Shuichi Fukuda *Editor*

Emotional Engineering

Service Development

Springer

Emotional Engineering

Shuichi Fukuda
Editor

Emotional Engineering

Service Development

Springer

Shuichi Fukuda, PhD
shufukuda@cdr.stanford.edu
shufukuda@gmail.com

ISBN 978-1-84996-422-7 e-ISBN 978-1-84996-423-4
DOI 10.1007/978-1-84996-423-4
Springer London Dordrecht Heidelberg New York

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

© Springer-Verlag London Limited 2011

Instron is a registered trademark of Instron, Coronation Road, High Wycombe, Bucks HP12 3SY, UK, www.instron.co.uk

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.
The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant laws and regulations and therefore free for general use.
The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Cover design: eStudioCalamar, Girona/Berlin

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Engineering is often thought to be very rational. But when we look way back into history, human beings made tools without the benefit of too much knowledge and experience. Instead, they were driven by their desire to explore new worlds by developing new tools. Indeed, *Homo faber* is the essence of engineering. We design and manufacture our products to develop new experiences. Engineering is more than just providing conveniences. If we remember that engineering started with emotion, it does not come as a surprise when we stop to consider the role of emotion in engineering.

The twenty-first century is an open world. In the twentieth century, technological advancements took place in relatively predictable directions, which meant we had enough knowledge and experience to make decisions. But today society is expanding rapidly, taking us beyond our traditional knowledge and experience. With such uncertainty, we need to look to our emotions to provide us with a sense of direction and a sense of balance, or what Aristotle described as "common sense". If emotion is interpreted in the sense of its original meaning, it is very much active, because it means to "move out", which comes from the Latin words *e* = *ex* = out, and *movere* (motion). It should also be noted that the word "motivation" has the same origin. In fact, emotion motivates our actions. It will take you higher and provide you with a better perspective. Emotion will play a very important role in rationalizing our actions or decision making in this age of increasing uncertainty.

Human beings are characterized in many ways, such as "*Homo faber*" (human makes a tool), "*Homo ludens*" (human plays), "*Homo loquens*" (human speaks), "*Homo mobens*" (human moves or travels), *etc.* They all describe how humans are motivated and they all lead to today's engineering. Our customers' expectations are quickly changing from better products to mental satisfactions. Therefore, more attention and effort should be focused upon human characteristics in our product design.

Traditional product development has been one way from the producer to the customer. Value was only thought about in terms of final product performance.

We forgot that our customers are not just passive consumers. They are very much active and creative. They invent many new ways of using our product, not only to meet their needs or preferences, but also to enjoy using them. They are creating value themselves by inventing a wide variety of experiences that the product designers never dreamed of. If we consider that they would also like to get involved in product development to exhibit their creativity more and to enjoy the experience, then our product development must be changed drastically to make this possible. Engineering can provide our customers with experiences not only as users but also as developers.

This book presents a wide variety of research topics on emotional engineering with contributions from leading experts in many different fields. Chapter 1 is an introduction to the topic. Chapters 2, 9, and 11 discuss emotional communication from different perspectives: Chapter 2 covers ambient information; Chapter 9, embodiment; and Chapter 11, human–computer interaction. Chapters 3, 4, 6, 9, and 13 study emotion in virtual environments: Chapter 3 looks at skill transfer; Chapter 4, skill and creativity, Chapter 6, emotional design; Chapter 9, embodiment; and Chapter 13, 3D shape evaluation. Chapter 4 discusses customer creativity and customer involvement in product development, which is not often discussed elsewhere. Chapters 5 and 16 are kansei engineering papers: Chapter 5 looks at driver emotion and head-up displays; and Chapter 16 discusses product design. The issue of driver emotion is also covered in Chapter 14. Chapter 16 deals with sound and describes kansei engineering approaches generally. Chapters 7 and 9 deal with emotion and robots. Chapters 4, 8, and 13 describe shape design. Chapter 8 discusses the relation of shape with $1/f$ fluctuations. Chapter 10 discusses affordance perception from an emotional standpoint. Chapter 12 describes how emotion can be utilized for rating TV programs. Chapters 15 and 19 both discuss footwear issues: Chapter 15 looks at the problem from a medical angle; whilst the development of sports shoes are covered in Chapter 19. Chapters 17 and 20 discuss emotion in engineering design teams: Chapter 17 describes how emotion can be captured and analyzed; and Chapter 20 looks at how better teams can be formed with due attention to emotion. Chapter 18 introduces how biological signals can be used to detect excitement. Finally, Chapter 21 discusses how emotional and creative motion can be generated.

I hope the reader will enjoy exploring the wide open world of emotional engineering. It will pave our way to wisdom engineering.

Finally, I would like to thank all the authors from the bottom of my heart and would also like to thank Mr. Anthony Doyle and Ms. Claire Protherough, both at Springer, UK, and also Mr. David Hemsley for editing the book.

Shuichi Fukuda

Core Ideas

1. Emotion is a compass. It provides us with a better, long-term perspective on the direction in which we should be moving. It is a prerequisite for strategic decision making under increasing uncertainties. But to make the best use of it we have to rationalize our emotions.
2. Emotion is motivation and expectation. It shows us where our customers are heading and what they would like to have. It leads to evolutionary and disruptive innovation and opens up new markets.
3. Emotion is reflective and can be a generator of new values. It creates values not only from our products but also from our processes. If we can get our customers involved in product development, they can create their own experience and enjoy telling their stories. In order to achieve this, our product development must be changed so that both producer and customer can enjoy working together.

Contents

1 Emotion: A Gateway to Wisdom Engineering 1
Shuichi Fukuda
1.1 Introduction 1
1.2 From Designer-centric to User-centric Design 2
1.3 Role of Emotion in Decision Making 2
1.4 The Creative Customer 3
1.5 Emotion = Move Out 4
1.6 Babies 5
1.7 Quality Function Deployment (QFD) 5
1.8 The Fluctuation Equation 5
1.9 The Emotion Equation 6
1.10 Rationalizing Our Actions Using Emotion 6
1.11 Closed and Open Worlds 7
1.12 Reflective Practice 8
1.13 *Homo Faber* 8
1.14 *Homo Ludens* 8
1.15 Pragmatism 9
1.16 Learning from Failure 10
1.17 Maslow's Hierarchy of Human Needs 11
1.18 Short-duration and Long-duration Emotions 13
1.19 Continuous Prototyping 14
1.20 Creating Markets 16
1.21 One Time Value and Lifetime Value 16
1.22 Concluding Remarks 18
References 19

2 SHOJI: A Communication Terminal for Sending and Receiving Ambient Information ... 21
Masaki Shuzo and Ichiro Yamada
2.1 Introduction ... 21
2.2 Ambient Information ... 22
2.2.1 Definition ... 22
2.2.2 Related Works ... 23
2.2.3 Our Approach ... 24
2.3 Target Usage ... 24
2.4 Technical Functional Requirements ... 25
2.5 Questionnaire ... 26
2.5.1 Overview ... 26
2.5.2 Results ... 27
2.6 Prototype Terminal ... 29
2.7 Emotion Extraction from Voice ... 31
2.7.1 Voice/Non-voice Recognition Process ... 31
2.7.2 Speaker Recognition Process ... 32
2.7.3 Emotion Recognition Process ... 33
2.8 Field Experiment ... 34
2.8.1 Method ... 34
2.8.2 Results ... 35
2.9 Conclusion ... 37
References ... 37

3 A Mixed Reality-based Emotional Interactions and Communications for Manufacturing Skills Training ... 39
Keiichi Watanuki
3.1 Introduction ... 40
3.2 Skills Transfer ... 41
3.2.1 Present Situation and Problem Points Regarding Methods of Skills Transfer ... 41
3.2.2 New Methods of Skills Transfer ... 43
3.2.3 Knowledge Creation and Skills Transfer by Sharing of "Ba" ... 44
3.2.4 Conversation Analysis and Extraction of Knowledge ... 46
3.2.5 Proficiency and Corporeality ... 47
3.3 Skills Transfer and Human Resource Development in the Foundry Industry ... 48
3.3.1 Knowledge and Skills Transfer of Casting Technology ... 48
3.3.2 New System for Skills Transfer ... 49
3.3.3 Skills Transfer System Using Multimedia Technology ... 49
3.3.4 Casting Skills Acquisition System Using VR Space ... 51
3.3.5 Skills Transfer and Human Resources Development Through Emotional Interaction in VR Space ... 53

3.3.6 Skills Transfer with an Annotation Display System 54
3.3.7 Work Training by OJT and Sharing a Ba in a Virtual Training System 55
3.4 Augmented Reality-based Training System 57
3.4.1 Augmented Reality 57
3.4.2 Training System for Pouring 58
3.5 Work Support for Elderly or Female Workers 59
3.5.1 Aging Society and Labor Problems 59
3.5.2 Work Support for the Work of the Elderly or Female Workers 60
3.5.3 Construction of a Teleworking Environment as a Work Form with Variety 60
3.6 Conclusions 61
References 61

4 Exploitation of Designers and Customers' Skills and Creativity in Product Design and Engineering 63
Monica Bordegoni
4.1 Introduction 63
4.2 One Product, many Customers 65
4.3 The Product Development Process 66
4.3.1 Designers and Engineers 68
4.3.2 Tools for Design and Engineering 70
4.3.3 Methods and Tools for Product Testing 71
4.3.4 Remarks About the Product Development Process 72
4.4 A Paradigm Shift in the Product Development Process 72
4.4.1 Product Customization and Participatory Design 73
4.4.2 PDP Scenario Based on User-centered Design 74
4.5 Design Tools for Product Designers and Customers 75
4.5.1 Enactive Interfaces 75
4.5.2 Design Tools Based on Enactive Interfaces 78
4.5.3 An Example: Enactive Interface for Shape Evaluation 79
4.5.4 Remarks 81
4.6 Product Prototyping and Testing 82
4.6.1 An Example: Mixed Reality Application for Product Evaluation 82
4.7 Conclusion 84
References 84

5 A Study on the Relationships Between Drivers' Emotion and HUD Image Designs 87
Shana Smith and Shih-Hang Fu
5.1 Introduction 87
5.2 Experiment Procedures 88
5.2.1 The First Experiment 89

5.2.2 The Second Experiment 96
5.3 Conclusions 100
References 100

6 Emotional Design in the Virtual Environment 103
Tetsuro Ogi
6.1 Introduction 103
6.2 Design in the Virtual Reality Environment 104
6.2.1 Interactive Visualization 104
6.2.2 Haptic Sensation 106
6.2.3 Application of VR to Design 107
6.3 Design Using an Augmented Reality Environment 108
6.3.1 Integration of Virtual Image with Real World 108
6.3.2 AR View 108
6.3.3 Application of AR to Design 110
6.4 Emotional Communication 111
6.4.1 Tele-immersion Environment 111
6.4.2 Video Avatar Technology 112
6.4.3 Video Avatar Communication 114
6.5 Education of Emotional Design 115
References 116

7 Emotional Robot Competence and Its Use in Robot Behavior Control 119
Natascha Esau and Lisa Kleinjohann
7.1 Introduction 119
7.2 Aspects of Emotional Competence in Human–Robot Interaction 121
7.3 Related Work 123
7.4 Modeling of Emotions 124
7.5 Emotion Recognition 126
7.5.1 Overview of VISBER 126
7.5.2 Overview of PROSBER 127
7.5.3 Fuzzy Emotion Classification 128
7.6 Dynamics of Emotions and Drives 129
7.6.1 Drives 131
7.6.2 Emotions 133
7.7 Control of Robot Behavior by Emotions and Drives 135
7.7.1 Behavior System 136
7.7.2 Control of the Behavior System by the Emotion Engine 137
7.8 Example of MEXI's Emotional Competence 138
7.9 Conclusions 141
References 141

8 Shape Design and Modeling with 1/*f* Fluctuation Characteristics 143
Katsuhiro Maekawa
8.1 Introduction 143
8.2 Shape Extraction from a Japanese Teacup 144
8.2.1 Experimental Method 144
8.2.2 Results and Discussion 148
8.3 Shape Design with 1/*f* Fluctuation Characteristics 150
8.3.1 Use of Measured Teacup Characteristics 150
8.3.2 Use of Mathematical Patterns 152
8.3.3 Use of Artificial Patterns 154
8.4 Modeling by Rapid Prototyping 156
8.5 Summary 158
References 159

9 Human-entrained Embodied Interaction and Communication Technology 161
Tomio Watanabe
9.1 Introduction 161
9.2 Embodied Virtual Communication System 163
9.3 Speech-driven Embodied Interaction System 165
9.4 Embodied Interaction and Communication Technology 170
9.5 Conclusions 176
References 176

10 Personal Cognitive Characteristics in Affordance Perception: Case Study in a Lobby 179
Y.S. Kim, J.Y. Jeong, M.K. Kim, S.W. Lee, and M. Kim
10.1 Introduction 180
10.2 Background 182
10.2.1 Affordance 182
10.2.2 Personal Creativity Modes 183
10.3 Case Study – Experiments 186
10.3.1 Participants and User Tasks 186
10.3.2 Activities and Related Affordances 188
10.3.3 Personal Creativity Modes 190
10.4 Case Study – Analysis on Personal Creativity Modes and Actives 193
10.4.1 Task: Waiting (T-1) 193
10.4.2 Task: Eating a Snack and Reading the Magazine (T-2) 195
10.4.3 Task: Sketching on Paper (T-3) 198
10.4.4 Task: Lacing Shoes and Shaking Sand (T-4) 202
10.4.5 Discussions 204
10.5 Conclusions 205
References 205

11 Emotion-oriented Human–Computer Interaction ... 207
Toshiyuki Yamashita, Ahmad Eibo, Takumi Ichimura, and Kazuya Mera
11.1 Introduction ... 207
11.2 Facial Expressions as a Human Interface ... 208
11.2.1 Face Selection by Fuzzy Reasoning Model ... 208
11.2.2 An Application of the Model to e-mail ... 210
11.2.3 Comparisons Among Faces ... 213
11.2.4 User's Faces ... 214
11.2.5 Noh Mask ... 216
11.3 Verbal Communication ... 219
11.4 Conclusions ... 220
References ... 221

12 BROAFERENCE: A Prototype of an Emotion-based TV Quality Rating System ... 223
Terumasa Aoki and Uwe Kowalik
12.1 Current TV Rating System and Its Problems ... 223
12.2 System Overview: the BROAFERENCE System ... 225
12.3 Face Detection, Face Feature Point Detection, and Tracking ... 226
12.3.1 Abstract of Face Detection, Face Feature Point Detection, and Tracking ... 226
12.3.2 Emotional Parameters ... 227
12.3.3 Interest Parameters ... 229
12.4 Global Head Motion Compensation ... 229
12.4.1 How to Compensate Global Head Motion ... 229
12.4.2 Evaluation ... 231
12.5 Action Unit Detection with Artificial Neural Networks ... 233
12.5.1 How to Detect Action Units ... 233
12.5.2 Classifier Training ... 235
12.5.3 Evaluation ... 236
12.6 Gaze Tracking Approach ... 237
12.6.1 How to Estimate Gaze Information ... 237
12.6.2 Mapping to 3D World Coordinates ... 237
12.6.3 Evaluation ... 240
12.7 Other Applications ... 241
12.7.1 Content Indexing Based on the Emotional Component ... 241
12.7.2 Remote Health Care Prevention and Life Support ... 242
12.7.3 Advanced Human–Computer Interface ... 243
12.8 Conclusion ... 243
References ... 244

13 Shape Evaluation Properties in Real Space and Virtual Space 247
Masashi Okubo
13.1 Introduction 248
13.2 System Evaluation 249
13.2.1 3D Models and Objects Used in the Experiment 249
13.2.2 Sensory Evaluation 251
13.2.3 Experimental Results 251
13.3 Influence of Viewpoint on Shape Evaluation in Virtual Space 254
13.3.1 Three Types of Viewpoint 254
13.3.2 Influence of Avatars' Forearms on 3D Shape Evaluation 255
13.3.3 Experimental Results 255
13.3.4 Influence of Avatar's Upper Body Image 258
13.4 Conclusions 260
References 261

14 Affective Driving 263
Yingzi Lin
14.1 Introduction 263
14.2 Natural-contact Biosensor Method 266
14.3 Inference of Human Emotion 266
14.4 Emotion Versus Performance in Driving 270
14.5 Issues and Future Research Directions 272
14.6 Conclusions 273
References 273

15 Approximate Multi-objective Optimization of Medical Foot Support 275
Masao Arakawa
15.1 Introduction 275
15.2 Approximate Multi-objective Optimization Using Convolute RBF 277
15.2.1 Convolute RBF 277
15.2.2 Satisficing Method 277
15.2.3 Data Addition 278
15.3 Arch Support 279
15.4 Model of an Arch Support 279
15.4.1 Spline Curve 279
15.4.2 How to Make a Foot Support Model 280
15.5 Design of Personal Medical Foot Support 281
15.5.1 Initial Pattern 282
15.5.2 Ideal Position of Arch 282
15.5.3 Detection from X-ray Digital Data 282
15.5.4 Procedure of the Proposed Method 283

15.6 Validation Test of Personal Medical Foot Support 284
15.6.1 Myoelectric Potential Measurement 285
15.6.2 Experiment 285
15.6.3 Results 286
15.7 Conclusions 286
References 287

16 Kansei Quality in Product Design 289
Hideyoshi Yanagisawa
16.1 What Is Kansei Quality? 289
16.2 Approaches Towards Kansei Quality in Product Design 290
16.3 Towards Diverse and Latent Kansei 291
16.4 A Method for Extraction of Diverse and Latent Evaluation Criteria of Kansei Quality 292
16.4.1 Extraction of Multiple Kansei Scales Considering Diverse Kansei Qualities 293
16.4.2 Strategies of Setting Design Feature Values 295
16.5 Case Study: Product Sound Quality 296
16.5.1 Sound Quality Metric as Design Parameters 297
16.5.2 Sensory Test Using Existing Samples (First Experiment) 298
16.5.3 Extracting Patterns of Personal Kansei Scales 299
16.5.4 Multiple Scales of a Target Kansei Quality 301
16.5.5 Unexplored Design Area in Feature Space 302
16.5.6 Gradient of a Target Kansei Quality in Feature Space 302
16.5.7 Setting Feature Values and Creating New Sounds 304
16.5.8 Sensory Test Using Composite Samples (Second Sensory Test) 304
16.5.9 Comparison of SD Scales Obtained from First and Second Experiment Data 305
16.5.10 Finding Kansei Factors 306
References 309

17 Emotion in Engineering Design Teams 311
Neeraj Sonalkar, Malte Jung, Ade Mabogunje
17.1 Introduction 311
17.2 Why Study Emotion in Engineering Design? 312
17.3 Perspectives in Emotion Research 313
17.3.1 Assigning the Label “Emotion” 315
17.3.2 How Does an Emotion Come About? 316
17.3.3 Distinguishing Between Different Emotions 317
17.3.4 Emotion in the Context of Current Studies 317
17.3.5 Measuring Emotion 318
17.4 Research Study 1 – The Role of Emotion in Group Ideation in Engineering Design Teams 319

17.5 Research Study 2 – Team Conflict in Engineering Design Teams ... 322
17.6 The Cultural Context of Emotion ... 324
17.7 Looking Forward ... 324
17.8 Conclusion ... 325
References ... 325

18 Measurement of Wakuwaku Feeling of Interactive Systems Using Biological Signals ... 327
Michiko Ohkura, Masahide Hamano, Hiroyuki Watanabe, and Tetsuro Aoto
18.1 Introduction ... 327
18.2 Indexes of the Wakuwaku Model ... 329
18.3 Evaluation Experiment ... 329
18.3.1 Construction of Evaluation Systems ... 329
18.3.2 Experiment to Evaluate the Systems ... 332
18.3.3 Experimental Results ... 333
18.3.4 Derivation of Wakuwaku Model ... 335
18.3.5 Confirmation of Wakuwaku Model ... 336
18.4 Experiment 2 ... 338
18.4.1 System Construction ... 338
18.4.2 Experiment to Evaluate the Systems ... 340
18.4.3 Discussion ... 341
18.5 Conclusion ... 342
References ... 343

19 Footwear Fitting Design ... 345
Tsuyoshi Nishiwaki
19.1 Introduction ... 345
19.1.1 Running Shoe Construction ... 346
19.1.2 Property Requirements ... 347
19.1.3 Running Shoe Design Flow ... 348
19.2 Fitting ... 349
19.3 Optical Measuring Instrument ... 351
19.3.1 Current Technique ... 351
19.3.2 New Technique ... 351
19.4 Evaluation Method of Shoes Fitting ... 354
19.4.1 Definition of Measurement Areas ... 354
19.4.2 Measurement ... 356
19.4.3 Measurement Results ... 356
19.4.4 Fitting Quantitative Evaluation Method ... 357
19.4.5 Validity Checking ... 358
19.5 Design Examples ... 360
19.6 Summary ... 361
References ... 362

20 Psychological Teamology, Emotional Engineering, and the Myers–Briggs Type Indicator (MBTI) 365
Douglass J. Wilde
20.1 Introduction 365
20.2 Underlying Psychological Theory 366
20.3 Cognitive Modes 367
20.3.1 Relative Scores and Coupled Attitudes 369
20.3.2 Graphical Interpretation 370
20.4 Team Construction 371
20.5 Team Organization 372
20.6 Conclusions 374
References 374

21 Designing Emotional and Creative Motion 377
Toshiharu Taura and Yukari Nagai
21.1 Introduction 377
21.1.1 Emotion and Design 377
21.1.2 Designing the Emotional and Creative Motion 378
21.2 Method of Designing Emotional and Creative Motion 379
21.2.1 Analogy to Natural Objects 379
21.2.2 Emphatic Blending of Motions 380
21.2.3 Outline of Proposed Method 380
21.3 Detailed Method and Example 381
21.3.1 Designing Emotional and Creative Motion by Focusing on Rhythmic Features 381
21.3.2 Example of Designing an Emotional and Creative Motion Using the Proposed Method 383
21.4 Summary and Conclusions 386
References 387

About the Editor 389

Index 391

Chapter 1
Emotion: A Gateway to Wisdom Engineering

Shuichi Fukuda[1]

Abstract Our society is changing rapidly from a closed world with clearly defined boundaries, to an open world with no boundaries and with frequent and extensive changes. At the same time, design is changing from designer-centric to user-centric, and humans and machines must work together as a team. Therefore, satisfying individual human needs rather than mass needs is becoming increasingly important. Such changes call for cyclic and reflective engineering. Emotion will play an important role in providing us with motives for interaction and with guidelines for rationalizing our actions. In the future, wisdom will be crucial in engineering, and emotion will play a very important role in realizing this wisdom-based engineering.

1.1 Introduction

The topic of emotion is now attracting considerable attention in many areas. However, it has rarely been discussed from an engineering perspective.

Our society is changing from a closed world with clearly defined boundaries, to an open world with no boundaries and with frequent and extensive changes. At the same time, design is changing from designer-centric to user-centric, and humans and machines must work together as a team to respond to a changing situation, adaptively and flexibly.

Therefore, individual human needs are becoming more important than mass needs. Such changes call for cyclic and reflective engineering. In the future, wisdom will be crucial in engineering, and emotion will play an important role in providing us with motives for interaction and with guidelines for rationalizing our actions.

---

[1] S. Fukuda
Stanford University, Department of Mechanical Engineering, Stanford, California, 94305, USA
e-mail: shufukuda@cdr.stanford.edu

1.2 From Designer-centric to User-centric Design

In the twentieth century, situations did not change appreciably – the world we lived in was closed. Twentieth-century engineering was context-independent and static. It was rationality-based. Designers could foresee the operating conditions of machines and systems, which were designed to work regardless of how situations changed. Machines were just tools which worked as users instructed. Users were supposed to follow the designers' instructions. In fact, most so-called human failures took place because users did not behave as designers expected.

As we enter the twenty-first century, however, situations are changing extensively and often. It is becoming increasingly difficult to foresee the conditions under which machinery and systems will be expected to operate. It is now the users who know what is happening, and no one else, but only they can respond to the situation. Therefore, design has changed from being designer-centric to user-centric. The role of a machine is now not only to respond flexibly and adaptively to the diverse requirements or commands of users, but also to sense and provide necessary information for them to correctly understand the situation in order to make appropriate decisions. Thus, sensing becomes very important for a machine, and users and machines must work together as a team to respond to changing situations. Human–machine relations were static in the twentieth century, but are set to become much more dynamic and interactive in the twenty-first century.

1.3 Role of Emotion in Decision Making

How users judge the situation and make decisions are very important considerations for user-centric design. Emotion seriously affects human judgment and action in emergencies such as fire, earthquake, *etc.*, and these areas have received a lot of research. However, emotion also plays a very important role in our daily decision-making and action.

For example, in economics, Keynes pointed out that economic agents act rationally in short-term expectations, because there is ample information to allow them to predict the near future with reasonable certainty. But in long-term expectations, emotion plays a very important role in their decision-making and action, because there is great deal of uncertainty in the far future [1, 2]. The term "expectation" in economics is nothing more than a motive for an economic agent to take action. Today, uncertainty is increasing rapidly. The situation that we now find ourselves in corresponds to the situation of Keynes' long-term expectations.

Later in 1947, Simon, the famous novel laureate in economics, pointed out that economic agents' actions are not completely rational, and their rationality is bounded, and that the existing mathematical economics could not describe true economic states [3]. He pointed out that the more complex and diverse an economy, the greater the number of factors that have to be considered, which decreases computability. To compensate for this, emotion then plays an important role in decision-making and action. When he proposed this idea, society was in the chaotic confusion that followed World War II.

Keynes and Simon only discussed economics. But the same argument holds in engineering, because our world is becoming more and more open with boundaries disappearing, and there are frequent and extensive changes. We have not discussed emotion because we deal with artifacts and up to now have considered humans only as consumers.

1.4 The Creative Customer

Emotion is often interpreted in the passive sense. For example, most kansei engineering researchers [4], tend to give a passive impression – they discuss what emotion a product will stir up in customers. This is similar to the concept of Affordance [5], which maintains that our external world provides both stimulus and motives for us to act.

Most of the current discussions about emotion in engineering are focused upon how we can incorporate emotion into product design. Functional improvements do not appeal to customers as much as it did in the past, so emotion is attracting attention as a new way of adding value. But unfortunately this way of thinking is still producer-centric. It does not acknowledge that design is changing rapidly.

The word "consumer" demonstrates conclusively how designer-centric this approach is. It regards users only as passive people, who just "consume" what they provide. Even the word "users" is sometimes used in the same way as the phrase "end users", which indicates that designers are here on this side and users are far away on the other side.

We will use the word "users" here, but our users are very active and creative. We will use the word in the same sense as "customers", who "customize" our products to their tastes and needs. That is why they are called "customers". In fact, there are many examples which demonstrate why our users are "customers", and they are very active and creative.

For example, kids use a slide as instructed for the first few times. But soon they invent new ways of sliding down (see Figure 1.1).

Figure 1.1 Creative kids

Youngsters rip holes in their jeans and wear them. They know jeans are not simply something to wear but also something that has a story. Stories are important to develop feelings of attachment (see Figure 1.2).

Oh, no.

You got holes there!

It's the fashion,

Grandma

Figure 1.2 Creative youngsters

Although creativity has recently drawn wide attention in engineering, most research looks only at how designers can be creative. We fail to understand that our users would also like to be creative and do not consider how we can satisfy that desire.

It must be added that in addition to satisfying such basic desires of our customers, our customers have to customize machines in order to work together better as a team.

1.5 Emotion = Move Out

Emotion is very active, and far from passive. The word "emotion" originates from the Latin *e* = out and *movere* = motion, *i.e.*, "move out". The word "motivation" also comes from the Latin *movere*. Therefore, emotion is very closely associated with action. Affordance is deeply associated with emotion because it maintains that our outside world supplies stimuli and provokes motives for us to take actions. Emotions have often been interpreted as passive and static, but it should be remembered that they are in fact very much dynamic.

1.6 Babies

Recently, in their quest to make robots more intelligent, robotics researchers have been studying the behavior of babies in order to understand their cognitive development. Developmental psychologists observe that babies move their bodies in random or haphazard ways for the first 4 months or so. Then they do not move too much for a while. After this period, they move their arms, legs, *etc.*, in a coordinated manner.

Piaget's cognitive development theory [6, 7] may provide an interpretation. To babies, the world around them is a wide open world without boundaries. They move their bodies and interact with the outside world in various ways in order to build up a model that will enable them to understand their environment. Once a model is formed, then they move based upon it.

Developmental psychology does not make clear what triggers babies to act or move. It may be because the outside world supplies stimuli and motives to them, as the theory of Affordance maintains. Then their movements are triggered by emotions. And just as with babies, emotion is considered to play a crucial role in the way that adults interact with the outside world.

1.7 Quality Function Deployment (QFD)

Quality function deployment (QFD) plays an important role in designing artifacts, by providing a basic tool for realizing customers' requirements as product functions. Functions are motions, so QFD describes the required functions using verbs. Attributes, such as surface conditions, may not seem to be associated with verbs. But they are – they are deeply associated with touching, or haptics. Therefore, attributes are also describable by verbs.

QFD is a tool used to clarify requirements rather than a tool to analyze motions, *i.e.*, it is not a tool for problem-solving, but a tool for setting a goal or for problem formulation. Verbs are accompanied by objectives. Therefore, if motions are described using verbs, the objectives of actions will become clear.

Baby development may be compared to QFD development. The period when babies move randomly and then stay motionless may be interpreted as a time when they are preparing QFD. Babies are relating their motions to their objectives and, when their modeling is complete, they start to move in a coordinated manner to solve the problems.

1.8 The Fluctuation Equation

Moving on from the discussion of whether these seemingly random behaviors of newborns are truly random or chaotic, let us consider the problem of fluctuations.

Fluctuations are discussed by many theories, such as random theory, chaos, complex systems, *etc.*, but here we will just mention that fluctuations may play a very important role when considering emotions.

A research group at Osaka University in Japan is currently carrying out the *Yuragi* (fluctuations) project [8]. Living bodies are so complex that deterministic control is impossible. The group at Osaka maintain that living bodies deal with their complexities by utilizing fluctuations. Yanagida *et al.* made it clear that myosins move along actin filaments utilizing Brownian motion [9], and proposed the following fluctuation equation:

$$\mathrm{d}x/\mathrm{d}t = (\mathrm{d}u(x)/\mathrm{d}t) \times \text{activity} + \eta,$$

where $u(x)$ = object function and η = noise.

When the environment is very complex and changes dynamically, $u(x)$ might be trapped in a local minimum and may not reach the global minimum, if the traditional control approach is applied. Ishiguro *et al.* demonstrated by using robots that the global optimum (minimum) solution can be obtained by changing "activity" in their fluctuation equation [10].

1.9 The Emotion Equation

Ishiguro's approach is very similar to the simulated annealing approach in the global optimization theory. In fact, the idea is the same, in that noise is used to prevent the solution from being trapped in a local minimum, thereby allowing the global minimum to be reached. There are similar equations in other areas, too.

Let us consider risk engineering. There are many definitions of risk. One of them is:

Risk = Probability of the occurrence of an event × Loss by that event.

As probability is a function $f(x)$, so the loss by the event corresponds to activity in the fluctuation equation.

Utility, which is important in economics, is defined as follows:

$$\text{Utility} = \text{Prob}(x) \times \text{Option},$$

where Prob(x) = probability of the occurrence of the event and Option = Utility of the event selected.

If we note that utility is a measure of satisfaction, and risk is its opposite, we can possibly define the emotion equation in a similar way to the fluctuation equation in the previous section.

1.10 Rationalizing Our Actions Using Emotion

If we regard emotions in an active sense as motives for action, then just as babies learn how to move, we can rationalize our action using emotion.

This is not an attempt to mechanize humans or simplify human behavior. It is important for us to work well with machines as a team, and to fully satisfy our expectations by interacting with the outside world in a more rewarding way.

Although economics emphasize utility, it is engineering which creates utility. Therefore, the maximization of utility is a basic issue in engineering. Short- and long-term expectations in Keynes' discussion correspond to the issues of local and global optimization in engineering. Keynes' statement that emotion plays an important role in long-term expectations implies that emotion is deeply associated with the problem of global optimization or the problem of an open world.

1.11 Closed and Open Worlds

In order to consider emotion from the standpoint of global optimization, let us look at the changes taking place in our society.

The twentieth century was the age of the closed world, where boundaries could be clearly defined. It could be compared to a railroad or to agriculture. Railroads have tracks so we know where we are going and don't have to worry about possible destinations. We can decide on our destination before we get on the train – what is important is whether we can get there faster or cheaper. In the case of agriculture, farmers accumulate experience by cultivating their farmlands, so they can reap a better harvest as their experience increases. Both of these worlds are closed. Therefore, it is straightforward to develop a model, leaving us to concentrate on problem solving. In other words, tactics are more important than strategy.

The twenty-first century, on the other hand, is an open world with no boundaries. It can be compared to a voyage or to hunting. A voyage is affected by many unpredictable factors, such as weather, sea conditions, *etc.* When a storm arises, we may have to anchor at a port we never anticipated. In the worst case, we may not be able to reach our planned destination. In such a case, we have to ask ourselves why we set sail. In hunting, we can never totally predict what animal we will come across. Judgment is more important than knowledge and experience. If it is stronger than you, you risk being eaten! It is crucial to judge whether it is stronger than you or not. The strongest survives.

Thus, in an open world, it is extremely difficult to build up a universal model, so we have to keep changing our models to adapt to the current situation. Therefore, problem formulation becomes very important and strategy is more important than tactics.

Closed and open worlds may be explained in connection with multimodal optimization problems. If the world is very small, there is only one peak, and the optimization is very easy. But when the world expands, there will be many peaks, which forces us to decide which peak we should climb. Thus, global optimization becomes necessary.

As discussed in the previous section, we can use emotions to formulate a problem appropriately and develop a workable model in order to rationalize our actions. This leads to better decision making in engineering, management, *etc.*

1.12 Reflective Practice

Schon pointed out that in management, design, and medicine, rationality does not hold and the professionals in these areas think as they act. He called them "reflective practitioners" [11]. Reflective practice may be compared to the law of action–reaction (Newton's third law) in physics. In the physical world, if you push something, it will push back. Similarly, in the world of stimulus–response, if we take some action in the outside world, we get a reaction – and this excites our emotions.

Although kansei engineering may give a passive and static impression, as mentioned earlier, because it does not explicitly consider such a stimulus–response (reaction) cycle (Norman called such an approach "visceral design" [12]), it is not passive or static. For example, Stark discovered that our visual perception is realized by the very active movements of our eyes [13]. Therefore, both kansei engineering and visceral design are also very dynamic. The only difference is that they do not explicitly discuss the stimulus–response (reaction) cycle.

1.13 *Homo Faber*

Engineering is fundamentally reflective. The word "*Homo faber*" demonstrates it best.

Why do human beings make tools? If it is just to use the tool, even animals can do that. But humans also take the trouble to develop the tool. As discussed above, the world used to be very open. Situations changed frequently and widely. So we could not make our tools in a straightforward manner. We used trial and error to adapt tools to our own situation and to our own needs. We were not just solving a problem, we were formulating the problem and developing a model and applying it to solve the problem. If it did not work, then we came up with another model until the solution was found. Reflective practice is the intrinsic quality of human beings.

1.14 *Homo Ludens*

Huizinga defined human as "*Homo ludens*" or "human plays" [14]. Why is play an intrinsic quality of human beings? It is not play if we do something as instructed. Play is a behavior that seemingly has no particular objectives. Play is not problem solving, but it is problem formulation. That is why play characterizes human beings.

Kids are called "genius of play". They invent many new ways of playing. Why are they called creative? Because they are not solving a problem, but instead are formulating a problem. They are not interested in how they can play better, but they are interested in formulating a problem and inventing new ways of playing.

Play is important to human beings, because they are fundamentally oriented toward the open world. For humans, problem formulation is more important than problem solving. That is why human beings have prospered on Earth, as we see today.

1.15 Pragmatism

As our society is quickly moving toward an open world, let us consider Pragmatism, which is a typical philosophy applied to an open world. Pragmatism originated in the UK and was developed in the USA. Now it is the philosophy representing the USA. The UK was a seafaring country and the USA explored the West. Both countries have histories linked to the open world.

Peirce, one of the initiators of Pragmatism, proposed Abduction [15]. Abduction literally means, to find some workable hypothesis somewhere and apply it to the current problem. If it works, then you solve the problem. If not, find another hypothesis which might work. Repeat this process until your problem is solved.

In a closed world, the boundaries are clearly defined so you can accumulate your experiences and structure them into knowledge (induction). Once a structured knowledge is formed, then you can apply it to solve each particular problem (deduction). That is, in a closed world, you can make a prediction. But when the world becomes open, the concept of a set, which works in a closed world, will not be valid any more. Namely, induction and deduction logics do not work and we have to keep modifying our model to adapt it to the changing situations.

Schon called such activities “reflective practice”, as mentioned earlier. The 1980s, when Schon proposed this idea, was a time when great social changes, such as the collapse of the Soviet Union, the fall of the Berlin Wall, *etc.*, took place one after another. Engineering also changed considerably during this period. Computers changed from central to distributed systems, which accelerated the spread of the Internet. Rapid prototyping emerged. In all fields, people had to think in action.

According to the Spanish diplomat and journalist, de Madriaga [16], “Englishmen think while they walk. Frenchmen run after they think. Spaniards think after they run.” This remark becomes much more interesting when we note that the UK used to be a seafaring country and that France used to be a country of agriculture.

Schon’s reflective practice corresponds to the famous Shewhart Cycle of quality engineering. Shewhart was also an earnest advocate of Pragmatism. In fact, the Shewhart Cycle (Figure 1.3) is nothing other than Abduction if we interpret the diagram as Plan = formulate a model, Do = apply it, Study = study it to see if it works, Act = solve the problem, if it works. Otherwise look for another model.

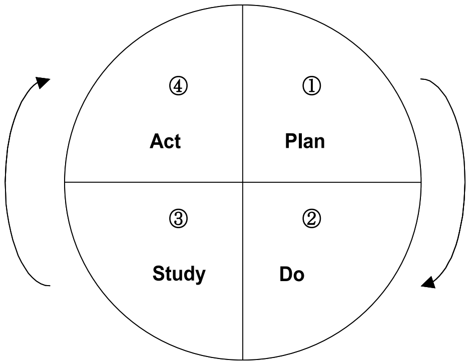


Figure 1.3 The Shewhart Cycle

The concept of the reflective cycle is common to Peirce, Shewhart, and Schon. Their way of thinking is not linear, as in induction or deduction, but is cyclic and reflective. If we quote from Shakespeare, their way of thinking is "All's well that ends well". This indicates that in an open world strategy is more important than tactics, or to put it in another way, problem formulation is more important than problem solving.

1.16 Learning from Failure

The reflective cycle is just the method of trial and error. Learning from failure is emphasized in the US education system. However, this does not mean "failure" in the usual sense, but means failure in the sense of "fail one's expectations", *i.e.*, learning from failure is learning how we respond to a situation when the result fails our expectations. In other words, learning how we should change our model dynamically to adapt to a changing situation to solve a problem, instead of developing a fixed model and applying it. This is very similar to what we observe in the cognitive development of newborns. It reflects the US history of exploration.

Braden, a famous tennis coach in the US [17], does not teach how to recover from misses, he just teaches the basic strokes and praises learners for their good shots. He explains why: muscles vary from person to person. Therefore, we must learn by ourselves how we recover from our misses. This indicates how important problem formulation is and how important it is to be able to change a model adaptively.

Recent progress of NIRS technology has made it possible to observe the changes of blood and internal secretions so that the relationship between our muscles and our emotions are being explained. Muscles, and therefore emotions, are deeply associated with how we recover from failure.

1.17 Maslow's Hierarchy of Human Needs

Expectations are deeply associated with human needs and they are deeply tied up with emotions. The twentieth century was the century of materials, but the twenty-first century is the century of the heart. In the century of materials, the primary engineering activity was to deliver products, so design was static and the flow was one way from producer to customer. But in the century of the heart, engineering is becoming interactive and reflective, so design needs to be dynamic. Therefore, we have to take the characteristics of human beings into consideration.

Maslow classified human needs into five levels [18] (Figure 1.4). The first, or bottom, level is physiological needs – we cannot survive unless we eat. The second is the need for safety and security – once our appetite is satisfied, we want to protect ourselves from rains, animals, *etc.* Recently, social uneasiness due to terrorism, *etc.*, and uneasiness caused by the increasing complexities of machines or systems, has been drawing increasing attention. These are also problems of emotion.

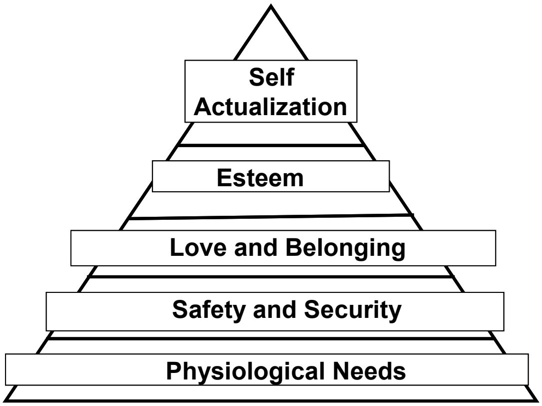


Figure 1.4 Maslow's hierarchy of human needs

Safety is related strongly to reliability. Traditionally, reliability was rational and discussions were made based upon numbers. Humans were also modeled as rational beings. Emotion was ignored, except in special applications. But recently, many cases are emerging which cannot be solved by such rational approaches. For example, in the world of the Internet, it is very important that the information you receive is truly reliable or trustworthy. This problem is very different from the problem of network reliability. Network reliability is a technical problem and can be solved rationally. But whether you can trust a piece of information on the Internet is the same problem as whether you can trust a person you have never met before. These problems cannot be solved rationally.

This problem is associated with the problem of understanding intent in communication. It should also be added that if we could transmit only intent without transmitting an enormous amount of data, we could reduce the amount of energy required for information transmission considerably. In understanding intent, emotion plays an important role. Why do we have to think EQ or emotional quotient [19]? Trust is associated with the problem of why humans can transmit a lot of information using only a small amount of energy. This is also another interesting problem of emotion.

The third level is the need for community – humans cannot live alone. In fact, cities historically developed around a “common”. The need for community was behind the development of Web 2.0. Common, community, and communication originate from the same word. Traditional product-centered engineering considered only the existing geographical markets. But in the twenty-first century when the world becomes open, market creation will become a big issue in engineering. Creation of a community where people can share values will lead to the creation of a market. Emotion deeply affects how people evaluate value.

The fourth level is the need for respect – people like to be respected. This is of course emotional. This need is deeply associated with brands. Brand has been a big issue in fields such as cosmetics, but it has not been discussed much in engineering. Any discussions were about product brands. Corporate brands were rarely discussed.

In the twentieth century, 20% of customers bought new products as soon as they were introduced. The other 80% observed how these products worked and how useful they were, and then made their decision to buy. But as we entered into the twenty-first century, many new never-before-experienced products emerged. And the life cycle of these products is getting shorter and shorter, so there is no time for the 80% to observe how the new products are used by the other 20%, who have led the market until now. Therefore, people have to know how much trust they can put in the company which produces such a product. Thus, the issue of corporate trust or brand is increasing rapidly in importance.

Corporate trust or brand is nothing other than how well a company can respond to customers’ expectations. Trust in a company and trust in Internet information are the same, and they are both deeply associated with emotion. Today, trust or brand building is becoming increasingly important. A product’s brand is inductive and it is developed through the experience of using each particular product. But the twenty-first century is an age of hunting. Customers do not know what animals (*i.e.*, products) they will come across. In such situations, we have to build trust in the system. This is the same idea of quality as the ISO 9000 series, which visualized a quality system instead of assuring the quality of each product.

It must also be added that the issue of corporate brand is associated with confidence or self-trust – asking ourselves the question, “I have bagged good games up to now so why not this time?”. Confidence provides a motive for going hunting. This is truly an outward motion (ex-motion), *i.e.*, emotion. Providing customers with such motives is very important for companies of tomorrow.

The fifth level, or the highest level of human need, is self-actualization. All humans would like to actualize themselves. Why do artists never give up their art, even though most of them are not well off while they are living? Because art is the world of self-actualization. Self-actualization may be expressed as a challenge. Challenge is the core and the mainspring of all human activities: "Why do you climb?". "Because there is a mountain.". Animals see the same mountains, but aren't driven to climb them. A human being climbs a mountain because it is a challenge. Games become popular because they provide a challenge. And challenge or self-actualization is of course a problem of emotion.

Historically, industries climbed up Maslow's pyramid of human needs. They developed from primary, to secondary, and then to tertiary. If we look at Maslow's pyramid from another angle, we will notice that as we climb up, it changes from mass to individual, and from material to heart (Figure 1.5). Therefore, not only is our real world changing rapidly from closed to open, but we can see that human beings by their very nature are driven to head for an open world.

Another definition of human, "*Homo mobens*" or "human moves", substantiates this argument. From the standpoint of the optimization of energy, it is best for us not to move except when eating, just as lions do. But human beings always challenge the unknown. That is our intrinsic nature and that is why human beings have spread all over the globe. But our physical world has boundaries, and when we cannot explore any more, we head toward a non-physical or mental world, where there are no boundaries for self-actualizations or challenges.

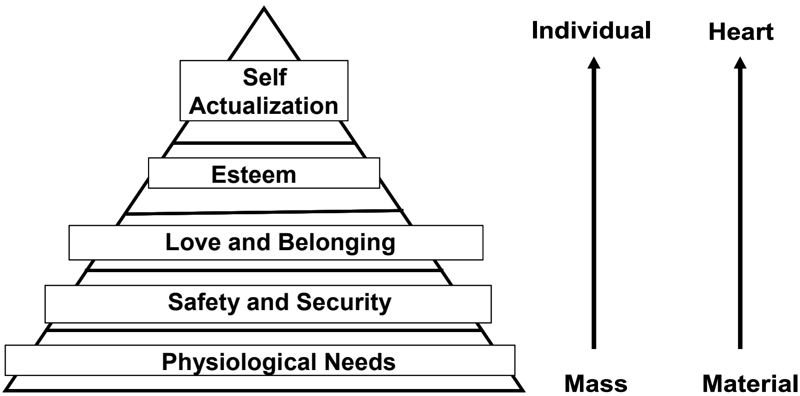


Figure 1.5 Changes as we go up

1.18 Short-duration and Long-duration Emotions

Just as there are short-term and long-term expectations, there are short-duration and long-duration emotions, and we must consider their dynamics separately.

Short-duration emotions are surprise, anger, *etc.*, and are sometimes called non-cognitive or instinctual emotions. Such short-duration emotions trigger an interrupt to the human brain and reduce its capability for information processing [20]. Although its overall capability may be reduced, we may interpret this as showing that our brain concentrates its capabilities on executing the necessary processes faster and more effectively. Such emotional problems as evacuating a building in a fire are short term or instinctual, and many researchers have studied them.

There is, however, another type of emotion, which has a long duration. This type of emotion is sometimes called cognitive emotion. As this emotion is associated with cognition, it varies widely from person to person, making it a greater challenge to analyze than a short-duration one. However, for the purposes of design, this long-duration emotion should receive more study.

For example, most research into human–machine interfaces has been or is being carried out on how a machine can be operated more easily from a designer-centric perspective. But in user-centric design, we should design a machine as a partner and take into consideration how a user will understand the situation and how he/she will respond to it.

Users respond to changing situations in many different ways. Therefore, machines have to understand the intent of a user and to respond and support adaptively and flexibly. Furthermore, some users might prefer to challenge, rather than to solve the problem in a safer and more secure way. This is very much associated with how a user recognizes the situation and with his/her personality.

1.19 Continuous Prototyping

Hardware and software development styles differ very markedly. Hardware is developed with fixed functions (Figure 1.6).

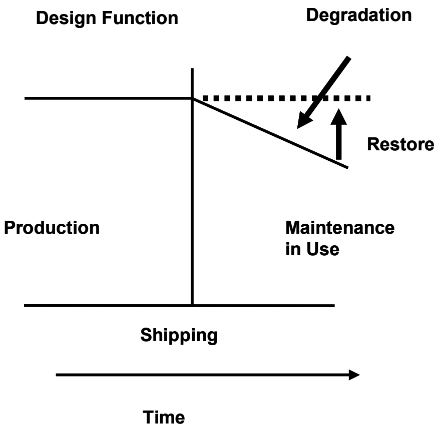


Figure 1.6 Hardware development

Software, on the other hand, is developed with growing or evolving functions (Figure 1.7). Artificial Intelligence (AI) introduced the concept of continuous prototyping into software development. It first provides very basic functions and then lets the functions grow with time and the accumulation of user's experience, and in response to their needs.

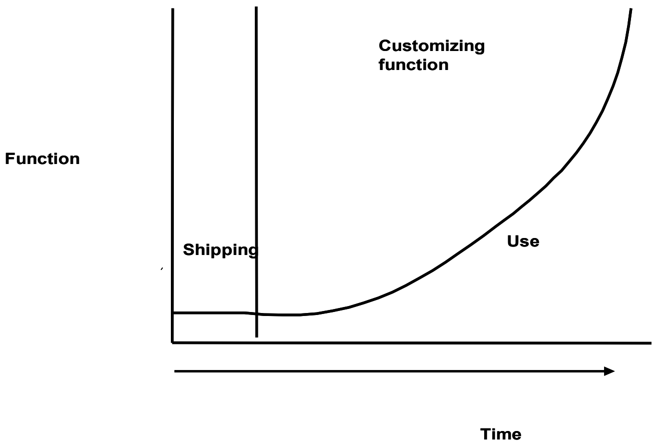


Figure 1.7 Software development

Until then, software programs had become much too complex and complicated, so that users needed a great deal of time to get used to them – and to make matters worse they could not grasp the whole picture and lost confidence. But if simple functions are provided first, users do not need to spend too much time learning the basics of the program, and can gain confidence through using the system as they wish. Thus, step by step, as functions grow, users' confidence and their trust in the system grow. In fact, the software development curve looks very similar to the learning curve. The more we learn, the more confident we become.

Hardware is physical so that it deteriorates once shipped. But if we turn the deterioration curve upside down, it also looks similar to the learning curve. Current maintenance aims primarily to restore the degraded functions back to their design level. But customers would like to be creative and would like to customize our products. This is self-actualization. Customers prefer to use products that have been tailored to their own needs and preferences. In other words, they prefer repair to maintenance. The word "repair" comes from "pare", which is associated with "prepare". That is, repair means to remodel or remake products so that they fit to the current environment and situation.

If we can get our customers involved in repair, their sense of involvement increases and their feelings of attachment will grow. We should move toward such service development beyond mere product delivery. Engineering tomorrow will be more consultation than production. The sale of hardware results in a win-lose situation; but service is non-physical so it is far easier to achieve a win-win situation. This also serves to establish a lifelong relationship with our customers.

1.20 Creating Markets

A marketing strategy such as BoP (Bottom of the Pyramid) can be regarded as continuous prototyping for hardware development. BoP refers to the poorest class or the poorest countries. In the market economics that have held sway up to now, such market segments have been ignored, and helping them has been considered charity rather than business. C.K. Prahalad at the University of Michigan first pointed out the possibility of BoP as a future market [21]. He is also famous as a proponent of value co-creation, which is described later. Certainly there are strong needs at BoP, so they could be markets. But the problem is whether they are capable of buying. Without any purchasing power, they cannot be markets. Therefore, new moves are emerging to give them purchasing power and thus make them new markets.

As indicated by the law of diminishing returns, the greater the number of products, the lower the financial returns or profits, when the market is closed or bounded. If we consider only existing markets, our market competitiveness quickly decreases. This can be explained from another angle by using Weber–Fechner's law from psychology. This law says that the larger the stimulus becomes, the larger the increment we need to perceive a change. For example, if someone with a quiet voice raises his/her voice a little, everybody would hear the difference, but if someone with a loud voice raises his/her voice a little, the difference would be difficult to hear.

Customers perceive an improvement when the initial product quality is low, but as quality becomes more and more improved, they find it more and more difficult to perceive the change. Therefore, in advanced countries, it becomes increasingly difficult to secure returns or profits as more products are put into the market, and we cannot increase the number of customers even if we make considerable efforts to improve the quality.

Products with lower functions can satisfy the needs of developing countries. These products increase their buying power so that eventually they can afford products with higher functions. This is true of the history of every hardware product. Such history-conscious product development must be explored. What is most important in a BoP market strategy is that it is not just about product delivery, but about seeing that products are developed to meet the emotional expectations of developing countries.

1.21 One Time Value and Lifetime Value

Most of the current discussions about industrial products consider value at a fixed point in time and pay attention to market share. But even though market share may be low, there is another strategy to secure customers and to have profits in the long run. BoP is one such example. If a lifetime value can be secured, industries can last for a long time and the investment risk can be reduced. Customers who share the same sense of value will not move to another company easily. Maslow's hierarchy of human needs indicates that industries will gradually move from mass to individual. Therefore, it is of the utmost importance for the industries of the twenty-first century to consider how they can cater for the needs of the individual.

Value is defined by traditional value engineering [22] as

$$\text{Value} = \text{Performance (Function)/Cost}$$

Value here is the profit to the producer. Traditional product development has been product-oriented so that performance here denotes the functions of the final product. Therefore, industries attempt to increase performance to increase value. But as the Weber–Fechner law teaches us, it becomes more and more difficult to get our customers to perceive quality improvements as the quality increases. That is why industries head toward cost reduction to increase value, because cost reduction is more visible and easier to estimate.

Traditionally, producers had much more information than customers (Figure 1.8). This asymmetry of information brought profits to producers, because if they produced a product that would level off the difference, customers were happy. This is the framework behind traditional value engineering.

But we are now an information society, so this asymmetry is quickly disappearing. In fact, it is not unusual for customers to have more information than producers. To cope with such situations, Prahalad and Ramaswamy proposed "value co-creation" [23]. In this, producers and customers work together to raise the level of information (Figure 1.9). The raising of the water level means profits to the producers and value or satisfaction to the customers. In traditional producer-centric product development, the customer's value or satisfaction was not considered explicitly. The chief attention was on how much profit the producers can yield. There was no value other than the producer's profit. But in value co-creation, the customer's value is now explicitly considered, and they can get involved in the process of realizing products that would be valuable to them.

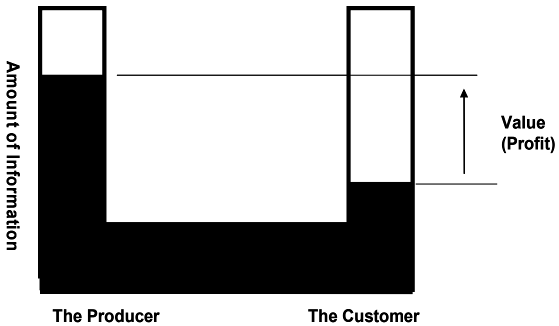


Figure 1.8 Asymmetry of information

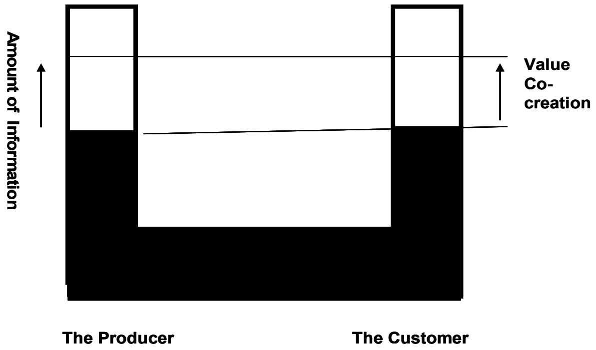


Figure 1.9 Value co-creation

Toffler proposed the prosumer (producer + consumer) system [24]. But in his book, the customer's desire for creativity is not considered, as can be understood from the fact that he used the word "consumer". He did not discuss explicitly the customers' desire to create their own values. Value co-creation is one step forward from this in that it points out this desire.

These discussions, however, are focused upon product value. But we must note that products alone do not contribute to value enhancement. Processes also yield values. It was pointed out earlier that repair would be one such value-creating activity.

Process values are closely associated with stories and experiences. Their importance has been recently pointed out by behavior economics. But economists can only discuss the stories or experiences when our customers use our products. We, engineers, can provide stories or experience throughout the whole product life cycle, from design to manufacturing and to use. Web 2.0, for example, illustrates how our customers would like to get involved in the process beyond just receiving the product.

1.22 Concluding Remarks

The following remark by de Madriaga [16] is very interesting:

Fair play is a term of sport. Let us note this: sport; pure action. Fair play means the perfect adaptation of the player to the game as a whole. It regulates the player's relations with his team-partners but also with his adversaries. This is already wisdom. For good relations with our allies are but reason. Now wisdom is something more than reason. It is a vision of the whole, an intuition of all as one single game, and of opposition as a form of co-operation. Fair play implies an effacing of the individual before the team, and of the team before the game. But this effacing does not mean annihilation. Far from it. It provides better conditions for the efficiency of the individual, since it makes his actions fit in with the actions of others in a perfect system of co-operation. This intuitive and instantaneous sense of balance between the individual and the community is the true essence of fair play.

Fair play cannot be put into formulas. It soars over all regulations, a living spirit. Elusive, yet precise; supple, yet exacting; it fits the mobile forms of life as closely as the glove in the hand. As a living spirit, it manifests itself in concrete actions. It is inseparable from action, undefinable without action. It is a way of doing things. In fact, *fair play is action*.

This implies that we should move away from traditional static and one-way product development and instead design our systems or machines in such a way that they can interact deeply with the outside world and change dynamically in response to their changes. And we have to work closely with these machines or systems.

In other words, how we can take reasonable or rationalized action under conditions of great uncertainty will be a big issue in the twenty-first century. Emotion will play an important role in providing guidelines or directions for wiser actions, because interactions bring forth wisdom, and emotions govern our motives for actions. Emotions are believed to govern our actions at the meta-level.

References

1. Hayes M.: The Economics of Keynes: A New Guide to the General Theory. Edward Elgar, Cheltenham (2008)
2. Dillard D.P.: The Economics of John Maynard Keynes: The Theory of Monetary Economy. Kessinger, Whitefish, MT (2005)
3. Simon H.A.: Administrative Behavior. Free Press, New York (1997)
4. http://www.jske.org
5. Gibson J.J.: The Ecological Approach to Visual Perception. Lawrence Erlbaum Associates (1987)
6. http://www.edpsycinteractive.org/topics/cogsys/piaget.html
7. Piaget J.: The Child's Conception of the World. Bowman and Littlefield, New York (2007)
8. Japanese Ministry of Education, Culture, Sports, Science and Technology: Yuragi – Biological fluctuation approach to develop intellectual artifacts and information systems. http://www.yuragi.osaka-u.ac.jp/introduction/yuragi.html (in Japanese) (2006)
9. Yanagida T., Iwaki M., Ishii Y.: Single molecule measurements and molecular motors. Phil. Trans. R. Soc. B **363**(1500):123–134 (2008)
10. Nurzaman S.G., Matsumoto Y., Nakamura Y., Koizumi S., Ishiguro H.: Yuragi-based adaptive searching behavior in mobile robot: from bacterial chemotaxis to levy walk. In: Proceedings of the IEEE International Conference of Robotics and Biomimetics 2008 (ROBIO 2008), pp. 806–811 (2008)
11. Schon D.A.: The Reflective Practitioner: How Professionals Think in Action. Basic Books, New York (1984)
12. Norman D.A.: Emotional Design: Why We Love (or Hate) Everyday Things. Basic Books, New York (2004)
13. http://berkeley.edu/news/media/releases/2004/10/27_stark.shtml
14. Huizinga J.: Homo Ludens: A Study of the Play Element in Culture. Maurice Temple Smith, London (1971)

15. Peirce C.S.: Studies in Logic. BiblioLife, Charleston, SC (2009)
16. de Madriaga S.: Englishmen, Frenchmen & Spaniards. Oxford University Press, Oxford (1929)
17. Braden V., Bruns B.: Vic Braden's Mental Tennis: How to Psych Yourself to a Winning Game. Little Brown & Co., London (1994)
18. Maslow A.H.: Motivation and Personality, 2nd edition. Harper and Row, New York (1970)
19. Goleman D.: Emotional Intelligence: Why it Can Matter More Than IQ. Bloomsbury, London (2005)
20. Kaiho H. (ed.): Warm Cognition Psychology. Kaneko Shobo, Tokyo (1997) (in Japanese)
21. Prahalad C.K.: The Fortune at the Bottom of the Pyramid: Eradicating Poverty through Profits. Wharton School Publishing, Indianapolis, IN (2006)
22. http://www.valuefoundation.org/
23. Prahalad C.K., Ramaswamy V.: The Future of Competition: Co-creating Unique Value with Customers. Harvard Business School Press, Harvard, MA (2004)
24. Toffler A.: The Third Wave: The Classic Study of Tomorrow. Bantam, New York (1984)

Chapter 2
SHOJI: A Communication Terminal for Sending and Receiving Ambient Information

Masaki Shuzo[1] and Ichiro Yamada[2]

Abstract In modern Japan, increased human mobility has resulted in many people being geographically separated from their families and friends. Thus, there is a need for communication devices that can provide a link between geographically separated family members and friends. Although there are two types of communication – instrumental and consummatory – few studies have been conducted on the latter. We have developed a communication terminal that uses the exchange of ambient information as a means to promote consummatory communication. A concept for effectively communicating ambient information was derived from data collected from questionnaires. This concept was used to develop a communication terminal called "SHOJI" that can send and receive ambient information such as the temperature, illumination, light color temperature, and noise level as well as information about the presence or absence of individuals, their movements, and their emotions. We evaluated it experimentally, in which the participants were parents and children living apart. They judged that the information was sufficiently expressed, which indicates that the terminal is useful for exchanging ambient information.

2.1 Introduction

In modern Japan, many people live apart from those with whom they have a close relationship, such as family members and romantic partners. Since they are geographically separated, they cannot often meet and communicate directly. Although

[1] M. Shuzo
Department of Mechanical Engineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
e-mail: shuzo@lelab.t.u-tokyo.ac.jp
[2] I. Yamada
Department of Mechanical Engineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
e-mail: yamada@mech.t.u-tokyo.ac.jp

computer network-mediated media communication tools will enable people living apart to interact more frequently, there are various obstacles to efficient media communication. One goal of media communication is to realize *kansei* communication, which conveys human emotions and/or the ambience of their locations.

Human group communication is often classified as either instrumental or consummatory communication [1]. The purpose of instrumental communication is to change the receiver's cognition, emotion, and/or action by communication. Examples of instrumental communication are providing knowledge to others, persuading others, and changing others' actions. The purpose of consummatory communication is communication itself, namely, to share one's experiences, emotions, knowledge, and/or opinions. Examples of consummatory communication are joking, complaining, and talking about mundane things. It is said that most daily communication is consummatory [2].

For both instrumental and consummatory communication to be effective, it is necessary to share communication assumptions between the two parties (Figure 2.1). However, most media communication does not convey nonverbal information such as gestures, eye movement, and facial expression and social information such as age, gender, and job. Most previous studies have focused on instrumental communication; few have focused on ambient information such as human unconscious information and environmental information. To convey the ambient information that tends to be missing in ordinary media communication would be an effective way to improve consummatory communication, which often requires the expression of emotion.

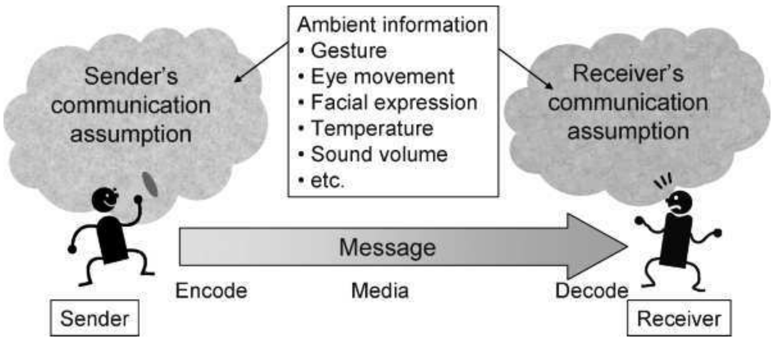


Figure 2.1 Communication model

2.2 Ambient Information

2.2.1 Definition

In this chapter, ambient information is defined as synthetic information consists of human unconscious information and environmental information. We classify information related to consummatory communication as either explicit or ambient. Explicit information is the information that the communication sender consciously wants to share with the receiver in consummatory communication. Examples of explicit information are content, the manner of speaking, and gestures used when joking. Ambient information, on the other hand, is implicit information that the communication sender does not want to consciously share with the receiver – for example, talking in a casual manner or muttering to oneself. Moreover, environmental information such as the balmy weather and birds twittering in the morning is also included in the ambient information.

We have investigated the effectiveness of ambient information in supporting consummatory communication. One reason we focused on ambient information is that it is often and easily lost in media communication. The language content of explicit information is easily conveyed by telephone or e-mail. Nonverbal information, such as gestures and facial expressions, can be conveyed using video chat. On the other, motions without conscious intention and soliloquy cannot be conveyed because the sender does not have the intention to do so. In addition, ambient information has rarely been conveyed because conventional media communication tools cannot sufficiently convey this information.

Another reason we focused on ambient information is that it is an extremely effective way to share communication assumptions. On the one hand, explicit information is conveyed only when the sender consciously conveys it. On the other, since ambient information involves many kinds of information that is automatically shared when the sender and receiver are in the same place, people unconsciously send and share such ambient information. This information promotes the sharing of communication assumptions. When most communication assumptions are shared, communication is more effective. Thus, we assume that the sharing of ambient information is useful for smooth consummatory communication.

2.2.2 Related Works

We can classify existing research on ambient information into two areas: "information that people unconsciously produce", such as emotion, presence, and motion, and "environmental information", such as temperature, illumination, and noise.

A previous study related to ambient information focused on *tsunagari* communication [3], which is communication "aimed at fostering a feeling of connection between people living away and maintaining their social relationships". To validate the concept of *tsunagari* communication, a prototype "Family Planter" terminal was developed. It uses an infrared sensor, an ultrasonic sensor, and a touch sensor to detect users' presence and motion information. These detected information is conveyed to other Family Planter terminals. Although the effectiveness of *tsunagari* communication had already been examined in experiments with family members and students as participants, only human presence and motion information was conveyed.

There have been previous studies in which ambient information other than presence and motion information have been communicated. For example, Lovelet [4] communicates human temperature information; a digital family portrait [5] conveys information regarding human activity (*e.g.*, opening or closing doors); Lumi Touch [6] transmits information regarding human touch; Peak-A-Drawer [7] is a virtual shared drawer; the Lover's cup [8] is a coffee cup that exchanges the time of drinking between two people in different places; and Meeting Pot [9] is a coffee maker that expresses coffee aroma to a remote office while the device is used.

Other studies have focused on emotion, which is an important communication assumption for consummatory communication. These, however, treated emotional information separately from other ambient information. For example, ComSlipper [10] is a pair of slippers with sensors that express the emotion of the other party; and TCON [11] is a tactile facial form that expresses the emotion of the other party. These studies communicate only a small portion of the various types of ambient information.

2.2.3 Our Approach

We propose expressing various types of ambient information by using a unified metaphor rather than focusing on specific portions of ambient information. We developed a method for expressing ambient information that helps two parties communicating to intuitively understand the situation of the other party. They can select the portion of ambient information to be transmitted depending on the time and circumstances. This method is expected to protect each party's privacy through the use of a metaphor.

We have now developed a communication terminal for conveying various types of ambient information that we call "SHOJI" (symbolic hosting online jog instrument).

2.3 Target Usage

SHOJI can be used for both private and public communication:

- private communication between:
  - family and husband or wife working at a distant location;
  - parents and children;
  - hospital patient and family;
  - romantic partners;
- public communication between:
  - head office and distant plant;
  - office and home of a teleworker.

Private communication is mainly between people who have a close relationship and cannot communicate easily due to their physical separation. For example, when a husband or wife assigned to work at a distant location returns to his or her temporary residence after work, he or she might want to know the living situation back home. Ambient information might enable him or her to feel at ease or motivate them to pick up a telephone and call home.

In public communication, people who work at a distant location could use ambient information to maintain their relationship. Since most communication in work situations is instrumental communication, people contact those working at distant locations by telephone, e-mail, or video conference only when necessary. However, to perform their jobs effectively, it is important for work group members to maintain their relationships by conducting a moderate amount of consummatory communication. Thus, creating communication cues is an important function for a terminal designed to communicate ambient information.

2.4 Technical Functional Requirements

Given the situations discussed above, we identified three technical functional requirements:

1. the ability to convey various types of ambient information that are effective in enabling each party to imagine the environment, motion, and state of mind of the other party;
2. the ability to protect the privacy of each party so that they do not feel uncomfortable;
3. the ability to enable people who have different lifestyles to communicate smoothly.

To meet the first technical functional requirement, we use the information that is part of the communication or communication assumption. This includes information about humans and their environments. Since the ambient information communication terminal is used in daily life situations, the information about the two parties should be detectable without imposing any physical restrictions on them. We thus use human presence and motion information. In addition, we attempt to detect emotion, which better represents the ambient information.

To detect emotion without physical restriction, we analyze the voice characteristics from voice information obtained with a microphone and extract the emotion. There are two reasons for detecting only emotion and not voice information itself. The first is to protect privacy. Although it is possible to communicate the voice information obtained from the microphone, users would certainly feel uncomfortable if they thought that their privacy was being invaded.

The second reason is that all the voice information is not always necessary for smooth consummatory communication. It is possible that conveying all the voice information would be bothersome for the users. In addition, one of the features of consummatory communication is sharing emotion. This means that extracting and conveying only the emotional information rather than all the voice information is more effective for consummatory communication.

In the architectural field, there have been several studies on the effect of environmental information, *i.e.*, room color and illumination. On the basis of these findings, we use sensible temperature (a compound function of temperature and humidity), illumination, noise, and color temperature. These kinds of environmental information are easy to detect, and helpful for users to imagine the other side because they change in a relatively short time. Table 2.1 lists the information detected by the SHOJI ambient information communication terminal.

Table 2.1 Information detected by the SHOJI terminal

Information	Type
Environment	Sensible temperature Illumination Noise Color temperature (light color)
Human	Presence Motion Emotion (extracted from voice)

Our solution to the second technical functional requirement, privacy protection, is to enable bi-directional communication of ambient information in an abstract manner. Since the digital expression of acquired information can lead users to feel that their privacy is being invaded, we express the acquired ambient information in a sensuous manner using light and color.

The third technical functional requirement is that people who have different lifestyles should be able to communicate smoothly. One of the potentially disruptive factors in telecommunication is the difference in lifestyles between the two parties. People in remote locations may not spend much time in front of a communication terminal, so they may feel and appear uncomfortable. Thus, the communication terminals must be able to cope with differences in lifestyles.

Our solution is to express not only the present ambient information but also prior ambient information. This means that even though one party may have already gone to bed, the other party can better understand his or her situation from the previous illumination and sensible temperature information.

2.5 Questionnaire

2.5.1 Overview

As mentioned above, the ambient information sent using the communication terminal should be expressed so that users do not feel that their privacy is being invaded. Our approach is to express information in an abstract and sensuous manner using light and color.

To determine the most appropriate expression method, we conducted a survey by questionnaire. We looked at four methods of expressing ambient information: using light form, using light color, using light brightness, and using motion. We asked participants about the validity of each method. The participants in the first survey were 63 adults (33 men and 30 women). They completed the questionnaire while watching demonstrations of some of the expression methods.

The participants were first told, “Please imagine a situation in which the terminal is placed in your living room and that your parents (or your child) are living apart. An identical terminal is placed in their living room. The terminals detect each room’s environmental information such as room temperature, noise level, and illumination level and also detect information about the people in the room such as their entering/leaving and emotions. The terminals exchange this information.”. The participants completed the questionnaire without further information about the terminal.

2.5.2 Results

First, the participants answered questions about the adequateness of four expression methods representing seven types of ambient information. The description of human presence information was “a measurement that represents whether a person is in the other living room”, that of human motion information was “a measurement that represents the speed of human motion in the other living room”, and that of emotional information was “a measurement that represents the emotional status of the person in the other living room”. The participants then chose the detail of the expression about the expression method that they rated “appropriate”.

The questions about appropriateness had a seven-point scale, from “not appropriate” to “appropriate”. The participants were asked questions about the further detail of the expression if they answered “appropriate”.

Figure 2.2 shows the percentage of respondents who answered “appropriate” for each type of information. The results showed that it was appropriate to express sensible temperature, color temperature, and emotion using color and that it was appropriate to express presence and motion information using light brightness and motion. Color expression was judged appropriate for illumination information by 79% of the participants. Light brightness and motion expression were judged appropriate for illumination information by 73%. There was thus little difference between the two expression methods. There was also little difference between using sound (79%) and light brightness and motion (75%) for noise information.

Next we describe the details of each expression method. The results of information for which only one expression method was judged “appropriate”, such as sensible temperature, color temperature, presence, motion, and emotion, showed the explicit tendency in each expression method. Following those results, we determined the detail expression.

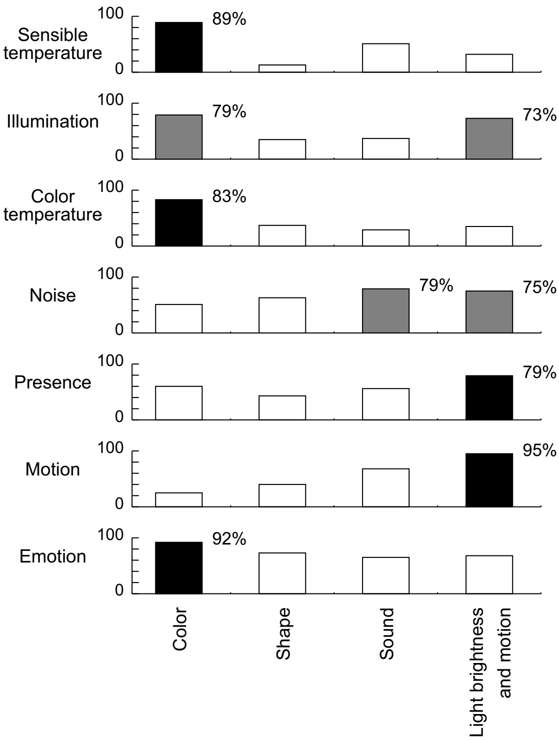


Figure 2.2 Percentage of respondents who answered “appropriate”

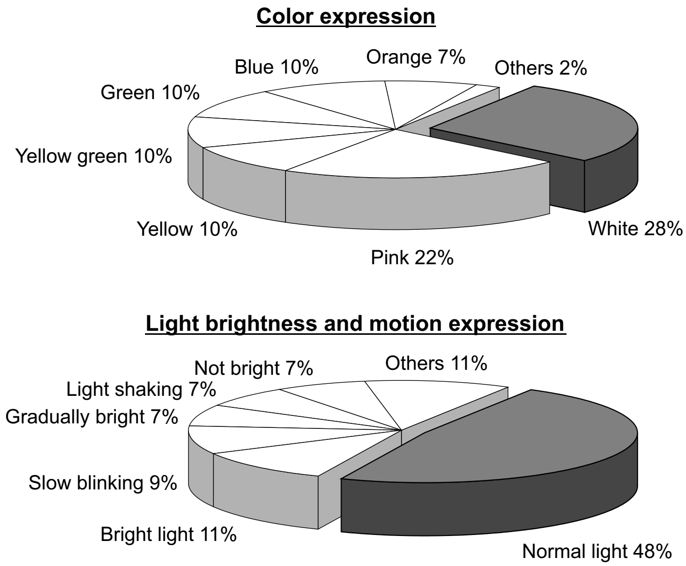


Figure 2.3 Ways of representing illumination information

The results for illumination and noise information were more complex. Figure 2.3 shows the results for color expression and light brightness and motion expression methods about the illumination information. Half of the participants answered "normal bright" for light brightness and motion expression. In contrast, an explicit tendency was not observed for color expression. Thus, we decided to express illumination information using light brightness and motion with normal light. In the same way, since we could observe a more explicit tendency in light and its motion than sound (data not shown), we also expressed noise by using light brightness and motion.

Some participants also stated that the expression of noise information by sound was troublesome in their daily lives. Thus, we decided not to use sound to express noise information.

2.6 Prototype Terminal

To validate the concept of ambient information communication in which users can feel close to the other party, we developed an experimental terminal. The terminal concept to express ambient information in abstract and sensuous ways is "communication over *shoji*". A *shoji* is a semitransparent room divider that has been traditionally used in Japan. People can be aware of the ambient information in the next room indirectly from the lights and sounds filtering through the *shoji* screen.

The *shoji* metaphor offers two advantages. The first is that it can be an effective way for people to comprehend an ensemble of information. Since the *shoji* is a common room divider in Japan, we assume that people will intuitively understand the *shoji* metaphor, *i.e.*, a person on one side of the divider will be aware of the environment and activities on the other side.

The second advantage concerns the invasion of privacy. Exchanging large volumes of mundane information often invades a person's privacy. This often occurs when the transmitted information is one-sided or blatant. Since a *shoji* screen is semitransparent when seen from either side, the exchange of information is reciprocal, and information is transmitted naturally, without any specific effort required to do so.

On the basis of this discussion, we developed our experimental communication terminal, "SHOJI" (symbolic hosting online jog instrument), which is shown in Figure 2.4. The figure also shows how emotional and other cue information is expressed through the terminal. Figure 2.5 shows a schematic of the terminal. It has a microphone, a temperature and humidity sensor, an illumination sensor, three infrared sensors, and an ultrasonic sensor. The terminal is connected to the network through a local area network port. The detected information is transmitted to a predesignated terminal that displays the received information using full-color light-emitting diodes (LEDs). The lights in the center of the terminal present the information about human activities, the lights in the upper portion of the terminal present information about the environment, and the lights in the lower portion show previous information about the environment.

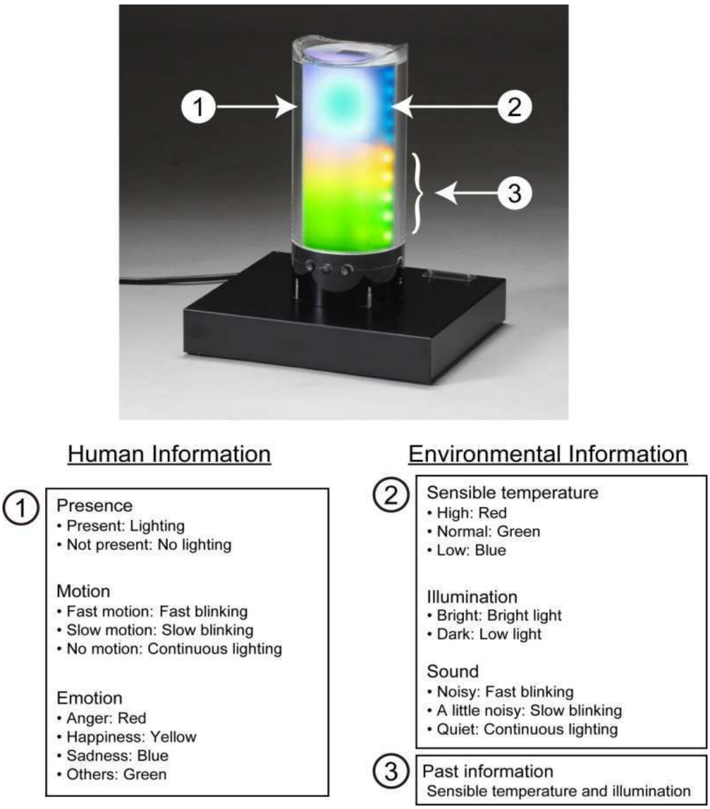


Figure 2.4 Appearance of the developed terminal "SHOJI" and adequate representation methods found in questionnaires

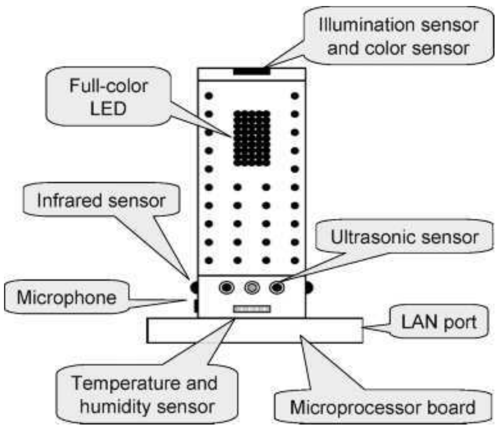


Figure 2.5 Schematic of the SHOJI terminal

2.7 Emotion Extraction from Voice

In this section, we describe an algorithm for emotion extraction under a real environment. Figure 2.6 shows a flowchart of the proposed algorithm, which consists of three serial recognition processes: voice/non-voice recognition, speaker recognition, and emotion recognition.

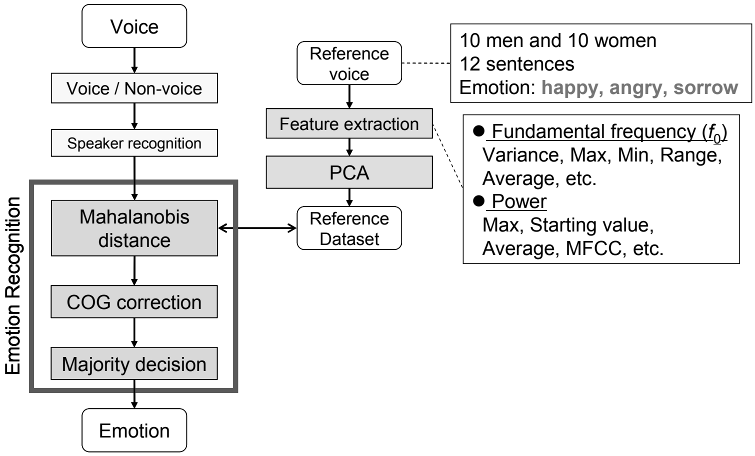


Figure 2.6 Flowchart for the extraction of emotion from voice data

2.7.1 Voice/Non-voice Recognition Process

Since the sounds from the microphone include various kinds of voice, noise, and environmental sounds, the determination of voice/non-voice is necessary prior to the emotion recognition. For the voice/non-voice recognition analysis, we used three datasets (one for voice, two for non-voice). The first dataset was voice data from ten adult men and ten adult women. The second dataset was the domestic sounds database [12]. The third dataset was the Real World Computer Partnership (RWCP) sound scene database in real acoustic environments [13].

In a discriminant analysis, the discrimination function with the modified Mahalanobis distance ($b = 2$) was used [14]. The fundamental frequency (f_0), variance and starting value of the power, and the average and variance of the 12-dimension mel-frequency cepstrum coefficient (MFCC) were selected as sound feature values. Figure 2.7 shows the computed Mahalanobis distance between voice and non-voice data. This result suggests that voice data can be discriminated from non-voice data at a distance larger than 80.

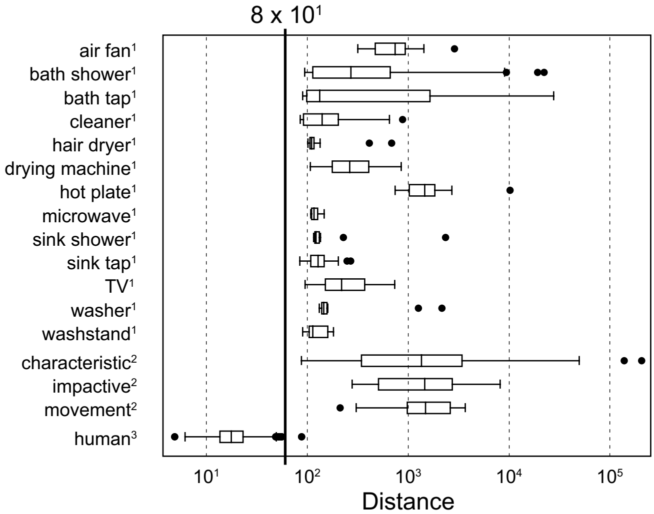


Figure 2.7 Mahalanobis distance between voice and non-voice data (1: AIST database of domestic sounds[3]; 2: RWCP sound scene database[4]; 3: human voice data)

2.7.2 Speaker Recognition Process

Since voice features show large natural variations between individuals, it is difficult to recognize the emotion of a specific person when several people are speaking simultaneously. Thus, speaker recognition is also necessary in order to recognize emotion with high accuracy. In a real environment, there will not be any prior reference data for a speaker's voice; therefore, the standard discrimination analysis, which requires reference data, would not be an appropriate method.

In this process, to discriminate between different speakers, we implemented a hierarchical cluster analysis method to accumulate and retain up-to-date reference data. We used Ward's method for the clustering analysis with the cosine distance. The voice data and feature values were the same as in the previous process. The results of clustering for up to ten men or ten women are shown in Figure 2.8. These results suggest that the recognition rate is about 80% when the number of speakers is three or fewer, which is reasonable in a normal-sized family room.

---

[3] JIS TR S 0001: 2002.
[4] http://tosa.mri.co.jp/sounddb/indexe.htm.
"Impactive" is sound caused by the impact of objects such as wood panel [kon, poku] and metal panel [kan, kin]. "Movement" is sound caused by human actions such as clapping hands [pan, pon] and snapping chopsticks together, opening bottle cap [baki, puchi]. "Characteristic" is sounds whose tone describes the type of source, such as bells, coins [chirin, shan], falling books, tearing paper [basa, bi-i].

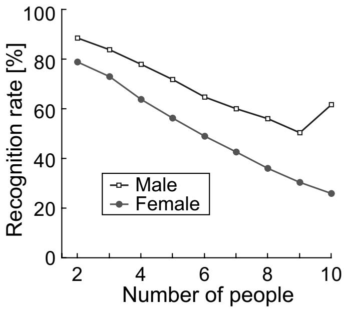


Figure 2.8 Speaker recognition rate obtained by cluster analysis as function of number of people

2.7.3 Emotion Recognition Process

Consequently, emotion recognition was conducted. The voice data from the 20 adults (ten men and ten women) mentioned above was treated as reference data. Each person spoke 12 sentences using three types of emotion (happy, angry, and sad). Three minutes of continuous voice data spoken by 11 other adults (five men and six women) was treated as target data.

In this process, the discrimination function with the modified Mahalanobis distance ($b = 2$) was also used, and the fundamental frequency and power were used as feature values. As a result, the emotion recognition rate was about 70% for men, and about 40% for women (data not shown).

Since human emotions generally do not change significantly over a short time scale, the majority decision process can be applied to this discriminant analysis output. Also, since the center of gravity (COG) in the voice feature space is different for each individual, it is effective to correct the COG once every minute in the discriminant analysis process. Even when starting from an initially low rate, the

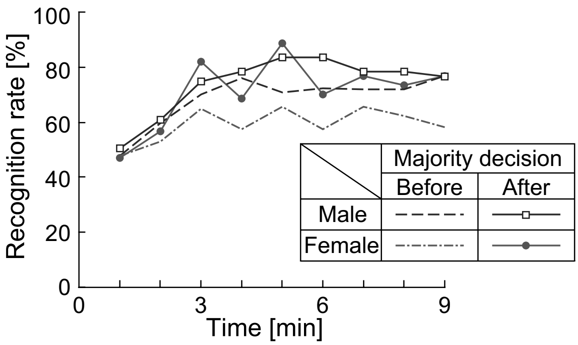


Figure 2.9 Emotion recognition rate using discriminant analysis with majority decision and COG correction

emotion recognition rate improved after a few minutes with the majority decision process and COG correction, resulting in a success rate of about 80% for both men and women in our study (Figure 2.9).

2.8 Field Experiment

2.8.1 Method

To examine the effectiveness of the concept, we conducted a field experiment using the SHOJI terminal with two couples (Figures 2.10 and 2.11). We set one terminal in each couple's living room where they mainly stayed. The terminals were then connected through the Internet. One couple was aware of the information detected in the other couple's terminal.

Participants were a married couple in their 20s and 30s, and a couple in their 50s who were parents of the wife of the younger couple. The senior couple lived on their own, and the young couple lived with their 6-year-old daughter. The distance between the living room of the young couple and that of the senior couple was about 30 min by train. They were instructed not to be particularly conscious of the terminal and to live as usual for a week.

Interviews were conducted before and after the experiment. In the pre-survey interview, they were asked about their lifestyles and communication frequency with the corresponding couple. In the post-survey interview, they were asked to evaluate the design of the terminal and the appropriateness of the ambient information expression method. They also answered questions about its effect on consummatory communication.

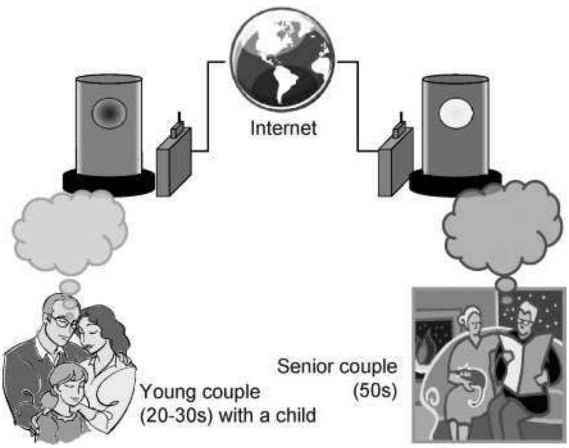


Figure 2.10 Field experiment using "SHOJI"

Figure 2.11 Sample photo of SHOJI terminal in living room used for field experiment

2.8.2 Results

Since the two couples did not live near each other, they did not meet frequently. The telephone was the most commonly used communication tool, and they talked around 5 h/month by telephone. Most communication was between the two wives.

The evaluation of the design of the terminal and the appropriateness of the ambient information expression method in the post-survey interview are summarized below in the form of responses:

- "I could grasp an overview of the ambient information without close attention to the terminal."
- "Although at first I could not understand the color expression of emotional information, within a day or so I could figure out the meanings. The expression of temperature information was intuitively easy to understand. Seven types of information were not too much to comprehend."
- "About the ambient information expressed on the terminal, once I become accustomed to the terminal, I figured out the information at once. However, in the current circumstances, I often forgot the expression method."
- "The expression by LED was good. The size of the terminal was appropriate. If the display region were smaller, the expression would be harder to understand."
- "The light of the terminal was too bright at night. I sometimes mistakenly thought that I had forgotten to turn off a light. It would be better to control the light volume of the terminal at night by detecting the illumination of the room."
- "The terminal was interesting because it was like an *andon* [Japanese indirect illumination made from washi paper, a wooden frame, and a candle]. The color was so beautiful. Washi was good because it was like a night-light."
- "The change in emotional information was a useful cue to talk with my daughter about the condition of her grandparents."

The design of the terminal was favorably evaluated by both young and senior couples. Then, although the young couple could easily understand the expression method of overall ambient information, the senior couple pointed out that there was too much information expressed on the terminal. Thus, the evaluation of overall expression method was affected by the participant's age. On the other hands, focusing on seven types of ambient information, both groups had good or not-so-bad evaluations to each expression method. During this experiment, the participants sometimes focused on various types of information depending on their situation. In one instance, they focused on the temperature information, and in another, they took notice of the emotional information.

The evaluation of the effect on consummatory communication is summarized below in the form of responses:

- "The terminal increased the chance to imagine the other side. Though the terminal was a 'machine,' I could feel connected to the other side."
- "The chances to communicate increased. My wife sometimes made a phone call to her mother when she watched the terminal."
- "Since the terminal informed us of the presence of the other side, I could easily contact them. In contrast, since I could judge the condition of my parents by watching the terminal, the number of times we contacted them to determine their condition decreased."
- "Since I could determine whether they were present, it was easy to contact. When I had to contact them, the information on the terminal helped me decide when to phone."
- "When our family ate dinner, we often talked about the other side. Since the terminal gave us information about them, we could not only talk about my family's events but also about the events of the other side."
- "I felt uncomfortable that the other side knew when I got up late. Additionally, I didn't want to convey information when my family was upset in the morning."

All the participants favorably evaluated the effect of the information sharing, which is the purpose of consummatory communication. In particular, the results of the interview showed several advantages, such as "it helps me to contact them by providing information about the situation of the other side", "ambient information leads users to imagine the other side", and "by watching the information on the terminal, users can carry on a pleasant conversation about the other side".

The effect of the terminal was ambiguous concerning the frequency of communication. On the one hand, thinking about the situation of the other side leads to contacting them. On the other hand, since users know the condition of the other side, they may not bother contacting them. Nevertheless, the results also suggest that communication became smoother even if the frequency of communication decreased.

Though the information that users want to convey increased communication effectiveness, there were problems caused by sending information that users did not want to convey. An increase in the information conveyed inevitably leads to an increase in not only the information that they want to convey but also in the information they do not want to convey. This issue requires further examination.

2.9 Conclusion

We have developed a communication terminal called "SHOJI" that conveys ambient information for the purpose of promoting fluent consummatory communication. The concept on which it is based was derived from data collected from questionnaires. SHOJI can send and receive such ambient information as the sensible temperature, illumination, light color temperature, and noise level as well as information about the presence or absence of individuals, their movements, and their emotions. The participants in an experimental evaluation judged that the information was sufficiently expressed, which indicates that the terminal is useful for exchanging ambient information.

We plan to conduct a longer-term field experiment based on the knowledge acquired in this study. In addition, on the basis of an evaluation of the effectiveness of each type of information, we will identify the information that is most effective for conveying ambient information.

Acknowledgements The authors would like to thank Mr. Masatoshi Kajimura and Mr. Fumihiko Takeishi of GS Yuasa Power Supply Ltd. for fabrication of the prototype terminal "SHOJI". Part of this research was supported by JST, CREST.

References

1. Dewey J.: Experience and Nature. Dover, New York (1958)
2. Rubin R.B., Perse E.M., Barbato C.A.: Conceptualization and measurement of interpersonal communication motives. Hum. Comm. Res. **14**:602–628 (1988)
3. Itoh Y., Miyajima A., Watanabe T.: "TSUNAGARI" communication: fostering a feeling of connection between family members. In: CHI '02 Extended Abstracts on Human factors in Computing Systems, pp. 810–811. ACM Press, New York (2002)
4. Fujita H., Nishimoto K.: Lovelet: a heartwarming communication tool for intimate people by constantly conveying situation data. In: CHI '04 Extended Abstracts on Human factors in Computing Systems, pp. 1553. ACM Press, New York (2004)
5. Mynatt E.D., Rowan J., Craighill S., Jacobs A.: Digital family portraits: providing peace of mind for extended family members. In: CHI '01 Extended Abstracts on Human factors in Computing Systems, pp. 333–340. ACM Press, New York (2001)
6. Chang A., Koerner B., Resner B., Wang X.: Lumi Touch: an emotional communication device. In: CHI '02 Extended Abstracts on Human factors in Computing Systems, pp. 313–314 ACM Press, New York (2002)
7. Siio I., Rawan J., Mynatt E.: Peek-A-Drawer: communication by furniture. In: CHI '02 Extended Abstracts on Human factors in Computing Systems, pp. 582–583. ACM Press, New York (2002)
8. Chung H., Lee C.H.J., Selker T.: Lover's cups: drinking interfaces as new communication channels. In: CHI '06 Extended Abstracts on Human factors in Computing Systems, pp. 375–380. ACM Press, New York (2006)

9. Siio I., Mima N.: Meeting Pot: coffee aroma transmitter. In: UbiComp'01 – International Conference on Ubiquitous Computing (2001)
10. Chen C.Y., Forlizzi J., Jennings P.: ComSlipper: an expressive design to support awareness and availability. In CHI '06 Extended Abstracts on Human factors in Computing Systems, pp. 369–380. ACM Press (2006)
11. Heesook S., Junyoung L., Junseok P., Youngjae K., Hyunjoo O., Taehwa L.: A tactile emotional interface for instant messenger chat. In: Smith M.J., Salvendy G. (eds) Human Interface and the Management of Information. Interacting in Information Environments, pp. 166–175. Springer, Heidelberg (2007)
12. Kurakata K., Matsushita K., Kuchinomachi Y.: Database of domestic sounds for evaluation of auditory-signal audibility: JIS/TR S 0001. Acoust. Sci. Tech. **24**:23–26 (2003)
13. Nakamura S., Hiyane K., Asano F., Endo T.: Sound scene data collection in real acoustical environments. J. Acoust. Soc. Japan E **20**:225–231 (1999)
14. Kato N., Abe M., Nemoto Y.: A handwritten character recognition system using modified Mahalanobis distance. Syst. Comput. Japan **28**:46–55 (1997)

Chapter 3
A Mixed Reality-based Emotional Interactions and Communications for Manufacturing Skills Training

Keiichi Watanuki[1]

Abstract In recent years, Japanese manufacturing industry has entered a very competitive era due to globalization and to the concept of manufacturing the right thing at the right place. At the same time, manufacturing in Japan is changing from a conventional mass production and perpendicular work division system to a high-precision, small-lot, multi-variety manufacturing system. In addition to this, the need and opportunity for face-to-face contact between colleagues has been drastically reduced as a result of factory and office automation. This means that human functions that are a result of human contact and relationships are being replaced by automated systems. This transfer of relations from human to machine systems causes individuals to be isolated in the process of their work. Therefore, appropriate staff-training programs with communications based on a long-term viewpoint are needed to create and transfer the skill/technology. However, due to the financial, manpower, and time constraints caused by the present tough economic conditions, it is difficult for small and medium-sized industries to organize good training programs. In reality, more than half of small and medium-sized industries implement on-the-job training, video libraries, technical documents, and the like in their training programs. Nevertheless, these time-consuming and inefficient methods do not make it easy to obtain the high-level skill necessary to be a skilled technician. This chapter considers some reasons for difficulties in skills transfer and human resource development, and addresses a new mixed reality-based job training and human resource development for skilled foundry workers. This chapter also looks at tacit knowledge, coherence in knowledge, and action in manufacturing.

[1] K. Watanuki
Department of Mechanical Engineering, Graduate School of Science and Engineering, Saitama University & Brain Science Institute, Saitama University,
255 Shimo-okubo, Sakura-ku, Saitama-shi, Saitama 338-8570 Japan
e-mail: watanuki@mech.saitama-u.ac.jp

3.1 Introduction

In the Japanese manufacturing industry, the hollowing out of industry due to the transfer of production centers overseas is having a serious effect on areas where industry is concentrated. In particular, in the core manufacturing technology industries which have supported the manufacturing industry to date, the deterioration in skilled labor is a concern. Also, with an aging population and young people avoiding going into the manufacturing industry, finding new staff/apprentices is becoming difficult, and consequently the Japanese manufacturing industry is facing a crisis. In order to continue the design and manufacture of high-added-value products in the future, the passing down of core manufacturing technology and skilled labor as well as the creation of knowledge is indispensable.

As an example, in the Japanese casting and product forming industry, a high level of international competitiveness has been maintained by rationalizing design and production processes, improving productivity and yield rates, and achieving shorter delivery times, as well as passing on technology and skills, and solving problems such as environmental issues. However, the product forming industry is almost totally comprised of small and medium-sized companies, and sub-contracting transactions are frequent, so at present the management situation is weak. In order to strengthen the product forming industry in Japan in the future, it is important that the technology and skills accumulated to date are further expanded by research and development, unique technologies are established and strengthened, technologies and skills are transferred, human resources are trained to use information and communication technology (ICT), and so on.

Therefore, this chapter describes the acquisition of manufacturing knowledge and human resources training by sharing a "ba" (place) for the application of multimedia technology, virtual reality (VR) technology, *etc.* [1, 2]. In particular, a system is described that uses a combination of a three-dimensional visualization system, created using VR/AR technology, and a force-sensing display device, created using robot technology, to enable not only the visual effect to be obtained, but also to obtain a sense of the force required in casting. It does this in combination with the communication between the engineer and skilled operative. By adding annotation and so on to the space in which a model is displayed in VR space, technical information can be shared, communication between engineer and skilled operatives promoted, and embodied knowledge can be acquired. Also, new experiments in operational support for elderly and female workers using VR technology are described.

3.2 Skills Transfer

3.2.1 Present Situation and Problem Points Regarding Methods of Skills Transfer

As the pace of technical change becomes faster and competition between companies becomes more severe, it is necessary for each company to try to learn faster than the other. Therefore, the number of companies undertaking knowledge management, such as knowledge sharing based on organizational learning theory, knowledge transfer, etc., is increasing, advanced technologies are being recorded in manuals, and information databases using ICT are being introduced. However, as the division of labor in manufacturing sites increases, obstacles to information sharing arise, such as the attitude “as long as it follows the manual, this is accurate enough for my work,” and not much thought is given even though accidents occur or defective products are produced. Therefore, there is a danger that knowledge management, which aims to share knowledge and skills among members of an organization in order to improve the rate of learning, can have the opposite effect and actually hinder the nurturing of skilled operatives. Also, if too much reliance is placed on knowledge accumulated in databases within the company, the opportunities for problem solving on one’s own reduce. Therefore, by making knowledge sharing easier in order to quickly develop human resources capable of raising efficiency, learning by experience is hindered, so when problems occur it is not possible to quickly respond to them. This leads to the dilemma that knowledge sharing can result in higher costs [3].

A skilled worker is one that within a specific field has both specialist training and practical experience, has acquired special skills and knowledge, has outstanding “structured knowledge” in that specific field based on their experience and training, understands problems deeply, can accurately and quickly solve problems, and has excellent self-monitoring skills. To develop such a skilled worker, practical on-the-job training (OJT) and so on is carried out, but various problems with this system are evident. Examples include the fact that learning through failure is difficult structurally, there are time restrictions, depending on the circumstances appropriate guidance is not necessarily provided, individuals become satisfied with their level of learning at a certain stage, and thus the knowledge at the site becomes fixed, hindering adaptation to a new environment, *etc.*

Also, although the necessity of knowledge transfer by engineers and skilled operatives in the core manufacturing technology industries has already been pointed out, the sense of crisis in qualitative and quantitative shortages was strengthened with the “year 2007 problem”. From 2007, the baby boomer generation, which was responsible for industry during the era of high growth rate in Japan, started to retire in large numbers, raising many concerns. A number of initiatives have been tried to solve the year 2007 problem, but there have been many drawbacks, for example, the skills transfer and know-how took time and did not proceed smoothly; the human resources systems providing guidance were insufficient; and other issues. The main background to these problems is that during the period of high economic growth (the period from roughly 1955 to 1975) new employees were employed in large numbers. Technology and skills were passed down through in-company training. However, after the oil-shock of 1973, energy efficiency, automation, and new technologies were rapidly introduced, and the traditional transfer of knowledge and skills tended to be neglected, which meant that the passing on of technology and skills to the younger generation slowed down. Also, from the economics viewpoint, due to the technology transfer associated with the relocation of factories abroad, mainly to South-East Asia, and with outsourcing, the hollowing-out of technology and skills within Japan has become serious. Following the collapse of the bubble economy in 1991, companies within Japan restructured, and the resultant cutbacks in capable technical and skilled staff, made the transferring of technology and skills even more difficult. Furthermore, as a result of the decline in the birthrate and the dislike for hard, dirty, or dangerous workplaces, it has become difficult to secure young engineers and skilled operatives. This has resulted in a large age gap between the baby boomer generation and the younger generation, so technology and skills are not smoothly passed down. As a result of this situation, measures to pass on technology and skills came into full swing.

At present, transferring skill is typified by technical documents, video libraries, and OJT. Creating technical documents and video libraries requires a lot of labor and cost, and despite their many merits they are not widely used due to poor ease of use. At present, most factories concentrate on OJT for transferring skills.

The problem points with conventional transfer skill methods are as follows. Technical documents are ideal for describing technologies, but they have the disadvantage that it is not possible to describe the movements and skills of operatives very well. Also, video libraries are good at recording and preserving skills, but a lot depends on the knowledge and capability of the person watching the video. Also, in the normal video format it is difficult to immediately watch the part you want to see. In contrast, OJT as shown in Figure 3.1 is an extremely good method of skills transfer, as the whole body can experience the training using a person's five senses of seeing, hearing, taste, smell, and touch. However, this requires much time, and there are problem points; for example, the opportunities for experience are limited by the production of many products in small quantities; and it also very much depends on the educational ability of the person being instructed.

In this way, each of the methods of skills transfer has many advantages, but also has a limit to transfer skills. In addition, looking at these methods from the viewpoint of knowledge conversion, with OJT and video libraries tacit knowledge is transmitted as tacit knowledge, but with technical documents tacit knowledge is converted into explicit knowledge for transmission. Here, explicit knowledge can be defined as "knowledge that can be represented in the language of documents and diagrams, *etc.*" and "knowledge based on objective general laws", and so-called "technology" corresponds to explicit knowledge. On the other hand, tacit

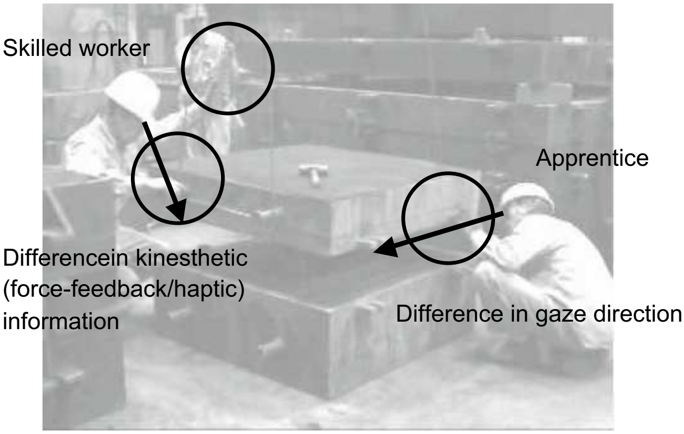


Figure 3.1 Difficulties of OJT

knowledge [4] can be defined as "knowledge that is difficult to represent clearly in the language of documents and diagrams, *etc.*" or "knowledge based on subjective or individual experience", and so-called "skills" correspond to tacit knowledge. Using these methods on their own, where tacit knowledge is transmitted as tacit knowledge or tacit knowledge is converted into explicit knowledge by omitting a part of the tacit knowledge, problems are experienced with the quality, quantity, or efficiency of the transmitted knowledge. In other words, with the main conventional methods of passing down skills, namely technical documents, video libraries, and OJT, effective skills transfer is not possible at present.

3.2.2 New Methods of Skills Transfer

As stated previously, although the conventional methods of skills transfer have many advantages, they also have various problems. Any future new system of skills transfer has the following requirements:

1. It must be possible to easily search for the knowledge that the user requires from amongst the large quantity of knowledge in the system.
2. It must not require a large amount of labor and a high cost for preparation of documents, but must be expandable and flexible enough to be able to re-use documents prepared previously.
3. It must provide a form for easily transmitting tacit knowledge.
4. The quality and quantity of knowledge that can be acquired from images is greatly affected by the knowledge and experience of the individual, so a form must be provided that does not depend on the capacity of the individual. In other words, the large amount of knowledge contained in images must be clearly presented without depending on the capability of the individual.

5. In order that a fictitious, virtual “place (ba)” can be shared, time, spatial, and scale restrictions that interfere with sharing of the “place (ba)” must be overcome at least partially.
6. What is required is not the skills transfer that concentrates on OJT, but an improvement in the quality and efficiency of the knowledge transmitted, using technical documents and video libraries, and so on, and utilizing their respective merits.

3.2.3 Knowledge Creation and Skills Transfer by Sharing of “Ba”

One of the models used for knowledge management is the SECI model (Socialization, Externalization, Combination, Internalization: organizational knowledge creation) advocated by Nonaka and others. This is a model in which knowledge is created by the interaction of tacit knowledge and explicit knowledge through four knowledge conversion processes: (1) socialization; (2) externalization; (3) combination; and (4) internalization [5]. SECI is obtained from the first letter of the four knowledge conversion processes. If these processes are applied to the process of acquiring skills at a foundry, then (1) socialization is the process in which a skilled operative clarifies awareness of problems and recognizes tacit knowledge through OJT or work on site; (2) externalization is the process in which explicit knowledge of casting techniques is acquired from internal documents, technical documents, *etc.*; (3) combination is the process by which a skilled operative generalizes various types of knowledge obtained in the externalization process, and forms their own image of the skill that is to be acquired; and (4) internalization is the process in which the casting skills are made their own by actions and learning based on the created image.

Nonaka and others introduced the concept of “ba” in the field of knowledge management, where ba is defined as a “shared context in which knowledge is created, shared, and utilized” [6]. Context is necessary for the creation of knowledge. The key concept for understanding ba is interaction. Knowledge is not created by the activities of an individual carrying out activities by him- or herself, but by interaction among individuals or interaction between individuals and the environment. Interaction occurs in actual space, virtual space, or a combination of the two. In particular, in socialization and externalization direct face-to-face interaction at the same time and in the same space is important. Ba is the context shared by people that are interacting. Both ba and its participants co-evolve as a result of the self-transcendence and knowledge creation through this interaction.

Itami defines ba as “the framework of circumstances in which people participate, mutually observe whether consciously or unconsciously, communicate, mutually understand, mutually approach, and mutually psychologically stimulate each other” [7]. Further, strictly from the viewpoint of the elements that comprise ba, it is defined as “the framework of circumstances in which a highly dense information interaction is continuously created in various forms through the sharing by more than a certain minimum amount by the participants of a ba of four basic elements that constitute ba, namely agenda (what the information relates to), interpretation code (how the information should be interpreted), information carrier (medium or media conveying the information), and desire for association".

"Ba" can be thought as one of the things to be considered in order to pass down skills. As used here, ba does not just refer to a so-called physical space, as stated previously, but means the space of a specific time, space, or "relationships". In other words, ba is the "interaction relationship". This ba, in other words how this interaction ought to be, may be classified into four types: originating ba, dialoging ba, systematizing ba, and exercising ba. These are defined as follows: (1) originating ba – ba defined by individual and direct interaction; (2) dialoging ba – ba defined by group and direct interaction; (3) systematizing ba – ba defined by group and indirect interaction; (4) exercising ba – ba defined by individual and indirect interaction [8].

In skills transfer, it is extremely important that ba is shared in a specific time and space. Specifically, the case of a highly skilled and experienced operative and a person accompanying the operative at work can be given. By sharing ba in this way, not only can knowledge obtained from the five senses be obtained, but also valuable knowledge of "information that cannot be understood unless in that ba" can be acquired. Hence it can be said that sharing of ba is very beneficial for the skills transfer.

However, the skills transfer by sharing of ba in this way has the following problems:

1. There are time, space, and scale restrictions on the sharing of ba, so there is a limitation making it not possible to efficiently transfer skills.
2. Too much depends on the capacity of the person transferring or receiving the skills, and a large amount of time is required. Under the present circumstances in which skills must be transferred under restrictions of time, cost, and personnel, there are some difficulties aspects of knowledge and skills transfer.

By using an immersion type virtual shared environment system that combines the display of visual information and force information, as described below, it is possible to partially overcome these problems. It is possible to promote the sharing of a fictitious virtual ba, which is considered to be very beneficial for the skills transfer.

One of the work improvement activities in the manufacturing industry is the PDSA/PDCA cycle (PDSA/PDCA cycle: Plan–Do–Study/Check–Act cycle). This PDSA/PDCA cycle is a management method for promoting improvements in production control and quality maintenance, and continuous improvement or work activities in the manufacturing and other industries. This was advocated by W.A. Shewhart and W.E. Deming, who founded quality control in the 1950s, and is also called the Shewhart Cycle or the Deming Cycle (Deming Wheel) [9, 10]. The PDSA/PDCA cycle includes the following four steps: (1) Plan – set a target, and design the process to achieve that target; (2) Do – implement the plan, and measure the results; (3) Study/Check – study and evaluate the measured results, and analyze the results by comparing with the target, *etc.*; (4) Act – identify the points that need to be changed in order to improve the process. These four stages are carried out in succession, and the final "Act" leads to the next PDSA/PDCA cycle, so that work is continuously improved in a spiralling up motion. In this way, continuous and effective knowledge acquisition and implementation by the sharing of ba are necessary for knowledge creation, skills transfer, production control, quality control, and continuous improvement activities in the manufacturing industry. In addition to the conventional skills transfer and work improvement activities, new activities through the sharing of ba and communication are necessary.

3.2.4 Conversation Analysis and Extraction of Knowledge

Ethnomethodology is a theoretical framework for analyzing the process of how people who are active in a real social situation can understand things by talking about those things in a common style and take this understanding to be their reality. Founded by H. Garfinkel, ethnomethodology uses conversation analysis and other methods to focus on the mutual actions between person and person [11]. This conversation analysis is a method that considers everyday conversation and conversation in systematic situations to be a social phenomenon. Social interaction actions have ordered structure and patterns. Also, the participants in the conversation understand the development of the conversation within the sequence of the flow of time. Based on these conversational analyses as the fundamental methodological premise, naturally occurring situations of mutual interaction are recorded by video, and the conversation and the actions are analyzed using the detailed transcript as material [12].

The transcript contains the name of the speaker on the left-hand side of a column, followed by a colon, and the speaker's conversation on the right-hand side. The transcript proceeds as a time series from left to right, and from top to bottom. Four types of symbols are used: for sounds; for the sequence of the conversation; for the transcriber's comments; and for non-verbal actions.

In casting design, the conversation analysis starts from the careful observation of the conversation data, consisting of the sound and visual data and the transcript. The conversation data includes specific actions, conversational procedures, linguistic forms, *etc.* From among these, examples of specific phenomena are collected, and systematically compared and classified. At this time the opinions of several engineers and skilled operatives are obtained, and the collection is analyzed. Based on this data, a knowledge tree is prepared, and it is important that the knowledge is expressed in an intuitive and easy-to-understand manner.

3.2.5 Proficiency and Corporeality

In recent years, research into proficiency has been expanding rapidly. Kitano and others have classified proficient people into skillful proficient people and adaptable proficient people, according to the level of flexibility and adaptability of their knowledge and skills. A skillful proficient person becomes proficient by repeating the same procedure many times, and the speed and accuracy of execution of a skill is outstanding. Also, an adaptable proficient person is a person that has derived conceptual knowledge from executing a procedure, so they can flexibly adapt to the changing circumstances of a task and arrive at an appropriate solution. K.A. Ericsson and others have pointed out that in order to attain proficiency in a skill, well thought out training based on a clear sense of purpose is indispensible [13]. R.R. Burton and others have argued that for learning complex skills, providing a small world in which the complexity gradually increases is effective [14]. A certain skill may consist of several lower level skills, but separating and learning each lower level skill is difficult. However, for this purpose a "small world" is provided for learning appropriate combinations of the lower level skills. By gradually increasing the complexity of these small worlds, the learner is given the opportunity to recognize and correct mistakes that were allowed in the previous small world. By actively monitoring this experience, the formation of still deeper knowledge is enhanced.

Characteristics of a proficient person include learning lower level skills, knowledge acquisition for appropriate problem solving, and the acquisition of appropriate evaluation criteria. Regarding the learning of lower level skills, it is said that in order to reduce the load on the short-term memory, a proficient person makes comparison with long-term memory and remembers by chunking familiar knowledge. There is also research that indicates that by using the knowledge within the long-term memory, the formed chunk does not stay in the short-term memory, but is soon transmitted to the long-term memory, where it can be remembered for a long time. Also, it is said that a proficient person does not carry out controlled processes under conscious control, but carries out automatic processes. Skills are automated as if the person were one with their tools. Next, regarding knowledge acquisition for appropriate problem solving, it goes without saying that a proficient person has mostly structured knowledge, but it is said that this is due to mapping intention and image (*i. e.*, in the internal world) with the corresponding appropriate superficial characteristics (*i. e.*, in the external world). According to H.L. Dreyfus and S.E. Dreyfus, the overall situation in the ba is understood as a pattern by this mapping, which suggests the capability of responding as necessary to confined problems or dispersed problems in accordance with the circumstances [15]. Furthermore, regarding the acquisition of appropriate evaluation criteria, it is thought that a proficient person can respond flexibly to changes in circumstances by constantly monitoring their own state, and by flexibly adjusting their own state.

J.J. Gibson refers to the relationship between the outside world and a person as "affordance", the value and meaning that goes beyond the superficial characteristics and properties of each individual thing or phenomenon. According to Gibson, this is not created in the head of a person, but is already in existence and is directly perceived by that person [16]. Saeki states that a person becoming knowledgeable does not mean that the person possesses knowledge, but that the person becomes more open to the outside world. Saiki states that when an open person encounters the outside world, the prior relationship between the outside world and the person is realized as an interaction. For learning it is necessary that the learner acts in a participating and active manner in order to bring into view a meaningful world [17].

For knowledge and skills transfer, training such as OJT in an actual social setting is of course necessary, as are well thought out training based on a clear sense of purpose, mapping of intention and image with the corresponding superficial characteristics and actions, the learner acting in a participating and active manner, bringing into view a meaningful world, and so on.

3.3 Skills Transfer and Human Resource Development in the Foundry Industry

3.3.1 Knowledge and Skills Transfer of Casting Technology

Kawaguchi City in Saitama Prefecture is one of the major industry clusters of casting related companies in Japan. In the 1960s, there were almost 600 foundries in Kawaguchi, which employed 12,000 people producing 400,000 tons per year. At present, Kawaguchi is a forest of high-rise apartments and office buildings, having been greatly transformed by the effect of the three factory acts. However, there are still over 100 foundries in the area that make use of advanced casting technology and skills to produce high-added-value cast products to support the foundations of Japan's advanced technology industry.

A cast product is normally produced through a sequence of processes. In particular, the work at the foundry site, such as melting, molding, mold assembly, pouring, and shake-out, requires not only casting knowledge, but also sensitivity to weight, a sense of touch, sound, smell, color, temperature, and so on, so skills that make use of all five senses are required. Much of the work at a foundry requires experience and intuition, including ramming, coating, mold assembly, and pouring, and the existence of this tacit knowledge is one of the factors that makes the acquisition of skills difficult.

In order to transfer these casting technologies and skills as well as create still higher added-value products, efforts directed towards the skills transfer and knowledge creation are necessary. For this purpose most companies have activities in place for skills transfer which are centered around technical documents, video libraries, and OJT, as mentioned above. However, the present situation is that effective transfer of advanced skills is not possible by the simple use of these methods. Therefore, a new system for advanced skills transfer is necessary.

3.3.2 New System for Skills Transfer

In the casting process, besides explicit knowledge of the technology for making castings, tacit knowledge in the form of skills is necessary. In order to design and manufacture high-added-value products, it is necessary to obtain knowledge that skillfully combines this explicit knowledge and tacit knowledge, as well as master technologies and skills through OJT or the like.

In our study, we consider an example of the utilization of a skills transfer system and an immersive virtual environment system and their correspondence with the knowledge conversion processes of the SECI model. Here, (1) socialization is the process in which a user clarifies awareness of problems and becomes aware of tacit knowledge through OJT and work on site; (2) externalization is the process in which the user acquires explicit knowledge from documents, technical data, *etc.*, using the skills transfer system; (3) combination is the process in which the user acquires knowledge from images, *etc.*, presented in the form of a combination of explicit knowledge and tacit knowledge; and (4) internalization is the process in which the user makes the actions and learning their own, based on the knowledge obtained from the skills transfer system. When making knowledge one's own while using the skills transfer system in the externalization and combination processes and using the immersion type virtual shared environment system in the internalization process, acquisition of new tacit knowledge as well as the creation of new knowledge is possible.

The authors have developed a skills transfer system for efficient and effective knowledge acquisition using multimedia and VR technology in each of the knowledge conversion processes of the SECI model. This is an example where the utilization of the skills transfer system and the system using the VR technology correspond to the knowledge conversion processes of the model. Here, externalization may be substituted for the process in which a user obtains explicit knowledge from documents and technical data, *etc.*, using a multimedia skills transfer system; combination may be substituted for the process in which the user acquires knowledge presented in the form of a combination of explicit knowledge and tacit knowledge; and internalization may be substituted for the process in which the user makes the acquired skills their own by mock experience within the VR space, based on the knowledge obtained from the skills transfer system. It is when making knowledge one's own while using the skills transfer system in the externalization and combination processes, and using the VR space in the internalization process, that the creation of new knowledge is possible

3.3.3 Skills Transfer System Using Multimedia Technology

Figure 3.2 shows an outline of the skills transfer system. By searching for and visualizing the knowledge, a user can efficiently access the required knowledge.

Tacit knowledge is presented using Synchronized Multimedia Integration Language (SMIL), and explicit knowledge is presented using eXtensible Markup Language (XML). Links are provided to the knowledge associated with keywords, so when necessary the associated information and knowledge can be efficiently acquired. Also, parts that cannot be expressed in images are complemented with links to three-dimensional CAD data and CAE simulations. It is expected that the use of this system will have a large effect that cannot be obtained through conventional skills transfer through sharing of a ba or skills transfer using video libraries, *etc.*

The system assumes the use of a web-based system, and therefore uses languages suitable for use on the web, such as XML, SMIL, and XVL (eXtensible Virtual world description Language). To use the system, the user first searches or navigates through the system, and the target knowledge is accessed from the list of search results or by following the navigation. In the case of knowledge classified as tacit, a multimedia screen using SMIL is presented; in the case of knowledge classified as explicit, documents are presented using XML. Links are provided to knowledge as keywords, and when necessary a user can select a link to receive only the knowledge that is required. Also, for parts that cannot be explained solely using images, and parts that cannot be explained solely using documents, documents and images are mutually linked so that their respective advantages can be utilized. Also, parts that cannot be explained using static or two-dimensional images are supplemented with images that can be easily controlled by the person watching, using three-dimensional CAD data or CAE simulations.

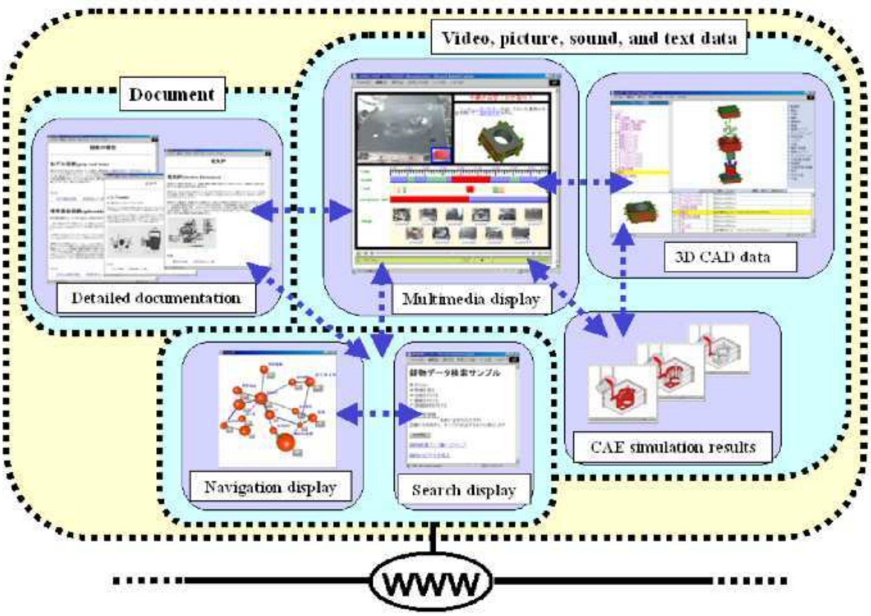


Figure 3.2 Knowledge transfer system

As described above, by linking tacit knowledge with explicit knowledge, tacit knowledge can be presented to a user in a form in which the knowledge is easily transmitted. Also, the required knowledge can be easily searched for, so the user can select the required information as necessary and learn it, so knowledge can be efficiently acquired.

3.3.4 Casting Skills Acquisition System Using VR Space

As shown in Figure 3.3, in the skills transfer system, (1) after acquisition of the explicit knowledge, (2) the explicit knowledge is linked with tacit knowledge. Then (3) using the skills training system based on VR technology, the knowledge obtained in the skills transfer system is internalized and corporeal knowledge is acquired. Finally, (4) through OJT on site the high-level skills are confirmed, and the new skills are acquired. By repeating this series of processes the skills are further developed, and new knowledge or higher level skills are acquired. Here the skills virtual training system using VR technology is called CAMRobot (Cyber Assist Meister Robot). This virtual training system combines a touch panel type computer, a three-dimensional visualization device, and a force sensing device, and supports education in design and manufacturing knowledge and knowledge creation for new high-added-value products. Using this system new people and people with insufficient accumulated experience can obtain experience that is

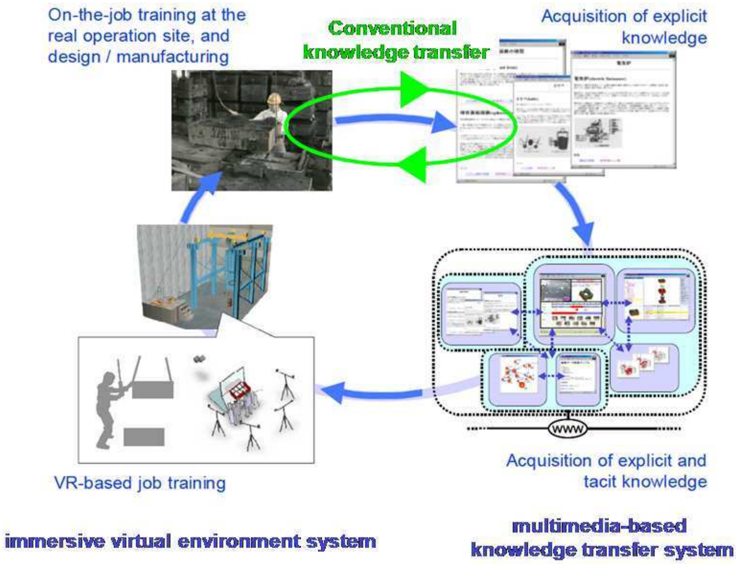


Figure 3.3 Process of knowledge acquisition and human resource development

close to repetitive on-site work. In this chapter, the annotation input and display system that can be used with the skills training system virtual training system is described. In this way annotation or similar can be shown in the space on a model displayed in the VR space, which promotes sharing of technical information, communication between engineer and skilled operative, and enables the acquisition of corporeal knowledge.

Figure 3.4 shows an outline of the immersive virtual environment system. This VR system combines a three-dimensional visualization device and a force feedback device, so a user can experience the hardness and weight of three-dimensional objects. By experiencing images displayed within the virtual environment, not only visually, but also by touch and by sensing force, the internalization of casting knowledge is promoted. The display of three-dimensional shapes in the three-dimensional visualization device and input control to the force feedback device are carried out in an integrated manner by a PC, so the three-dimensional images and the force sensing are synchronized. The position of the field of vision is fed back to the PC by a head tracking device fitted to the three-dimensional visualization glasses, and images are displayed in real time in accordance with the position of the field of vision. In the force feedback device, the position of the tip of the manipulator and the load on the manipulator are fed back to the PC, and the position and force of the manipulator are appropriately controlled taking into consideration the displayed three-dimensional shape. It is considered that by using this system to provide a visual and force sensing ba for the acquisition of skills, it is possible to partially overcome the problems associated with the acquisition of casting skills.

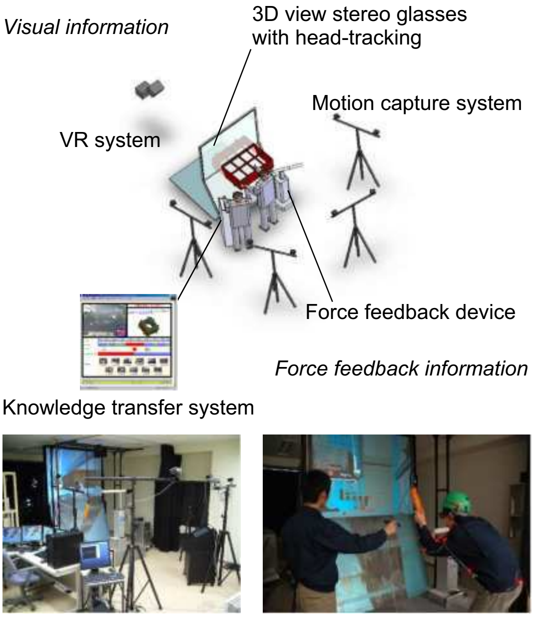


Figure 3.4 Immersive virtual environmental system

In the ramming process, foundry sand is rammed with a rammer while pouring the foundry sand into a metal form, so that the sand is evenly spread throughout. The knowledge to be acquired for this process is not only visual, but also the sensation of ramming with the rammer is important. The foundry sand frequently used today starts to harden soon after it is poured into the metal form because it contains hardening resin, so the operation must be quickly completed. Unskilled persons learn this work through OJT on site, but failure is normally not acceptable in product manufacture, so the opportunities for an inexperienced person to learn are limited. When the VR system is used, the sensation of the ramming process is provided by the three-dimensional visualization device and the force feedback device. Also, because a real product is not being manufactured, it is possible to repeat this experience any number of times. Also, by measuring in advance the work data of an experienced skilled person, this data can be taught using the force feedback device, so it is possible to directly experience the details of the work of a skilled person. By learning not just the visual information, but also data associated with the sensation of the force applied by the rammer in the ramming operation, it is possible to directly acquire tacit knowledge.

3.3.5 Skills Transfer and Human Resources Development Through Emotional Interaction in VR Space

Our proposed system is the framework for skills transfer and knowledge creation through communication between engineers and operatives in VR environment and the skills transfer system in accordance with this framework. A system has been developed which enables explicit knowledge from technical documents, *etc.*, to be efficiently acquired using multimedia, and tacit knowledge such as skills to be acquired through visual and force sensing experience of actual work on site using a combination of the three-dimensional visualization system using VR technology and force feedback device utilizing robotic technology. It is important to experience these with communication between several operatives mixed in. Our system is displayed on projection VR systems, which allow many people to simultaneously share the experience. Interaction between all the viewers present is generally very high. Most trainees in this experience are influenced by the experience, and immediately begin sharing their feelings with the others in this group. Also, most trainees can acquire tacit knowledge, embodied ***knowledge,*** and skills by sensory-motor and implicit emotional learning memory. An emotion is a mental and physiological state associated with a wide variety of feelings, thoughts, and behavior. This is based on the Latin *emovere*, where *e*- (variant of *ex*-) means “out” and *movere* means “move”. So, it is easy to acquire the skills transfer and human resources development through emotional interaction in VR space. In VR space an environment is created in which several engineers and operatives can participate, and while communicating cooperatively acquire design and manufacturing knowledge. In other words, a virtual environment is created in which effective OJT can be carried out with highly experienced skilled operatives. By doing this before the actual OJT, the skills transfer and human resources development can be carried out more effectively over a shorter time period.

3.3.6 Skills Transfer with an Annotation Display System

Figure 3.5 shows the structure of the annotation display system. The principle of this system is that a magnetic three-dimensional position sensor contained on the tip of the annotation display device carries out the writing operation, and in addition acquires positional information. At the same time the model for annotation is displayed within the VR space. By turning the display device switch on or off, it is possible to show text or graphics.

This annotation display system was applied to the mold assembly work. Tests were carried out on mold assembly work using VR with ten pairs of people, each pairing consisting of an experienced person (instructor) and an inexperienced person (operative). A questionnaire survey was carried out with the instructors regarding the ease of communication with and without the annotation display, and a questionnaire survey was carried out with the operatives regarding the ease of understanding with and without the annotation display. When the explanation times with and without annotation display are compared, no difference in explanation time was found. This is considered to be because the annotation display system did not take so much time to operate; it was easier to explain using a combination of words and annotation display than by words alone. For work training, explanation using words alone is difficult, and for pointing out a particular place or the like, it is considered to be more effective for the instructor to use the annotation display system. Also, from the results of the questionnaire regarding the ease

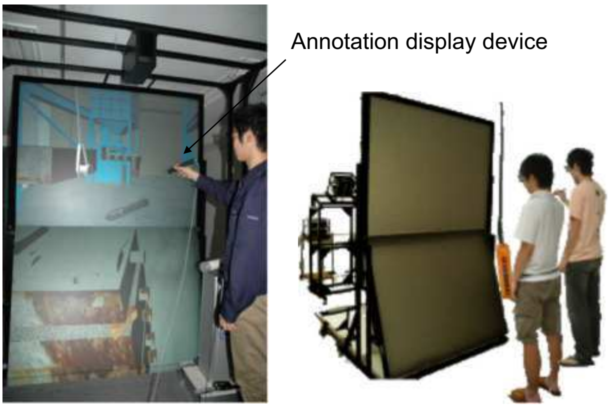


Figure 3.5 Skills transfer with an annotation display system

of communication with and without the annotation display system, all persons replied that explanation was easier using the system, demonstrating that the annotation display system is effective for communication.

3.3.7 Work Training by OJT and Sharing a Ba in a Virtual Training System

A new motion capture system was introduced into the virtual training system, and the new virtual training system for analyzing the bodily movements of the operative and analyzing the work are performed. In this system a high-accuracy motion capture system is added to virtual training system. The motion capture system has a resolution of 3600×3600 pixels, 480 Hz high-speed sampling, and LEDs each with their own ID, *etc.* It is capable of recording the motion history of the operative with high speed and high accuracy. Using our system and near-infrared spectroscopy, it is possible to study in detail the work motions and brain activities of skilled workers as shown in Figure 3.6.

Using the motion capture system, the motion histories and coordinating motions of operatives were recorded and studied. LEDs were fitted on various parts of the body of the operative, and the trajectory of each part was recorded. In the figure, one point each for the trajectory of the head, left hand, and right hand only are shown. Differences in the motion history of different operatives were observed. By determining the differences in motions of the operatives not only qualitatively, but also quantitatively, it is considered that in work training using VR it is possible to study the degree of proficiency of work or clearly see the differences from the work of a skilled operative in VR space.

We studied the results of a study of differences in the training effect for mold assembly training in VR space when visual information and force feedback information are linked, for the case of visual information only. Eight subjects whose knowledge of casting was poor carried out an operation to assemble a top mold on a bottom mold by operating the buttons on the crane operation panel in VR space. The time taken from start of assembly to completion and the positional error between the top mold and the bottom mold when the assembly was completed were measured. The positional error was found when the dowel in the top mold was placed into the dowel hole of the bottom mold when the eight subjects were provided with visual information only. Also, the positional errors using visual information and force feedback information were compared. It can be seen that the positioning accuracy is improved when force feedback information is provided compared with the case when using visual information only. We also studied the average number of times during the operation that the mold touched the core or dowel. In actual work, if the core or dowel is damaged, it is necessary to repair the damage, but here even in the case of contact, the work can be carried on to the end. As a result, it can be seen that for most subjects, the number of contacts with the dowel or core was less than when visual information was linked with the force feedback in-

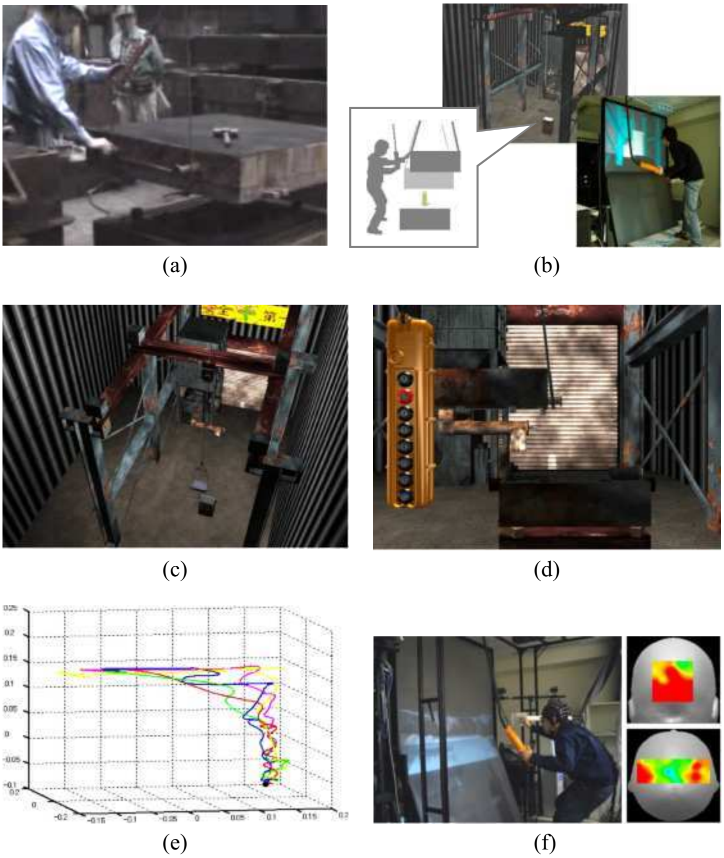


Figure 3.6 Mold assembly: (a) actual operation, (b) virtual operation, (c) and (d) virtual foundry, (e) result of virtual training, and (f) brain activity measurement

formation. The average amount of time required for each subject to complete the work was calculated. It can be seen that for all subjects, the operation time was shorter when visual information was linked with force feedback information.

The following is a description of the differences between work training by OJT and by the virtual training system. In work training using the VR system with high-accuracy motion capture, there were a number of differences compared with when using OJT on site. By using the two systems the operative and the skilled person can share a ba at a remote location, and work training can be carried out between the skilled person and the operative at the remote location. In this case, by capturing the motions of the work of the skilled person and displaying them superimposed on the same positions of the operative, the operative can stand in the viewpoint of the skilled person, and can trace the work motions. Compared with watching from a different viewpoint and then copying on site, more efficient work training is possible. Also, the display of annotation within the space, and display of associated information, images, *etc.*, within the space is possible, unlike in the real space. Therefore, in combination with motion capture, by displaying with an arrow the direction of movement of the arm or foot during work training, and displaying the amount of force applied using numerical information or image information, it is expected that work training can be made still more efficient.

3.4 Augmented Reality-based Training System

3.4.1 Augmented Reality

Augmented reality (AR) is the overlay of computer graphics onto the real world and has many potential applications in industrial and academic research [18]. One of the most difficult parts of developing an AR application is precisely calculating the location of the camera in real time so that the virtual image is exactly aligned with real world objects. The ARToolKit allows this to be done easily. The AR-ToolKit tracking works as shown in Figure 3.7.

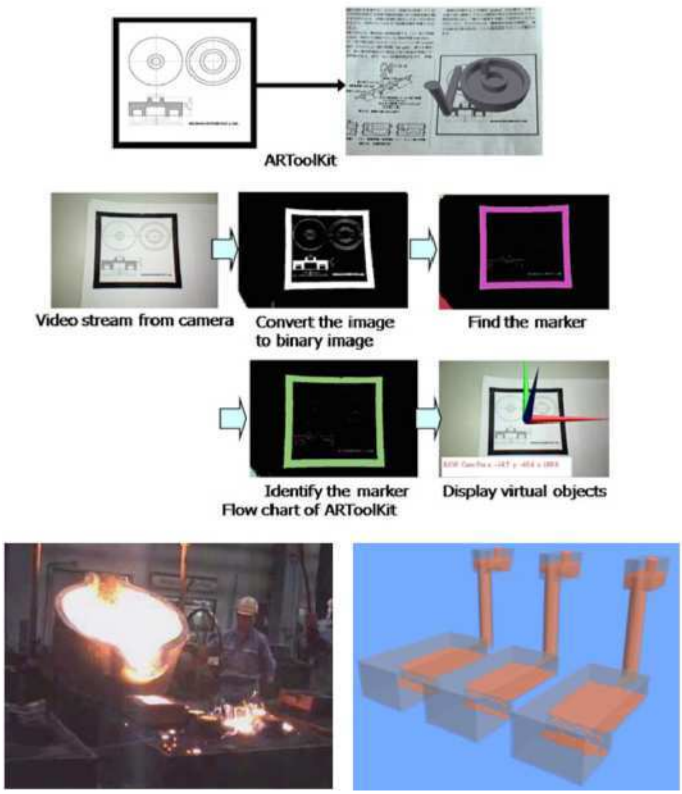


Figure 3.7 AR-based training system

ARToolKit is a software library for building AR applications. The ARToolKit video tracking libraries calculate the real camera position and orientation relative to physical markers in real time, enabling the development of a wide range of AR applications. Some of the features of the ARToolKit include single-camera position/orientation tracking, tracking code that uses simple black squares, the ability to use any square marker pattern, simple camera calibration code and sufficient speed for real-time AR applications. ARToolKit is a C/C++ language software library that allows programmers to easily develop AR applications. ARToolKit uses computer vision techniques to calculate the real camera position and orientation relative to marked cards, allowing the programmer to overlay virtual objects onto these cards. The fast precise tracking provided by ARToolKit is expected to enable the rapid development of many new and interesting AR applications.

3.4.2 Training System for Pouring

Casting is an operation whereby metal is shaped by pouring it in a liquid state into a mold, followed by solidification. Casting is also a method of detailing the metal as a result of pouring it into a mold.

Fluidity is the ability of metals to flow through a gating system to fill the cavity of a casting mold and conform to the shape of the mold. Fluidity is determined by the solidification interval, viscosity, and surface tension. Metal is cooled by heat transfer upon contact with the mold and the viscosity consequently increases. The metal solidifies upon further cooling, resulting in a misrun. To avoid misrun, engineers must pour the metal quickly. However, since the mold is made from sand, pouring too quickly can cause breakage or distort the shape of the mold. Therefore, it is critical to pour the metal at an appropriate rate. Engineers must consider both the shape of the mold and the metal's fluidity to determine the proper rate of pouring. However, this is a difficult task, especially for beginners. Therefore, engineers must be trained to pour metal properly in addition to studying its fluidity.

Figure 3.7 shows the AR-based training system for pouring. We built the system using the ARToolKit. The system displays fluidity that is simulated by simple physical models in real time.

The ARToolKit tracking works as follows. (1) The camera captures video of the real world and sends it to the computer. (2) The software on the computer searches through each video frame for any square shapes. (3) If a square is found, the software calculates the position of the camera relative to the black square. (4) Once the position of the camera is known, a computer graphics model is drawn in the same position. Thus, users can see virtual objects from every angle by moving a camera. In the proposed system, we display a mold and ladle on the marker and users operate a keyboard to rotate the ladle and pour metal.

The proposed system displays the fluidity in the mold for training purposes. It is important to display fluidity in real time, while considerable detail is not necessary, since the system is not for fluid analysis. The proposed system can also display the results of the fluid analysis that is preliminarily calculated in considerable detail if requested by the users. The proposed simulation, which uses simple physical laws, is considered to be useful for training.

3.5 Work Support for Elderly or Female Workers

Here, new experiments in lifestyle support for the elderly and work support for workers, *etc.*, using VR technology are described.

3.5.1 Aging Society and Labor Problems

The total population of Japan as of October 1, 2004 was 127.69 million, with 24.88 million people aged 65 years or older, or 19.5% of the total population. The elderly population will increase rapidly up to 2020, and in conjunction with the reduction in total population the percentage of elderly is predicted to reach 26.0% in 2015, and 35.7% in 2050. The working population aged 65 years or older in 2004 was 4.9 million, or 7.4% of the total population.

Currently in Japan, a wide range of measures for the aging society, such as measures for employment and income, health and welfare, learning and social participation, lifestyle environment, and survey research, are being taken, and the number is steadily increasing. In order to further maintain and expand the vitality of Japan, it is important to achieve a society in which everyone from young people to old people, in particular those responsible for the aging society, can use their capability and experience, not only at home but also to be active in society.

In the “Questionnaire survey of the promotion of participation of the elderly in society” prepared by the Japanese government, it was pointed out that it is necessary to create a system in which all workers, including the elderly, can work in accordance with their willingness and capability. It is necessary for problems such as wages and conditions, forms of work, retirement system, and so on, to be solved, and there is a strong need to support improvements in worker education and working capability.

For work support, it is very important to share a ba in a specific time and space. Specifically, it is important to carry out work with an instructor in an actual workplace. By sharing a ba in this way, not only can knowledge acquired through the five senses be obtained, but also valuable “information not obtainable unless in that ba” and corporeal knowledge can be obtained. Sharing a ba in this way can be said to be very beneficial for work support. Work support by sharing a ba in such a way places a large burden on the aptitude of the personnel carrying out the work guidance, and a lot of time is required. This has the problem that it is difficult to carry out work support under the present circumstances in which there are time, spatial, and financial restrictions.

Using VR technology in which visual information is combined with force feedback information, and where both are provided at the same time, these problems can be partially overcome. It is possible to promote the sharing of a mock virtual ba, and communication from person to person can be exchanged, so it is considered that this is very beneficial for work support.

3.5.2 Work Support for the Work of the Elderly or Female Workers

Based on the rapid spread of the effect of the low birth rate and aging in society, in June 2004 the law concerning stabilization of employment of elderly persons was partially revised with the aim of creating an employment environment in which the knowledge and experience of the elderly can be utilized. Also, female workers play a large role in the development of the economic society, and it is important in the future to provide an environment in which female workers can use their full capability. By using the VR system as described, it is possible to create a work environment that takes into account elderly and female workers, and use it for work training in that environment. Assuming various work environments and work contents, as shown in Figure 3.8, a work environment suitable for all workers, including the elderly and females, can be studied while providing mock experience within the environment, and work training can be supported.

Figure 3.8 Work support considering universal design

3.5.3 Construction of a Teleworking Environment as a Work Form with Variety

In recent years, attention has been focusing on teleworking (work at home, work at satellite offices, *etc.*) as a flexible form of work, unconstrained by place and time. For workers that need to care for the elderly, for children, or other nursing care, teleworking is a useful form of work that reduces the burden of commuting, and has few bodily or time restrictions. In the past, telework was mainly applicable only to programmers or desk workers, but by using a fusion of the three-dimensional visualization device, force and touch sensing, and the display device developed by the authors, the manufacture of shaped products, product development, and other creative work based on visual information and force and touch sensing information is becoming possible.

3.6 Conclusions

This chapter has described a new VR/AR-based skills transfer and human resources training taking casting as an example where high-level skills are required. Also, new experiments in work support for the elderly and females using VR technology have been described. In order to efficiently and reliably internalize work in manufacturing workplaces, effective skills transfer and human resources development can be carried out not just by knowledge alone, but by adding visual information, force feedback information, and emotional communication.

References

1. Watanuki K., Kojima K.: Knowledge acquisition and job training for advanced technical skills using immersive virtual environment. J. Adv. Mech. Des. Sys. Manuf. **1**(1):48–57 (2007)
2. Watanuki K.: Virtual reality-based job training and human resource development for foundry skilled workers. Int. J. Cast Metals Res. **21**(1–4):275–280 (2008)
3. Matsuo M., Easterby-Smith M.: Knowledge Sharing Dilemma: Knowledge and Knowing in Japanese Firms. Academy of Management Annual Meeting (2004)
4. Polanyi M.: The Tacit Dimension. Routledge & Kegan Paul, New York, pp. 3–25 (1966)
5. Nonaka I.: The knowledge-creating company. Harvard Bus. Rev. 96–104 (1991)
6. Nonaka I., Konno N.: The concept of ba: building a foundation for knowledge creation. California Management Rev. **40**(3):40–54 (1998)
7. Itami H.: Management of Ba. NTT Publishing, Tokyo (1999)
8. Nonaka I., Toyama R., Konno N.: SECI, ba and leadership: a unified model of dynamic knowledge creation. Long Range Planning **33**:5–34 (2000)
9. Shewhart W.A.: Statistical Method from the Viewpoint of Quality Control. Dover, New York (1939)
10. Deming W.E.: Out of the Crisis. MIT Center for Advanced Engineering Study (1986)
11. Garfinkel H.: Studies in Ethnomethdology. Prentice-Hall, Englewood Cliffs, NJ (1967)
12. Heritage J.: Garfinkel and Ethnomethodology. Polity Press, Cambridge (1984)
13. Ericsson K.A., Krampe R.T., Tesch-Roemer C.: The role of deliberate practice in the acquisition of expert performance. Psychol. Rev. **100**:363–406 (1993)
14. Burton R.R., Brown J.S., Fischer G.: Skiing as a Model of Instruction, Everyday Cognition: Its Development in Social Context. Harvard University Press, Harvard, MA (1984)
15. Dreyfus H.L., Dreyfus S.E.: Mind over Machine. Free Press, New York (1986)
16. Gibson J.J.: The Ecological Approach to Visual Perception. Houghton Mifflin, London (1979)
17. Saeki Y., Sasaki M.: Active Mind: Human Cognition in Action. University of Tokyo Press, Tokyo, pp. 25–54 (1990)
18. Kato H.: ARToolKit library for vision-based augmented reality. Technical Report of IEICE. PRMU, **101**(652):79–86 (2002)

Chapter 4
Exploitation of Designers and Customers' Skills and Creativity in Product Design and Engineering

Monica Bordegoni[1]

Abstract Nowadays, increasing importance is devoted to "product experience", which induces customers to choose a particular product over another. A shift of the current product development process is necessary to design, validate, and produce several variants of a product, in order to better satisfy customers' needs, preferences, and skills. This is possible if customers become a central part of the product development process and are more directly involved in the design and evaluation of new products, through the use of more natural and easy-to-use tools. Their involvement would lead to better products, which were more effective and in line with customers' aesthetic and functional expectations, and would contribute to the acceptance and success of these products.

4.1 Introduction

Functionality has always been considered an essential precondition for both product satisfaction and market success. Nowadays, increasing importance is devoted to "product experience", which induces customers to choose a particular product over another. Researchers from many disciplines, including industrial design and engineering, ergonomics, and psychology, as well as industries operating in various sectors, are responding to this need to incorporate users' emotional needs into product design.

Customers' needs, preferences, desires, and tastes are all aspects that are typically considered during the specification of products. Customers' emotions can be considered as an additional aspect that can influence product design. Unlike the other aspects, emotions are more complex to detect, since they are not explicitly

[1] M. Bordegoni
Dipartimento di Meccanica, Politecnico di Milano, Via La Masa, 1, 20156 Milano, Italy
e-mail: monica.bordegoni@polimi.it

expressed by customers, but may be important since they relate to the reasons for choosing a product, to instinctive choices of products, and to the establishment of effective relationships with products.

One issue is to understand how it is possible to capture the emotional needs of customers and incorporate them into the design process. For this reason, research on the methods and tools necessary to consider emotion in design is important in order to bring users closer to design and manufacturing. In this context, *emotional engineering* is a growing discipline, addressing the technology to design goods which appeal to emotion and sensibility by translating human sensibility and images into physical design factors. Emotional engineering is related to kansei engineering [1], which is a method invented by Nagamachi in 1970, and is now addressed by various researchers and conferences [2].

In addition, in the market analysis undertaken at the start of the product development process, customers are considered in terms of average needs, desires, tastes, *etc.* Actually, customers are individuals, each with their own skills and preferences due to the fact that they have different life experiences and different cultural backgrounds. For these reasons, the market should propose variations on products belonging to the same family, in order to offer products that best meet customers' needs and expectations, and are therefore successful.

This implies that design and engineering activities should be supported by tools that allow the design and verification of several variants of new products at low cost and in a short time. Specifically, while tools for engineering meet engineers' expectations, tools for industrial designers are still too technical. The translation of designers' creativity and emotional factors into a product shape is constrained by the overly technical interfaces of current computer-aided styling tools. Consequently, we need to develop tools that are more natural and in line with designers' typical ways of working. In addition, we also need to find design paradigms that involve the customers with the product design process. Their involvement would ensure that the products will be suitable for their intended purpose in the context in which they will be used.

This requires a paradigm shift of the product development process, where new tools for design and testing support a closer collaboration between designers and customers, and a allow a wider involvement of customers in the design process. In order to do that, we start by analyzing the current product development process, concentrating on the conceptual design and aesthetic aspects of products, and identifying the limits of the tools for design and testing and pointing out when and how customers are involved in the development process. Then, we define a new paradigm for product development, which is based on easy-to-use tools for design and user-centered design paradigms that are suitable for "common users", including customers. Customers will be able to co-design and evaluate their products together with designers, and ultimately, the most creative and capable customers, will also be able to directly design their products (Figure 4.1).

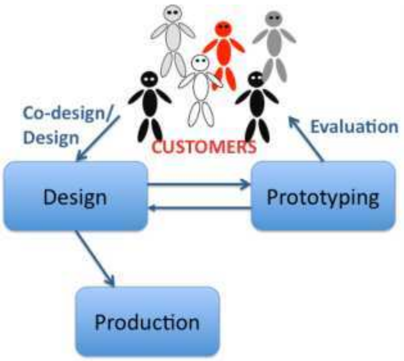


Figure 4.1 Role of customers in the product development process of the future

4.2 One Product, many Customers

Cars, tee-shirts, canned beverages, mobile phones … what do they have in common? They are objects which are widely distributed and commercially available, which are usually anonymously and uniformly manufactured in massive quantities. We generally call them *products*.

Products have forms and functions. Products are characterized by two aspects: products have a form, and products have a function. Therefore, two issues are important when thinking about a new product: the aesthetic aspects of the product, and the purpose the product is designed for, *i. e.*, the tasks it has to accomplish.

Products have customers. Products are made to be sold on the marker; they are made for several kinds of customers, who are the potential buyers. The way that customers choose products is variable: it could be by brand, by product attributes (function, color, material, shape), by use occasion, or simply by emotional response. Products are successful if sold: customers' appraisal is crucial.

Customers are individuals. Products are made for customers who have different know-how, skills, preferences, emotional reactions, and habits. It is well known that people from the East have tastes that are completely different from those of western people. Some of these characteristics depend on the customer's cultural background, and some on his or her experience:

- a person's skills, emotional reactions, habits relate to their cultural background;
- a person's know-how and preferences, and also skills and emotional reactions, relate to their experiences.

According to these differences, customers have different needs, tastes, and expectations with respect to products. For example, some customers are technologically oriented and therefore prefer and are more keen on using products which include more functions; people respond subjectively and differently to the aesthetics of an object – an object that appears ugly to one person may be considered cool and beautiful by another.

One product has several customers. Most of the time the relationship between products and customers is *one to n*, where a single product is required to satisfy the needs and expectations of several customers. In other cases, products can be customized by customers, minimally or significantly, in accordance to options that have been planned by the producers. There are products on the market today where customers can customize what they intend to buy – for example, cars and shoes.

Products have a context of use. Products are used within a context. The context is related to environment, situation, working conditions, social and cultural issues, and local needs. The context may be very variable and diversified depending on individual customers. The context of use may affect the use and the perception of the product, but unfortunately it can be very challenging for producers to consider during development all the possible contexts where the customers might use the product.

The relationships between product, customers, and context of use are summarized in Figure 4.2.

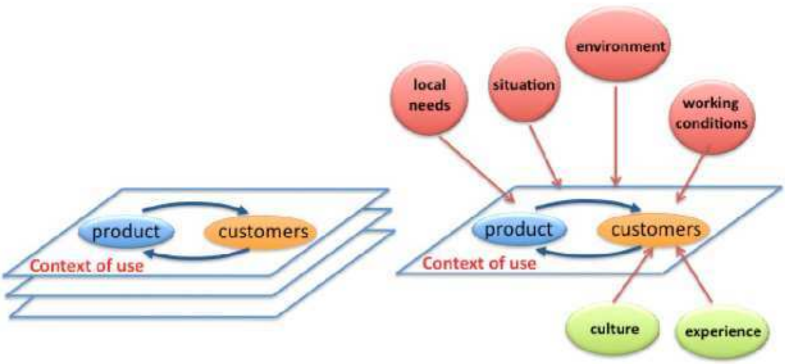


Figure 4.2 Relationship between a product and its customers distinguished by culture and experience within the context of use

4.3 The Product Development Process

Products characteristics are the results of the process of product design, which is the process according to which products are conceived and specified [3]. The *product development process* (PDP) is the full set of activities necessary for transforming an idea, a concept of a product that a designer has, into a product manufactured and distributed into the market.

The development of a product starts with a *market analysis*, where marketing people examine and analyze the profiles of potential customers' through interviews and focus groups, and finally define the needs and desires of what we call the "average customer". The average customer is a sort of stereotype, *i. e.*, a standardized description of a group of people that includes their average needs, preferences, *etc.* The analysis made by the marketing people is interpreted by designers in order to generate the product concept.

In the subsequent *design phase* the product is conceived and specified in its general shape and function. Product design is the phase where concepts and specifications are created and developed, and where the function, performance, value, and appearance of products are optimized [4]. Industrial designers study both function and form, and the relationships, interactions, and connections between the product and the user. In general, industrial designers deal with user research, comparative product research, initial sketching of the product shape, model making, and also the initial prototyping and testing of the product. All these activities aim at describing the overall architecture of the product, including some preliminary hypotheses about shape and functions. During the design activities, the product concepts are validated by the marketing people, tested through panels of users that represent the average customer, and finally passed as technical requirements and specifications to engineers.

In the *engineering phase* the product is specified in detail – dimensions, tolerances, materials, manufacturing process, costs, *etc.* – in order to produce it. Engineers are used to their problems being described in the form of clear and possibly quantified specifications and objectives, so they know when they have succeeded in specifying the product. Therefore, we can say that engineers develop their detailed design process well bounded by constraints. The activities that engineers mostly deal with are problem solving, rules of thumb, numerical tables and equations, books of regulations and companies' standard rules, comparative performances, and costs evaluation. Engineers produce a set of analyzed alternatives consistent with the technical specifications, which have to be evaluated by designers, and this happens in a closed loop. Then, the accepted alternatives are updated and finally fully detailed.

At the end of this process, the final solution accepted by designers and engineers, and also approved by the marketing people, consists of a unique solution that can be expanded as a family of variants of the selected solutions to satisfy, in the largest sense, the community of average customers. Therefore, the family of product variants is presented to the market and should satisfy, if successful, the needs and desires of the target market.

Figure 4.3 shows the three actors – marketing, designers, and engineers – taking part in the PDP, and their tasks and relationships.

In conclusion, we remark the fact that this approach is characterized by the initial synthesis of the multiple and different needs and desires of potential customers into a set of average needs and desires that imply the profile of an average customer. Then, designers and engineers work on the product, and the customers are only involved in the product testing phases.

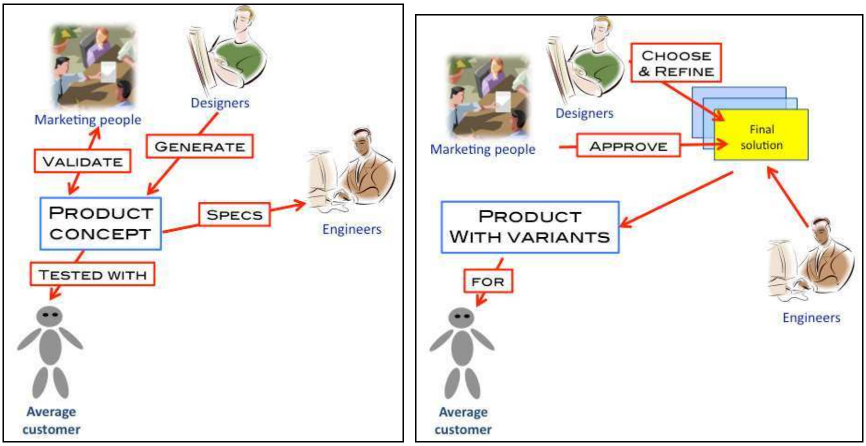


Figure 4.3 Product development process activities performed by marketing, designers, and engineers

4.3.1 Designers and Engineers

Industrial design and industrial engineering are two separate processes that work together in the development of a product, focusing and addressing different aspects of the product – the first is more aesthetic, the second, more technical – but both addressing from complementary points of view the product's functionalities [5]. There is a major difference between the design process in the engineering domain and in the aesthetic domain, which we represent through a spiral converging to a specific solution in the engineering domain (Figure 4.4 (a)), and a spiral that is evolving and moving forward from the initial idea in the aesthetic domain (Figure 4.4 (b)).

Both processes have a design space in which the designer moves with the aim of finding a solution to the given problem. The *engineering domain* is characterized by a well-defined solution space accompanied by metrics that allow us to measure the quality of the solution. We know the starting point and how far our

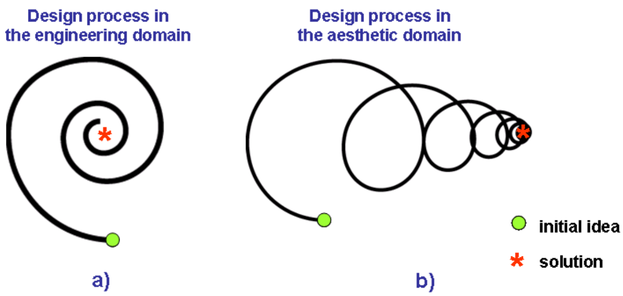


Figure 4.4 The design process: (a) in the engineering domain, and (b) in the aesthetic domain

solution has gone from this point. This can be quantitatively measured by considering aspects like product performance, and by verifying technically measurable features like structural rigidity, weight, efficiency, *etc.*

The *aesthetic domain* is definitely fuzzier in its boundaries and its objectives, since characteristics like elegance and beauty cannot be precisely and universally defined and perceived. Unlike the engineering domain, it is not possible to define *a priori* the final target of the solution from the initial idea, and it is also difficult to say how far we have moved from the starting point and how far we are from an ideal target. It is during the evolution of the initial rough solution that designers understand more clearly where they are going and how far they are from the initial point, and in which directions they could and could not go. Furthermore, the quality of the solution is not easily measurable; in fact, the evaluation of the result is very much subjective and related to human perception.

Let us analyze in more details the design and engineering process. The process starts from an initial product idea or concept and moves on to the representation of this idea through a modeling phase (Figure 4.5). This is followed by a verification and evaluation phase where prototypes are used to demonstrate, prove, and validate aspects of the product design. There are many levels of prototypes aimed at visualizing the product, representing its shape, its functional aspects, and how it works, *etc.* Therefore, prototypes are initially only visual, and then built physically. This is the point at which customers are usually involved.

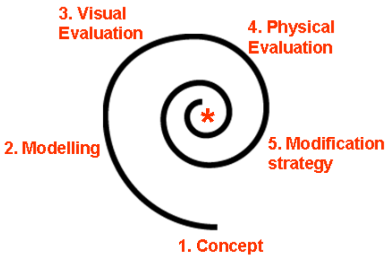


Figure 4.5 Main phases of the design and engineering process of a new product

The physical prototype (also called the physical mock-up, or PMU) necessary for the physical evaluation cannot be immediately derived from the digital model, but instead requires a production phase that can be manual or consist of a technological process including CAM (computer-aided manufacturing), milling, and finishing activities, or through rapid prototyping. In many industries, the products are quite complex and require several iterations of design, prototyping, and testing. Every time the design team considers that a PMU is necessary at that design stage, the design process breaks up in order to physically build the prototype. This activity requires long time that is at least an order of magnitude higher than the time required by the other product development phases. So, it often happens that in order to meet time constraints the number of PMUs is reduced and fewer solutions can be evaluated – to the disadvantage of product quality.

The problem faced by companies is to find a trade-off between obtaining the best result among the various identified solutions in the shortest time. As reported above, in the design process most time is dedicated to building PMUs. Ideally, the number of PMUs should be reduced while maintaining a high level of quality of the product. Since PMUs are important to the aim of assessing the product design, one of the interests of the research is in studying new design tools that allow a reduction in the number of PMUs, whilst maintaining and possibly increasing product quality, and at the same time exploiting the skill of their users, *i. e.*, the designers.

4.3.2 Tools for Design and Engineering

Turning now to tools that support the product design and engineering process, we note that there are today many computer-aided tools aimed at designers and engineers, and these generally speaking are software programs that assist in the design of products.

These design and engineering tools have been conceived and developed by computer scientists, and are typically made for a specific category of users. It is very typical that computer scientists are not experts in the problems addressed by the applications; therefore they are obliged to define the specifications of the software applications on the basis of an analysis of users' expectations and requirements, and then translate these into application functionalities. Due to this intermediation, the resulting applications may not be fully in line with users' expectations and modus operandi. In addition, most software applications are not easy to customize in order to meet a particular user's preferences.

Tools for engineers are technical tools oriented to technical users that allow the detailed definition of products (including dimensions, tolerances, manufacturing features, *etc.*). In general, engineers are satisfied with the current evolution of these tools, which adequately satisfy their needs.

Conversely, *tools for designers* are technical tools developed for creative users. From surveys of designers it comes out that they consider these tools too technical and too complex, and that they often interrupt their creative flow. This is also proved by the fact that designers often prefer using "non-digital" tools, *i. e.*, physical prototypes, sketches on paper.

If we consider the major features that distinguish the kinds of users of design tools, we notice that they are distinct and have several differences. In fact, designers and engineers use different methods, have different world views, have different histories, use different vocabularies, sometimes have differing prejudices – and they even have different demeanors and dress styles! For what concerns our research interest, we have observed in the previous section that they both have completely different approaches and techniques to design. In these diverse contexts, the role of technology has to be different.

Regarding the tools for industrial designers, we have said that they still are the more problematic. In fact, computer-aided styling (CAS) tools are based on interaction paradigms that are not in line with designers' skills, and in general they are barriers limiting the designers' creativity. The level of interaction of these tools is low, and consequently the capability of exploiting the designers' creativity is rather limited.

We need to develop tools that are more oriented to designers and more in line with their profile, which can be summarized as follows:

Designers are creative, non-technical people, who often use their hands and manual skills in their daily work.

If the design tools for creative people could be based on more designer-oriented, user friendly, and natural interaction modalities the designers could get the following benefits. They could:

- create more product variants in less time, which can be shown to customers;
- quickly modify proposed concepts of new products in order to better meet the customers' expectations;
- be more *emotionally involved* during the design process because they are less constrained by technological factors, and thus better able to express their creativity.

4.3.3 Methods and Tools for Product Testing

We have seen that the design and engineering loop may be a long process leading to the final product through the development and testing of various intermediate prototypes, some of which are subject to testing and evaluation by customers. The tests are performed with a sample of customers, appropriately selected for the purposes of the product analysis. The test results are used to feed back to the design and engineering activities.

In recent years companies have tried to optimize their product development process in order to reduce the overall time and costs. A major impact in terms of costs and time is related to the physical prototyping, as already discussed in Section 4.3.1. Companies are approaching the problem in two ways: by reducing the number of physical prototypes and by replacing physical prototypes with virtual ones. The first approach has the disadvantage that fewer solutions and variants can be tested, with a possible loss of final product quality. The second is potentially more beneficial, but requires research and study about the degree of realism of the testing performed through virtual prototypes that we can reach today using the available technologies, and about the consequent acceptance by customers of virtual testing as compared to real testing. The use of virtual prototyping for product testing during the process does not imply that physical prototypes are fully removed from the process, but rather that fewer prototypes, and ideally a final one, are built instead.

A final consideration concerns the testing protocols. In fact, during the testing sessions with customers, companies are already used to following well-defined protocols that have some important open issues:

1. how to effectively measure the customers' satisfaction and evoked emotions;
2. to what extent can the customers require changes in the product form or function that can be shortly, or even immediately, tested again?

4.3.4 Remarks About the Product Development Process

To conclude this section, we list some remarks that may be considered to be crucial points for future research and development of theories, methods, and tools for product development.

Product customization. The first issue concerns the possibility for customers to customize the products they intend to buy. This requires a paradigm shift in the product development process whereby customers are considered as individuals and not as average people, and activities concerning the design, engineering, and also manufacturing of products should be flexible enough to accept new and continuous specifications, also coming from customers, and consequently managing the following design and production activities.

Easy design tools for designers and also for customers. Tools for product design have to be more intuitive and designer-oriented so that, first, the creativity of designers is fully supported and exploited and their creative flow is not interrupted by the technicalities of the tools and technology, and products may be designed with more emotional involvement from designers; second, there is no need to build all the physical prototypes that are necessary today; and, finally, if tools are easy enough to use, we can also predict that in the near future customers may also be involved in the product conceptual process, and ultimately decide to be – to a certain extent – the designers of their products.

Virtual testing tools of new products. Testing of new products is a crucial aspect of the development process since the success of a product mainly depends on customers' acceptance. Major involvement of customers in the testing phases would reduce the risk of not meeting customers' requirements and not satisfying their needs and desires. More effective and less expensive tests may be achieved through the testing of virtual prototypes, which are flexible, versatile, and easily modifiable and configurable. In addition, we need to find more effective ways of measuring customers' satisfaction and emotions when faced with prototypes of new products.

4.4 A Paradigm Shift in the Product Development Process

As extensively discussed in the previous sections, products are made for average customers. In the context of mass-customized products, it would be more successful for companies and effective for customers if it was possible to make different products for different customers to use in different contexts and situations.

If we consider the variability of products versus the levels of various aspects related to products, such as acceptance, usability, comfort, emotion, and affection, we can easily see that these levels are quite low when the product variability is limited (Figure 4.6). This is the current situation. When the product variability increases, we would expect that the aspects more related to the various perception levels of the products would also increase. In fact, the variability of products leads to a better satisfaction of the expectations of a larger group of customers.

If today a company is producing a product for the average customer, we can expect in the future that a family of products will be made for many customers.

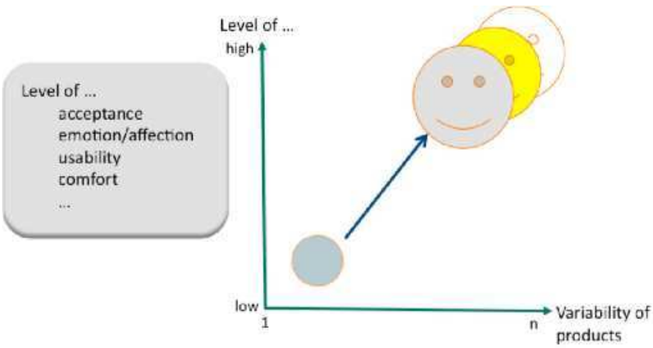


Figure 4.6 Variability of products versus levels of acceptance, usability, comfort, *etc.*

4.4.1 Product Customization and Participatory Design

Customization and personalization of products is one way of allowing customers to have products that closer satisfy their requirements, needs, and preferences [6]. A successful example is found in the automotive sector, where cars can be personalized by selecting interiors, wheel trims, *etc.*; another example is the shoes sector, where the colors and materials of the collar, sole, and shoelaces can be chosen from a palette [7].

In all these cases, users select from a variety of options to custom tailor their products. Form a technical point of view we talk about *product configuration*, where a standard product is customized by individual users to meet their specific needs and preferences. Each user can configure his or her own product by choosing from a pre-defined set of variants; the configuration tool uses a combinatorial approach that allows users to select one of the possible configurations. The manufacturability of each customized product is guaranteed: the product is assembled-to-order, made-to-order, engineered-to-order, but not designed-to-order.

An alternative and more powerful approach relates to user-centered design (UCD) and participatory design. UCD is a broad term to describe design process where users influence how a design takes shape. It was introduced in the 1980s by Donald Norman [8, 9]. Involving users in product design leads to better products, which are more effective and efficient, and contributes to their acceptance and success. Similarly, in participatory design the users are involved in the development of products [10].

The level of involvement of customers in the design process may vary: they can be co-designers working side by side with designers, or can be designers themselves. In the first case, we recognize that customers are not able to understand the language of designers; therefore, prototypes, mock-ups, and story boarding [11, 12] have been used to fill this gap. In the second case, the design is directly activated by the needs of the customer, who can define the specifications of the product he or she wants. Since customers are not designers or engineers, it is necessary to have appropriate design and engineering tools that are oriented to novice users. In order to support product design, the tools also have to encapsulate the knowledge and rules able to drive correctly the "why" and "how" of the decision process of design [13].

Because the product eventually has to be produced, it is necessary that the manufacturing is flexible so as to produce products not fully foreseen and planned in advance: rapid manufacturing may provide the technological solution to this problem in the future.

4.4.2 PDP Scenario Based on User-centered Design

In a new scenario where customers can design new products on the basis of a user-centered design approach, the role of marketing, designers, and engineers is different from the typical current ones that we described in Section 4.3.

Marketing people analyze, designers interpret, and engineers predict the customer's needs and desires, and identify a space/niche for potential products. Subsequently, marketing people define the price, market position, and potential market, designers generate concepts at large, and engineers define and identify technologies, constraints, and potential costs. They collectively build a conceptual view of responses to needs that is then validated through market testing.

The goal is to design and develop an easy system for customers in order to give them the possibility of "designing their own product". The system should be built around a specific "product archetype" supported by a "knowledge-based framework" (KB) and with a very easy and high-level interface that allows the customer to define and customize his or her own product (Figure 4.7).

The product archetype and the interface will be defined and generated by designers and engineers and the KB framework will be generated and populated by engineers.

Two major issues have to be faced in order to generate such a system:

- formalize the design and manufacturing knowledge and build the KB-driven framework;
- develop an easy-to-use interface for product design that is tailored to the customers' skills.

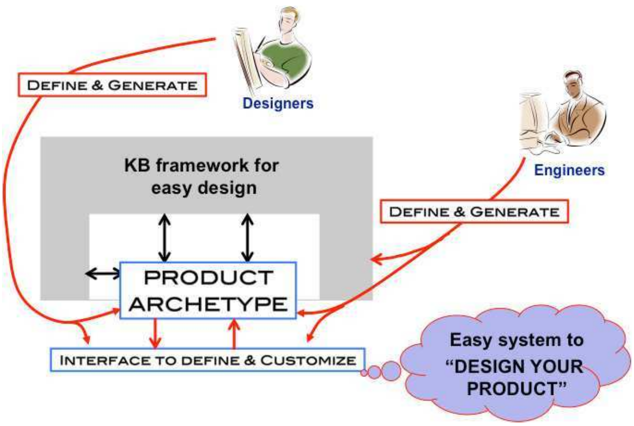


Figure 4.7 KB framework based on easy-to-use interface for product design

4.5 Design Tools for Product Designers and Customers

New design paradigms contribute to the development of *new technologies* and *new interaction modalities* more oriented to "common" users. Easy-to-use human–computer interfaces are based on multiple modalities (vision, sound, sense of touch) and perception–action interaction.

4.5.1 Enactive Interfaces

Enactive interfaces are a type of human–computer interfaces that are more natural and suitable to designers and common users. They are based on the concept of *enactive knowledge*. The concept of enactive knowledge is present in several application domains [14]. Enactive knowledge refers to the information that the user gains through *perception–action* interaction in the environment [15]. By doing things and performing actions, users learn how to perform and increase their knowledge about the environment and the way of doing things.

Enactive knowledge has several properties and features. It is *natural* in terms of:

- learning process, where users learn by doing; and
- the way it is applied in the environment, which happens through direct actions.

For knowledge as a general concept, enactive knowledge is:

- *physical* – it is related to physical interaction with objects;
- *personal* – it is very much dependant on the user's personality and habits;
- *experiential* – it relates to the way of doing, to abilities and skills, and depends on the user's experience;
- *cultural* – the way of doing and acting most of the time is related to social aspects, attitudes, values, practices, and legacy;
- *contextual* – both actions and perceptions depend on the context in which the user operates.

Enactive knowledge is inherently *multimodal* since it requires the coordination of the various senses that are involved in the perception–action interaction. Actions may be performed in various ways, including gestures, gaze, voice, and speech. The perception of the environment may happen through various and multiple modalities, such as vision, sense of touch, hearing, sense of smell.

Typically, enactive knowledge is related to manual tasks – humans' own knowledge on concrete actions in their hands, which are used for grasping, manipulation, handling, and shaping. An example of enactive knowledge and manual tasks is related to the field of industrial design, where a model maker uses his or her own knowledge on manually sculpting clay to naturally carve a piece of raw material into the pleasant object they have in their mind [5]. The carving action is natural for them, as they are used to using their hands to mold objects into shapes. Their ability depends on experience, and their creativity depends on personal and cultural issues. We can say that the evolution of the shape is driven by the perception of the various stages of the shape that the modeler gains through vision and through the sense of touch.

Enactive interfaces, defined as human–computer interfaces based on enactive knowledge, are capable of conveying and interpreting a user's input inserted in the forms of manual actions. The system interprets the input and provides an adequate response in perceptual terms. Responses may be provided in the form of images, haptic (kinesthetic and tactile) feedback, and sounds. Enactive interfaces are driven by the manual actions of users, and the users' actions are driven by the systems' responses perceived by the users. For example, the enactive approach is typical of assembly activities, where the users' actions are driven by the feedback obtained from the collisions of the parts.

Action for perception is an active form and involves active exploration of the world through touch. Through actively exploring the environment it is possible to perceive an object's shape, texture, hardness, temperature, size, and weight [16]. Unlike vision where the object can be perceived almost instantly, touch necessitates exploration over a period of time [17].

The users' perception is given by the integration of the various models perceived, and depends on the information included in each representation modality and in its coherence. In addition, the use of multiple modalities for representing information would allow us to overcome problems related to the perception threshold typical of some senses, or the limitations of some technological implementations (for example, haptic devices).

Enactive interfaces are characterized by a closed loop between the natural gestures of the user and the perceptual modalities activated. This improves the usability of systems since the closed loop decreases the humans' cognitive load traditionally required for interpreting and understanding the symbolic and iconic information provided by a computer-based system, thus allowing the user to quickly decide on the action to perform.

Figure 4.8 shows an interaction model of enactive interfaces based on the action–perception paradigm where the various modalities – vision, sound, and sense of touch – are integrated [15]. The mathematical model of an object is represented through different modalities, and therefore different models are generated, each dedicated to a particular communication channel: geometrical models dedicated to the visual output channel are converted by the human–computer interface (HCI) into images, and haptic models dedicated to the haptic output channel are converted by HCI into force feedback. The various models may possibly vary the fidelity and accuracy with which they represent the mathematical model. The kind of communication related to image and sound is passive, since the user receives information without specifically requiring it. Haptic perception is different, since the system exerts a force reaction only after a user's input. The perception is reactive in this case.

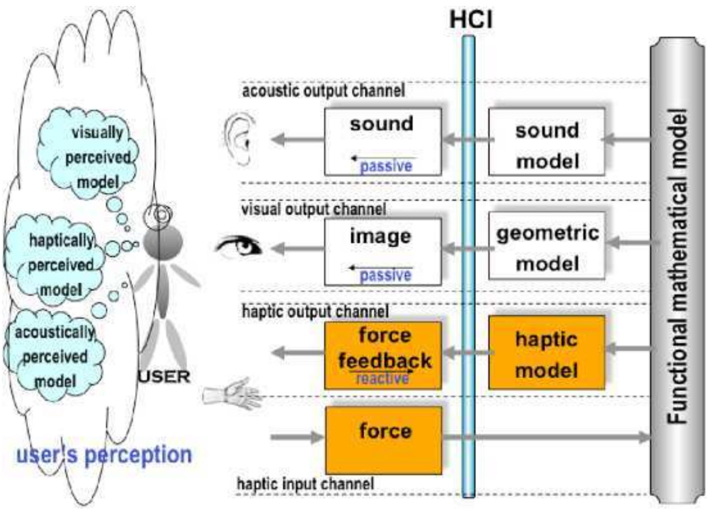


Figure 4.8 Action–perception model of enactive interfaces

Bruner [18] states that the traditional interaction where the information is mediated by a computer is mostly based on symbolic or iconic knowledge, and not on Enactive knowledge. In fact, information is presented in the form of words and mathematical symbols that have a conventional meaning (symbolic knowledge) or in the form of visual images that have a cultural significance (iconic knowledge).

Human–computer interfaces based on this type of knowledge are quite poor, limited and do not allow a full exploitation of users' skills. These interfaces are predictable, deterministic, and deductive, but leave little room for creativity.

Human–computer interfaces based on enactive knowledge, named *enactive interfaces*, would be more natural and effective ways of interacting with a system. The physical interaction supported by enactive interfaces generates emotions and can suggest to the users new ideas in a non-deductive way, supporting reasoning that starts from the particular and evolves towards more general solutions. The development of such interfaces requires the creation of a common vision and integration of different research areas such as haptic-based interaction and sound processing together with geometric modeling, computer graphics, and rendering. In addition, more focus should be placed on the manual aspects of interaction.

4.5.2 Design Tools Based on Enactive Interfaces

There are several situations where users are required to do things they usually do manually – with their hands – using a computer-based system. Obviously, they are not able to act in an intuitive way as well as they cannot exploit their skills, and they also have to use their time for learning a new way of working that is often inefficient and unsatisfactory. A typical example concerns the design of aesthetic products where stylists and designers are required to use computer-aided design tools strongly based on their internal geometric modeling engine with the pretensions to obtaining the same results. It is indeed well known that the ways of working are not comparable – users' satisfaction is not the same, creativity is limited, and the quality of the results may not be the same.

An interesting objective of research is investigating if and how enactive interfaces can be exploited in the tools for the product design context. New user interfaces based on enactive knowledge would support richer interaction that would allow users to operate with hands and to interact with three-dimensional models of products more easily and naturally, and also with the possibility of choosing the most appropriate and favorite modalities and senses during the interaction.

It is interesting to investigate how enactive interfaces can be used and exploited to increase the level of interaction in product design tools, which are presently still based on traditional ways of interaction. In order to do that we have to investigate and understand the following:

- interaction modalities, including visualization, sound, sense of touch;
- multiple modalities that can be used to enhance information, complement information, accommodate the perception thresholds typical of some senses, and the limitations of some technological implementations;
- the relationship between perception and action.

4.5.3 An Example: Enactive Interface for Shape Evaluation

The results of the reasoning about enactive interfaces presented in the previous section has been used as theoretical background for the implementation of a system based on haptic interfaces for the evaluation of aesthetic shapes of virtual objects, oriented towards the conceptual product design. The system has been developed within the context of the SATIN (Sound And Tangible Interfaces for Novel product design) research project, funded by the European Commission (www.satin-project.eu).

The motivation underlying the study of this system is presented below. It is common practice after creating a new shape in a digital modeling environment for the designer to evaluate the quality of the shape from an aesthetic point of view. This evaluation is a complex and personal mixture of global and local characteristics that are not easily formalized in precise terms. Usually, it is possible to translate in precise "mathematical terms" the local characteristics of shapes and also what can be considered as bad attributes of shapes; but it is very difficult to formalize global characteristics to deal with relative proportions, positions, and relations. As part of the design process, professionals in this field are experienced in the ability to "touch & feel" surfaces of products in order to check and evaluate their quality. Therefore, in order to build tools that enable the same experience in a virtual environment we need to render the geometric features of shapes, such as their curvature or discontinuities, in a "touch & feel" way.

The possibility of immediately touching a curve during the conceptual phase of a new product would add an additional and innovative modality for shape evaluation where the perception of the quality of the curve is not only based on mathematic representation and visualization, but is also based on other modalities, such as touch. For this reason, within the context of the research project SATIN we have implemented a haptic linear strip conforming to the curve.

The haptic strip that we can build on the basis of current know-how and technology is limited; specifically we are not capable of representing all curvature domain and geometric features through those interfaces [19]. Since the virtual representations are in general not as good as reality, several studies have demonstrated that multimodality helps in compensating, integrating, and enhancing information. The aim of our system is not to develop an environment that is an exact copy of the real one, but rather the intention is to provide the possibility of performing the same tasks in an intuitive way, exploiting the users' skill as much as possible.

The user interface developed aims at supporting designers' enactive knowledge and also includes interactive sound as an integral component into this shape evaluation system [20]. We have investigated how sound can be used for providing information to the user about the curvature that cannot be rendered, given the actual technological limitations, through other modalities that stimulate the haptic and the visual perceptual system. Those curve features, such as discontinuities, that cannot be represented haptically are rendered through sound. Specifically, metaphoric sounds are used to convey information about curve characteristics. In summary, features of the geometrical model are converted into force feedback, visual rendering, and sounds (Figure 4.9).

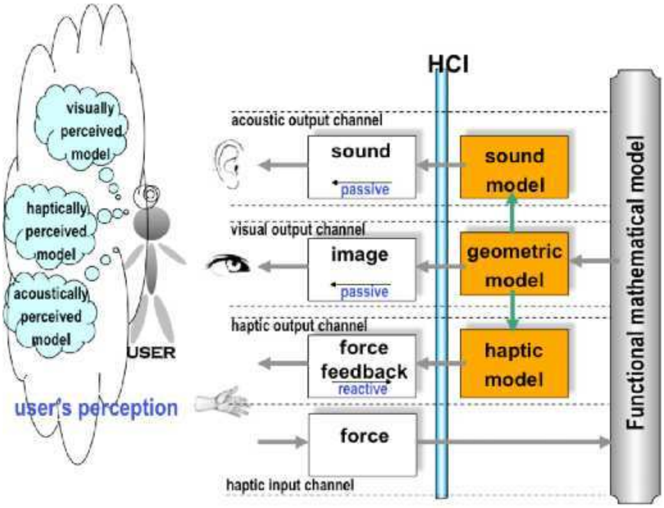


Figure 4.9 Enactive interfaces for virtual shape exploration based on metaphoric sounds

As mentioned earlier, the auditory system is highly sensitive in distinguishing sounds with different frequencies and/or levels [21]. It is, therefore, possible to use sound signals to represent qualitative and also quantitative information about specific shape properties of an object that the user is evaluating. This is particularly interesting for shape properties that are difficult to explore via the haptic modality, like the variation of, and discontinuities in, shape parameters such as curvature or tangency. We have thought of rendering sounds in such a way that the users experience the close relationship between their own actions, *i. e.*, exploring different positions along a surface by moving a finger, and the changing properties of the interactively generated sound (Figure 4.10). In order to test the concept of enactive interface based on a haptic–sound interface we have performed some experiments [15, 22], which have demonstrated the effectiveness of using a double channel, haptic and sound, for rendering information to users. It was detected that due to the close temporal relationship between the action of exploring and the sound feedback, the user can easily make the link between action and perception. Sonification thus means that by listening to the sound, the end-user learns something about the interaction that otherwise would remain hidden. All users agreed that this haptic–sound modality used for representing the mathematical characteristics of the shape is a much more natural and intuitive way of interacting and feeling curves in respect to the traditional methods based on pure vision used by CAD tools. In addition, they all felt more emotionally engaged and participatory during the shape evaluation tasks in respect to traditional modalities.

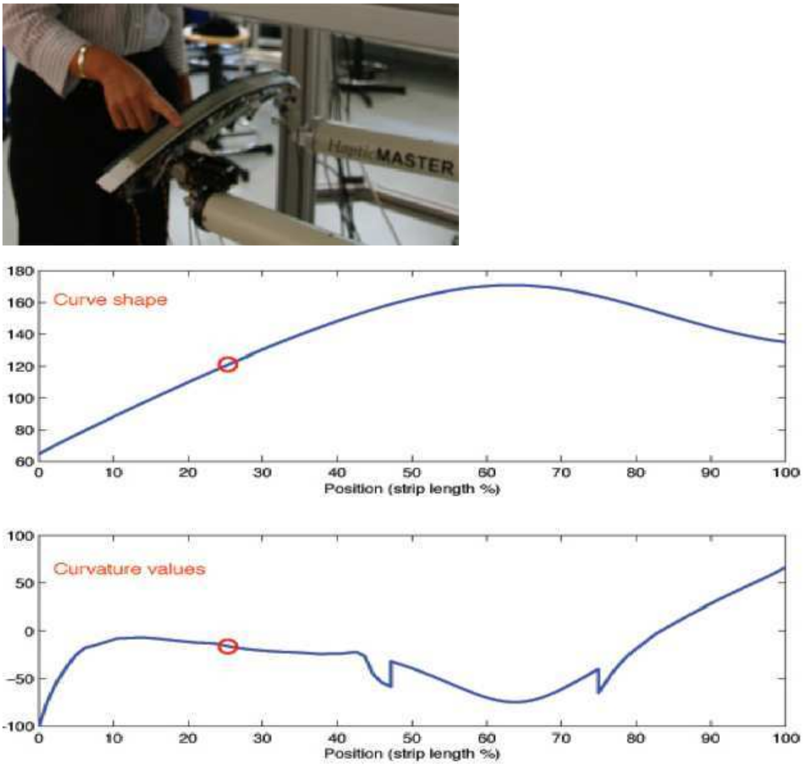


Figure 4.10 User's finger exploring the geometrical properties of a curve (SATIN project)

4.5.4 Remarks

To conclude this section, we list the benefits of the new design tools based on enactive interfaces that can be easily used by designers and also by customers.

Designers can better exploit their creativity. By using more natural ways of creating shapes, designers can express their creativity in a more natural manner and they may feel more emotionally involved in the design process; the typical technological aspects of design tools no longer represent barriers and do not interrupt the designers' creative flow. The new way of interacting with applications based on enactive interfaces would reduce the reaction and decision time during the designers' performance of the activities.

Customers can be co-designers. If easy-to-use design tools are available, we can foresee that customers will have the possibility of participating in the design activities with the designers. This will be possible even if customers do not fully understand the language spoken by the designers, since the "conversation" and exchange of opinions and requirements would be performed by means of a natural interaction with the virtual representations of the future products. Customers will also have the possibility of experimenting and testing the virtual prototypes of new products and showing what they desire through the direct use and modification of the products.

4.6 Product Prototyping and Testing

Prototyping is the design test and verification phase of product development aimed at demonstrating, proving, and validating aspects of product design. There are several types of prototypes that can be used for visualization, representation, functional verification, *etc.*

Typical prototyping methods include clay mock-up, milling machined prototypes, rapid prototypes, and so on. These are all physical prototypes that have several drawbacks: they are static, not easily modifiable, and can represent few variants.

In addition, in order to be effective, tests should be performed with an appropriate number of customers, should be done as realistically as possible, and should be easily modifiable.

In the last years, we have assisted with the advent and diffusion of practices based on virtual prototyping. More recent trends tend to use mixed reality technologies for product prototyping, where physical components of the testing environment are fused with product components that have been constructed only digitally [7, 23, 24]. Applications based on mixed reality integrate haptic interfaces in order to perform force feedback and tactile response to users' interaction with digital parts [25, 26].

Mixed prototyping has several benefits: there is no need to re-model parts of the environment that exist already and of which we do not have a digital model, and we can perform ergonomic tests that would not be possible simply using virtual prototyping. In addition, several customers can test products whose configuration can be easily and quickly changed. Therefore, virtual testing based on mixed reality is a more flexible and effective way of directly validating several aspects of new products with a group of customers.

4.6.1 An Example: Mixed Reality Application for Product Evaluation

We have developed a mixed reality application for the evaluation new of a video storage device. We needed to test and validate the initial concept of the device developed by the designers in respect to its ergonomic aspects (device handling), the aesthetic quality of its shape, and the graphical user interface and use of functions. The fully virtual prototype of the device did not allow us to perform all the tests we required. For example, we were not able to test with customers the comfort of the handling. Conversely, the fully physical prototype was static, not animated, and not interactive: we were not able to test the user interface and the use of the functions for storing and accessing videos on the device.

Therefore, we decided to build a more complex prototype where physical and virtual parts coexisted. The physical prototype of the device was produced by means of a rapid prototyping manufacturing process. The prototype was integrated with a pattern-based tracking system. The user wears a head-mounted display and handles the physical prototype: the tracking system registers the virtual representation of the device and its graphical user interface (GUI) with the physical prototype. In order to detect the selections of the GUI functions the physical prototype has been equipped with a touch panel: when the user physically interacts with the GUI, the system detects the selection and changes the GUI screen accordingly. The GUI was developed using Adobe Flash software and is controlled by the user using a touch panel; it is integrated as a texture map on the digital model.

Figure 4.11 shows a user holding the video storage device with the right hand and interacting with the device GUI using a pen. Through the helmet the user sees the virtual device and the GUI developed with the RTT RealView module (www.rtt.ag) and the real environment including his hands (shown as real thanks to an algorithm based on chroma key). The user can evaluate the ergonomic aspects of the device just by physically handling it. Any request for device shape modification can be addressed by manufacturing a new rapid prototype and modifying the virtual model of the device. The GUI "look and feel" can be changed through a rapid modification of the graphical appearance of the functions and of their activation sequence.

We were able to involve several potential customers in the evaluation phase. Customers were able to use a preview of the future video storage device, to physically handle and operate it in a very realistic way. Customers' requests for modifications could be handled and satisfied quickly, acting both on the physical and virtual components of the video storage device prototype.

Figure 4.11 Mixed reality application for video storage device evaluation and testing

4.7 Conclusion

A major trend concerns the design and production of product variability so that they can better satisfy the needs, preferences, and skills of individual customers. This can be obtained by means of a product development process where the design is based on a user-centered approach. This approach tries to satisfy customers' needs and wants, and to optimize how people can, want, or need to work, rather than forcing the customers to accept, adapt, and change the way they work just to accommodate the product design and engineering approach.

An extensive involvement of customers ensures that the product will be suitable for its intended purpose in the context in which it will be used. On the one hand, this approach leads to the development of products that are more effective, and helps designers in handling customers' expectations about a new product. On the other hand, customers are more satisfied, as they feel that their preferences are taken into consideration in the design phase, and they find them in the products. In addition, the sense of ownership and affection of customers for the final products would be definitely increased.

In a product-consumer society the relationship between consumers and manufacturers is linear and simple. If consumers are heavily involved in the product development, they become *stakeholders*: they affect or can be affected by the actions of the business.

Finally, we can state that by using a user-centered design approach and easy-to-use design tools, *creativity* is within everyone's reach: products that have never been made before, but which are also useful, beautiful, and impressive, can be made by each of us or with our contribution, even if we are not designers [27].

References

1. Nagamachi M.: Kansei Engineering. Kaibundo, Tokyo (1989)
2. Fukuda S.: Emotional engineering session. Proc. IDETC/CIE ASME 2009 Conference, San Diego, CA (2009)
3. Ulrich K.T., Eppinger S.D.: Product Design and Development. McGraw-Hill, New York (2004)
4. Eder W.E., Hosnedl S.: Design Engineering – A Manual for Enhanced Creativity. CRC Press, New York (2008)
5. Bordegoni M., Cugini U.: A conceptual design tool based on innovative haptic interfaces. Proc. DETC'06 ASME 2006, Philadelphia, PA (2006)
6. Hvam L., Mortensen N.H., Riis J.: Product Customization. Springer, Berlin (2008)
7. Bimber O., Raskar R.: Spatial Augmented Reality: Merging Real and Virtual Worlds. A K Peters (2005)
8. Norman D.A.: The Design of Everyday Things. Doubleday, New York (1988)
9. Norman D.A., Draper S.W. (eds.): User-Centered System Design: New Perspectives on Human-Computer Interaction. Lawrence Earlbaum Associates, Hillsdale, NJ (1986)
10. Muller M.J.: Participatory design: The third space in HCI (revised). In: Jacko J., Sears A. (eds.), Handbook of HCI, 2nd edition. Lawrence Erlbaum, Mahway, NJ (2007)

11. Preece J., Rogers Y., Sharp H.: Interaction Design: Beyond Human-computer Interaction. Wiley, New York (2002)
12. Tudor L.G.: A participatory design technique for high-level task analysis, critique and redesign: the CARD method. In: Proceedings of the Human Factors and Ergonomics Society, Seattle (1993)
13. Schreiber G., Akkermans H., Anjerwierden A. *et al.*: Knowledge Engineering and Management – The CommonKADS Methodology. MIT Press, Cambridge, MA (1999)
14. www.enactivenetwork.org. Accessed October 2009
15. Bordegoni M., Cugini U.: Multimodal perception-action interaction. In: Proceedings of the IDETC/CIE ASME 2009 Conference, San Diego, CA (2009)
16. Wolfe J.M., Kluender K.R. *et al.*: Sensation and Perception. Sinauer Associates, Sunderland, MA (2006)
17. Loomis J.M., Klatzky R.L. *et al.*: Similarity of tactual and visual picture recognition with limited field of view. Perception **20**:167–177 (1991)
18. Bruner J.: Toward a Theory of Instruction. Belknap Press, Cambridge, MA (1966)
19. Bordegoni M., Cugini U., Antolini M., Covarrubias M.: Geodesic haptic device for surface rendering. In: Hirose M., Schmalstieg D., Wingrave C. A., Nishimura K. (eds.) Joint Virtual Reality Conference of EGVE–ICAT–EuroVR, Lyon, France (2009)
20. Bordegoni M., Ferrise F., Shelley S., Alonso M.A., Hermes D.: Sound and Tangible Interface for shape Evaluation and Modification. In: HAVE 2008 – IEEE International Workshop on Haptic Audio Visual Environments and their Applications, Ottawa, Canada (2008)
21. Hermann T., Hunt A.: An introduction to interactive sonification. IEEE Multimedia **12**(2): 20–24 (2005)
22. Sharples S., Hollowood J., Lawson G. *et al.*: Evaluation of multimodal Interaction Design Tool. In: CREATE Conference, Computer Society, Covent Garden, London (2008)
23. Burdea G., Coiffet P.: Virtual Reality Technology, 2nd edition. Wiley, New Jersey (2003)
24. Burdea G.C.: Haptic feedback for virtual reality. Int. J. Design Innovation Res. Special issue on Virtual Prototyping **2**(1):17–29 (2000)
25. Haller M., Billinghurst M., Bruce T.: Emerging Technologies of Augmented Reality – Interfaces and Design. Idea Group, Hershey, PA (2007)
26. Hayward V., Ashley O., Hernandez M.C. *et al.*: Haptic interfaces and devices. Sensor Rev. **24:**16–29 (2004)
27. Fukuda S.: Developing evolving machines for creative customers. In: Proceedings of the IDETC/CIE ASME 2009 Conference, San Diego, CA (2009)

Chapter 5
A Study on the Relationships Between Drivers' Emotion and HUD Image Designs

Shana Smith[1] and Shih-Hang Fu

Abstract This study explores the relationships between drivers and head-up display (HUD) presentation image designs using kansei engineering. There are two major experiments in this study. The objective of the first experiment is to cull the representative kansei words from a large pool of kansei words which were associated with HUD physical design elements, using factor analysis and cluster analysis. In the second experiment, the relationships between the representative kansei words and HUD presentation images are found using the Quantification Theory Type I (QT1) method. The results of the study can be used as a reference for future HUD image design.

5.1 Introduction

Research shows that many traffic accidents are caused by human errors in longitudinal or lateral control [1]. Head-up display (HUD) technology was originally developed to help pilots operate aircraft, particularly military aircraft, more safely. In 1988, HUD technology was also introduced in the automobile industry. For example, BMW's M series and X5 were both equipped with HUDs. Automobile HUDs present important information to drivers, such as speed, warnings, gas level, gears position, radio setting, and position navigation. As a result, with an HUD, drivers spend less time lowering their heads to check important driving information. Instead, they can view information from presentation images which are projected on the windshield and in the drivers' line of sight. Prior research shows that the response time to an urgent event is faster with an HUD than an

[1] Shana Smith
Department of Mechanical Engineering, National Taiwan University, Taipei, Taiwan
e-mail: ssmith@ntu.edu.tw

HDD (head-down display), and speed control is also more consistent with an HUD. Moreover, using an HUD causes less mental stress for drivers and they are easier for first time users to use [2, 3]. Therefore, in the future, HUDs might become an indispensable device for drivers, since they can create a much safer driving environment.

There has been significant prior research into the hardware structure of HUDs, but little research so far on HUD presentation image design. However, HUD presentation image design is an important factor which can enhance the effectiveness of HUDs. To design an HUD image, some aspects should be taken into consideration. For example, Yoo *et al.* studied the effect of HUD warning location on driver response and driver performance enhancement [4]. In addition, HUD image quality is also an important factor which affects consumer purchasing. Most drivers prefer consumer-oriented HUDs, which are more attractive, useful, and sensitive to their personalities and their emotions.

This study aims to explore the relationships between HUD presentation image designs and drivers' emotions using kansei engineering. Kansei, which is a Japanese word, means the customers' psychological feelings. Kansei engineering aims to produce a new product which is more acceptable based upon consumers' feelings and demands [5, 6, 7]. Kansei engineering can be used to determine consumers' emotions from a psychological basis. With kansei engineering, statistical methods are used to analyze the kansei and translate the analyzed data into the design domain [8]. For example, Mazda used a method of category classification to develop a sports car, and Nissan utilized a hybrid kansei engineering system to design a new steering wheel for a passenger car [9].

There are three types of kansei engineering: category classification, kansei engineering systems, and hybrid engineering systems [10]. In this study, category classification was used to find relationships between the kansei and the design specifications. Through kansei engineering, a collection of kansei words were found which describe the main factors associated with HUD image design elements.

5.2 Experiment Procedures

The major objective of this study is to find the relationships between kansei words and HUD presentation image design elements. The kansei words are used to express customers' emotions and demands, and the design elements are used to extract the physical design importance for HUD presentation images. The experiment flowchart is shown in Figure 5.1.

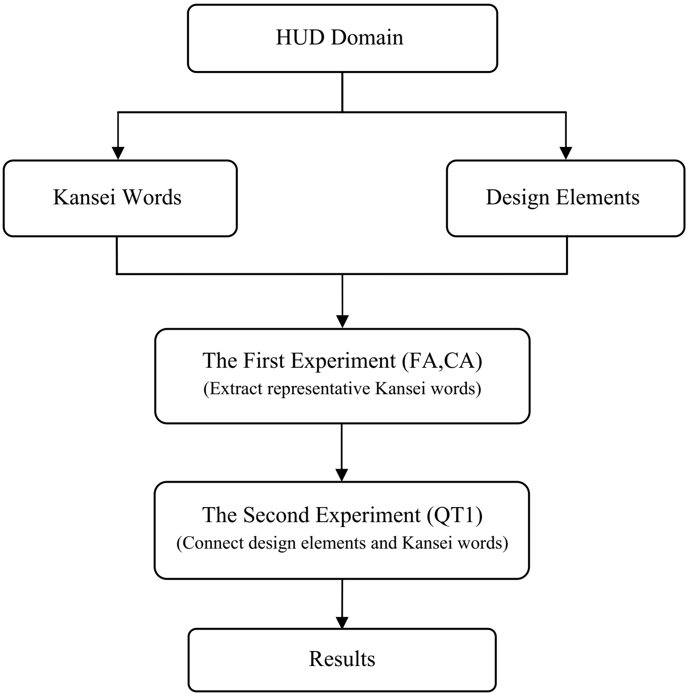


Figure 5.1 Experiment flowchart

5.2.1 The First Experiment

5.2.1.1 Collect HUD Images

In the first experiment, existing HUD presentation images were collected, as shown in Figure 5.2. Most existing HUD presentation images from automobiles, such as the Chevrolet Corvette, GMC Arcadia, and Pontiac Grand Prix, include speed, gear position, engine speed, outside air temperature, and oil level. For all of the given vehicles, HUD presentation images are green in color and three inches by five inches in size. The images are projected so that they appear to be in front of the windshield at a distance of about 1 m. Some HUD presentation images from video games or website were also collected. It can be seen from the examples that existing HUD presentation image designs are quite diverse.

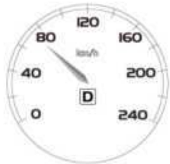
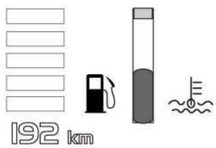
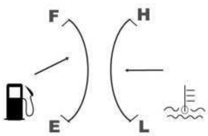
5.2.1.2 Extract Image Design Elements

Next, new HUD presentation images were designed based on the existing commercial HUD presentation images which were collected. However, this study focused on the effectiveness of different image designs on drivers' emotions, rather than on the overall design patterns themselves. The aim of this step was to list all the design elements from the existing images. For the given study, enumerating image design elements, such as fonts, colors, the presentation methods for speed and gas level (illustration or meter), and so on, was equivalent to spanning the space of properties in kansei engineering [11].

Figure 5.2 Existing commercial HUD images

Each design element was evaluated, and important elements were selected for further study. Seven design elements were extracted: image content, forms for major content, forms for secondary content, image location, ratio of major content to secondary content, image size, and color. Each element had several categories or levels, as shown in Table 5.1. For example, color had three levels.

Table 5.1 HUD presentation image design elements of first experiment

Design elements	Design categories		
Form	 A	 B	 C
	 D	 E	 F
Content	1–2 items • Speedometer and gear position • Rev counter	3–4 items • Speedometer • Rev counter • Gear position • Fuel gauge	Over 5 items • Speedometer • Rev counter • Fuel gauge • Temperature gauge Odometer
Secondary content form	 Illustration	 Meter	
Location	Left	Center	Right
Ratio	1:1	2:1	
Size	Large	Medium	Small
Color	Orange	Green	Blue

5.2.1.3 Collect Kansei Words

In this step, as many kansei words as possible were collected from customers, automobile magazines, and the Internet. Each kansei word was arranged as a pair with opposite meanings for semantic evaluation using Osgood's original semantic differential (SD) scale or Nagamachi's SD scale [12] (Figure 5.3). Osgood's SD scale uses synonyms and antonyms for spanning the range of ratings [13]. With the method used by Nagamachi and most Japanese researchers, an extreme adjective is put on the left-hand side of the scale and "not at all" is added to the same adjective on the right-hand side of the scale [14, 15]. For the given study, 86 pairs of kansei words were collected, and every pair of kansei words had opposite meanings to span the range of ratings.

Figure 5.3 There are two different semantic differential scales: (a) Osgood's SD scale, and (b) Nagamichi's SD scale

5.2.1.4 Select Kansei Words

In this step, important kansei words were culled from the large pool of kansei words collected in the previous step through discussion. Three of the subjects had engineering backgrounds, while the other three had design backgrounds. For each of the HUD presentation images collected in the previous step, each study subject chose pairs of kansei words from the pool and marked them with check marks. Kansei words with more than three checks were retained, and others were dropped. From a first-round evaluation, 45 pairs had more than three checks. After a second round evaluation, only 32 pairs of kansei words remained, as shown in Table 5.2.

Table 5.2 32 pairs of kansei words

fashionable – ordinary	outgoing – reserved	modern – retro	young – mature
stylish – conservative	passionate – dull	amiable – strange	comfortable – uncomfortable
elegant – ugly (unattractive)	vigorous – lethargic	with a sense of technology – without a sense of technology	of high quality – not of high quality
sporty – not sporty	innovative – non-innovative	brand-named – not brand-named	soft – harsh
female – male	gentle – manly	relaxing – anxious	aristocratic (noble) – common
vivid – gloomy	moderate – freezing	organized – scattered	spacious – crowded
tidy – messy	concrete – abstract	easy to understand – esoteric	explicit – ambiguous
human – not human	considerate – not considerate	secure – dangerous	tired – not tired

5.2.1.5 Create New Images Using the Taguchi Method

Testing all possible combinations in Table 5.1 would require evaluating 1944 possibilities. Therefore, the Taguchi method was used to conduct a sufficient but fewer number of experimental evaluations [16, 17]. The Taguchi method uses orthogonal arrays to find balanced design solutions in an experimental space. Each row in the Taguchi table represents each experimental setup. In this study, an L18 Taguchi table was chosen and used, with 18 experimental trials extracted to replace the original 1944 possibilities. Next, six representative samples were chosen for factor analysis. The samples which were chosen are shown in Figure 5.4.

Figure 5.4 Representative HUD images

5.2.1.6 Questionnaire

The questionnaire of the first experiment was designed based on the six HUD image samples and the 32 pairs of kansei words. Every pair of kansei words was scored from 1 to 7, using a semantic differential scale [18, 19]. The semantic differential scale is the most common type of rating scale [12]. Subjects in the experiment consisted of 30 people, whose ages range between 18 and 65. In the questionnaire, the 32 pairs of kansei words were arranged into seven major questions so that the participants could be guided into accomplishing the objectives of the study. Moreover, Guilford [20] showed that randomizing the order of the kansei words could lead to better survey results. Therefore, in this study, the orders in which both image samples and kansei words were presented to the subjects randomly.

5.2.1.7 Representative Kansei Words

Factor Analysis

The survey data was first analyzed using explorative factor analysis. Factor analysis was conducted using an extraction method based on the principle component analysis. Components with eigenvalues greater than 1 were extracted and the resulting data was rotated using the varimax rotation method. Five factors were extracted to explain the total variance. The cumulative contribution of the five factors was 100%. The results after varimax rotation are shown in Table 5.3.

Table 5.3 Component matrix of the 32 pairs of kansei words

	Component				
	1	2	3	4	5
Stylish – conservative	.983	–.069	–.092	.140	–.004
With a sense of technology – without a sense of technology	.977	–.112	–.140	–.090	.075
Relaxing – anxious	.945	.184	.137	.135	.188
Innovative – non-innovative	.943	–.023	.094	.301	.103
Fashionable – ordinary	.939	–.328	.076	.051	.054
Modern – retro	.933	.296	.183	–.088	.023
Sporty – not sporty	.930	–.333	.133	–.037	–.068
Vigorous – lethargic	.915	–.248	.196	.229	.100
Young – mature	.894	.319	.275	–.155	.015
Considerate – not considerate	.817	–.256	–.140	.484	.117
Comfortable – uncomfortable	.794	–.089	.554	–.219	–.085
Of high quality – not of high quality	.760	–.589	.263	–.059	–.061
Soft – harsh	.726	–.480	.443	.115	.183
Elegant – ugly	.712	–.660	.114	–.124	.169
Brand-named – not brand-named	.695	–.631	–.153	.280	–.130
Gentle – manly	.610	–.007	.360	.447	.546
Spacious – crowded	–.117	.981	–.138	.076	–.005
Explicit – ambiguous	–.241	.970	–.029	.004	.007
Tidy – messy	–.234	.964	–.058	.087	–.070
Easy to understand – esoteric	–.039	.930	–.304	–.089	.181
Vivid – gloomy	.019	.910	.211	.357	–.022
Organized – scattered	.289	.873	.271	.013	.283
Concrete – abstract	–.406	.828	.043	–.225	.313
Aristocratic – common	.696	–.710	–.056	–.090	–.005
Human – not human	.675	.705	.008	.216	–.024
Outgoing – reserved	.478	.695	–.044	.213	–.491
Amiable – strange	.230	–.275	.920	–.033	.154
Tired – not tired	–.158	–.459	–.873	–.040	.039
Female – male	–.143	.441	–.802	.203	.316
Moderate – freezing	–.253	.610	.696	.249	.130
Passionate – dull	.193	.459	–.089	.835	–.220
Secure – dangerous	.292	.262	–.090	–.152	.903

Cluster Analysis

Cluster analysis was used to divide the kansei words into five groups based on the rotated component matrix. There are five centers in the five groups. The distance between each kansei word and the center of its group can be used to estimate the importance of the kansei word in its group, as shown in Table 5.4. Finally, five representative pairs of kansei words were extracted: "*modern – retro*", "*masculine – feminine*", "*relaxing – anxious*", "*soft – harsh*", and "*explicit – ambiguous*".

Table 5.4 The distance between each kansei word and their group centers

	1st	2nd	3rd	4th	5th	Distance
Fashionable – ordinary	.923	–.361	.083	.076	.071	0.22633
Of high quality – not of high quality	.919	–.365	.140	–.009	–.053	0.29834
With a sense of technology – without a sense of technology	.739	–.615	.267	–.044	–.047	0.29322
Elegant – ugly	.690	–.681	.118	–.115	.181	1.70683
Brand-named – not brand-named	.697	–.507	.449	.125	.199	1.78321
Relaxing – anxious	.675	–.732	–.053	–.079	.008	1.66617
Innovative – non-innovative	.665	–.661	–.147	.296	–.110	1.57465
Human – not human	.788	–.292	–.130	.506	.141	1.40133
Female – male	.582	–.034	.368	.454	.565	1.49388
1st group center	.742	–.472	.122	.134	.106	
Soft – harsh	–.133	.446	–.801	.206	.316	0.65736
Explicit – ambiguous	.691	.676	.017	.255	–.013	0.66862
Passionate – dull	.181	.435	–.08	.855	–.202	0.73836
Outgoing – reserved	.501	.670	–.037	.255	–.483	0.76893
Considerate – not considerate	.294	.260	–.089	–.155	.902	0.92779
Secure – dangerous	–.166	–.450	–.875	–.05	.038	1.08487
2nd group center	.228	.340	–.311	.228	.093	
Gentle – manly	.942	.149	.146	.168	.204	0.22516
Modern – retro	.943	.264	.191	–.05	.034	0.24557
Stylish – conservative	.975	–.106	–.083	.175	.014	0.31664
Sporty – not sporty	.928	–.061	.103	.332	.124	0.32265
Young – mature	.907	.289	.282	–.118	.023	0.32324
Vigorous – lethargic	.975	–.143	–.133	–.057	.089	0.35626
Comfortable – uncomfortable	.795	–.114	.558	–.194	–.077	0.52590
3rd group center	.924	.040	.152	.037	.059	
Amiable – strange	.212	–.284	.921	–.04	.158	0.53592
Vivid – gloomy	–.249	.613	.697	.244	.128	0.53592
Tired – not tired	.478	.695	–.044	.213	–.491	1.25845

Table 5.4 Continued

	1st	2nd	3rd	4th	5th	Distance
4th group center	–.019	.165	.809	.102	.143	
Easy to understand – esoteric	–.207	.978	–.029	.014	–.002	0.17600
Organized – scattered	–.084	.983	–.136	.093	–.01	0.18704
Spacious – crowded	–.202	.970	–.057	.099	–.078	0.23135
Concrete – abstract	–.004	.934	–.302	–.074	.173	0.34257
Aristocratic – common	.036	.901	.215	.374	–.019	0.42398
Tidy – messy	–.373	.847	.04	–.23	.297	0.46252
Moderate – freezing	.313	.863	.276	.032	.282	0.51564
5th group center	–.074	.925	.001	.044	.092	

5.2.2 The Second Experiment

In order to find the relationships between the five representative kansei words and the HUD design elements, the Quantification Theory Type 1 (QT1) method was used.

5.2.2.1 Create New Images Using the Taguchi Method

The HUD design elements in the second experiment are shown in Table 5.5. Since it is impossible to test all the design combinations, the Taguchi method was used to conduct a sufficient number of experimental evaluations. An L18 orthogonal array was used. In addition, three existing commercial HUD images were also added in the second experiment. Therefore, there are 21 HUD image samples, as shown in Figure 5.5.

Table 5.5 HUD presentation image design elements in the second experiment

Design elements	Design categories		
Forms	Illustration	Meter	Illustration and meter (Mix)
Secondary content forms	Illustration	Meter	
Content	1 item • Speedometer	3 items • Speedometer • Gear position • Fuel gauge	5 items • Speedometer • Gear position • Fuel gauge • Time • Navigation
Color	Orange	Green	Blue
Location	Left	Center	Right
Font	Arial	Electronic	
Scenario	Day time	Night	

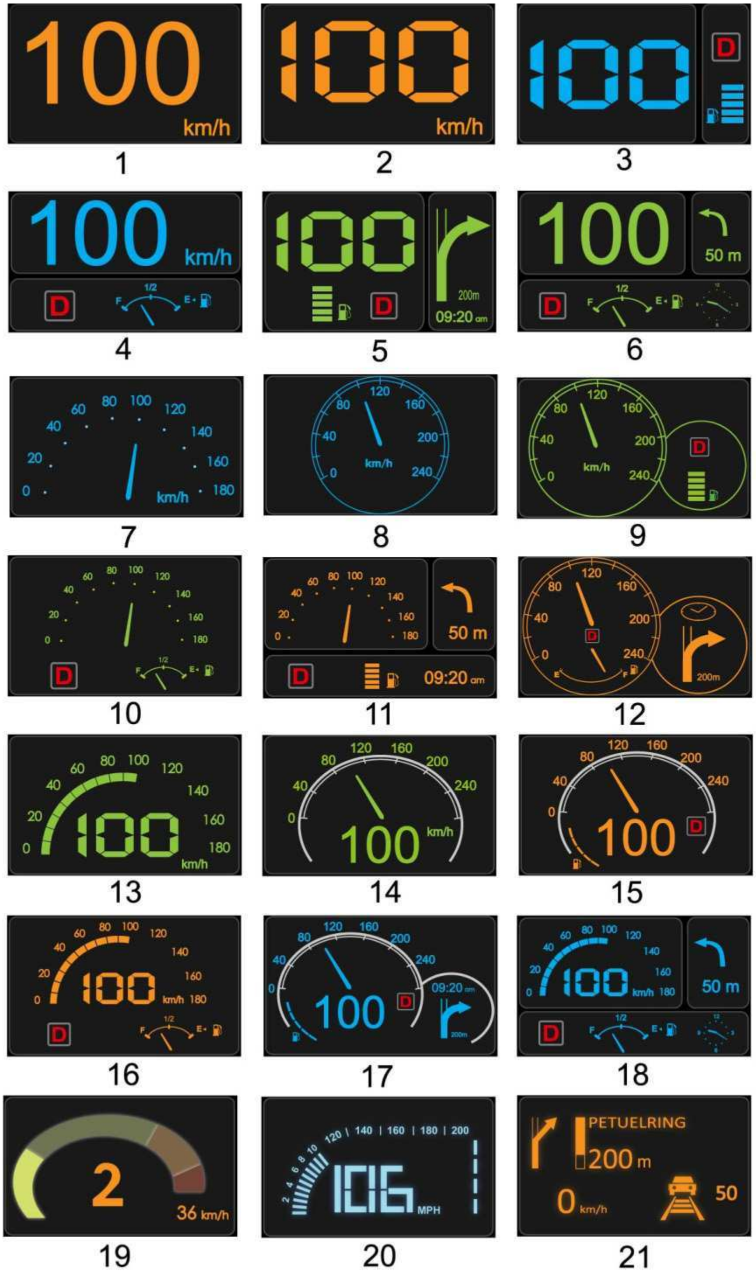


Figure 5.5 21 HUD presentation images

5.2.2.2 Questionnaire

A questionnaire in the second experiment was designed based on the 21 HUD image samples and the five pairs of representative kansei words. Each pair of kansei words was scored from 1 to 7, using the semantic differential scale.

5.2.2.3 Experiment and Data Analysis

In the second experiment, a driving simulator and an HUD device were used to simulate the actual driving environment, as shown in Figures 5.6 and 5.7. There were ten subjects in the second experiment, with ages ranging between 18 and 65. A survey was conducted after the test drive, and the data was analyzed using the QT1 method.

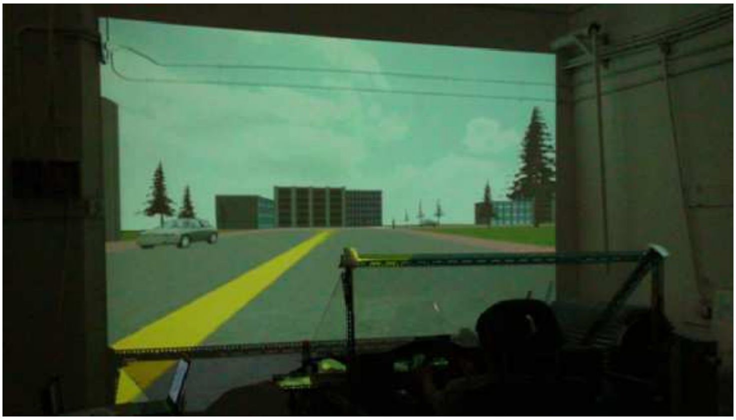


Figure 5.6 Driving simulator

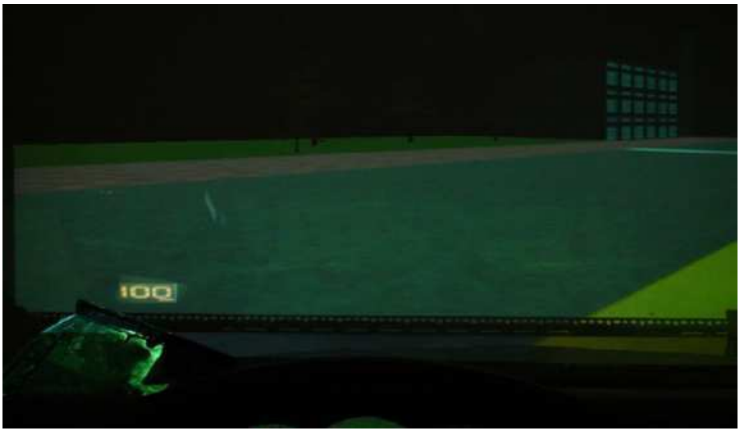


Figure 5.7 HUD presentation image projected onto a windshield

Table 5.6 Results of the QT1 analysis for "modern" and "retro"

Modern – Retro			Modern – Retro
Design element	Category	PCC	–1 –0.5 0 0.5 1
Form	Illustration	0.715	–0.18
	Meter		0.79
	Mix		–0.46
Secondary content form	Illustration	0.576	–0.31
	Meter		0.41
Content	1 item	0.634	0.65
	3 items		–0.24
	5 items		–0.28
Color	Orange	0.497	–0.31
	Green		0.01
	Blue		0.41
Location	Left	0.485	0.2
	Middle		0.19
	Right		–0.39
Font	Arial	0.426	0.22
	Electronic		–0.25
Scenario	Daytime	0.465	0.27
	Night		–0.25
Multiple correlation coefficient 0.89			

5.2.2.4 Quantification Theory Type 1 (QT1)

This study attempted to quantitatively relate each kansei word to the HUD design elements, using the Quantification Theory Type 1 method. Table 5.6 shows the results of the analysis concerning kansei words “modern” and “retro”. Since the multiple correlation coefficient was 0.89, the correlation between the kansei words and the design elements was 89%. The partial correlation coefficients in this table are expressed in the form of a bar chart, which indicates the influence of each design category on the emotion of “modern” and “retro”. In each item, the category on the left-hand side of the bar chart shows “modern”, and on the right-hand side shows “retro”. Therefore, the HUD presentation image was interpreted as being “retro” when the Form is in meter and “modern” when the Form is in mix.

5.3 Conclusions

This study analyzes the relationships between drivers’ emotions and HUD physical design elements using kansei engineering. The results can be used to evaluate and enhance future HUD presentation image designs. Results of the study show that drivers’ kansei responses consist of five emotional factors. The study also shows that certain aspects should be considered when designing an HUD presentation image. First, the presentation image of an HUD should be simple so that drivers can obtain the information quickly and correctly. Second, to minimize the influence of the external environment, the HUD presentation image should not contain too many colors. In addition, according to the attributes of an HUD, some information might be better shown in illustrations, and others might be better displayed in the form of meters. Furthermore, the study results can also be applied to normal dashboard design, especially for digital dashboards.

References

1. Wierwille W.W.: Development of an initial model relating driver in-vehicle visual demends to accident rate. In: Proceedings of Third Annual Mid-Atlantic Human Factors Conference, Virginia Polytechnic Institute and State University, Blacksburg, VA (1995)
2. Liu Y.C.: Effect of using head-up display in automobile context on attention demand and driving performance. Display **24**:157–165 (2003)
3. Liu Y.C., Wen M.H.: Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. Int. J. Human-Comp. Studies **6**:679–697 (2004)
4. Yoo H., Tsimhoni O., Green P.: The effect of HUD warning location on driver responses. In: International Transportation Systems World Congress, pp. 1–10 (1999)
5. Chen K., Chiu S.C., Lin F.C.: Kansei design with cross cultural perspectives. In: Aykin N. (ed.) Usability and Internationalization, Part I, HCII 2007, LNCS 4559, pp. 47–56. Springer, Berlin/Heidelberg (2007)

6. Nagamachi M.: Kansei Engineering: a new ergonomic consumer-oriented technology for product development. Int. J. Industrial Ergonomics **15**:3–11 (1995)
7. Nagamachi M.: Kansei engineering: the framework and methods. In: Nagamachi M. (ed.) Kansei Engineering 1. Kaibundo, Kure, pp. 1–9 (1997)
8. Nagamachi M.: Perspectives and the new trend of Kansei/affective engineering. TQM J. **20**(4):290–298 (2008)
9. Nagamachi M.: Kansei engineering as a powerful consumer-oriented technology for product development. Appl. Ergonomics **33**(3):289–294 (2002)
10. Nagamachi M., Okazaki Y., Ishikawa M.: Kansei engineering and application of the rough sets model. Proc. IMechE Part I: J. Syst. Control Eng. **220**(8):763–768 (2006)
11. Grimsæth K.: Kansei Engineering: Linking emotions and product features. Department of Product Design, Norwegian University of Science and Technology (2005)
12. Schütte S., Eklund J.: Product Development for Heart and Soul. Linköping University, Department for Quality and Human Systems Engineering, Linköping (2003)
13. Osgood C.E., Suci G.J., Tannenbaum P.H.: The Measurement of Meaning. University of Illinois Press, Illinois (1957)
14. Ishihara S.: Kansei engineering procedure and statistical analysis. Workshop at International Conference on Affective Human Factors Design, Singapore (2001)
15. Nagamachi M.: Workshop 2 on kansei engineering. In: Proceedings of International Conference on Affective Human Factors Design, Singapore (2001)
16. Antony F., Antony FF.: Teaching the Taguchi method to industrial engineers. Work Study, **50**(4):141–149 (2001)
17. Taguchi. G.: Introduction to Quality Engineering. Asian Productivity Organization, Tokyo (1986)
18. Küller R.: Semantisk Miljö Beskrivning (SMB). Psykologiförlaget AB Liber Tryck, Stockholm, Stockholm (1975)
19. Osgood C.E, Suci G.J., Tannenbaum P.H.: Measurement of meaning. In: Osgood C.E., Snider J.G. (eds.) Semantic Differential Technique – A Source Book. Aldine, Chicago, IL, pp. 56–82 (1969)
20. Guilford J.P.: Psychology Methods. McGraw-Hill Publishing Company, New York (1971)

Chapter 6
Emotional Design in the Virtual Environment

Tetsuro Ogi[1]

Abstract In order to realize human-centric design, the design specification should be based not only on the functional aspects of the product but also on an emotional evaluation, such as the impression gained from the shape or the size of the designed model. A virtual environment in which the designed model is visualized interactively using a full-scale stereo image can be used effectively as a method of emotional evaluation. Design support and design evaluation using virtual reality, augmented reality, and tele-immersion technologies are discussed. The collaborative design between designers or between designer and customer is achieved using video avatar technology in the tele-immersion environment. In addition, a design education program that includes the emotional evaluation in the virtual environment is also introduced.

6.1 Introduction

Recently, human-centric design has been become an important concept in various fields. The specification for products designed for humans should be based not only on their functional aspects, such as their size and strength, but also on emotional qualities, including the impression the user derives from the shape and size of the designed product. Although rapid prototyping is often used to evaluate the shape of a design based on kansei and the emotional response of the user, it is difficult to make a full-sized model. Thus, there is a need for an evaluation method in which the user can visualize an actual sized model in the virtual environment using virtual reality technology [1]. This chapter discusses the current technology

[1] Tetsuro Ogi
Graduate School of System Design and Management, Keio University
4-1-1 Hiyoshi, Kouhoku-ku, Yokohama 223-8526, Japan
e-mail: ogi@sdm.keio.ac.jp

trends, which include recent research in design support, design evaluation, and collaborative design using virtual reality, augmented reality, and tele-immersion technologies.

6.2 Design in the Virtual Reality Environment

6.2.1 Interactive Visualization

Virtual reality is a simulation technology that realizes an interactive virtual experience using high-presence images such as stereo or wide field of view images as well as sound and haptic sensation information. In order to achieve a virtual reality experience, special devices are being used. For example, interactive stereo images are represented using a goggle-type HMD (head mounted display) system that the user wears on his or her head or an IPT (immersive projection technology) system that uses projectors and multiple screens [2].

Figure 6.1 shows K-Cave, the CAVE-type immersive projection display developed at Keio University [3]. In Figure 6.2 the configuration of this system is shown. This display consists of multiple screens placed to the front, left, and right of the user, and also on the floor. A three-dimensional image space can be generated by projecting synchronized stereo images onto each screen. A stereo image is generated as follows: two DLP projectors with circular-polarizing filters are used for each screen to project passive stereo images. The user wears stereo glasses with the same polarizing filters in order to experience the stereo images, based on the binocular parallax by seeing the right eye image and the left eye image separately. In addition, an electromagnetic sensor that measures three-dimensional position and orientation information in space is attached to the glasses, which enables the images corresponding to the user's view position to be generated interactively by tracking the user's head position.

Figure 6.1 Immersive projection display K-Cave

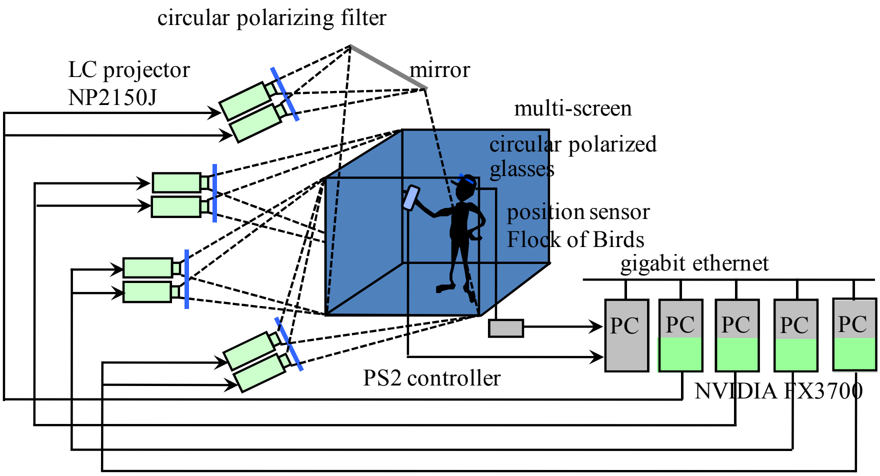


Figure 6.2 System configuration of the K-Cave system

In this kind of virtual reality system, since the user can feel a sense of depth and obtain a sense of the size of the displayed image using the effect of binocular parallax, he or she can evaluate the size of the displayed object based on a sensation of the actual size. Moreover, since the image seen from the user's viewpoint is represented using tracking data, the displayed image can be seen from various directions interactively, enabling the user to evaluate its three-dimensional shape. Thus, it is expected that the virtual reality environment can be used as a support environment for kansei and emotional evaluation, in which the user can evaluate whether the sense of size or shape experienced from the model design satisfies his or her demands. Figure 6.3 shows an example where the user is looking at a three-dimensional full-sized model in the immersive virtual environment.

Figure 6.3 Visualization of a three-dimensional model in the CAVE system

6.2.2 Haptic Sensation

Virtual reality technology utilizes not only visual sensation but also other sensations. When we recognize the shape and surface quality of an object, we usually touch the object and hold it in our own hands to get a feeling for it. Therefore, it is very important to use haptic sensation in the virtual reality environment together with visual information. Haptic sensation can be roughly classified into force feedback sensation and tactile sensation.

The force feedback device generates a reaction force from the object to the user's hand or fingertips so that the user can recognize the shape of the virtual objects. As a means of generating a reaction force, several methods have been proposed, including the use of a motor, string, master arm, or exoskeleton structure for the arm or finger. For example, commercial devices such as PHANToM, where the reaction force is represented through a pen, and CyberGrasp, which uses a hand exoskeleton structure, have been developed [4, 5].

Figure 6.4 shows a user wearing the force feedback device HapticGEAR in the CAVE system [6]. In this device, four strings are attached to the pen held in the user's hand, and the reaction force is generated by winding the strings from the corners of the frame on the shoulders using DC motors. This device has the advantage of being a wearable force display that can be used in the CAVE-type immersive projection display, though the direction of the generated reaction force is limited to the frustum shaped by the four strings. By using this device, the user can feel the reaction force from the visualized model and recognize its three-dimensional shape.

Figure 6.4 Wearable force feedback device HapticGEAR

On the other hand, tactile sensation includes the sense of touch, the sense of temperature, and the feeling of the texture, such as the smoothness or roughness that we feel by tracing the surface of an object with our fingertip. The tactile display represents these kinds of sensations based on computer simulation. In order to generate a tactile sensation, several methods have been proposed and used, such as utilizing air pressure, a pager motor, a micro pin array, wind pressure, and electrical stimulation. Whereas the pager motor only represents the sense of touch, the micro pin array can represent the feeling of texture by controlling the stimulation pattern according to the movement of the user's fingertip.

Figure 6.5 shows a tactile data glove, named the VibroGlove [7]. In this device, 17 pager motors are attached to the data glove, on the palm and on each side of the fingers. This device can represent the sense of touch by vibrating the pager motor at the point touching a virtual object. In addition, since it can represent the movement of tactile stimulation by controlling the vibrating part, a flow field can be generated using tactile sensation. In this figure, the user is investigating the air flow around the model of an airplane using tactile and visual information, and the result of the evaluation can be used as feedback for the airplane design. Thus, the haptic sensation can be used effectively for designs based on an emotional evaluation.

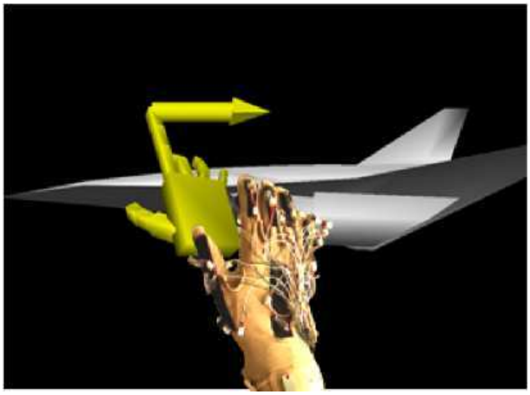


Figure 6.5 Tactile data glove VibroGlove

6.2.3 Application of VR to Design

Virtual reality technology using three-dimensional space has been applied to product design in various research fields.

For example, in the field of architectural design, the user can evaluate a model of a building or office by visualizing it as a full-scale stereo image. Though the architectural strength or the seismic isolation structure can be evaluated using computer simulation, it is impossible to evaluate the sensation derived from the size or height of the building by looking at design data only. These sensations can be evaluated by experiencing an actual-sized model visualized in the virtual environment.

The virtual space can also be effectively used for evaluating furniture or equipment in a room. For example, when a kitchen is customized, the user can evaluate the position of a faucet or the height of a shelf based on his or her body size through their experience in the virtual environment.

Recently, in the car industry, a large-screen immersive projection display has often been used for evaluating designs and presenting them to customers. At the design stage, the immersive environment can be used to evaluate the position of the steering wheel or to design the front panel by visualizing them with full-scale models based on CAD data. In this case, the interaction, in which the position and color of the car parts can easily be changed in real time while the user is evaluating them sitting on a virtual driver's seat, is very valuable. Moreover, since it is also possible to present the car to the customer with a high-quality presence by visualizing a full-scale model of it, the virtual car can be exhibited with the various options available, creating a virtual showroom.

6.3 Design Using an Augmented Reality Environment

6.3.1 Integration of Virtual Image with Real World

Augmented reality is a technology that integrates the virtual image with the real world, though the virtual reality generates the complete virtual world without the real scene. For typical augmented reality display systems, HMD and IPT are used. HMD systems for augmented reality are classified into optical see-through systems and video see-through systems. In an optical see-through system, the user can see the augmented reality world where the computer graphics image of the virtual world is integrated with the real world scene using a half mirror. On the other hand, in a video see-through system, the real scene is captured by a video camera attached to the HMD and the user sees the augmented reality image generated by superimposing a computer graphics image onto the captured video image.

The IPT system for augmented reality generally uses the optical see-through method by placing a large half mirror in the room, and the user can see an integrated scene of the real world and a reflected computer graphics image projected onto the floor or ceiling screen. This method is often called spatial augmented reality technology since it uses the actual space around the user as the augmented reality world [8].

6.3.2 AR View

Figure 6.6 and Figure 6.7 show the appearance and system configuration of the spatial augmented reality display named AR View [9]. In this system, a large semi-transparent 2.0 m by 2.6 m mirror film is placed on the floor at an angle of 45 degrees, and an active stereo image with a wide field of view is projected onto the floor screen using two DLP projectors fixed to the ceiling. The user can experience the augmented reality world where the real object behind the half mirror film and the computer graphics image are integrated, by wearing LC shutter glasses with an electromagnetic position sensor.

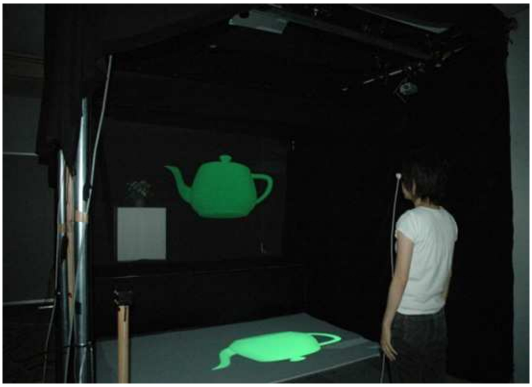


Figure 6.6 Spatial augmented reality display AR View

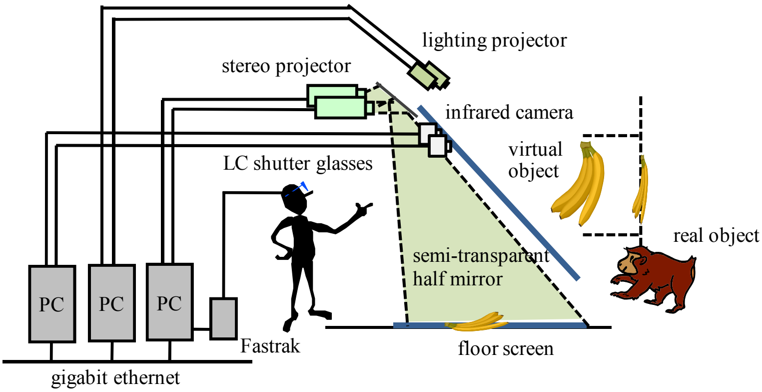


Figure 6.7 System configuration of AR View

It is generally difficult for an augmented reality display using a semi-transparent mirror to represent the correct occlusion when virtual objects overlap real objects. When virtual objects are located behind real objects, the occlusion effect can be represented by rendering a black shadow on the virtual objects. On the other hand, when real objects are located behind virtual objects, parts of the real objects should be occluded. In this system, the correct occlusion effect is realized by using a lighting projector technique, as shown in Figure 6.8 [10].

This technique is used to illuminate the real object using projectors instead of standard light bulbs. In this system, low brightness LED projectors of 15 lumens are used. Since the real object is illuminated by a projector using white light, an occlusion shadow can be created on the real object where it overlaps the virtual object without illuminating it. In order to implement this method, it is necessary to create models of the shape of the virtual and real objects in the computer and calculate the shadow according to the movement of the user's head. Therefore, in cases where the shape or location of the real object is changed, an infrared multi-camera system is used to measure the shape of the real object and to realize the correct occlusion effect using the lighting projector.

By using this technique, an integrated scene of the designed model and real objects can be correctly represented.

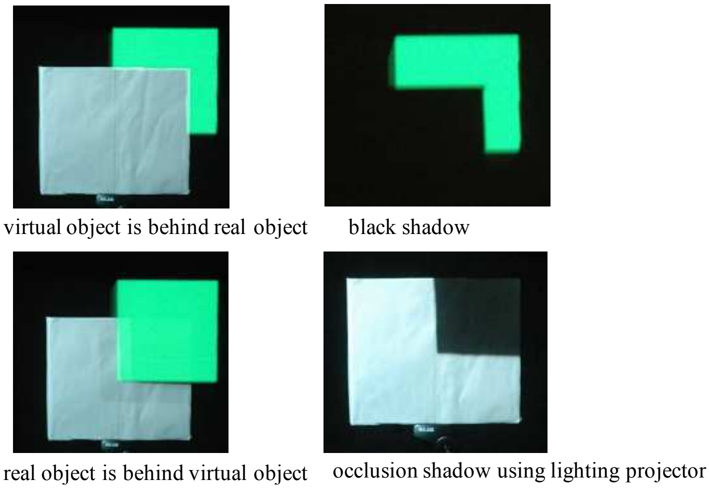


Figure 6.8 Correct occlusion effect using the lighting projector technique

6.3.3 Application of AR to Design

When augmented reality technology is applied to design support or design evaluation, a three-dimensional design model can be displayed together with a real world scene. In this case, there are various forms of integrating real objects and virtual objects, such as visualizing a three-dimensional object in real space or visualizing the design space around real objects.

For example, when a user is planning to buy furniture or an electrical appliance, he or she can evaluate whether these objects are suitable for their room, by looking at the visualized images of these products overlapping with their real room using the augmented reality devices. And when the user is thinking of buying a new house or a building under construction, it becomes possible to judge the ease of use of the room based on kansei and an emotional evaluation by visualizing the virtual house or building as a background to his or her real furniture or appliances.

6.4 Emotional Communication

In order to realize human-centric design, it is important to transmit emotional information between humans. When large-scale products or systems are developed, it is impossible for them to be completed by one person, and many people have to collaborate to carry out the work. Moreover, communication between the designer and the customer is also necessary to design easy to use products. In these cases, it is important to realize emotional communication, whereby the remote user can share not only design data but also kansei and emotional information such as what he is looking at or what he is thinking about.

In order to realize high-presence communication remotely, we are constructing a tele-immersion environment by connecting together several CAVE-like immersive projection displays and developing video avatar communication technology. The details and current status of the video avatar communication technology in the tele-immersion environment are discussed in the following sections.

6.4.1 Tele-immersion Environment

In this research, several virtual reality displays were connected as shown in Figure 6.9 through the JGN2plus (Japan Gigabit Network 2 plus), which has a bandwidth of 20 Gbps, as the backbone for constructing the tele-immersion environment [11]. CAVE-type immersive projection displays at Keio University, The University of Tokyo, University of Tsukuba, and Kyoto University were used. In addition, super high-definition tiled displays at Keio University, Iwate Prefectural University, Kyoto University, and Osaka University were used [12].

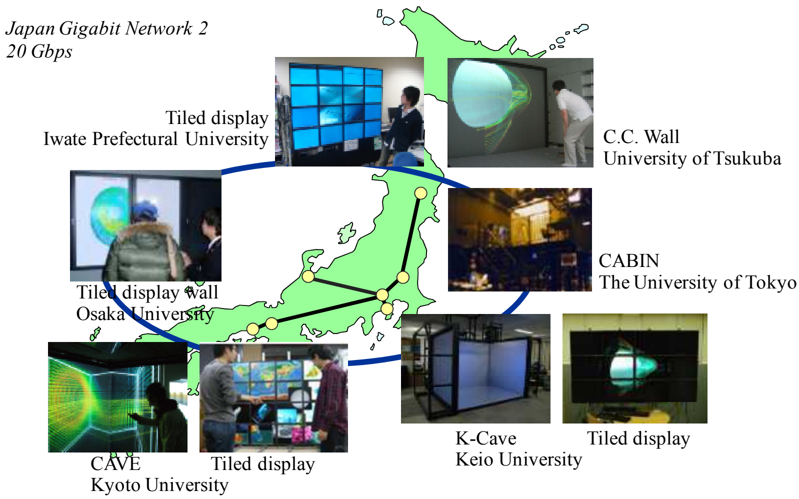


Figure 6.9 Tele-immersion environment using the JGN2plus

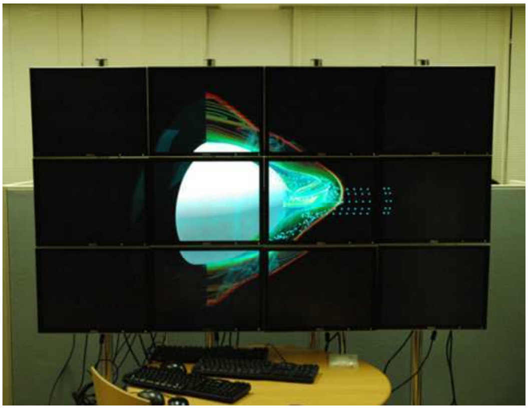


Figure 6.10 Super-high-definition tiled display

In the networked CAVE display, visualization data can be shared in the immersive virtual environment with a wide field of view. On the other hand, the tiled display, which is constructed by arranging several LC monitors, can be used as a super-high-definition display environment with a large display area. Figure 6.10 shows the tiled display system at Keio University. Though the current tiled display does not support stereo images, it can be used effectively to display a large amount of information, including text and image data. This network environment can be used effectively for communication sharing of large-scale visualization data.

6.4.2 Video Avatar Technology

For communication in the tele-immersion environment, video avatar technology was developed [13]. The video avatar can be used to achieve high-presence and emotional communication in the networked virtual environment. In this method, the user's figure is captured by a video camera and this image is transmitted to another site. At the receiving site, the user's image is segmented from the background and synthesized as a video avatar in the shared virtual world. By performing this process mutually between remote sites in real time, the video avatar can be used as a communication method.

Figure 6.11 shows the method of generating a video avatar. In this method, the user's image is captured by an IEEE1394 camera at 15 Hz. The background subtraction method that calculates the difference between the user's image and the background image captured beforehand is used to segment the user's figure from the background. In this case, since the view position of the user is measured by the position sensor at each site and is transmitted to the other site together with the user's image, the video avatar image can be placed at the position where the remote user is. The video avatar can be synthesized in the shared virtual world in real time by texture mapping the video image onto the video avatar model.

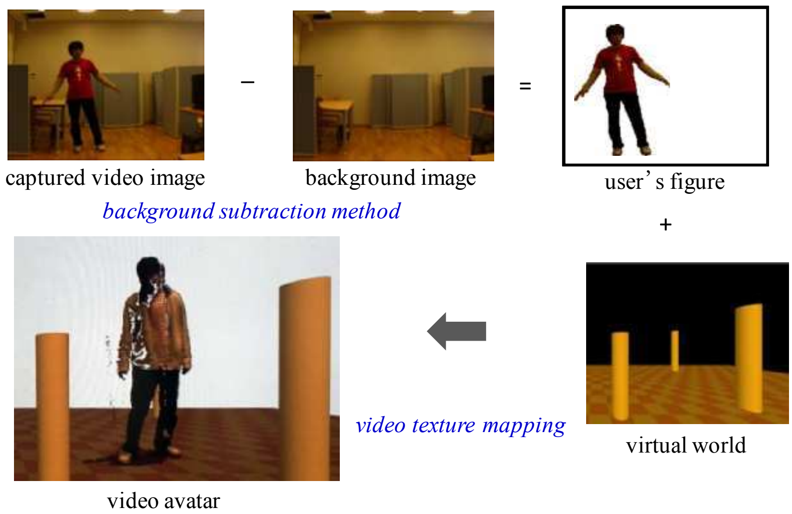


Figure 6.11 Video avatar generation method

Various methods are used to create the video avatar model, such as a simple two-dimensional billboard model, a 2.5-dimensional surface model generated from a stereo camera image, and a three-dimensional voxel model reconstructed from multi-camera images. Figure 6.12 shows the construction methods of various video avatar models [14].

In the case of a two-dimensional video avatar (Figure 6.11 (a)), although the video avatar is placed at the three-dimensional position, the avatar itself is represented as a two-dimensional image, so that gestures such as pointing at objects with the index finger cannot be accurately transmitted to the other user. When the 2.5-dimensional video avatar is used (Figure 6.11 (b)), it can represent the three-dimensional gesture of the user because it has depth information. However, when observed from the side, the user sees a defective model of the avatar because it has

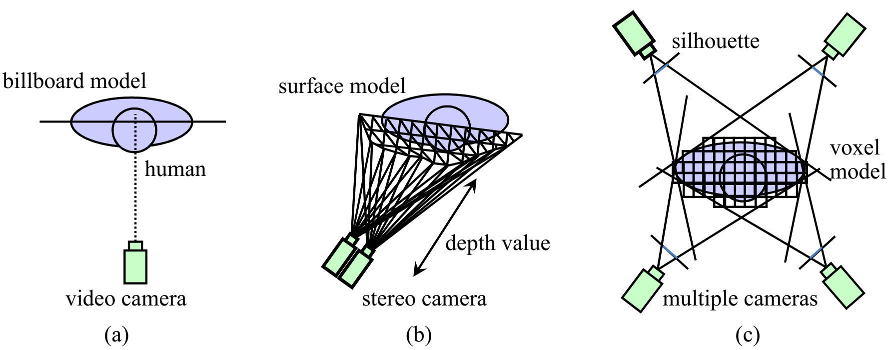


Figure 6.12 Construction methods of the video avatar model: (a) 2D video avatar, (b) 2.5D video avatar, and (c) 3D video avatar

a shape model only for the camera direction. Therefore, when the representation of an accurate three-dimensional gesture is necessary for communication between multiple users, the only effective method is to generate the video avatar using the three-dimensional model.

Figure 6.11 (c) shows the method of constructing the three-dimensional video avatar using shapes from the silhouette method. This method can reconstruct three-dimensional voxel shapes from the intersection of back-projected cones of silhouette images that have been captured by multiple cameras. Though the representation of a voxel shape has large errors when few video cameras are used, the accuracy of the approximate shape model would be much higher if the number of cameras were to be increased. By selecting the video image captured from the direction closest to the other user's view position and mapping it onto the voxel model as video texture, a three-dimensional video avatar can be generated. Though the three-dimensional video avatar is superior to the two-dimensional or 2.5-dimensional video avatar in its ability for expression, it is inferior to them in image quality, since it depends on the accuracy of the voxel data, the calculation time for reconstructing the three-dimensional model, and the amount of transmission data. Therefore, several problems must be solved to represent a high-quality image of a three-dimensional video avatar in real time.

6.4.3 Video Avatar Communication

Emotional communication in the shared virtual world is expected to be used effectively for the product design in various fields. For example, although a lot of people such as designers, experimental engineers, and computer scientists are engaged in the design phase, these people are not usually in the same location. Therefore, it is desirable that remote users can have discussions with each other while sharing visualized design data between remote places. In this case, the high-presence communication techniques that transmit emotional information are expected to be very effective.

In addition, communication is important not only between designers but also between designers and customers. The customer usually cannot evaluate whether the designed product satisfies his or her requirements before seeing the completed product. Moreover, the customer's requirements often change after he or she has seen the completed model. Therefore, this type of communication is expected to be an effective way for the designer and customer to discuss designs in the design phase while looking at visualized full-scale three-dimensional models in the shared space.

Figure 6.13 shows a remote user having a discussion with another user's video avatar while sharing the three-dimensional model in the networked CAVE environment. In this method, since each user can walk through freely in the shared virtual space, the positional relationship between them can be represented correctly, and they can discuss standing side by side looking at the visualized model from the same direction or stand face to face to look at the model from opposite directions. Thus, remote users are able to perform the kansei and emotional evaluation in the high-presence communication environment [15].

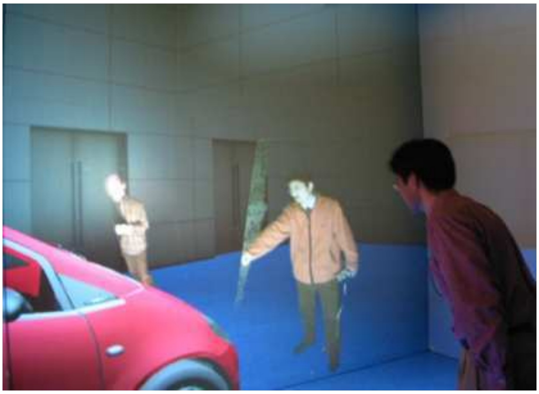


Figure 6.13 Video avatar communication sharing the designed model

6.5 Education of Emotional Design

In order to spread the concept of human-centric design based on emotional evaluation, it is necessary that the designers themselves acquire it as a practical design method, and not just as a demonstration or experiment. Therefore, design education that includes not only the functional specification but also human-centric emotional evaluation is necessary.

For such a purpose, the authors are experimentally assembling a design education program, which includes emotional evaluation in virtual space, for a class in the graduate school. In this class, about 20 students are taking a seminar which includes emotional design in virtual space as well as geometrical and functional design using CAD and CAE Tools. For example, when a chair is designed, they are given the design parameters of favorite shape and appropriate size for the user's body, as well as functional specifications such as the weight of the chair, the strength needed to support the user's weight, and the stability for tilting.

The students first make a rough design of their favorite shape using the CAD system CATIA and confirm the three-dimensional shape in the stereo vision display. In this phase, they can recognize the difference between the impressions felt from the two-dimensional image and the three-dimensional stereo image, and they often change their design concept. After designing the rough shape, they calculate the weight of the chair, the strength for the user, and the stability for tilting by using the structural analysis function of CATIA. They then decide the dimensions and select the materials. Moreover, at this stage they also evaluate the height and area of the seat by visualizing a full-scale model in the K-Cave immersive projec-

Figure 6.14 Evaluation of the designed chair using the K-Cave system

tion display environment. Figure 6.14 shows a student evaluating the shape and height of the seat by visualizing a stereo image in the K-Cave system. By conducting the functional and emotional evaluation repeatedly, the product that satisfies both functional and emotional specifications is designed. At the end of the class, each student has a presentation in the stereo display environment to explain the concept and the features of his or her design.

This class is considered to be useful to allow the students to gain an understanding of the importance of human-centric design based on an emotional evaluation, such as the user's impression of the size and shape of the product, as well as the functional design of its geometrical shape, strength, and structure. Some students actually repeatedly changed the sizes in the detailed model according to how they felt, or made big changes to the conceptual design after conducting an emotional evaluation in the CAVE environment.

In the design process, although the emotional evaluation is conducted in the latter half, since the current exercise is given according to the progress of the lecture, it should be conducted from the first stage of the rough design. Moreover, since the current CAD system for geometrical design and the VR system for emotional evaluation are separate systems, the data from the CAD system must be transmitted to the VR system to visualize it. In future, it will be necessary to develop and use an emotional design system that integrates the CAD environment and the visualization environment without the transmission of data.

References

1. Hollerbach J.M., Cohen E., Thompson T.V., Freier R., Johnson D.E. (eds.): Virtual prototyping for human-centric design. In: 2000 NSF Design and Manufacturing Research Conference (2000)
2. Bullinger H., Riedel O., Breining R.: Immersive projection technology – benefits for the industry. In: Proceedings of International Immersive Projection Technology Workshop, pp. 13–25 (1997)

3. Tateyama Y., Oonuki S., Sato S., Ogi.: K-Cave demonstration: seismic information visualization system using the OpenCABIN library. In: Proceedings of ICAT 2008, pp. 363–364 (2008)
4. Massie T., Salisbury T.: The PHANToM Haptic interface: a device of probing objects. ASME Winter Annual Meeting **55**(1):295–300 (1994)
5. http://www.cyberglovesystems.com
6. Hirose M., Hirota K., Ogi T., Yano H., Kakehi N., Saito M., Nakashige M.: HapticGEAR: the development of a wearable force display system. In: Proceedings of IEEE VR2001, pp. 123–129 (2001)
7. Yano H., Kakehi N., Ogi T., Hirose M.: Haptic interface for immersive projection display. In: Proceedings of HCI International '99, pp. 1030–1034 (1999)
8. Bimber O, Raskar R.: Spatial Augmented Reality. AK Peters, Natick, MA (2004)
9. Ogi T., Murase K., Saito K., Koyama T.: Development of immersive augmented reality display: AR View. In: Proceedings of AGIAGRAPH 2007, pp. 161–168 (2007)
10. Bimber O, Fröhlich B.: Occlusion shadows: using projected light to generate realistic occlusion effects for view-dependent optical see-through displays. In: Proceedings of ISMAR '02, pp. 186–195 (2002)
11. http://www.jgn.nict.go.jp/english/
12. Ebara Y., Kukimoto K., Koyamada K.: Evaluation experiment on eye-to-eye contact in remote communication with tiled displays environments. In: Proceedings of AINA Workshops 2008, pp. 1017–1022 (2008)
13. Ogi T., Yamada T., Tamagawa K., Kano M., Hirose M.: Immersive telecommunication using stereo video avatar. In: Proceedings of IEEE VR2001, pp. 45–51 (2001)
14. Ogi T., Yamada T., Kurita Y., Hattori Y., Hirose M.: Usage of video avatar technology for immersive communication. In: Proceedings of the First International Workshop on Language Understanding and Agents for Real World Interaction, pp. 24–31 (2003)
15. Ogi T.: Building networked immersive VR applications using the MVL toolkit. In: Proceedings of ICAT2004, pp. 202–209 (2004)

Chapter 7
Emotional Robot Competence and Its Use in Robot Behavior Control

Natascha Esau and Lisa Kleinjohann[1]

Abstract Emotional competence plays a crucial role in human communication and hence has also gained increasing attention for the design of interaction processes between humans and robots. Like humans, emotionally intelligent robots should be capable of coping with the emotions of their human counterpart as well as with their own artificial emotions, which requires some key competencies. Robots need the abilities to recognize and to understand human emotions in a certain situation, they have to be able to react adequately in order to regulate their own emotions as well as the emotions of their human counterpart, and they have to express their own emotions in an adequate way. In this paper we elaborate the concepts of emotional competence and show how artificial emotions and drives can be integrated into a robotic system to realize emotionally competent and proactive behavior. For this purpose we propose a fuzzy emotion model which is used as basis for human emotion recognition and for representing the static aspects of a robot's emotions. Subsequently, a dynamic model for artificial robotic emotions and drives that allows for adequate control of robotic behavior is described. Furthermore, the application of our concepts in the emotionally competent robot head MEXI is presented.

7.1 Introduction

Industrial robots are widely used today, for instance in manufacturing, assembly and packing, transport, or in the mass production of consumer and industrial goods. They usually perform dedicated tasks in dedicated environments and have a limited degree of autonomy. Their history reaches back to the middle of the

[1] N. Esau and L. Kleinjohann
University of Paderborn, Fürstenallee 11, 33102 Paderborn, Germany
e-mail: nesau@c-lab.de, lisa@c-lab.de

twentieth century. Since then there have been considerable advancements in their sensory abilities and motor skills, allowing the utilization of robots in “natural” environments not specifically tailored to robotic needs, such as cleaning, housekeeping, lawn mowing, inspection, surveillance, and guide or “edutainment” robots. At the end of the 1990s with the advent of affordable personal robots and “edutainment” robots, which frequently interact with their owner, a new research field called social robotics emerged.

The focus of social robotics is no longer merely the accomplishment of a specific task, as it has been for industrial and many service robots. But the interaction between robots and humans has come to the fore. Since human interactions usually involve not only the rational level but also the emotional level, emotional aspects have also gained increasing attention in robotic research. According to R. Picard, who coined the term *affective computing* [1], human–computer interaction can be considerably improved when technical systems are aware of their user’s feelings and can react adequately.

However, the aspects emotion recognition and adequate reaction are only part of what is called emotional competence [2, 3]. If we imagine more sophisticated machines like service robots that are required, for instance, to distribute meals in hospitals or support elderly persons, these machines should also be able handle and show their own (artificial) emotions, since this would create some kind of relatedness and increase the acceptance of their human users. This directly leads to two further aspects of emotional competence: emotion representation and emotion regulation, *i. e.*, adequate handling of own emotions in different situations. While emotion recognition and emotion representation are the focus of many researchers nowadays, realizations of emotion regulation and adequate emotional reaction mechanisms to control the robot behavior have rarely been reported. However, in order to show emotional competence in man–machine cooperation all four aspects have to be considered.

In this paper we first describe the psychological concepts of emotional competence more deeply and show how they can be integrated into a robotic architecture (Section 7.2). After summarizing related work in Section 7.3, we concentrate on modeling and recognizing human emotions (Sections 7.4 and 7.5). Subsequently, in Section 7.6 we describe how the dynamics of artificial emotions together with

Figure 7.1 MEXI – Machine with Emotionally eXtended Intelligence

robotic drives, which support a robot's pro-active behavior, are modeled to represent the robot's own emotionality. Furthermore, we present the robot head MEXI (see Figure 7.1) and show how its behavior-based architecture integrates the aspects of emotion recognition, representation, and regulation as well as the selection of appropriate behaviors (Section 7.7). Section 7.8 gives an example of how MEXI adequately reacts to emotions recognized from the face of its human counterpart. Finally, Section 7.9 gives an outlook on future work.

7.2 Aspects of Emotional Competence in Human–Robot Interaction

Since the beginning of the 1990s the notion *emotional competence* or *emotional intelligence* has gained major psychological interest. The term *emotional intelligence* was introduced by Salovey and Mayer in 1990, as a kind of social skill that incorporates the abilities of recognizing and supervising own and others' emotions, and how this information is used to control ones thoughts and actions [2]. In 1995, Daniel Goleman published his popular book *Emotional Intelligence* [4], emphasizing the importance of emotional and social factors in human interaction.

Emotional competence is the ability to cope with the emotions of oneself and others and to evaluate and regulate them in a specific context in order to avoid conflicts or stress. As already mentioned it covers the following four aspects:

- the ability to perceive the emotions of oneself and others;
- the ability to understand the emotions of oneself and others;
- the ability to regulate the emotions of oneself and others in various ways;
- the ability to express one's own emotions in an adequate way.

The concept of emotional competence can also be applied to the communication or interaction between humans and robotic systems, leading to man–machine interfaces that are able to cope with or even avoid conflicts. Accordingly a robot has to provide the following functionalities in order to show emotional competence:

- *Emotion recognition* – the robot has to be able to recognize the emotional state of its human counterpart based on the information received by its sensory inputs.
- *Emotion regulation* – the robot has to select a situation-adequate artificial emotional state for itself based on information about its human counterpart and its own state.
- *Behavior selection* – the robot has to select appropriate actions to avoid stress and conflicts.
- *Emotion representation* – the robot has to be able to show its own (artificial) emotions through appropriate multi-modal communication channels, *e. g.*, facial expression, prosody of natural language or gestures.

A robot can only be called emotionally competent if all four functionalities are available.

The definition of emotional competence implies the presence of an emotional state in robots as well as in humans. This emotional state can be influenced by external and internal factors. External factors are stimuli generated by the environment, while internal factors correspond to influences generated by the robot (or human) itself.

According to psychological investigations, drives are the most important internal factors influencing the emotional state of humans. The satisfaction of a drive is, for instance, often accompanied by a positive emotion reflecting the associated comfortable feeling. On the other hand, if a drive is not satisfied over a longer period of time, negative emotions can often arise.

Hence, emotions are the result of a complex information processing. This process involves not only the perception and cognitive evaluation of external stimuli, but also the evaluation of drives and their dynamics as internal stimuli influencing the emotional state.

Due to the prominent role of drives in human emotional processes, we also included a drive system into a robotic system. Thus, instead of the world model usually used to plan a robot's behavior, the drive system (together with the emotions) is used to determine a robot's pro-active, emotional behavior.

Figure 7.2 shows the principal architecture for realizing emotional competence in the feedback loop of human–robot communication.

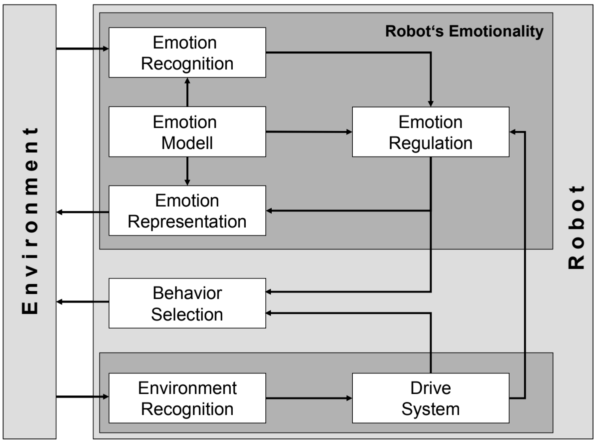


Figure 7.2 Emotional competence in human–robot communication

In the first step the robot perceives its environment using different sensors. It has to extract the relevant information about its human counterpart, *i.e.*, facial expressions, natural language utterances, or gestures, from these perceptions in order to recognize the respective emotion. This information about the human counterpart and the environment is then evaluated. Depending on the results of this evaluation the robot's drive state will be updated. Furthermore, the evaluation results together with the actual drive state will influence the robot's emotional state, which is further determined by intrinsic emotion regulation mechanisms.

The actual drive state together with the actual emotional state will determine the robot's behavior, including the expression of respective emotions by according facial expressions or natural language utterances. By executing this behavior using corresponding actor elements the robot is able to interact with its environment. Thus a feedback loop between robot and environment is realized, where on the one hand the environment acts upon the robot and on the other hand the robot influences the environment by its actions.

This view of robotic emotions and drives as an expression of dynamic processes involving several subsystems leads to real emotional competence in human–machine interaction. Thereby a more natural interaction between humans and robot can be reached, avoiding potential strain or conflicts.

The next section describes related work with special focus on how existing robots deal with the four aspects of emotional competence described above.

7.3 Related Work

Numerous architectures have been proposed covering one aspect or another of emotional competence. Several expressive face robots and android robots have been implemented, especially in Japan and the USA, where the focus has been on mechanical engineering and design, visual perception, and control. Examples are the android Repliee Q2 developed by Osaka University and Kokoro Corporation [5] or the humanlike robot Albert Hubo from Hanson Robotics Inc [6], which are constructed to resemble females or males and are even equipped with synthetic skin, teeth, and hair. They can only represent emotions and thereby try to imitate humans.

Emotion recognition from facial expressions and natural speech has been studied by many researchers, but often separately from the other elements of emotional competence. An overview can be found in [7], for instance. The humanoid robot WE4-RII [8] also determines how stimuli from the environment are evaluated in its current emotional state and how the robot reacts on them. Their model of emotion dynamics is inspired by the motion of a mass–spring system. Arkin *et al.* [9] discuss how ethological and componential emotion models influence the behavior of Sony's entertainment robots. The homeostasis regulation rule described in [10] is employed for action selection in Sony's AIBO and the humanoid robot SDR as well as in our approach. Canamero and Fredslund [11] created the humanoid robot Feelix based on an affective activation model that regulates emotions through stimulation levels. Part of it relying on Tomkins' idea that overstimulation causes negative emotions [12] is also adopted in our approach realized in the robot head MEXI. Velasquez proposes a framework consisting of five subsystems for handling perceptions, emotions, drives, behaviors, and actors [13]. They are represented as interconnected nonlinear computational units. Via nonlinear functions of so-called releasers and weights the activation of each basic system, *e. g.*, an emotion or behavior, is calculated. For emotions, an automatic decay and an influence on other emotions are also integrated, as in MEXI. However, the automatic excitation of drives as proposed to realize pro-active behavior in MEXI is not described.

Most similar to our work is the robotic head Kismet built at MIT [14]. The organization and operation of Kismet's motivation system is strongly inspired by various theories of emotions and drives in humans and animals. Kismet's target is to imitate the development and mechanisms of social interaction in humans. However, Kismet does not recognize emotions but only tries to extract a speaker's intention from natural language. Furthermore, Kismet does not reason about the emotional state of others [15], which is a crucial element of emotional competence in order to react adequately. Recently, MIT announced a development of Nexi that will not only be able to express different emotions but will also be able to detect the emotions of humans and act accordingly [16].

In contrast, our approach realized in the robot head MEXI supports all four elements of emotional competence. MEXI recognizes human emotions from speech and facial expressions and incorporates mechanisms to react adequately to the emotions recognized in its human counterpart. MEXI may, for instance calm down angry persons or encourage sad persons. For handling its own emotions, MEXI does not try to imitate humans, like Kismet and many other android robots. Rather, MEXI follows a constructive approach in order to realize the control of purely reactive behavior by its drives and artificial emotions, and shows this internal state using corresponding facial expressions and speech utterances.

7.4 Modeling of Emotions

Psychologists have tried to explain the nature of human emotions for decades or even centuries. Nevertheless no unique established emotion model exists. However, human emotions are now usually seen as dynamic processes that involve several modalities, such as gestures, physiological arousal, and subjective feeling [17].

Psychological literature describes two major types of emotion models, *i. e.*, models relying on basic emotions and dimensional emotion models that distinguish, for instance, the dimensions valence, potency, arousal, or intensity. In addition, mixtures of these models can be found.

Our work is built on emotion theories with basic emotions, since they are used by most technical systems for emotion recognition. Although such emotion models have existed for several decades, even now there is no general agreement among psychologists how many basic emotions exist and what they are.

Plutchik, for instance, describes eight basic emotions (acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise) that are used as building blocks for so-called secondary or even ternary emotions [18]. Other psychologists often rely on different smaller subsets of basic emotions (cf. [19]).

Due to this variety we decided to develop an emotion model that is adaptable to the set of basic emotions (cf. [20]). Following Plutchik's theory, blends of emotions and emotion intensities can also be modeled.

The emotion model proposed in this paper is based on fuzzy hypercubes [21]. Emotion intensities are mapped to values within the interval [0, 1]. In the remainder of this section we first define a fuzzy set corresponding to an emotional state and then show how it is represented in a fuzzy emotion hypercube.

Fuzzy set for emotional state. Let *BE* be a finite base set of *n* basic emotions $e_1, e_2, \ldots, e_n$, and $\mu_{FE_j} : BE \to [0,1], j = 1,2,\ldots\}$, an infinite set of fuzzy membership functions. Then each $FE_j := \{(e_i, \mu_{FE_j}(e_i) \mid e_i \in BE\}, j = 1,2,\ldots$ defines a fuzzy set corresponding to one emotional state E_j.

Fuzzy emotion hypercube. If *BE*, μ_{FE_j}, and FE_j are defined as described above, we shall use the membership vector

$$(\mu_{FE_j}(e_1), \mu_{FE_j}(e_2), \ldots, \mu_{FE_j}(e_n)) =: (\mu_{FE_j}(e_i))$$

to denote a point in an *n*-dimensional hypercube.

Each axis of the hypercube corresponds to one basic emotion e_i. Thus a membership vector $(\mu_{FE_j}(e_i))$ corresponds to one emotional state E_j and can be interpreted psychologically as vector of emotion intensities $(I_{e_j}) := (I_{e_1}, I_{e_2}, \ldots, I_{e_n})$.

The number *n* of distinguished emotions depends on the psychological theory or, in the case of technical systems, on the intended application. If, for instance, the three basic emotions happiness *h*, anger *a*, and surprise *s* shall be distinguished, a three-dimensional unit cube is needed for modeling emotional states.

The corners in the unit cube describe dual memberships (0 or 1) for all emotions; vertices describe dual memberships for two emotions and the third one varies from 0 to 1. For example, the point $E_1 = (1.0, 0.2, 0.3)$ corresponding to the fuzzy set $FE_1 = \{(h,1.0),(a,0.2),(s,0.3)\}$ represents a happy emotional state. The point $E_2 = (0.2,1.0,0.9)$ corresponding to $FE_2 = \{(h,0.2),(a,1.0),(s,0.9)\}$ certainly represents an emotional state for a derived emotion from anger and surprise. The point $(0,0,0)$ represents the entirely neutral state where no emotion is present.

Figure 7.3 shows how the unit cube could be further divided in order to represent basic emotions and their mixtures. In the subcubes denoted by a single emotion, the membership function of this emotion takes values in the interval [0.5, 1.0], whereas the membership values for the other emotions have intensities below 0.5. Therefore, it is reasonable to associate the subcube with this basic emotion. In the subcubes denoted with a sum of emotions (*e. g.*, *Surprise* + *Happiness*) memberships of these emotions are in the interval [0.5, 1.0], whereas the membership of the third emotion is below 0.5. Hence a derived emotion from these two basic emotions (*e. g.*, surprise and happiness) is assumed. The subcube where the membership values of all basic emotions are between 0.5 and 1.0 is denoted by the sum *Surprise* + *Anger* + *Happiness*.

If a general *n*-dimensional emotion hypercube is regarded certainly not all combinations of up to *n* emotions make sense. However, whether a combination is reasonable or not is certainly a psychological question.

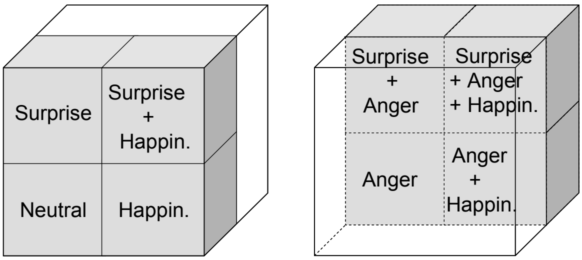


Figure 7.3 Subdivision of the unit cube representing basic and derived emotions

For the robot head MEXI we use this fuzzy emotion model to represent the output of the emotion recognition modules described in the next section. Furthermore, it is also useful for representing the state of MEXI's own emotions happiness, sadness, anger, and fear.

7.5 Emotion Recognition

The ability of perceiving and recognizing human emotions is a key feature to be realized by a robot to show emotional competence in its interactions with humans. Humans show their emotions by various means in their body language, speech, or facial expressions and usually also interpret these means all together in order to analyze the emotional state of others. However, facial expression and the prosody of natural speech are those representations of emotions that are most commonly analyzed by existing approaches in order to recognize the current emotional state of a human. An overview of current emotion recognition approaches can be found in [7].

In our approach we also rely on facial expressions and prosody for emotion recognition. For this purpose we realized the systems VISBER (Video Based Emotion Recognition) [22] and PROSBER (PROSody Based Emotion Recognition) [23] and integrated them into the robot head MEXI.

This section will give an overview of VISBER and PROSBER and the recognition results obtained by them. Furthermore, the principal architecture for a fuzzy emotion classification relying on the fuzzy emotion model described in Section 7.4, also pursued by VISBER and PROSBER, will be outlined.

7.5.1 Overview of VISBER

VISBER processes video sequences taken by a Philips webcam in the JPEG format with a frame rate of 30 frames per second.

Images have a size of 320 × 240 pixels and contain one face in frontal pose. We assume a homogeneous illumination for the video sequence but the background of the face may change. For classification of facial expressions we use a fuzzy rule-based approach that works on single images. It analyzes the deformation of a face using a size invariant, feature-based representation by a set of typical angles. The choice of angles for classification provides a size invariant classification and saves the effort for normalization. Automatic adaptation to the characteristics of individual human faces is achieved by an optional, short training phase that can be done before the emotion recognition starts. VISBER distinguishes the basic emotions happiness, sadness, anger, and fear.

In contrast to most existing approaches, blended emotions with varying intensities of these basic emotions as proposed by many psychologists can also be recognized and represented by the fuzzy emotion model described above (see also [20]).

Recognition results are shown in Table 7.1. They were obtained using at least three video sequences with a minimum of 100 frames for each emotion. The average recognition rate is 85%, which is comparable to existing systems where recognition rates from about 80 to 95% are reported.

Table 7.1 VISBER: confusion-matrix

Emotion	Happiness	Sadness	Anger	Fear	Neutral state	Sadness and fear
Happiness	75%	3%	1%	1%	20%	–
Sadness	0.5%	87%	–	–	12.5%	–
Anger	1%	1%	86%	–	12%	–
Fear	–	2%	–	86%	11%	1%
Neutral state	4%	5%	–		91%	–

7.5.2 Overview of PROSBER

PROSBER takes single sentences as input and, like VISBER, uses the basic emotions happiness, sadness, anger, and fear for classification. Speaker-dependent as well as speaker-independent emotion recognition is supported. Since the set of features relevant for a specific emotion differs considerably between different emotions, PROSBER autonomously selects up to six relevant speech features for each basic emotion from a set of 20 analyzed features. Among these analyzed parameters are, for instance, fundamental frequency, energy, speech/pause rate, speech speed, jitter, or shimmer of the input signal. The actual feature vector comprises statistical information about the dynamic course of some of these parameters. By the restriction to up to six features, on the one hand, over-fitting is avoided and, on the other hand, analysis effort is decreased to support real-time emotion recognition.

Table 7.2 shows the results obtained for speaker-independent recognition with six features. The best average recognition rates were achieved with a feature selection choosing six features. For speaker-dependent recognition the algorithm recognizes 84% of the test samples correctly with 5% ambiguity.

That is, for 5% of the test samples two or more fuzzy models reach the same maximum value. For speaker-independent recognition 60% of the test samples were recognized correctly without any ambiguous choices.

Table 7.2 PROSBER: confusion-matrix

Emotion	Happiness	Sadness	Anger	Fear	Neutral state
Happiness	23%	6%	38%	19%	4%
Sadness	0%	84%	4%	4%	8%
Anger	8%	0%	84%	4%	4%
Fear	8%	28%	12%	42%	12%
Neutral state	4%	24%	8%	4%	64%

7.5.3 Fuzzy Emotion Classification

Relying on the fuzzy emotion model presented in Section 7.4, a fuzzy rule-based approach for emotion classification is a quite natural choice. The principle structure of such a rule-based system is shown in Figure 7.4. For each basic emotion $e_i, i = 1,\dots,n$, to be recognized and hence distinguished in the fuzzy emotion model an independent fuzzy rule system is used. In our case four rule bases for the emotions anger, fear, happiness, and sadness are used by VISBER and PROSBER, respectively.

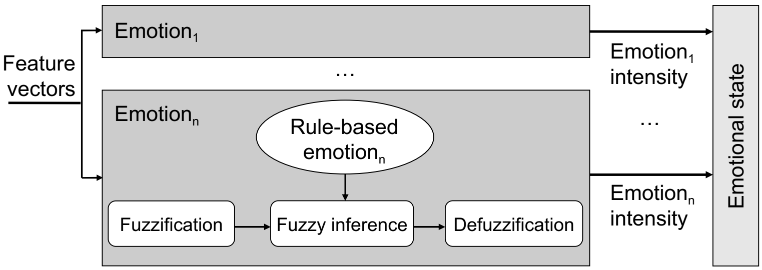


Figure 7.4 Principle structure of fuzzy emotion classification

Each rule takes fuzzified features $f_j, j = 1,\dots K, K \le 6,$ as input and produces a fuzzy emotion value I_{e_i} as output.

As already mentioned, VISBER uses six typical angles characterizing human facial expressions and PROSBER autonomously determines up to six features from a set of 20 selected speech features.

The rule system for VISBER [20] was designed based on psychological findings and empirical studies while the rule system for PROSBER was automatically generated from a set of training samples using a fuzzy grid method (cf. [23]).

We represent each feature f_j and emotion intensity I_{e_i} by three to five triangular or trapezoidal membership functions from the set *very low*, *low*, *medium*, *high*, and *very high*. However, the actual start and end coordinates as well as the maximum coordinates are generated automatically during a training phase for VISBER and PROSBER. This representation is simple enough to support real-time emotion recognition, yet allows to distinguish degrees to which a feature or emotion is present in the current input sentence. Furthermore, it is in line with the approaches used by psychologists, who often use two up to ten levels for characterizing psychological phenomena such as emotion intensities.

For the defuzzification of emotion values we use the center of gravity (COG) method. By projecting the COG onto the x-axis we calculate the corresponding emotion intensity. Hence a four-dimensional vector

$$(I_h, I_s, I_a, I_f) = (\mu_E(h), \mu_E(s), \mu_E(a), \mu_E(f))$$

containing the intensities of the four basic emotions happiness h, sadness s, anger a, and fear f is generated. This vector represents the membership values for each emotion and hence determines a point in a four-dimensional emotion hypercube (see Section 7.4).

The representation of an emotional state in a fuzzy emotion hypercube focuses on the static aspects of an emotion. However, emotional competence also includes the aspect emotion regulation. For this aspect, the dynamic behavior of emotions and drives also has to be considered as described in the next section.

7.6 Dynamics of Emotions and Drives

Emotions and drives, also called activating components in psychology, belong to the major factors determining human actions, besides rational elements. The motivation of human actions results from the interaction between these activating components and cognitive processes, which are responsible for deciding on the targets and plans of human actions.

Drives can be interpreted as subjective processes indicating a deficiency or a demand that causes the desire to satisfy this demand or to overcome the deficiency. Such a demand usually motivates and evokes human action for its satisfaction. An example is hunger or thirst, which causes the human action of walking to the breadbox or refrigerator. In psychology, drives are often viewed as homeostatic processes that motivate actions in order to reach and keep a certain balance.

Figure 7.5 shows how different behaviors, namely *coping behaviors*, are activated in a cyclical process of drive emergence and satisfaction (lower part). Furthermore, it shows the influence a drive and its state may have on emotions (upper part) and corresponding so called *expressive behaviors*. While expressive behavior is spontaneous and cannot be controlled consciously (*e. g.*, gestures, facial expressions, *etc.*), coping behavior is rationally motivated and tries to reach a certain goal (*e. g.*, drive satisfaction, work, *etc.*) [24].

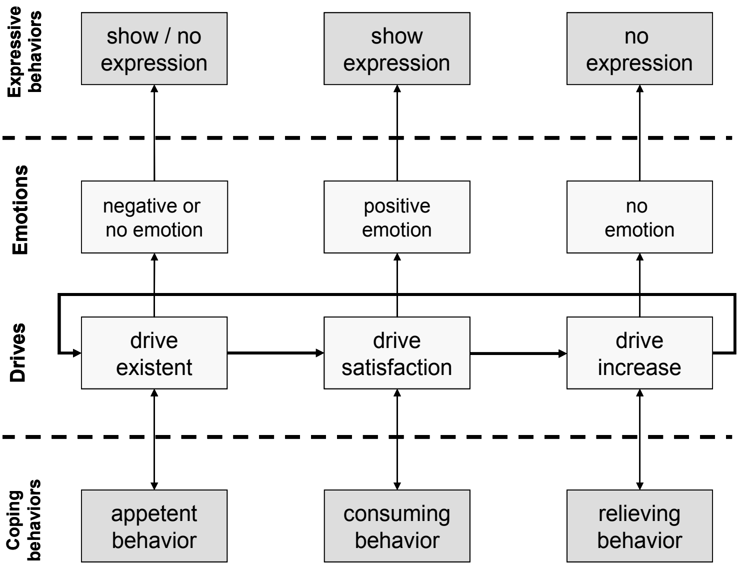


Figure 7.5 Drives and emotions as activating factors of action

If a drive is existent a corresponding appetent behavior is activated, *e. g.*, a search for food in the case of hunger. In order to satisfy the drive a consuming behavior (*e. g.*, eating) is performed when a perception or stimulus from the environment (*e. g.*, food) indicates this. This is followed by a relieving behavior (*e. g.*, sleeping) and a phase where the drive increases again. This cyclical increase and decrease of a drive also influences the emotional system. If a drive is existent over a longer time, this will certainly cause negative emotions, while the satisfaction of a drive is accompanied by a positive feeling and corresponding expressive behavior. When the drive is balanced or in its homeostatic area again, it is accompanied by a neutral emotional state.

Our robot head MEXI is endowed with a communication drive, a playing drive, and an exploration drive to realize pro-active behavior. Other drives may be useful, depending on the application domain. If MEXI's communication drive is very high, it looks for persons in its environment (appetent behavior). When it sees a person, the communication drive is satisfied by communicating with him or her and by following his or her face with its view (consuming behavior). This is accompanied by the emotion happiness, shown in MEXI's facial expression by activating a smiling behavior. Afterwards when the communication drive is satisfied, MEXI averts its gaze from the person and simply looks around (relieving behavior).

Accordingly, MEXI's basic behaviors can also be classified into expressive behaviors, resulting in different facial expressions or prosody of MEXI's utterances, and coping behaviors such as talking, playing, following a face, *etc.* While the expressive behaviors mainly depend on MEXI's emotional state, the coping behaviors are activated in response to its drive state.

Since a robot does not really *feel* emotions or drives, they have to be synthesized by a piece of software. In the following sections we show how we realized such artificial drives and emotions including their dynamics in the robot head MEXI, whose main purpose is to interact with people by communicating. These mechanisms are encapsulated in a specific software component, which we call *Emotion Engine* in the remainder of this chapter.

The principal modeling approach, however, is not restricted to MEXI but can be adapted to other application domains as well by introducing corresponding drives and emotions.

7.6.1 Drives

In order to realize a simple control mechanism for MEXI we mainly consider cyclical homeostatic drives as described above. Homeostatic drives motivate behavior in order to reach a certain level of homeostasis. Violation of the homeostasis is most likely to cause an according behavior the more the discrepancy between homeostasis and the actual value increases.

Each drive d_i is represented by a strength value ranging from –1 to 1. The value 1 represents a very large drive striving for its satisfaction and –1 expresses that a drive was overly satisfied. Drives have an upper and lower threshold th_u and th_l. Within the interval $]th_l, th_u[$ a drive is in homeostasis. Only when a drive's value is outside this interval, it strives to dominate the robot's behavior (the shaded areas in Figure 7.6) by configuring respective behaviors (see Section 7.7.2). This also avoids MEXI's behavior oscillating in an attempt to satisfy competing drives. In order to realize pro-active behavior MEXI's drives follow an internal regulation mechanism by automatically increasing and decreasing in a cyclical manner, even if no external stimulation is present.

The course of a drive $d_i(t)$ over time t is determined by the following equation, where $\Delta d_i(t)$ denotes the change between time points $t-1$ and t:

$$d_i(t) = d_i(t-1) + \Delta d_i(t), \text{ with } \Delta d_i(t) = c_{d_i} \cdot k_{d_i} \cdot \delta_{d_i}(t)\,.$$

c_{d_i} are positive values from the interval $]0, 0.5[$ that determine the gradient of a drive due to intrinsic regulation. They were determined experimentally. This intrinsic regulation happens even, if MEXI receives no perceptions or executes no behavior that could influence the respective drive d_i.

The acceleration factors $k_{d_i}(t)$ may accelerate or slow down the intrinsic regulation if their absolute value is >1 (acceleration) or <1 (slow down). They determine the influence of external stimuli or of MEXI's own behavior regarding a specific drive. How these factors are determined is described more closely below.

For a drive $d_i(t)$ the factor $k_{d_i}(t)$ can be calculated as follows:

$$k_{d_i}(t) = \begin{cases} k, \text{ if } p_{d_i}(t) = 1 \text{ and } \delta_{d_i}(t) = 1 \text{ or} \\ \qquad p_{d_i}(t) = -1 \text{ and } \delta_{d_i}(t) = -1, \\ \frac{1}{k}, \text{ if } p_{d_i}(t) = \frac{1}{2} \text{ and } \delta_{d_i}(t) = 1 \text{ or} \\ \qquad p_{d_i}(t) = -\frac{1}{2} \text{ and } \delta_{d_i}(t) = -1, \\ 1, \text{ otherwise,} \end{cases}$$

where $1 < k \leq \frac{1}{c_{d_i}}$ and $p_{d_i}(t) \in -1, -\frac{1}{2}, \frac{1}{2}, 1$ denotes the classification of the stimuli at the time t. The value -1 indicates stimuli that cause a steeper decrease, $-\frac{1}{2}$ or $\frac{1}{2}$ denote stimuli that cause a decelerated decrease or increase, and 1 indicates stimuli that accelerate a drive's increase.

For drives an additional factor $\delta_{d_i}(t) \in \{-1, 1\}$ determines the direction of their course. It is dependent on the previous values $\Delta d_i(t-1)$ and $d_i(t-1)$. When, for example, a drive becomes 1 at the time $t-1$ $(d_i(t-1) = 1)$ it starts decreasing due to MEXI's internal regulation mechanisms and $\delta_{d_i}(t)$ becomes -1, meaning that the drive is going to be satisfied. When a drive becomes -1 at $t-1$ $(d_i(t-1) = -1)$ it automatically starts increasing and $\delta_{d_i}(t)$ becomes 1. This factor enables the cyclical behavior of drives.

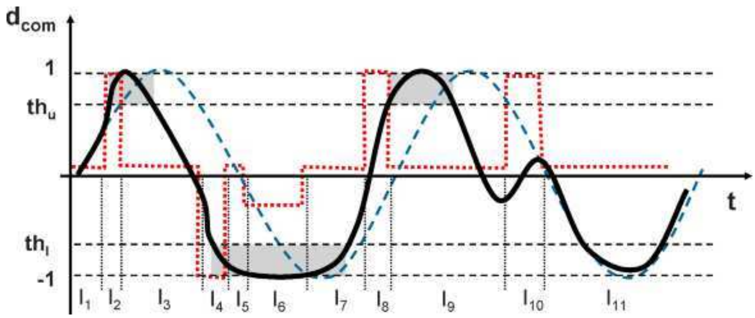


Figure 7.6 Example course of a communication drive

The principle development of a drive over time is shown in Figure 7.6. The solid curve shows the course of a drive d_i. In order to realize pro-active behavior MEXI's drives increase and decrease in a cyclical manner. As a default excitation MEXI's drives would follow a sine wave (dashed curve) and incorporate only internal regulation mechanisms. This is realized by assigning $\delta_{d_i}(t)$ the values 1 and –1 alternately. Stimuli, *i. e.*, perceptions and own behavior, influencing the drive are depicted as dotted line in Figure 7.6. Stimuli may accelerate a drive's increase (in the intervals I_2, I_8 is the stimuli $p_{d_i}(t) = 1$). The stimuli that satisfy the drive cause a steeper decrease (the interval I_4, $p_{d_i}(t) = -1$). In these intervals ($t \in \{I_2, I_4, I_8\}$) the acceleration factor is greater than 1 ($k_{d_i}(t) > 1$) and accelerates a drive's increase (I_2, I_8) or its decrease (I_4). Other stimuli may cause a slower increase or decrease of the drive. In this case the factor $k_{d_i}(t)$ has a value between 0 and 1 ($0 < k_{d_i}(t) < 1$). The slower increase is shown in interval I_6 ($p_{d_i}(t) = -\frac{1}{2}$), where a negative stimulus indicates its over-satisfaction. In some timing intervals MEXI's perceptions or their absence do not influence the depicted drive (indicated by a zero line of the stimuli, *e. g.*, intervals $I_1, I_3, I_7, \ldots$). In these cases the factor $k_{d_i}(t)$ is equal to 1 and the course of a drive runs in parallel with the excitation function.

Imagine, for example, that MEXI sees a human face in a state where its communication drive d_{com} is decreasing ($\delta_{\text{com}} = -1$). Since the current perception signals a potential communication partner for MEXI, its communication drive should be satisfied faster. This is reached by setting the acceleration factor $k_{\text{com}}(t)$ to a value $k > 1$. If no stimuli, that may influence the communication drive, are recognized $k_{\text{com}}(t)$ remains at 1 and does not accelerate the normal internal regulation of the communication drive. If, for instance, the person disappears (the face becomes smaller) the communication drive might decrease slower, and hence the value of $k_{\text{com}}(t)$ should be in the interval $]0, 1[$.

7.6.2 Emotions

For MEXI we distinguish a small set of basic emotions that can be represented easily by MEXI's facial expression and audio output, *i. e.*, anger, happiness, sadness, and fear. Happiness is a positive emotion, which MEXI strives for, while the others are negative ones that MEXI tries to avoid.

Each emotion e_i is represented by a strength value between 0 and 1. For each emotion a threshold *th* defines when the robot's behavior will be configured to show this emotion, *e. g.* by corresponding facial expressions (the shaded areas in Figure 7.7). This is done by increasing the gain values of the respective desired behavior(s) (see Section 7.7.2 and Figure 7.8).

Figure 7.7 shows as a solid curve the development over time of the positive emotion happiness. The dotted curve shows the duration and evaluation of MEXI's current perceptions. For positively evaluated perceptions the curve is above the time axis, for negative ones it is below the time axis.

In contrast to drives, for each emotion only one threshold th defines when the robot's behavior will be configured to show this emotion, *e. g.*, by corresponding facial expressions (shaded area) (see Section 7.7.2). Furthermore, emotions only have an automatic decay but do not automatically increase like drives.

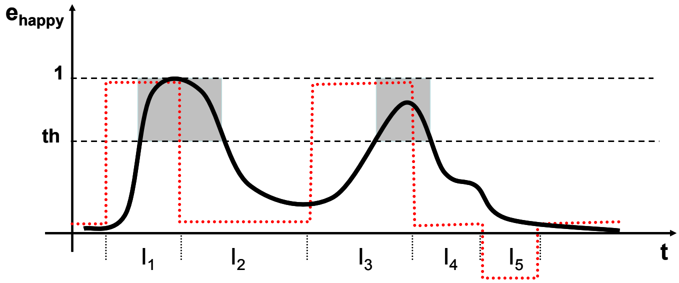


Figure 7.7 Example course of the emotion happiness

The course of an emotion $e_i(t)$ over time t is determined by the following equation, where $\Delta e_i(t)$ denotes the change between time points $t-1$ and t:

$$e_i(t) = e_i(t-1) + \Delta e_i(t), \text{ with } \Delta e_i(t) = c_{e_i} \cdot k_{e_i}(t),$$

where c_{e_i} are positive values in the interval $]0, 0.5[$ that determine the gradient of an emotion due to intrinsic regulation. They were determined experimentally. This intrinsic regulation happens even, if MEXI receives no perceptions or executes no behavior that could influence the respective emotion e_i.

The acceleration factors $k_{e_i}(t)$ may accelerate or slow down the intrinsic regulation if their absolute value is >1 (acceleration) or <1 (slow down). They determine the influence of external stimuli or of MEXI's own behavior regarding a specific emotion.

The acceleration factor $k_{e_i}(t)$ for an emotion e_i depends on the previous value of the emotion $e_i(t-1)$, the evaluation of the current perception and also on the current drive state. If a drive d_j concerning an emotion e_i increases, MEXI reacts in a neutral way, *i. e.*, $k_{e_i}(t) = 0$ and hence also $e_i(t) = 0$. If that drive starts decreasing ($\delta_j(t)$ changes its value from 1 to –1 and hence Δd_j becomes <0) then the emotion e_i increases very rapidly and $k_{e_i}(t) \geq 1$. The decrease of an emotion could also be accelerated by setting $k_{e_i}(t)$ to a value $k \leq -1$.

The increase of a positive emotion may be caused by positive perceptions of its environment (intervals I_1, I_3) and by the drives (their fulfillment). Hence, for $t \in \{I_1, I_3\}$ the acceleration factor $k_{e_i}(t)$ is larger than 1 ($k_{e_i}(t) > 1$). The decrease happens automatically with a certain adaptable amount per time unit, if the positive stimulus has disappeared (intervals I_2, I_4). In this case the acceleration factor $k_{e_i}(t)$ is equal to –1. By a negative stimulus the decrease of $e_i(t)$ is accelerated by setting $k_{e_i}(t) < -1$ (interval I_5). How these factors are determined is described by an example below.

Imagine for instance the communication drive d_com and its impact on the emotion happiness e_happy. Assume that it is initially not satisfied ($d_\text{com} > th_\text{u}$ and $\Delta d_\text{com} < 0$). Then happiness may be increased by a positive perception perhaps a human face. This should result in an accelerated increase of e_happy, which is reached by setting the respective acceleration factor $k_\text{happy}(t) > 1$. If the face suddenly disappears, happiness decreases ($k_\text{happy}(t) = -1$). If according stimuli are recognized, happiness could decrease even faster resulting in $k_\text{happy}(t) < -1$. For $\Delta d_\text{com}(t) > 0$ the factor $k_\text{happy}(t)$ can be calculated as follows

$$k_\text{happy}(t) = \begin{cases} -1, & \text{if } e_\text{happy}(t-1) \in\ ]0,1], \\ 0, & \text{if } e_\text{happy}(t-1) = 0, \end{cases}$$

and for $\Delta d_\text{com}(t) < 0$

$$k_\text{happy}(t) = \begin{cases} 1, & \text{if } e_\text{happy}(t-1) \in [0,1[\text{ and } p_{e_\text{happy}}(t) = 1, \\ k, & \text{if } e_\text{happy}(t-1) \in [0,1[\text{ and } p_{e_\text{happy}}(t) = \frac{1}{2}, \\ -k, & \text{if } e_\text{happy}(t-1) \in\ ]0,1] \text{ and } p_{e_\text{happy}}(t) = -\frac{1}{2}, \\ -1, & \text{if } e_\text{happy}(t-1) \in\ ]0,1] \text{ and } p_{e_\text{happy}}(t) = -1, \\ 0, & \text{if } e_\text{happy}(t-1) = 1 \quad \text{and } p_{e_\text{happy}}(t) > 0 \text{ or} \\ & \quad e_\text{happy}(t-1) = 0 \quad \text{and } p_{e_\text{happy}}(t) < 0, \end{cases}$$

where $1 < k \le \frac{1}{c_{e_\text{happy}}}$ und $p_{e_\text{happy}}(t) \in -1, -\frac{1}{2}, \frac{1}{2}, 1$ denotes the classification of the stimuli at the time t. The value -1 indicates stimuli that cause a steeper decrease, $-\frac{1}{2}$ or $\frac{1}{2}$, denotes stimuli that cause a decelerated decrease or increase and 1 indicates stimuli that accelerate an emotion's increase.

Since MEXI's emotions are modeled by strength values between 0 and 1, a certain emotional state can also be represented graphically in the fuzzy emotion model proposed in Section 7.4 as a point in a four-dimensional fuzzy hypercube.

7.7 Control of Robot Behavior by Emotions and Drives

In this section we show how artificial drives and emotions as proposed above are used to control the behavior of a robot in order to show emotional competence. We realized the concepts in our robot head MEXI as a feedback loop between *Emotion Engine*, *Behavior System*, and the environment as depicted in Figure 7.8.

MEXI perceives inputs from the environment via two cameras and a microphone and processes them in the *Perception* component. MEXI's output to the

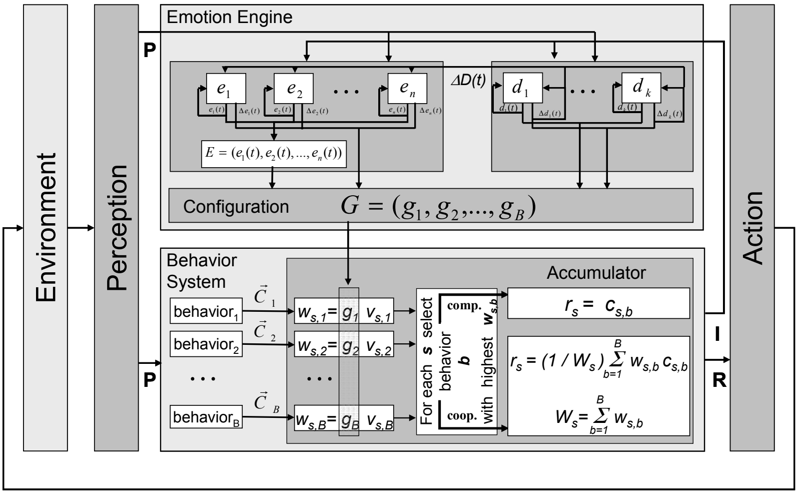


Figure 7.8 Control of the behavior system by the emotion engine

environment is generated by the *Action* component, which controls MEXI's movements realized via 15 model craft servo motors and speech utterances via a built-in loudspeaker. The Behavior System combines appropriate basic behaviors and produces the corresponding nominal values for controlling MEXI's actors represented by the Action component. Without control by the Emotion Engine, the Behavior System would select behaviors in a purely reactive fashion depending on its current perceptions, which also includes, *e. g.*, the emotions of its human counterpart. An overview of the realization concepts of the Behavior System is given below. In order to show emotional competence, MEXI has to select behavior that on the one hand is adequate with respect to the emotions of its human counterpart and on the other hand also maintains and regulates MEXI's own emotional state in such a way that MEXI's drives are kept in the homeostatic area, and negative emotions are avoided while positive ones are reinforced. For this purpose the Behavior System is configured by the Emotion Engine in a feedback loop as described below. A more detailed description of this architecture can be found in [25].

7.7.1 Behavior System

For realizing MEXI's *Behavior System* the paradigm of *Behavior Based Programming* developed by Arkin [9] is applied.

So-called basic behaviors b like *Smile*, *LookAround*, *FollowFace*, *AvertGaze*, or *Sulk*, which may be either cooperative or competitive, are mixed by an *Accumulator* to compute the nominal vector R for the actor system from the fixed-sized vectors C_b . A C_b generated by the basic behavior b contains 15 triplets $(c_{s,b}, v_{s,b}, m_{s,b})$ (one for each servo motor s realizing MEXI's movements) consisting of the nominal value $c_{s,b}$, a vote $v_{s,b}$ for each of the nominal values, and a mode flag $m_{s,b}$ (cooperative versus competitive) (see Figure 7.8).

Using the vote values a behavior can signal whether the output is of high importance $v_{s,b} = 1$ or should have no influence at all $v_{s,b} = 0$. Via this mechanism the influence of MEXI's external perceptions, *e. g.*, a human face or its toys, on its behavior is realized. Apart from that also its internal emotional state may influence the actual behavior by setting appropriate *gain* values as described below. The accumulator combines external votes and internal gains of a behavior b by calculating a weight $w_{s,b} = g_{s,b} \cdot v_{s,b}$ as their product. This weight is used in order to rank the behaviors. If the behavior b with highest weight is competitive, a winner-takes-all strategy is used and the resulting nominal value for servo s is $r_s = c_{s,b}$. Else a weighted median of all cooperative values is calculated. The vector I shows which behavior is really executed and to what extent it influenced the actually set nominal values for the servo motors. These calculations are described in more detail in [25].

7.7.2 Control of the Behavior System by the Emotion Engine

The *Emotion Engine* maintains MEXI's artificial drives and emotions via the dynamic processes described in the previous section. Depending on the current state of emotions and drives the behavior of the Behavior System has to be configured in such a way that appropriate emotion regulation is achieved. This is done via the *Configuration* component.

Via this component the Emotion Engine determines for each basic behavior $b_i, i = 1,2,...,B,$ a so-called *gain* value $g_i, i = 1,2,...,B,$ that indicates how adequate the behavior is for MEXI's current emotional state and hence how important it is that the Behavior System selects this behavior for execution at a certain point in time t (cf. Figure 7.8). For the calculation of gains we distinguish between *expressive* and *coping* behaviors [24]. As already stated, an expressive behavior is usually spontaneous and unconscious and does not try to do anything (like smiling to express happiness), while coping behavior tries to reach some goal as for instance working actions or satisfaction of drives.

MEXI's *expressive behaviors* comprise behaviors involved in the generation of its facial expressions or emotional prosody of speech utterances like *Smile* or *Sulk*. They are used to represent MEXI's emotional state and mainly depend on the current values of emotions. If a perception causes a certain emotion in MEXI the emotion strength is increased. If a certain threshold is reached the gains for according behaviors, *e. g.*, those that generate the corresponding facial expression are set to 1 and an according speech output and prosody are produced.

Coping behaviors are those that are used to satisfy MEXI's drives, such as communication with humans (*FollowFace*) or playing with its toy (*FollowToy*). Via these behaviors MEXI tries to keep its drives at a homeostatic level. Therefore, they mainly depend on the current strength values of MEXI's drives. If a drive increases the gain for the *consuming* behaviors, that will satisfy that drive, is increased by a certain (variable) amount per time. If a drive decreases the gains for the consuming behaviors are decreased respectively. Conflicts between drives may be solved, for instance, by a fixed priority.

This dependency between expressive behaviors and emotions on the one hand and coping behaviors and drives on the other hand is also reflected in the calculation of gain values for behaviors of these two categories shown below.

The following formula shows an example for the calculation of the gain values g_{Sm} for the expressive behavior *Smile*. The formula is also applied for the example session described in Section 7.8.

$$g_{\mathrm{Sm}} = \begin{cases} 1, & \text{if } th \le e_{\mathrm{happy}} \le 1, \\ g_{\mathrm{Sm}} + c_{\mathrm{Sm}}\,\Delta e_{\mathrm{happy}}, & \text{if } 0 < e_{\mathrm{happy}} < th, \\ 0, & \text{otherwise,} \end{cases}$$

where $c_{\mathrm{Sm}} > 0$ is a constant and is determined experimentally. The gain g_{Sm} depends on the emotion happiness e_{happy}, its actual change $\Delta e_{\mathrm{happy}}$, and on the threshold th . It has to be noted that $\Delta e_{\mathrm{happy}} \le 0$ since the excitation function for happiness realizes an automatic decrease as soon as the maximum emotion strength of 1 is reached.

The formulas for calculating the gain g_{FF} of the coping behavior *FollowFace* are shown below:

$$g_{\mathrm{FF}} = \begin{cases} 1, & \text{if } \; 0 < d_{\mathrm{com}} < th_{\mathrm{u}} \text{ and } \Delta d_{\mathrm{com}} < 0, \\ & \text{or } \; 1 \ge d_{\mathrm{com}} \ge th_{\mathrm{u}}, \\ g_{\mathrm{FF}} + c_{\mathrm{FF}}\,\Delta d_{\mathrm{com}}, & \text{if } -1 \le d_{\mathrm{com}} \le 0 \text{ and } \Delta d_{\mathrm{com}} < 0, \\ & \text{or } \; 0 < d_{\mathrm{com}} < th_{\mathrm{u}} \text{ and } \Delta d_{\mathrm{com}} > 0, \\ g_{\mathrm{FF}} - c_{\mathrm{FF}}\,\Delta d_{\mathrm{com}}, & \text{if } -1 \le d_{\mathrm{com}} \le 0 \text{ and } \Delta d_{\mathrm{com}} \ge 0, \end{cases}$$

where $c_{\mathrm{FF}} > 0$ is a constant and is determined experimentally.

The gain g_{FF} depends on the communication drive d_{com}, the direction of its change Δd_{com} ($\Delta d_{\mathrm{com}} < 0$ indicates decrease and $\Delta d_{\mathrm{com}} > 0$ an increase of d_{com}) and from the upper and lower thresholds of the drive th_{u} and th_{l} .

MEXI's behavior configuration using gain values calculated by the formulas above is illustrated in the next section.

7.8 Example of MEXI's Emotional Competence

For an example of how MEXI reacts adequately to human emotions, look at the diagrams in Figure 7.9 that show how MEXI behaves when at time t_0 a person appears and MEXI sees a human face. At t_1 MEXI recognizes that the person has a sad facial expression and tries to distract her from that feeling by playing (MEXI plays by getting shown its toy and tracking it). At t_3 MEXI detects the toy and plays with it until it becomes boring and MEXI looks at the human face again at t_4.

The upper two diagrams show the votes $v_{s_h,b}$ and $v_{s_m,b}$ for two groups of servo motors involved in the execution of different sets of behaviors. Motors for head and eye movements $s_h \in \{1,...,7\}$ are for instance needed for the behaviors *FollowFace, FollowToy*, or *LookAround*, which are competitive, but not for *Smile* or *Sulk*, which are cooperative behaviors. On the other hand, motors for movements of the mouth corners, $s_m \in \{8,...,11\}$, are needed for *Smile* or *Sulk* and not for the other behaviors. For behaviors not involving a group of servos the corresponding votes are constantly set to 0. The votes of the other behaviors depend on MEXI's current perceptions and somehow reflect its current action tendency due to these perceptions not taking into account its internal state.

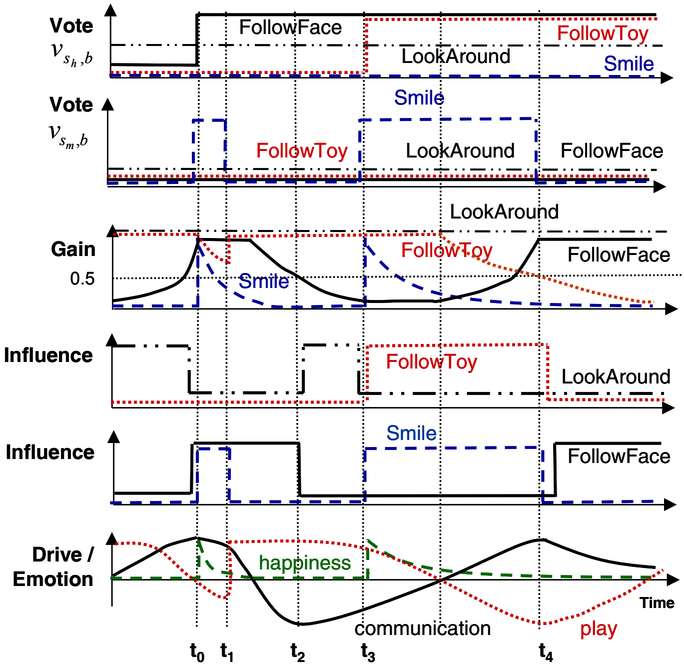


Figure 7.9 Example session

The votes $v_{s_h,b}$ for the behavior $b = FollowFace$ and for $b = FollowToy$ are set to 1 when MEXI sees a human face (t_0 and afterwards) or its toy respectively (t_3 and afterwards). The behavior *LookAround*, which lets MEXI look for interesting things (*e.g.*, human faces), corresponds to MEXI's exploratory drive. It describes a kind of default behavior for the head movement motors. Therefore, the votes $v_{s_h,b}$ for $b = LookAround$ are constantly set to 0.5 and the gain is set to 1.0.

The behavior *Smile* only involves the motors for mouth movements. Since MEXI by default should be friendly, the corresponding votes $v_{s_m,b}$ for b = *Smile* are set to a high value of 0.8 when MEXI perceives a human face (starting at t_0). When MEXI classifies the human's face as sad at t_1, it adapts its own facial expression to a neutral expression and no longer intends to smile. Hence, the votes $v_{s_m,b}$ for *Smile* are set to 0 and remain there until MEXI sees its toy at t_3 and wants to smile again due to this perception ($v_{s_m,b} = 0.8$ at t_3).

The gain values reflect MEXI's action tendency due to its current internal state. The weight combines gains and votes, *i. e.*, internal and external action tendencies, and the influence shows which behavior is really executed (and to what extent in the case of cooperative behaviors). The influence of *LookAround* is 1 until t_0 and MEXI is looking around since nothing else is seen. When MEXI's communication drive d_{com} increases the gain for *FollowFace* (g_{FF}) also rises until d_{com} reaches its upper threshold ($th_{\text{u}} = 1$) and also g_{FF} is set to 1 at t_0. Since MEXI sees a face at t_0, the weight of *FollowFace* also becomes 1 and *FollowFace* is executed (its influence is 1) because its weight is now higher than that of *Look-Around*. At t_1 MEXI detects the sad face of its human counterpart and stops smiling (vote of *Smile* is set to 0, hence the weight and influence become 0 too). Since MEXI now wants to distract the human from her sadness by playing with its toy, its own *play* drive increases to 1, causing the gain for *FollowToy* to increase as well. Since MEXI now wants to play, the communication drive starts decreasing faster until it becomes 0 at t_2. Now the weight of *LookAround* is lager than that of *FollowFace* and MEXI looks around (influence of *LookAround* = 1 from t_2 to t_3) until at t_3 the weight of *FollowToy* is larger than that of *LookAround*, since MEXI detects its toy at t_3 and the corresponding vote $v_{s_h,FollowToy}$ becomes 1. Then the pla*y* drive and due to this the gain of *FollowToy* start decreasing until the drive is –1 ($th_{\text{l}} = -1$) at t_4. Since at t_4 the communication drive has reached its upper threshold ($th_{\text{u}} = 1$), the gain of *FollowFace* is set to 1 and its weight is the highest again resulting in execution of *FollowFace* after t_4 (influence $= 1$).

The emotion *happiness* is set to 1 each time the influence of *FollowFace* or *FollowToy* starts rising (t_0, t_3), *i. e.*, the communication drive or the playing drive starts being satisfied (δ_{com} is set to –1). As result of the emotion *happiness* rising (threshold = 1) the gain value of the behavior *Smile* also is set to 1 and MEXI starts smiling, because the vote is also 1 (t_0 and t_3). *Happiness* automatically decreases caused by the respective excitation function (*e. g.*, from t_3 until t_4). This decrease becomes faster by setting $k_{\text{happy}} < -1$, when MEXI receives corresponding perceptions such as a sad human face (t_1). MEXI smiles until the weight of *Smile* becomes 0 because of MEXI's perception ($v_{s_m,Smile} = 0$ at t_1 and t_4). Another reason to stop smiling may be that the gain of *Smile* has decreased to 0 because of a corresponding decrease of *happiness* due to its excitation function e_{happy}.

This example illustrates how artificial emotions and drives can be used to dynamically control the selection of adequate emotional robot behavior.

7.9 Conclusions

In this chapter we presented concepts for integrating emotional competence into a robotic system and showed how the behavior-based architecture of the robot head MEXI supports these concepts. The chapter concentrates on two important aspects of emotional competence: how internal regulation mechanisms for emotions and drives can be realized and how adequate behaviors for external emotion regulation as well as for reaction on human emotions can be incorporated. MEXI also integrates the components VISBER and PROSBER for emotion recognition from facial expression and natural speech as outlined above. Hence all four aspects of emotional competence are realized. Based on these building blocks MEXI can communicate with humans in real time. We presented MEXI's software architecture and showed how it is used to realize its actions without any explicit world model and goal representation. Instead, MEXI's artificial emotions and drives maintained by the Emotion Engine are used to evaluate its perceptions and control its future actions in a feedback loop. The underlying Behavior System and the Emotion Engine are based on the behavior-based programming paradigm extending Arkin's motor schemes to a multidimensional model of reactive control. This architecture supports a constructive approach for synthesizing and representing MEXI's artificial emotions and drives rather than emulating human ways of *feeling*. Our experiences with MEXI at different public exhibitions and fairs show that MEXI, although realizing only a restricted set of emotions and drives, attracts human spectators and maintains their communication interest. For future work it would be interesting to psychologically evaluate how different humans react to MEXI. Also an extension of MEXI's communication behavior to other tasks by integrating appropriate actor facilities and basic behaviors into the Behavior System and the Emotion Engine could be envisaged.

References

1. Picard R.W.: Affective Computing. MIT Press, Cambridge, MA (2001)
2. Salovey P., Mayer J.D.: Emotional intelligence. Imagination, Cognition, and Personality **9**:185–211 (1990)
3. Seidel W.: Emotionale Kompetenz. Gehirnforschung und Lebenskunst. Elsevier-Spektrum Akademischer, Munich (2004)
4. Goleman D.: Emotional Intelligence: Why it Can Matter More than IQ. Bantam Books, New York (1996)
5. Ishiguro H., Nishio S.: Building artificial humans to understand humans. J. Artificial Organs **10**(3):133–142 (2007)
6. Hanson D.: Exploring the aesthetic range for humanoid robots. In: Proceedings of the Cognitive Science (CogSci 2006) Workshop on Android Science, Vancouver, Canada, pp. 16–20 (2006)
7. Zeng Z., Pantic M., Roisman G.I., Huang T.S.: A survey of affect recognition methods: audio, visual, and spontaneous expressions. IEEE Trans. Pattern Anal. Mach. Intell. **31**(1):39–58 (2009)

8. Miwa H., Itoh K., Matsumoto M. *et al.*: Effective emotional expressions with emotion expression humanoid robot WE-4RII. In: Intelligent Robots and Systems, 2004 (IROS 2004). Proceedings of the 2004 IEEE/RSJ International Conference, pp. 2203–2208 (2004)
9. Arkin R., Fujita M., Takagi T., Hasekawa T.: An ethological and emotional basis for human–robot interaction. Robotics Autonomous Sys. **42**(3–4):191–201 (2003)
10. Arkin R.: Homeostatic control for a mobile robot, dynamic replanning in hazardous environments. In: Proceedings of the SPIE Conference on Mobile Robots, Cambridge, MA, pp. 240–249 (1988)
11. Canamero L., Fredslund J.: I show you how I like you – can you read it in my face? IEEE Trans. Sys. Man Cybernet. Part A: Sys. Humans **31**(5):454–459 (2001)
12. Tomkins S.S.: Affect theory. In: Scherer K.R., Ekman P. (eds.) Approaches to Emotion, pp. 163–195. Lawrence Erlbaum, Hillsdale, NJ (1984)
13. Velasquez JD.: Modelling emotion-based decision-making. In: Proceedings of the 1998 AAAI Fall Symposium Emotional and Intelligent: The Tangled Knot of Cognition, Orlando, FL, pp. 164–169 (1998)
14. Breazeal C.: Affective interaction between humans and robots. In: Kelemen J., Sosk P. (eds.), Proceedings of the ECAL 01, pp. 582–591. Springer, Prague (2001)
15. Breazeal C.: Emotion and sociable humanoid robots. Int. J. Human Comp. Stud. **59**(1–2):119–155 (2003)
16. MIT builds Nexi the emo robot. http://www.robotworldnews.com/100531.html (2008). Accessed April 4, 2008
17. Lazarus RS.: Emotion and Adaptation. Oxford University Press, New York (1991)
18. Plutchik R.: The Emotions. University Press of America (1991)
19. Ortony A., Turner W.: What's basic about basic emotions? Psychological Rev. **97**(3):315–331(1990)
20. Esau N., Wetzel E., Kleinjohann L., Kleinjohann B.: A fuzzy emotion model and its application in facial expression recognition. In: Proceedings of the ASME 2007 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC/CIE 2007) (2007)
21. Kosko B.: Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence. Prentice Hall, Englewood Cliffs, NJ (1992)
22. Esau N., Wetzel E., Kleinjohann L., Kleinjohann B.: Real-time facial expression recognition using a fuzzy emotion model. In: Fuzzy Systems Conference, 2007 (FUZZ-IEEE 2007, pp. 1–6). IEEE International (2007)
23. Austermann A., Esau N., Kleinjohann B., Kleinjohann L.: Prosody based emotion recognition for MEXI, In: Intelligent Robots and Systems, 2005 (IROS 2005). Proceedings, 2005 IEEE/RSJ International Conference, pp. 1138–1144 (2005)
24. Maslow A.: A theory of human motivation. Psychological Rev. **50**: 370–396 (1943)
25. Esau N., Kleinjohann L., Kleinjohann B.: Emotional communication with the robot head MEXI. In: Proceedings of the International Conference on Control, Automation, Robotics and Computer Vision (ICARCV 2006) (2006)

Chapter 8
Shape Design and Modeling with 1/*f* Fluctuation Characteristics

Katsuhiro Maekawa[1]

Abstract This chapter first describes the extraction and analysis of 1/*f* fluctuation characteristics involved in the shape of ceramics. The surface topography of various teacups is measured using X-ray CT or laser metrology, and the wave patterns are then analyzed using the FFT method to clarify power spectrum–spatial frequency characteristics. Next, a couple of methods for shape design that utilize the power spectrum characteristics are explained. Finally, the modelling of aesthetically pleasing objects is demonstrated using the rapid prototyping method.

8.1 Introduction

In general, the value of a commercial product is defined as performance divided by cost, but consisting of many other factors, including functional capability, usability, maker brand, and eco-friendliness. Design to please the eye is another factor that adds value to a product. As new products having a high value reflect this trend, we tend to select a product by relying on our aesthetic feeling. The aesthetics of a product are influenced by such elements as shape, texture, color, pattern, and visual harmony with its environment. These qualities are related to individual sensitivity or kansei [1], and to the individual's cultural background, so it is difficult to characterize them quantitatively [2].

Fluctuation is a spatial and time change that is difficult to predict. Figure 8.1 shows typical patterns of the power spectrum as a function of frequency in a double logarithmic plot. The spectrum is defined as the root-mean-square wave amplitude contained in every frequency. White noise is not dependent on frequency, whereas the spectrums of $1/f$ and $1/f^2$ fluctuations change with frequency; and the inclinations of these characteristic lines are –1 and –2, respectively. It is worth

[1] K. Maekawa
Faculty of Engineering, Ibaraki University,
4-12-1 Nakanarusawa, Hitachi, Ibaraki 316-8511, Japan
e-mail: mae@mx.ibaraki.ac.jp

pointing out that these three fluctuations tend to give human beings different impressions: disorder with white noise, aesthetic and affective satisfaction with the $1/f$ fluctuation, and blandness and monotony with the $1/f^2$ fluctuation [3].

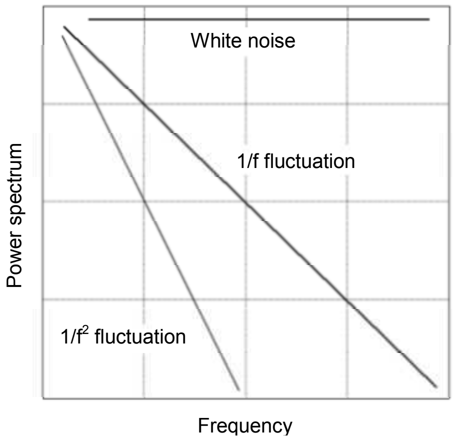


Figure 8.1 Logarithmic diagram of power spectrum with typical fluctuations

We can see $1/f$ fluctuation characteristics in various natural phenomena, such as breezes, streams, candle flames, and the luminous patterns of fireflies [4]. The $1/f$ fluctuation is also introduced into many industrial products, such as the control of air-conditioning equipment, and the control of the oscillations of massage tools [1]. However, no application to shape design has been reported.

The design and modeling of aesthetic forms is a matter of concern in this chapter. First, shape fluctuation is investigated using noncontact metrology and analyzed by Fourier transformation in order to highlight eye-pleasing characteristics. Next, a computer-based design of an affectively satisfying shape with kansei features is described. Incorporated into the computer-aided design, rapid prototyping technology is utilized to generate complicated plastic models with satisfactory dimensional accuracy. Finally, the possibility of using artificial patterns with $1/f$ characteristics is discussed to envisage the kansei design.

8.2 Shape Extraction from a Japanese Teacup

8.2.1 Experimental Method

Objects to be measured are Japanese teacups as shown in Figure 8.2: both handmade and mass-produced ones. In order to extract features contained in their outer surfaces, we tried two kinds of methodology: X-ray CT measurement and laser-CCD displacement sensing.

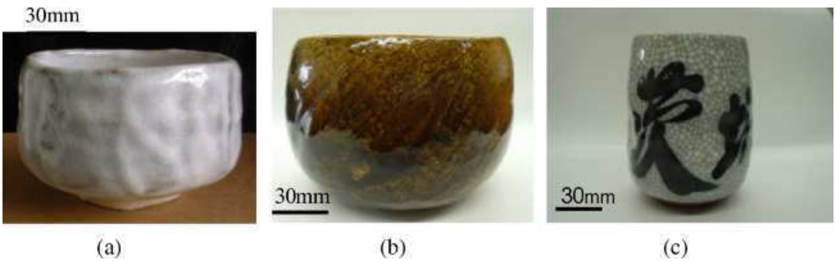


Figure 8.2 Japanese teacups used for measurement: (a) teacup A, (b) teacup B, and (b) mass-produced teacup

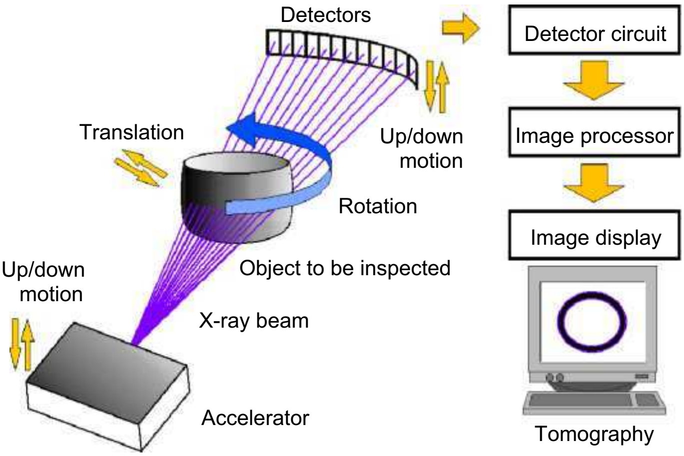


Figure 8.3 Schematic of X-rayCT measurement

Figure 8.3 illustrates the X-ray CT equipment used for contour measurement. The computer-assisted tomography scanner can create sectional images; X-rays that penetrate an object to be inspected are detected, then the signals are filtered and processed, and finally the image is displayed on a computer screen. Figure 8.4 shows the image of a cross-section at a distance of 26.6 mm from the bottom. As many as 235 such images are generated at vertical intervals of 350 μm. One pixel of the picture corresponds to 350 μm and has 8-bit or 256 gradations.

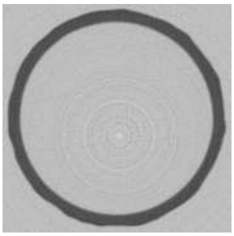
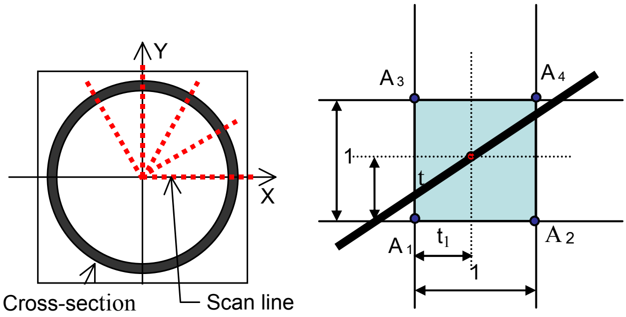


Figure 8.4 X-ray CT image of teacup (ϕ120 mm) cross-section

It is necessary to construct contour data of the outer surface from each bit-mapped X-ray CT. As shown in Figure 8.5 (a), the center of the round image is regarded as the origin from which 1024 scan lines are spaced radially and equally. When scanning toward the origin, an intersection between the outer surface image and the line gives the coordinates of the outer surface. The intersection is determined by the luminance of each pixel. For example, as shown in Figure 8.5 (b), the luminance L_x at point X along the scan line can be calculated by

$$L_x = \{L_{A1} + (L_{A2} - L_{A1})t_1\}(1 - t_2) + \{L_{A3} + (L_{A4} - L_{A3})t_1\}t_2, \tag{8.1}$$



(a) (b)

Figure 8.5 Logarithmic diagram of power spectrum with typical fluctuations: (a) scan lines, and (b) calculation of luminance

where L_{A1}, L_{A2}, L_{A3}, L_{A4} denote the luminance at the four corners, and t_1 and t_2 are parameters ranging from 0 to 1. Concerning the judgment of the outer contour, it is necessary to set a luminance threshold.

The contour data thus obtained from each cross-section are filtered to reduce noise, finally being composed to define polygonal surfaces with triangular patches. The data covers an area of 27.3–71.75 mm from the bottom of the teacup. The fast Fourier transformation (FFT) is applied to the frequency analysis, in which each power spectrum is added at every contour data and the average is finally taken. Note that the spatial frequency is defined as the number of wave periods per unit distance. In this case the unit distance is taken as the circumference of a teacup.

The other three-dimensional shape measurement system is shown in Figure 8.6, consisting of a z-axis stage, a turntable, a laser-CCD displacement sensor, a stage controller, and a personal computer for system control. A teacup is placed on the stage for rotation to sample the surface topography at a rate of 2048 points per rotation. The measurement is repeated by 143 layers with a vertical interval of 210 μm, so that the whole cross-section data for contour shape is obtained. Taking the shape of the teacup and the measurement range of the sensor, a range of 45 mm to 75 mm from the bottom was measured. Table 8.1 summarizes the laser-CCD displacement sensing conditions.

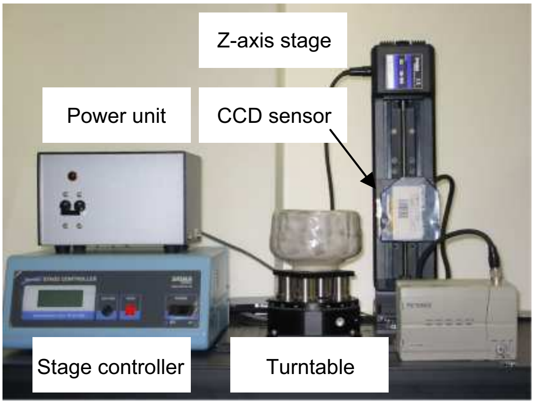


Figure 8.6 Laser-CCD displacement sensing apparatus

Table 8.1 Laser-CCD displacement sensing conditions

Photoflood lamp	Semiconductor laser with 670 nm wavelength
Photo acceptance	CCD
Measurement distance	30 mm
Measurement range	Standard distance±5 mm
Power input	100 V AC
Output	0.95 mW maximum
Laser beam spot	φ30 μm at standard distance
z-axis resolution	1 μm
Sampling cycle	512 μs

When the contour of a teacup is assumed to be a perfect circle, the system measures not only fine concavities and convexities on the surface but also deviations from the circular form. These data are regarded as waves involved on the perfect circle. In the power spectrum analysis, the FFT method enables us to decompose complicated waves into frequency components. A power spectrum of these components yields the topographical features of a teacup on the spatial frequency axis.

However, it is difficult to align the center of the turntable with that of the teacup since the object to be measured is somehow out of round. The influence of this eccentricity may appear in the measured data. Figure 8.7 (a) shows an off-centered circle with the radius r' and the deviation a. When θ is defined as the angle between the scan line and the x-axis, the distance between the intersection and the origin can be expressed as a function of θ:

$$r(\theta) = a\cos(\theta+\varphi) + \sqrt{a^2\cos^2(\theta+\varphi) - (a^2 - r'^2)}, \tag{8.2}$$

where φ is the angle between the line OO′ and the x-axis.

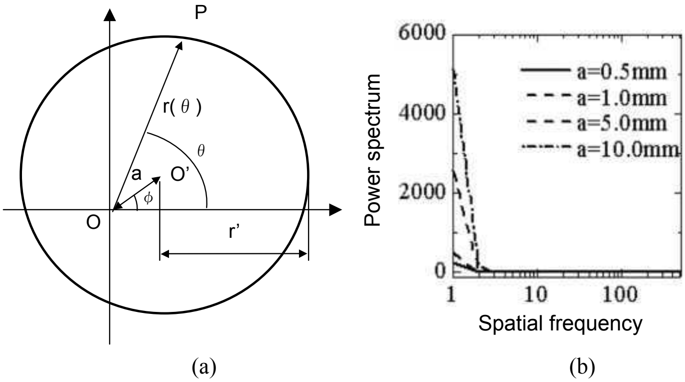


Figure 8.7 Laser-CCD displacement sensing apparatus: (a) eccentric circle, and (b) power spectrum characteristics

For the 1024 scan lines, $r(\theta)$ was calculated by Equation 8.2 when r' is equal to 60 mm and a is assumed to be 0.5, 1.0, 5.0, or 10.0 mm. Figure 8.7 (b) shows the result of the FFT analysis. The power spectrum increases with increasing deviation at a low frequency of less than 3. Since the higher frequency region is a matter of concern, as will be shown later, the influence of eccentricity can be neglected.

8.2.2 Results and Discussion

In terms of the X-ray CT measurement of Teacup A shown in Figure 8.2 (a), Figure 8.8 shows the calculated contours with three different luminance thresholds: 96, 135, and 157. It is found that the contour line appears closer to the center when the threshold is decreased: a diameter difference of about 1% is involved.

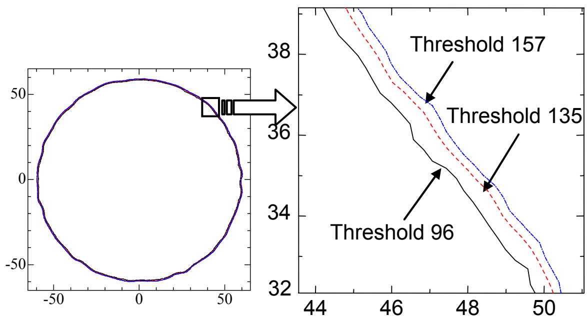


Figure 8.8 Calculated contours with varying luminance thresholds

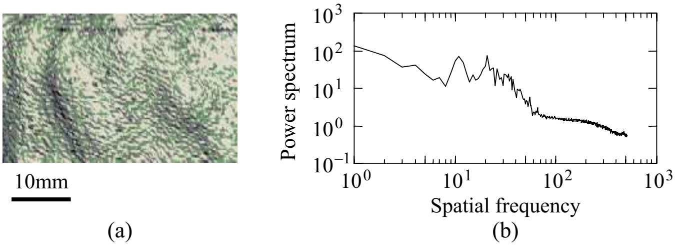


Figure 8.9 Reconstructed surface topography and power spectrum characteristics: (a) CG image, and (b) power spectrum characteristics

Figure 8.9 shows the teacup surface created on the basis of the contour data and its power spectrum as a function of the spatial frequency when the luminance threshold is set at 135. With the threshold varying from 96 to 157, the power spectrum was similar in the low-frequency region of less than 70, whereas the inclination of the characteristic line is slightly different in the higher region: –0.80 for a threshold of 96, –0.97 for 135, and –0.81 for 157. In the high-frequency region of more than 70, the slope has a value of –1, which means that the surface can be characterized by the 1/*f* fluctuation. Note that the spatial frequency 70 is equivalent to a wavelength of 5.38 mm.

The surface texture reconstructed using the contour data yields a rough topography. This feature corresponds with the power spectrum in the low-frequency region of less than 70. A fine, uneven texture is duplicated at a threshold of 135, where the characteristic line slope has a value of –0.97.

For the laser-CCD displacement sensing, on the other hand, Figure 8.10 (a) shows the FFT results of Teacup B shown in Figure 8.2 (b), while Figure 8.10 (b) depicts that of the mass-produced teacup shown in Figure 8.2 (c). In addition to the power spectrum as a function of spatial frequency, wavelength is also plotted on the upper abscissa; the spatial frequency is dependent on the diameter of the teacup. The range of some frequencies is not shown in these figures: low frequencies of less than 3 and high ones with a wavelength of longer than 380 μm (the limit of resolution). The point where the largest change occurs in the frequency

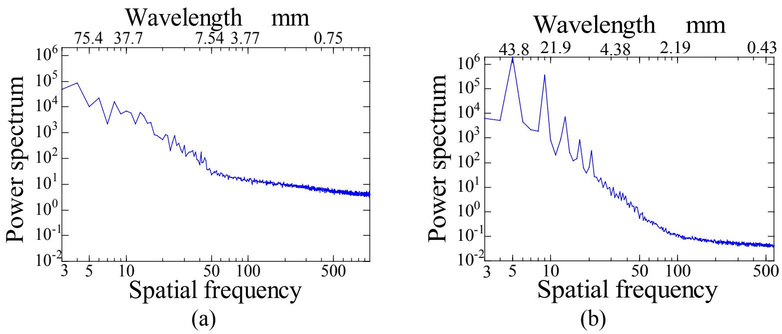


Figure 8.10 Power spectrum characteristics of teacups: (a) teacup B, and (b) mass-produced

range of 30 to 500 is defined as the "intersection". Table 8.2 compares the frequency and inclinations at the intersection, as well as an average inclination and power spectrum over the whole frequency range. In the case of the handmade Teacup B, the change in slope appears at a spatial frequency of around 50, which corresponds to a wavelength of 7.54 mm as the outer diameter of the teacup is set at 120 mm. The average inclination of the characteristic line is close to –1 or the $1/f$ fluctuation. The surface topography of the teacup contains a moderate combination of regularity and irregularity, as shown in Figure 8.2 (b). According to the aesthetic criterion adopted here, it can be said that the handmade object provides an eye-pleasing impression.

Table 8.2 Features abstracted from power spectrum characteristics

Object measured	Average inclination	Average power spectrum	Intersection frequency	Inclinations around intersection	
Teacup B	–1.06	237	50	–2.75	–0.63
Mass-produced	–1.86	55	120	–3.92	–0.43

On the other hand, the mass-produced teacup yields a higher inclination (–1.86) in the lower frequency range of less than 100. The magnitude of power spectrum is lower than that of the handmade one. This means that the surface of the machine-made teacup is relatively smooth or monotonous. Some spikes in the power spectrum stem from a periodical wave of asperities. Thus, some fine irregularity with a moderate combination of spatial frequency features the $1/f$ fluctuation characteristics.

As far as the sensibility of the surface topography is concerned, it seems that the laser-CCD displacement sensing is superior to the X-ray CT measurement.

8.3 Shape Design with 1/*f* Fluctuation Characteristics

8.3.1 Use of Measured Teacup Characteristics

Computer-based modeling is applied to shape design with $1/f$ fluctuation characteristics. Figure 8.11 shows the modeling process. First, a reference shape with a smooth surface is approximated by polygons using triangle patches, and then each cross-section where asperities are copied is defined and sliced at even intervals from the base. Finally, a spiral flow line is defined on the circumferential plane, and fluctuation is added along this line, *i. e.*, after the intersection between the sliced plane and the flow line is calculated, it is moved in the radial direction by the specified amount. This is repeated from the bottom to the top. A flow line is used because it mimics the making of a shape on a potter's wheel, *i. e.*, the base clay rotating on the potter's lathe is spirally modified by a tool or fingers, which involves some fluctuation.

The base shape is 120 mm round, 75 mm high, and 5 mm thick; and the plateau at the bottom is 5 mm high, assuming a real teacup. The fluctuation characteristics extracted from the teacup are used to duplicate pattern on the circumferential surface.

Figure 8.12 illustrates the wire-frame model used for the shape design: (a) the reference shape, and (b) the shape after modification. Note that the number of meshes is greatly reduced. The number of the sliced cross-sections is 140, and the number of the intersections in a layer is 2048. Thus, the resolutions in the axial and circumferential directions are 36 µm and 184 µm, respectively.

Figure 8.13 shows the CG image of the model duplicated using the frequency data of Teacup B, while Figure 8.14 is an enlarged version of the surface texture for visual effects: (a) the model with the 1/*f* fluctuation magnified 3.16 times, and (b) its power spectrum characteristics. Although the level of the power spectrum is increased over the whole frequency range by the enlargement, the profile of waviness and the inclination are precisely duplicated.

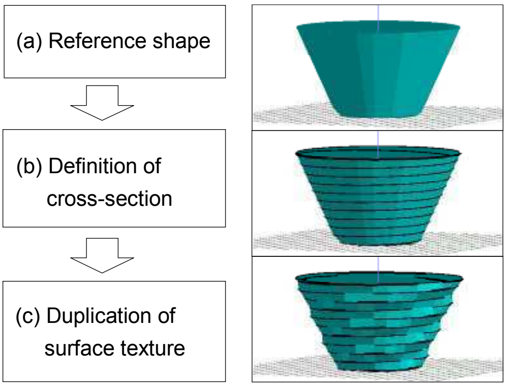


Figure 8.11 Schematic of shape design using 1/*f* characteristics

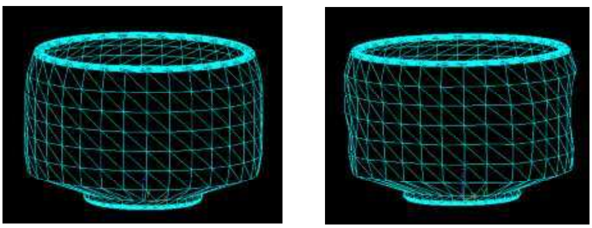


Figure 8.12 Wire-frame model of teacup: (a) reference shape, and (b) after duplication

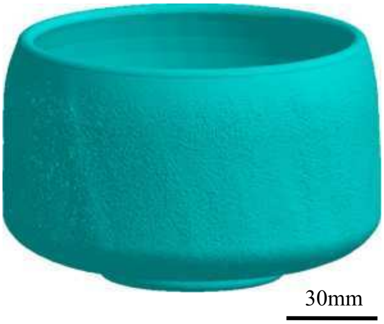


Figure 8.13 CG image of duplicated teacup with original fluctuation data

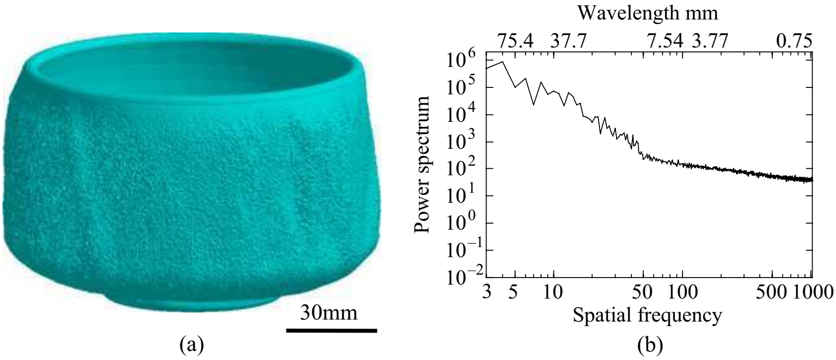


Figure 8.14 CG image of duplicated teacup with a magnification of 3.16: (a) CG image, and (b) power spectrum characteristics

8.3.2 Use of Mathematical Patterns

So far, the application of the measured 1/*f* fluctuation to shape design has been described. From an engineering point of view, it is inadvisable to take samplings of kansei characteristics from the natural world, including Japanese teacups; it takes time and needs high-capacity storage. One of the solutions to the generation of 1/*f* fluctuation characteristics is to use a mathematical approach. For instance, the following equation derived from a chaos model by one-dimensional mapping [5] generates 1/*f* fluctuation features:

$$\begin{aligned} x_{t+1} &= x_t + 2x_t^2 \qquad \left(0 \le x_t \le 0.5\right), \\ x_{t+1} &= x_t - 2\left(1 - x_t\right)^2 \quad \left(0.5 < x_t \le 1\right). \end{aligned} \tag{8.3}$$

Figure 8.15 (a) shows the waveform given by Equation 8.3, while Figure 8.15 (b) plots its power spectrum, postulating an initial value of 0.8. The 1/*f* waves expressed by Equation 8.3 give rise to sudden changes in waveform, which may lead to uncomfortably rough surfaces. In order to obtain some fine irregularity with a moderate combination of spatial frequency, a smoothing function defined by

$$\left.\begin{aligned} r &= (1.0-s)r_1 + sr_2 \\ s &= (1.0-\cos\pi t)/2.0 \end{aligned}\right\}(0<t<1) \tag{8.4}$$

can be used, where r, r_1, and r_2 are the radial coordinates of the polar coordinate system.

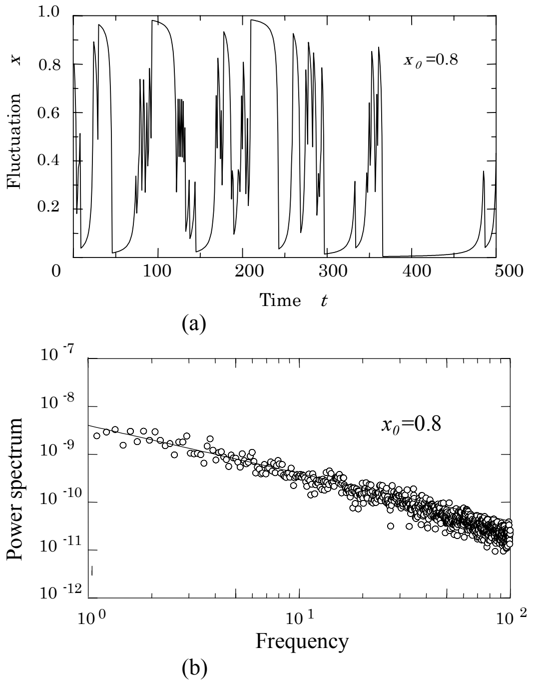


Figure 8.15 Generation of 1/*f* fluctuation characteristic with Equation 8.3: (a) waveform, and (b) power spectrum characteristics

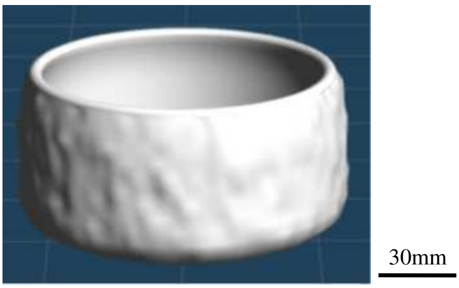


Figure 8.16 CG image of duplicated teacup with mathematical fluctuation data

Next, this artificial waveform is copied in a similar way as described in the previous section. Figure 8.16 shows the result of a CG image to be compared with Teacup A in Figure 8.2. Its power spectrum characteristics have been analyzed to have the $1/f$ property on the circumferential surface [6]. In the modeling, the roughness of maximum height was set at ±2% of the radius of the reference shape. Compared with the real teacup, an outstanding similarity of the surface texture can be recognized.

8.3.3 Use of Artificial Patterns

Another way of incorporating $1/f$ fluctuation characteristics is the two-dimensional mapping of artificial patterns. As shown in Figure 8.17, a reference shape without fluctuations is first defined: an equally spaced array of quadrangular pyramids in this case. The apex of each pyramid is translated by dx, dy, and dz in the x, y, and z directions, respectively. Note that the displacements given by Equation 8.3 are composed as a three-dimensional vector quantity.

Regarding the use of artificial $1/f$ waves, on the other hand, a 120 × 120 mm square sample of wallpaper is selected. A three-dimensional stylus instrument with 50 μm horizontal and 25 μm vertical resolutions measures the unevenness of the surface. The number of 1024 × 128 roughness data are collected from the rectangu-

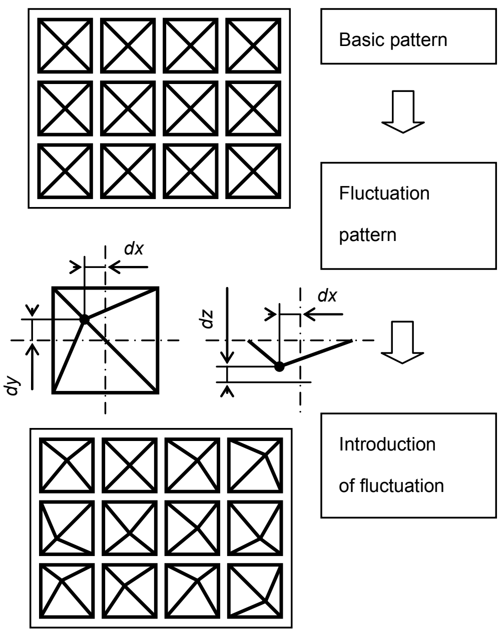


Figure 8.17 Schematic of shape design using two-dimensional mapping of $1/f$ characteristics

lar sample area of 102.4 mm by 12.7 mm, where the measured points are as small as 100 μm in the x and y directions. Unexpected discontinuities in unevenness often appear as the stylus traverses the edges of the wallpaper. This may give rise to a high spectrum at a specific frequency. To avoid this singularity, the measured data are multiplied by a window function [7]. The Hann window can be used:

$$w(n) = 0.5 - 0.5\cos\left(\frac{2\pi n}{N}\right), \tag{8.5}$$

where n is the number of measured points and N is the total data per cycle.

Figure 8.18 shows the topographical characteristics of the wallpaper: (a) the real CCD image, and (b) the power spectrum as a function of spatial frequency. The frequency or the wave number involved in the length of the wallpaper sample is 102.4 mm in this case. In the power spectrum, an inflection point can be recognized at a frequency of around 25, or a wavelength of 4 mm. The inclination of the characteristic line is approximately –1 at high frequencies above the inflection point.

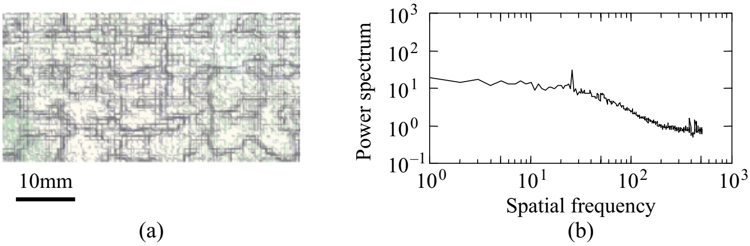


Figure 8.18 Wallpaper sample and its power spectrum characteristics: (a) CCD image, and (b) power spectrum characteristics

When the topographical data sampled from the wallpaper is applied to the outer surface of the reference shape, the problem is how to connect the edges or boundaries of the patches smoothly; a discontinuity or a large gap disturbs the 1/f fluctuation characteristics there. Figure 8.19 illustrates one of the solutions: selective duplication, where m is the length of the wallpaper referred to, and n is the length of the part used. A "continuous part" of the 1/f fluctuation data is selected and pasted onto the periphery of the reference shape. The continuous part means that the sides of the patch have a similar wave pattern. The same wallpaper as shown in Figure 8.18 can be used, *i. e.*, 12.8 × 12.8 mm square patches are pasted on the outer surface of a teacup.

Figure 8.20 (a) shows the graphical image of the teacup on which the surface unevenness of the wallpaper has been copied. The dimensions of the teacup are 120 mm in diameter, 77.5 mm in height and 5 mm in average wall thickness. The maximum peak-to-valley roughness is set at 1 mm. Figure 8.20 (b) depicts the power spectrum of the virtual teacup as a function of spatial frequency. Compared with the real surface texture shown in Figure 8.18, fluctuated patterns have been produced with satisfactory accuracy.

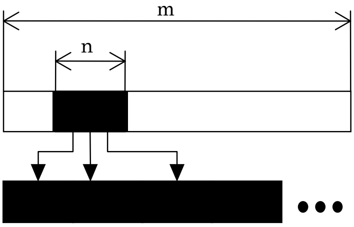


Figure 8.19 Generation of continuous data from patches

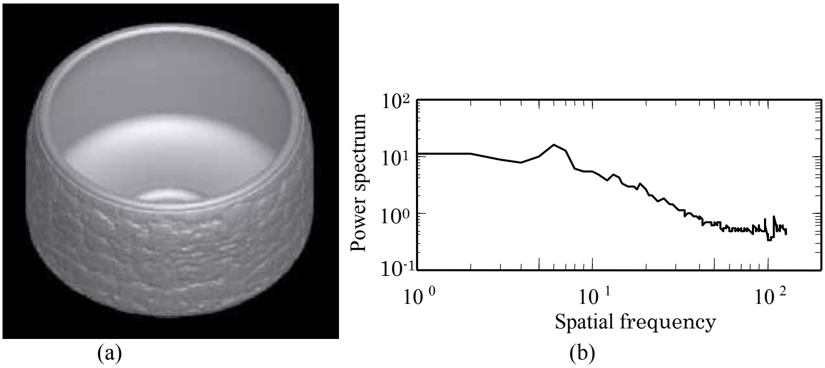


Figure 8.20 CG image of duplicated teacup with wallpaper patterns: (a) CCD image, and (b) power spectrum characteristics

8.4 Modeling by Rapid Prototyping

Objects with 1/*f* fluctuation characteristics have a high degree of complexity, so that it is difficult to fabricate them by conventional methods such as machining. Layered manufacturing technology [8], which enables the fabrication of three-dimensional objects from CAD data, has developed apace recently as a tool for the rapid prototyping (RP) of industrial products. The concept is that a model or a component is first modeled by a CAD system. Then, it is converted into an STL file format that can approximate the model surface by polygons. The STL model is sliced into cross-sections on a computer. The cross-sections are systematically recreated layer by layer using various materials, such as photo-curable resin, polymer powder, paper sheet, and metal powder. The integration of these layers makes it possible to form a three-dimensional object in a short time.

As mentioned in the previous section, the sliced model with 1/*f* fluctuations has already been completed. Accordingly, the use of the RP technology is suitable for this purpose. We can employ the Selective Laser Sintering (SLS) method [9], as shown schematically in Figure 8.21. A thin layer of heat-fusible powder is deposited onto the powder bed in the process chamber. The cross-section layer of the object under fabrication is selectively scanned on the powder layer by a CO_2 laser

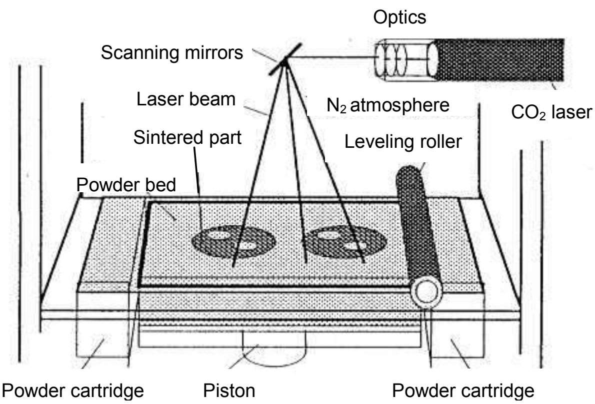


Figure 8.21 Schematic of rapid prototyping by SLS using Nylon powder [9]

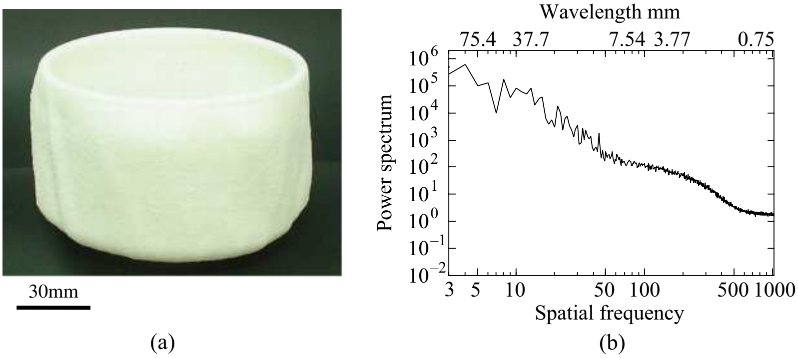
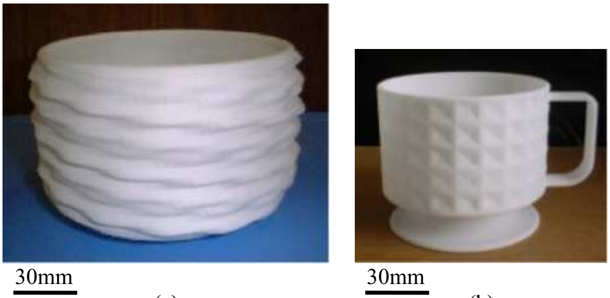


Figure 8.22 Rapid prototyping of duplicated teacup with a magnification of 3.16: (a) CCD image, and (b) power spectrum characteristics

beam. The irradiation of the beam elevates the powder temperature to the melting point, fusing the powder particles to form a solid mass. When the cross-section is completely scanned, powder is added by the roller on top of the previously scanned layer. These processes are repeated until the part is completed. The SLS process has the advantage of being able to use a wide variety of powders for sintering, including wax, Nylon, stainless steel, and casting sand, but the dimensional accuracy of the SLS part is around 100 μm. A Nylon powder with an average diameter of 50 μm can be used on the Sinterstation 2500 PLUS.

Figure 8.22 shows the result of a duplicated teacup with a magnification of 3.16 by the RP method: (a) the Nylon model, and (b) its power spectrum characteristics. The power spectrum is in accordance with that in Figure 8.10 (a) in the range of spatial frequency below 300 or wavelength above 1.51 mm. The poor reproducibility in the high-frequency range is probably due to the powder size used.

Finally, other models with 1/*f* fluctuation characteristics are shown in Figure 8.23: (a) a teacup with 5% fluctuations and three orbits, and (b) a teacup with arrays of artificial patterns. Thus, the SLS method together with the design method stated in the previous section can generate any complicated models with 1/*f* fluctuation characteristics in a relatively short time.



(a) (b)

Figure 8.23 Rapid prototyping of teacups with various 1/*f* fluctuation characteristics: (a) 5% fluctuation and three orbits, and (b) fluctuated array patterns

8.5 Summary

An attempt has been made to develop a method of introducing 1/*f* fluctuation characteristics into shape design and modeling of aesthetic objects.

First, topographic features of various Japanese teacups have been measured using X-ray CT or laser metrology, and their wave patterns have been analyzed using the FFT method and characterized using power spectrums. The outer shape of the handmade teacup has an inclination close to –1 on the spatial frequency-power spectrum diagram, which means that 1/*f* fluctuation characteristics are involved in it. A mass-produced one, on the other hand, tends to have an inclination of nearly –2, which gives us a monotonous or dull impression.

Next, the 1/*f* fluctuation property thus measured has been applied to shape design with computer-based modeling. Other approaches using a mathematical model and a wallpaper pattern that can generate 1/*f* fluctuation characteristics have been demonstrated. Finally, rapid prototyping technology has been utilized to prove the CAD model. The design and modeling method with the kansei feature can be used not only in the duplication of handmade ceramics but also in the creation of original aesthetic objects.

As far as the realization of the CG model with pottery material is concerned, it is neither simple nor straightforward; rapid prototyping technology with ceramic materials has not yet been established. The plastic RP model can be used as a mold for evaporative pattern casting [10], or alternatively can be pressed onto a half-dried clay object that has been formed on a potter's wheel. This challenging issue is another topic of discussion.

Acknowledgments The author would like to express sincere thanks to Mr. Kazuo Nakamura for providing the beautiful teacups for measurement, and Mr. Tomohiro Nishii, former graduate student, Ibaraki University for his dedication to the computational work. This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Exploratory Research, 14658110 in 2002 and Sekisui Chemical, Co., Ltd., Research Grant in 2003.

References

1. Nagamachi M.: A new ergonomic consumer-oriented technology for product development. Int. J. Industrial Ergonomics **15**(1):3–11 (1995)
2. Childs T.H.C.: An essay on affective design – applied science or art? In: 2nd International Conference on Affective Human Factors Design, Singapore (2001)
3. Musha T.: Yuragi no Hassou, Ch. 2. NHK Syuppan, Tokyo (in Japanese) (1998)
4. Inagaki T., Inuzuka K., Agu M., Akabane H., Abe N.: $1/f^n$ fluctuation phenomena in luminous patterns of fireflies and their healing effect. J. JSME **67**(657):1583–1590 (in Japanese) (2001)
5. Akabane H., Nakaya K., Agu M.: Asymmetrical property of $1/f$ resistance fluctuation for time inversion. Jpn. J. Appl. Phys. **39**:2865–2869 (2000)
6. Maekawa K., Nishii T., Hayashi T., Akabane H., Agu M.: Laser fabrication of affective 3D objects with $1/f$ fluctuation. JSME Int. J. (A) **46**(3):460–466 (2003)
7. Kivinukk A., Tamberg G.: On sampling operators defined by the Hann window and some of their extensions. Sampl. Theor. Signal Image Proc. **2**(3):235–257 (2003)
8. Kruth J.-P., Leu M.C., Nakagawa T.: Progress in additive manufacturing and rapid prototyping. Annals CIRP **47** (2):525–540 (1998)
9. Lü L., Fuh J., Wong Y.-S.: Laser-Induced Materials and Processes for Rapid Prototyping. Kluwer, Singapore (2001)
10. Kalpakjian S., Schmid S.R.: Manufacturing Engineering and Technology, 4th edition, pp. 274–276. Prentice Hall, Englewood Cliffs, NJ (2001)

Chapter 9
Human-entrained Embodied Interaction and Communication Technology

Tomio Watanabe[1]

Abstract An embodied communication system for mind connection (E-COSMIC) has been developed by applying the entrainment mechanism of the embodied rhythms of nodding and body movements to physical robots and CG characters in verbal communication. E-COSMIC comprises an embodied virtual communication system for human interaction analysis by synthesis and a speech-driven embodied interaction system for supporting essential human interaction and communication based on the analysis that uses the embodied virtual communication system. A human-entrained embodied interaction and communication technology for an advanced media society is introduced through some applications of E-COSMIC. A generation and control technology of human-entrained embodied media is also introduced.

9.1 Introduction

In human face-to-face conversation, a listener's movements such as nodding and body motions are interactively synchronized with the speaker's speech. Embodied rhythms between voice and movement are mutually synchronized not only between talkers but also in a talker. The phenomenon is observed in an infant's movements in response to the mother's speech as a primitive form of communication [1, 2]. This synchrony of embodied rhythms in communication, referred to as entrainment, generates the sharing of embodiment in human interaction, which plays an important role in human interaction and communication. Entrainment in communication is also observed in physiological indices such as respiration and

[1] T. Watanabe
Faculty of Computer Science and Systems Engineering, Okayama Prefectural University
CREST of Japan Science and Technology Agency
111 Kuboki, Soja, Okayama 719-1197, Japan
e-mail: watanabe@cse.oka-pu.ac.jp

heart rate variability [3]. This embodied communication, which is closely related to behavioral and physiological entrainment, is an essential form of communication that forms the basis of interaction between talkers through mutual embodiment. Hence, the introduction of this mechanism to a human interface is indispensable to the realization of human-centered essential interaction and communication systems.

In this chapter, by focusing on the embodied entrainment, the human-entrained embodied interaction and communication technology through the development of the embodied communication system for mind connection (E-COSMIC) is introduced for supporting human interaction and communication [4]. E-COSMIC mainly comprises an embodied virtual face-to-face communication system and a speech-driven embodied interaction system, as shown in Figure 9.1. The former is developed for human interaction analysis by synthesis and the latter, for supporting human interaction and communication based on the analysis that uses the former. The effectiveness of the system is demonstrated by some actual applications on robot/CG and human interactive communications. With the aim of creating embodied media that unify performers and audiences for supporting the creation of digital media arts for entertainment and education, the generation and control technology of human-entrained embodied media is also introduced.

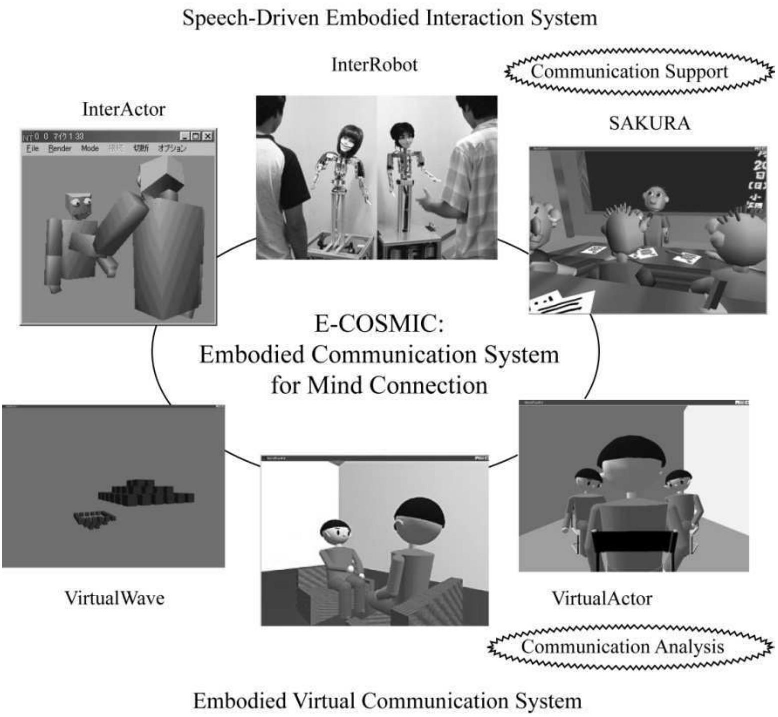


Figure 9.1 E-COSMIC

9.2 Embodied Virtual Communication System

The concept of an embodied virtual face-to-face communication system is illustrated in Figure 9.2. The figure presents a VirtualActor (VA), an interactive avatar, that represents the talker's interactive behavior such as gestures, nodding, blinking, facial color and expressions, paralanguage, respiration, *etc.*, based on one's verbal and nonverbal information as well as physiological information in a virtual face-to-face communication environment. Figure 9.3 provides an example of a virtual face-to-face scene with two VAs from the diagonal backward viewpoint of one's own VA. The motions of the head, arms, and body for each VA are represented based on the positions and angles measured by four magnetic sensors that are placed on the top of the talker's head, both wrists, and the back of the body [5]. Two remote talkers can communicate through their VAs and become aware of the interaction through the embodied interaction of VAs in the same virtual communication environment from any viewpoint. The analysis by synthesis for interaction in communication is performed by processing the behavior of VAs, such as cutting or delaying the motion and voice of VAs in various spatial relations and positions (Figure 9.4). For example, to examine the effects of only nodding on interaction, it is possible for a VA to represent just nodding without body motion even if the talker nods with body motion. Thus, the characteristics and relations of the embodied interaction between talkers are systematically clarified through the analysis by synthesis of interaction in communication by using the system in which talkers are the observers of interactions as well as the operators of interaction through their VAs. Further, physiological measurements such as respiration, heart rate variability, and facial skin temperature, as indices of emotional states in communication are utilized not only for quantitatively evaluating the interaction but also for transmitting the change in talkers' emotions through the VA affect display in which facial color and expressions are synthesized based on the measurement. An embodied virtual group communication system was also developed for three human interaction supports and analyses by synthesis, as indicated in Figure 9.1 [6].

Not only we have created a VA that represents human behavior precisely, we have also created an abstract avatar of a wave (VirtualWave, or VW) that is constructed from 6×6 cubes, as shown in Figure 9.5. Here, the communication function of the VA is simplified as a function of the motion of the VW in order to clarify an essential role of interaction. The rhythm of the VW is used to characterize the interactive rhythm, and this behavior is represented using only the motion of the head; the motion is measured by a magnetic sensor placed on top of the talker's head. This is because the head motion performs the essential function of regulating the flow of conversation, for example, nodding, by which each talker discriminates one's VW from the partner's VW and shares their interactions. A vertical shift such as nodding is expressed by an up-and-down displacement of cubes in which the wave approaches a quadrangular pyramid in shape, as shown in Figure 9.5. A horizontal shift such as back-and-forth is expressed by the parallel displacement of cubes in proportion to the shift. VW is represented by a frame rate of 30 f/s. The effects of the head motion on the interaction of VAs have already been demonstrated [7].

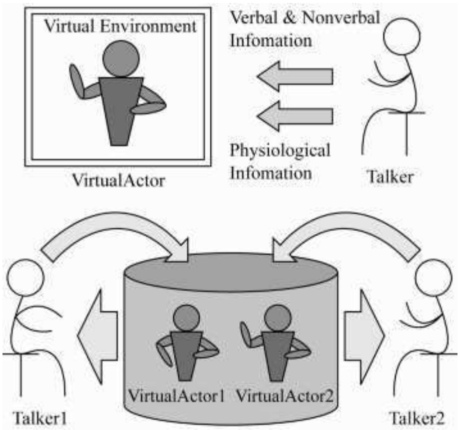


Figure 9.2 Concept of the embodied virtual communication system

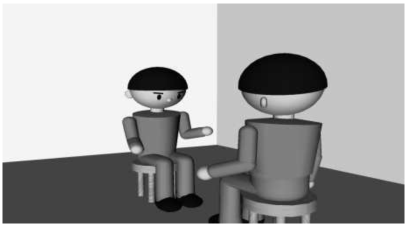


Figure 9.3 Example of a virtual face-to-face scene with two VirtualActors representing the talker and his/her partner

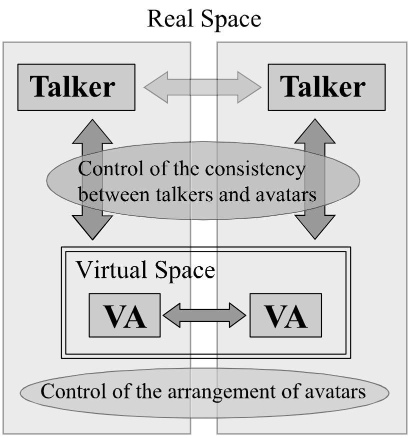


Figure 9.4 Analysis model for human interaction via embodied avatars

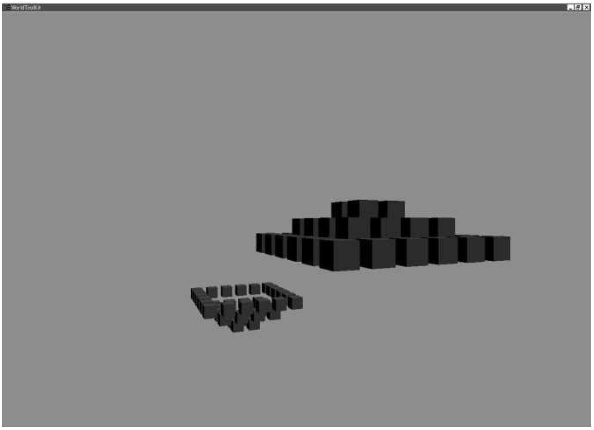


Figure 9.5 Example of VirtualWave's (VW) communication scene

9.3 Speech-driven Embodied Interaction System

Based on the human interaction analysis that uses the embodied virtual communication system, a speech-driven embodied interaction system is developed for supporting human interaction by generating the communicative motions of a physical robot referred to as InterRobot or a CG character known as InterActor; these communicative motions are coherently related to speech input [8]. The concept is presented in Figure 9.6. The system comprises two InterRobots (or InterActors) that function as both speaker and listener based on speech input. When Talker 1 speaks to InterRobot 2, InterRobot 2 responds to Talker 1's utterance with an appropriate timing through its entire body motions, including nodding, blinking, and actions, in a manner similar to the body motions of a listener. Thus, Talker 1 can talk smoothly and naturally. Subsequently, the speech is transmitted via a network to the remote InterRobot 1. InterRobot 1 can effectively transmit Talker 1's message to Talker 2 by generating the body motions similar to those of the speaker based on the time series of the speech and by simultaneously presenting both the speech and the entrained body motions. This time, Talker 2 in the role of a speaker achieves communication in the same way by transmitting his/her speech via InterRobot 1 as a listener and InterRobot 2 as the one talking to Talker 1. Thus, in this manner, two remote talkers can enjoy a conversation via InterRobot. The information transmitted and received by this system is only through speech. Of significance is the fact that it is a human who transmits and receives the information; the InterRobot merely generates the entrained communicative movements and actions based on speech input and supports the sharing of mutual embodiment in communication.

With regard to a listener's interaction model, the nodding reaction model from a speech ON–OFF pattern and the body reaction model from the nodding reaction model are introduced (Figure 9.7). When $Mu(i)$ exceeds a threshold value, nodding

M(*i*) is estimated as the weighted sum of the binary speech signal *V*(*i*). The body movements are related to the speech input by operating both the neck and one of the wrists, elbows, arms, or waists at the timing over the body threshold. The threshold is set lower than that of the nodding prediction of the MA (moving average) model, which is expressed as the weighted sum of the binary speech signal to nodding. In other words, for the generation of body movements when InterActor functions as a listener, the time relationships between nodding and other movements are realized by varying the threshold values of the nodding estimation.

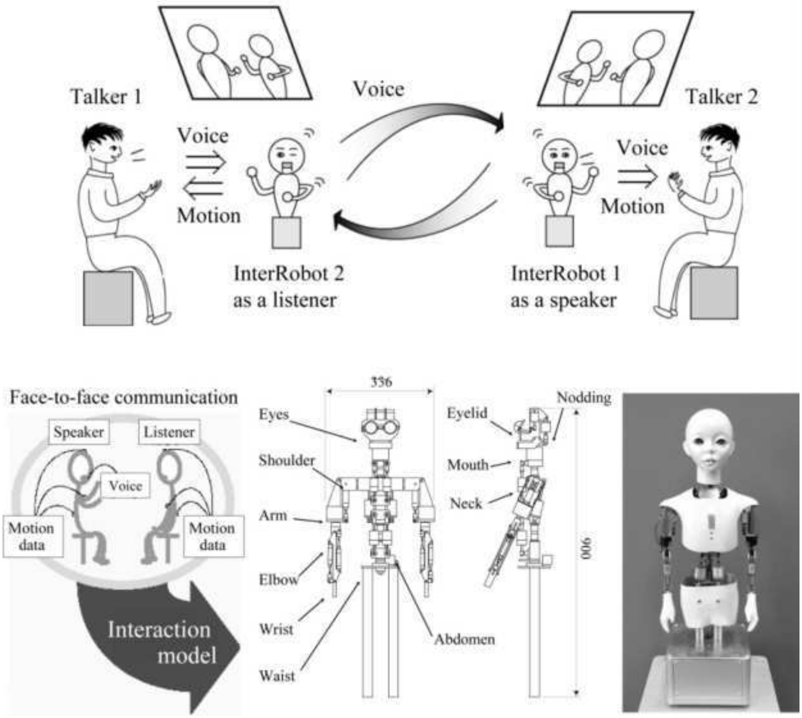


Figure 9.6 Concept of the speech-driven embodied interaction system

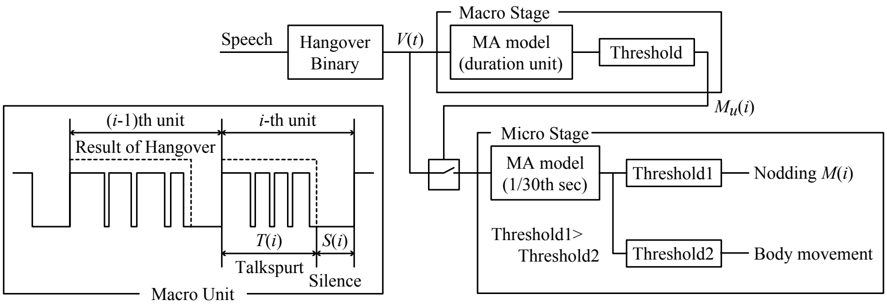


Figure 9.7 Interaction model

$$M_u(i) = \sum_{j=1}^{J} a(j)R(i-j) + u(i) \tag{9.1}$$

$$R(i) = \frac{T(i)}{T(i) + S(i)} \tag{9.2}$$

$a(j)$: linear prediction coefficient

$T(i)$: talksqurt duration in the ith duration unit

$S(i)$: silence duration in the ith duration unit

$u(i)$: noise

$$M(i) = \sum_{j=1}^{K} b(j)V(i-j) + w(i) \tag{9.3}$$

$b(j)$: linear prediction coefficient

$V(i)$: voice

$w(i)$: noise

The body movements as a speaker are also related to the speech input by operating both the neck and one of the other body actions at the timing over the threshold, which is estimated by the speaker's interaction model as its own MA model of the burst-pause of speech to the entire body motion. Because speech and arm movements are related at a relatively high threshold value, one of the arm actions in the preset multiple patterns is selected for operation when the power of speech is over the threshold. The expressive actions of InterActor are shown Figure 9.8.

We developed a system superimposed on a nodding response model for the analysis by synthesis of embodied communication under the expected conditions of promoted interaction. In addition, for this system, we performed experiments

(a)

(b)

Figure 9.8 Expressive actions of InterActor: (a) listener's action, and (b) speaker's action for the analysis by synthesis of embodied communication by examining the sensory evaluation and the voice–movement analysis while inconsistently adding nodding responses in VA. We found that the cross-correlation between the talker's voice and the listener's head movement in the inconsistently activated condition increases at a significance level of 1% compared to that observed under normal conditions. The result also demonstrates that the system superimposed over the nodding response promoted interaction in embodied communication [9]. Furthermore, we have also developed a speech-driven embodied entrainment system called "InterWall" in which interactive CG objects behave as listeners on the basis of the speech input of a talker (Figure 9.9). This system can support human interaction and communication by producing embodied entrainment movements such as nodding on the basis of the speech input of a talker. We confirmed the importance of providing a communication environment in which not only avatars but also CG objects placed around the avatars are related to virtual communication.

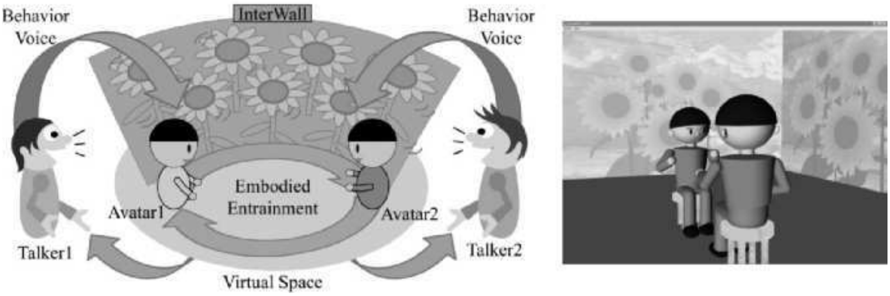


Figure 9.9 InterWall

Figure 9.10 illustrates a speech-driven embodied group-entrained communication system referred to as SAKURA [10]. SAKURA activates group communication in which InterActors are entrained to one another as a teacher and some students in the same virtual classroom. By using SAKURA, talkers can communicate with a sense of unity through the entrained InterActors by using only speech input via the network. Figure 9.11 depicts a physical version of SAKURA with InterRobots. Their entrained movements and actions based on speech can activate and assist human embodied interaction and communication. Figure 9.12 indicates another physical version of SAKURA with four InterRobots and one InterActor, which is exhibited in the National Museum of Emerging Science and Innovation where visitors can enjoy a dynamic experience of embodied communication. They perceive the effects of group-entrained communication environment intuitively and recognize the importance of embodied communication.

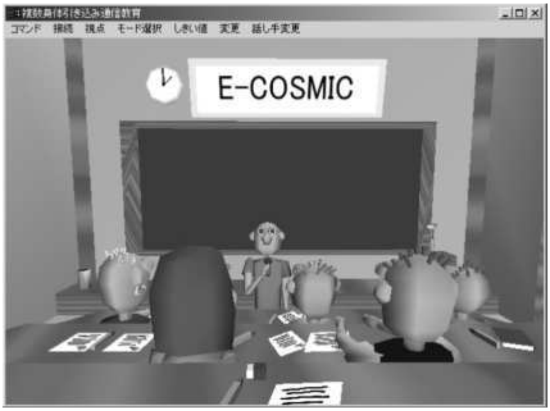


Figure 9.10 SAKURA: The speech-driven embodied group-entrained communication system

Figure 9.11 Physical version of SAKURA with InterRobots

Figure 9.12 Speech-driven embodied interaction system with InterRobots and an InterActor in the National Museum of Emerging Science and Innovation

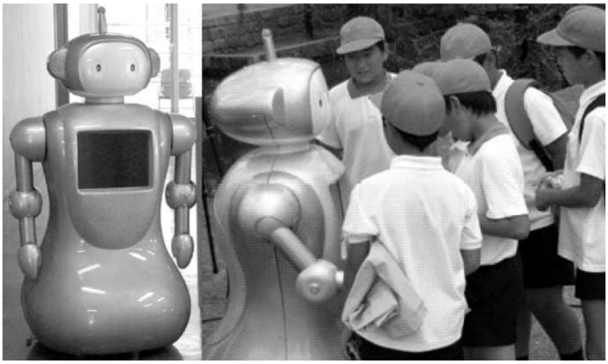


Figure 9.13 Interaction scene between an InterRobot and children

9.4 Embodied Interaction and Communication Technology

In this section, some actual applications of InterRobot/InterActor to human interface are introduced. Figure 9.13 depicts an interaction scene between an InterRobot and children. This InterRobot is commercially available and marketed for kindergarten use. Children enjoy and are excited about having conversations with the InterRobot, while the teacher standing behind the InterRobot enjoys talking and encouraging children in a new communication mode from a completely different standpoint, just changing to a friend and so forth. By focusing on an animal character, which is the type most preferred by children, an animal-type InterRobot/InterActor known as InterAnimal has been developed in order to encourage and cheer up children, as depicted in Figure 9.14. The bear-type InterAnimal shown in Figure 9.15 was a popular exhibit in the 2005 EXPO. Figure 9.16 illustrates a toy version of InterRobot with the function of a listener, which generates the listener's actions of nodding, tilting his/her head, and moving his/her arms up and down, based on speech input. The stuffed toy bear is eager to listen without ever uttering a word. It is commercially available and marketed under the name of Unazukikun.

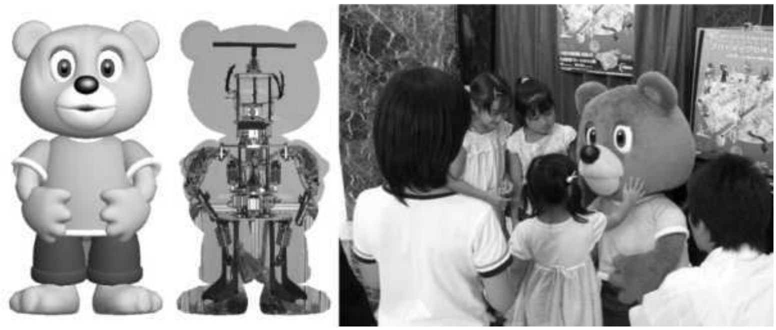


Figure 9.14 InterAnimal: Animal-type InterActor

The InterActor, as indicated in Figure 9.17, is also commercially available under the name of InterCaster through which news and media contents are effectively and cordially transmitted in a commercial program. By superimposing In-

Figure 9.15 InterAnimal in EXPO 2005

Figure 9.16 Toy version of InterRobot

Figure 9.17 InterCaster in an educational program

terActors as listeners on the video images of a lecture such as an education program, the InterActor-superimposed learning support system has been developed, as illustrated in Figure 9.18 [11]. The system provides group-entrained interaction effects for audiences who watch the video, in which two InterActors at the bottom of reduced-size images are entrained with the lecturer's speech.

The InterActor is a speech-driven CG-embodied interaction character that can generate communicative movements and actions for an entrained interaction. An InterPuppet, on the other hand, is an embodied interaction character that is driven by both speech input, similar to the InterActor, and hand motion input, like a puppet. Therefore, humans can use the InterPuppet to communicate effectively by using deliberate body movements as well as natural communicative movements and actions [12]. An advanced InterPuppet with a cellular phone-type device was developed as indicated in Figure 9.19. On the display, two characters – representing the talker and his/her partner – were arranged at an angle of 90° to each other, based on the finding of the conversation arrangement of the InterActors. The

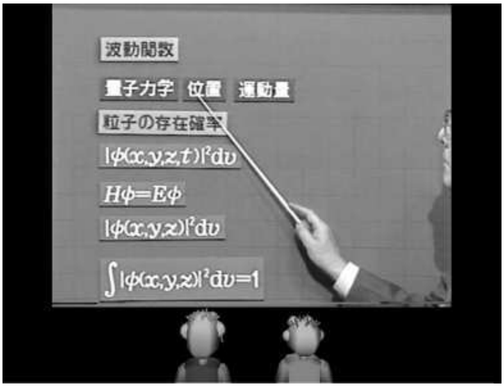
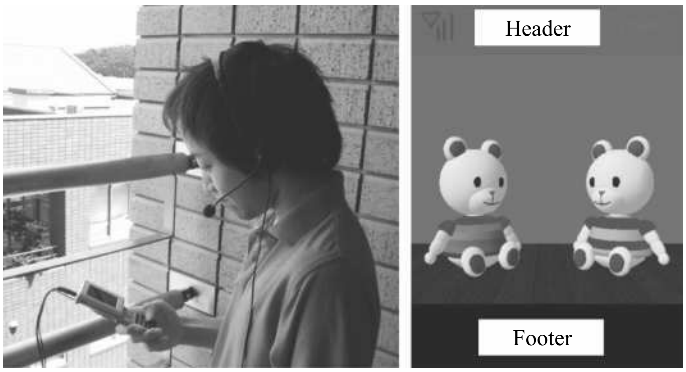


Figure 9.18 InterActor-superimposed learning support system

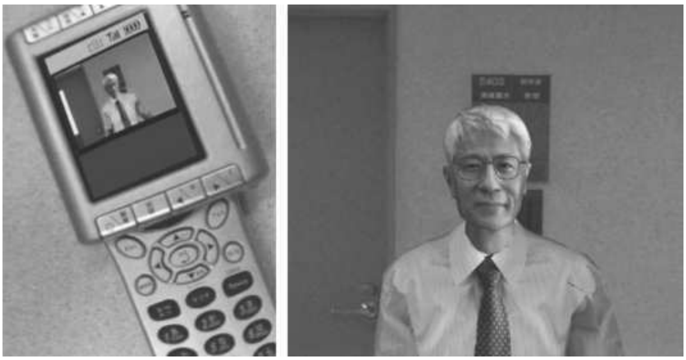


(a) (b)

Figure 9.19 InterPuppet with a cellular phone-type: (a) an image of someone using InterPuppet and (b) screen shot screen comprises installed headers and footers similar to those in cellular phones. The character resembled an animal type, which could be reminiscent of a doll. When the talker speaks to another user, the talker's InterPuppet behaves as the speaker and the partner's InterPuppet responds as a listener. Then, voice is transmitted via the network to the remote partner and the InterPuppet. If the user inputs his/her hand motion, the hand motion is converted to InterPuppet movement and is included in the movements of the InterActor [13].

Figure 9.20 shows the system appearance and the CG image of the system. In this system, communications using a videoconferencing telephone become possible by presenting a real-type CG character with the user's characteristic as well as an anonymous CG character. It is difficult for a videoconferencing cellular telephone to continuously capture the user with the camera when the user speaks while carrying it, because the camera is set in the body of the cellular phone. In addition, unexpected backgrounds and real-time telecasted images cannot be recognized by a videoconferencing telephone. If the InterPuppet characteristic is ideally employed, a real-type CG character can express intentional body movements that include intentions and emotions in addition to entrainment body motion. Therefore, remote face-to-face communications such as videoconferencing telephones will be enabled and involve only speaking and easy key operations, instead of needing to take care of the camera as well as the environment. In addition, the user can prepare his/her own face background and can freely set the conversation arrangement and the aspect.

Figure 9.21 shows the speech-driven embodied entrainment systems InterPointer and InterVibrator which support embodied interaction and communication during presentations [14]. InterPointer is a pointing device with a visualized response equivalent to the nodding response and its response time to the speech input is similar to that of a listener. InterVibrator is a vibration device with a vibratory response by nodding response in the same way. An integrated system of InterPointer and InterVibrator for supporting interactive presentation is developed.



(a) (b)

Figure 9.20 Application to video conferencing telephone: (a) system appearance, and (b) CG image

Figure 9.22 shows an example of a presentation using the system. It is expected to support interactive communication by synchronizing the embodied rhythms using a visualized response and a vibratory response.

A speech-driven embodied entrainment chair system called "InterChair" is being developed for supporting embodied interaction (Figure 9.23). The system generates bodily responses equivalent to nodding in the same timing as listener to speech input. The system naturally forces a user sitting on it to nod with backward–forward motions of 0.01 G. The nodding responses activate embodied interactions among a speaker and listeners around the user.

Furthermore, we have also developed an embodied entrainment character chat system called "InterChat" by introducing an enhanced interaction model that can create motions where the communicative motions and natural actions can be generated easily from both the typing rhythms and the meaning of the words of sending and receiving messages because the typing rhythms resemble the speech rhythms (Figure 9.24). The system is expected to be the basis for a new type of interactive chat communication with embodied entrained characters. The development expands the applications of the embodied interaction and communication technology from a voice input to a typing input.

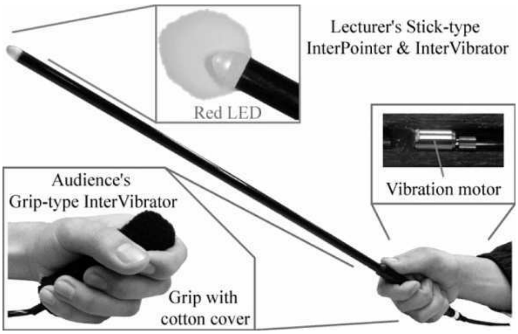


Figure 9.21 InterPointer and InterVibrator

Figure 9.22 Example scene of presentation using InterPointer and InterVibrator

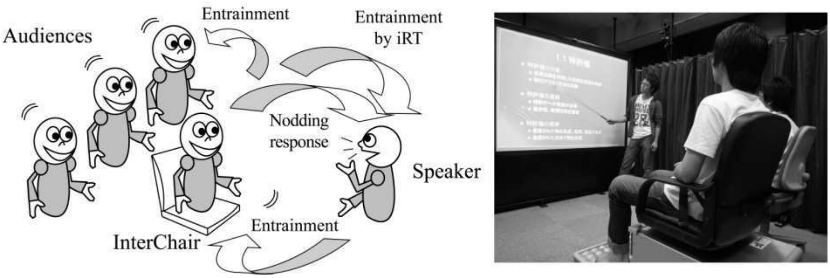


Figure 9.23 Concept of InterChair

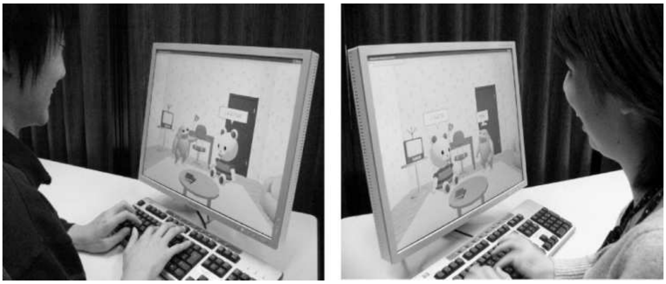


Figure 9.24 InterChat

Figure 9.25 Human-entrained embodied media

With the aim of creating embodied media that unify performers and audiences for supporting the creation of digital media arts for entertainment and education, we will develop a generation and control technology of human-entrained embodied media by developing and integrating the following three technologies: (1) "embodied entrainment media technology" to set embodied media alight with virtual audiences' entrained responses; (2) "embodied space and image media technology" to integrate and display special media with embodied audiences; and (3) "embodied acoustic media technology" to produce music and embodied acoustics from body motions. Figure 9.25 depicts our embodied media exhibited in the National Museum of Emerging Science and Innovation. What I want to convey with the media is the mystery and importance of embodied interaction and communication.

9.5 Conclusions

The human-entrained embodied interaction and communication technology for an advanced media society was proposed through the development of E-COSMIC for supporting essential human interactive communication based on the entrainment mechanism of the embodied rhythms between speech and body movements such as nodding. Some actual applications pertaining to robot/CG and human interactive communication were also demonstrated. In particular, the speech-driven embodied interaction system, such as InterRobot and InterActor, is a robust and practical communication support system for everyday living, which activates embodied interaction and communication in a new communication mode by using only speech input. The speech-driven entrainment technology for enhancing interaction and communication would be expected to form the foundation of mediated communication technologies as well as the methodology for the analysis and understanding of human interaction and communication, and allow the development of a new embodied communication industry for supporting essential human interactive communication.

References

1. Condon W.S., Sander L.W.: Neonate movement is synchronized with adult speech: interactional participation and language acquisition. Science **183**:99–101 (1974)
2. Kobayashi N., Ishii T., Watanabe T.: Quantitative evaluation of infant behavior and mother–infant interaction. Early Development and Parenting **1**:23–31 (1992)
3. Watanabe T., Okubo M.: Evaluation of the entrainment between a speaker's burst-pause of speech and respiration and a listener's respiration in face-to-face communication. In: Proceedings of the 6th IEEE International Workshop on Robot–Human Interactive Communication (RO-MAN'97), pp. 392–397 (1997)
4. Watanabe T.: E-COSMIC Embodied communication system for mind connection. In: Proceedings of the 13th IEEE International Workshop on Robot–Human Interactive Communication (RO-MAN 2004), pp. 1–6 (2004)

5. Watanabe T., Ogikubo M., Ishii Y.: Visualization of respiration in the embodied virtual communication system and its evaluation. Int. J. Human Comp. Interaction **17**(1):89–102 (2004)
6. Shintoku T., Watanabe T.: An embodied virtual communication system for three human interaction support and analysis by synthesis. In: Proceedings of the 5th IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA2003), pp. 211–216 (2003)
7. Watanabe T., Okubo M.: An embodied virtual communication system for human interaction sharing. In: Proceedings of the 1999 IEEE International Conference on Systems, Man, & Cybernetics, pp. 1060–1065 (1999)
8. Watanabe T., Okubo M., Nakashige M., Danbara R.: InterActor: speech-driven embodied interactive actor. Int. J. Human Comp. Interaction **17**(1):43–60 (2004)
9. Sejima Y., Watanabe T., Yamamoto M.: Analysis by synthesis of embodied communication via virtual actor with a nodding response model. In: Proceedings of Second International Symposium on Universal Communication (ISUC2008), pp. 225–230 (2008)
10. Watanabe T., Okubo M: SUKURA voice-driven embodied group-entrained communication system. In: Proceedings of HCI International 2003, vol. 2, pp. 558–562 (2003)
11. Watanabe T., Yamamoto M.: An embodied entrainment system with interactors superimposed on images. In: Proceedings of HCI International 2005, vol. 4, pp. 1–6 (2005)
12. Osaki K., Watanabe T., Yamamoto M.: Development of an embodied interaction system with interactor by speech and hand motion input. Trans. Human Interface Soc. **7**(3):89–98 (2005)
13. Osaki K., Watanabe T., Yamamoto M.: A cellular phone type of mobile system with speech-driven embodied entrainment characters by speech and key inputs. Trans. Human Interface Soc. **10**(4):73–83 (2008)
14. Nagai H., Watanabe T., Yamamoto M.: A speech-driven embodied entrainment system with visualized and vibratory nodding responses as listener. Trans. Jpn. Soc. Mech. Eng. **75**(755):163–171 (2009)

Chapter 10
Personal Cognitive Characteristics in Affordance Perception: Case Study in a Lobby

Y.S. Kim[1], J.Y. Jeong[2], M.K. Kim[3], S.W. Lee[4], and M. Kim[5]

Abstract User activities in performing tasks are influenced by the way the user perceives the related context and environment, and determined by the user making a judgment on their preferences. Structures in the physical environment afford user activities when they are properly perceived. This chapter addresses how user activities and perceived affordances are different and reflect personal creativity modes, which are determined by factual–intuitive perception inclination and subjective–objective decision preferences as well as the introverted–extroverted nature of the user. To enable the designing-in of various affordance features for diverse users in varying contexts, an understanding of the relationship between the personal characteristics of the user and affordance perception would be helpful. We conducted a case study in a public space used by many ordinary people. User activities and behaviors were analyzed in specific tasks given to 20 students in a building lobby they had never previously visited. The tasks were devised so that various affordance features would be relevant, while eliminating factors affecting the affordance perception (culture, intelligence, *etc.*) other than those due to the personal characteristics of the user. User activities can be classified into several

[1] Y.S. Kim
Creative Design Institute, Sungkyunkwan University, Suwon, Korea, 440-746
e-mail: yskim@skku.edu
[2] J.Y. Jeong
Creative Design Institute, Sungkyunkwan University, Suwon, Korea, 440-746
e-mail: jiyunjeong@skku.edu
[3] M.K. Kim
Creative Design Institute, Sungkyunkwan University, Suwon, Korea, 440-746
e-mail: skylord@skku.edu
[4] S.W. Lee
Creative Design Institute, Sungkyunkwan University, Suwon, Korea, 440-746
e-mail: sangwonl@skku.edu
[5] M. Kim
Creative Design Institute, Sungkyunkwan University, Suwon, Korea, 440-746
e-mail: meek@skku.edu

different groups for each task based on the affordance features involved in their activities. These differences are then compared with their personal creativity modes. For users of less common activities for some tasks, relevant personal cognitive characteristics have been identified.

10.1 Introduction

User activities in performing tasks are influenced by the way the user perceives the related context and environment, and determined with their preferences in judgement. Structures in the physical environment afford user activities when they are properly perceived. In addition, user activities reflect their emotional situation. Even the way users perceive the surrounding environment will be affected by their cognitive and emotional states as well as their cultural and social backgrounds.

Affordance is a concept that is highly related to human perception, judgment, and action. Thus it is to be used in designing various artifacts such as products, architectural structures, and spaces, as well as services. Norman mentioned "unarticulated needs", which cannot easily be known with simple interviews, focus group interviews, surveys, or questionnaires [1]. In reality, these unarticulated needs can be dug out by monitoring and analyzing users' activities and contexts. Since affordance is also highly associated with the activities of users, it is necessary to substantially consider affordance and its links with human emotions and personality.

Norman mentioned that designers could communicate with end-users via the system images of products by describing the designer's conceptual model and the user's mental model [1, 2]. He regarded affordance, which was one of main system images of the products, as a tool to understand users and designers and to create a bridge between them. In addition, the affordance is highly related to features of artifacts that drive certain user activities. However, users usually perceive the affordance associated with the features in diverse ways, and they may not recognize the designer's original intents embedded in the features [3]. In other words, users could perceive affordances from features that the designers did not originally intend. Therefore, it is necessary to consider the personal or emotional characteristics of users when designing artifacts.

Recently, the relationship between users and designers has been an important issue in the field of design. To properly address this issue, it is necessary to understand the emotional characteristics of users. Emotion has a significant influence on human decisions and activities. However, human emotion is very complicated, and can be expressed very diversely according to one's experience, personal characteristics, and context. Research efforts to try to understand users have recently been very active in the field of design with studies into the context of artifacts. By taking the context into consideration, designers can gain empathy with users, avoid the fixation on preset assumptions about the user or the product, and therefore create innovative concepts on how a product can be experienced [4].

There is also a similar research trend in the field of industry. Several methodologies to reflect users and context in the course of design have been devised and used, including usability analysis, scenario-based analysis, affordance-based design, context mapping, and so on.

At the Creative Design Institute, research into the interactions among users (consumers), designers, and products is being conducted so that various issues related to affordance are properly addressed in the design of products, spaces, and services [5, 6]. Figure 10.1 shows a diagram of the interactions among consumers, designers, and products representing the research philosophy of the Creative Design Institute. As can be seen in Figure 10.1, the cognitive and emotional aspects of consumers are very significant and should be properly addressed in the design of the products. The cognitive and emotional aspects of designers are seen as important factors for design processes and design creativity. In addition, the context – including social and cultural aspects – is very importance and its influences on consumers, designers, and products should be investigated. Consequently, if these aspects are properly managed then an enterprise can improve its competitiveness.

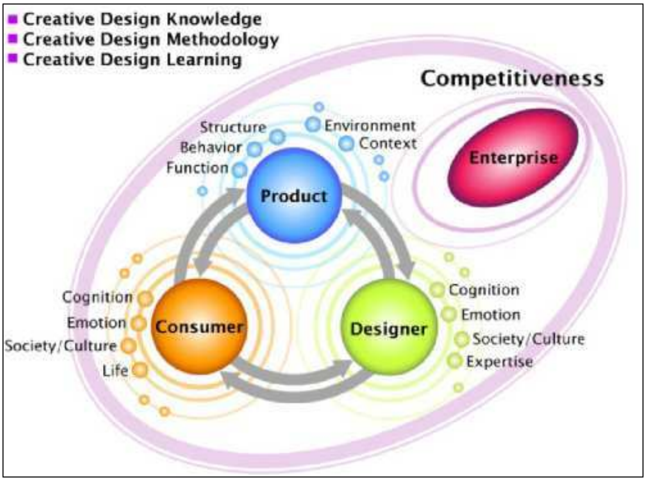


Figure 10.1 Interactions among consumers, designers, and products

This chapter addresses how user activities and perceived affordances are different and reflect personal creativity modes, which are determined by factual–intuitive perception inclination and subjective–objective decision preferences, as well as the introverted–extroverted nature of the user. We conducted a case study in a public space – a building lobby – used by many ordinary people. User activities and behaviors were analyzed in several specific tasks given to 20 students in the lobby of a building they had never previously visited. The tasks were devised so that various affordance features could be relevant, while eliminating factors affecting the affordance perception other than those due to the personal characteristics of the user. User activities can be classified into several different groups for each task based on the affordance features involved in their activities. These user activity differences are then compared with their personal creativity modes. For users of less common activities for some tasks, relevant personal cognitive characteristics have been identified.

10.2 Background

10.2.1 Affordance

Affordance was coined by perceptual psychologist James J. Gibson [7] as follows: "The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill.". It implies the interaction of the animal and the environment. Gibson's essential concept of affordance is that the relationship exists as a pair of animal and environment, and some parts of this relationship are the invariant features of the environment that permit the animal to do things. From an investigation of the affordances of everyday things such as doors, telephones, and so on, it was argued that the form of everyday things provides strong clues to their operation as a result of the mental interpretation of them, where the mental interpretation is based on people's past knowledge and experiences [8].

Instead of Gibson' affordance, Norman introduced *perceived affordance*, which is about characteristics in an object's appearance that give clues as to its proper operation. According to Norman, the Gibson's affordance is *real affordance*, which is about the physical characteristics of an object or interface that allow its operation [2]. Norman distinguished between two perspectives on products: design and emotion. He focused on two product factors – functionality and appearance (though there are a number of different constraints or considerations for designers and also users) – because these two are the most relevant for understanding the relationship between design and emotion. From the user's perspective, he focused on three kind of emotional response to products [1, 9]. Koutamanis mentioned that the affordances could be perceived by users differently from the designer's original intent [3]. As can be seen in Figure 10.2, the fence also provides the users with the affordance of sitting. This could be due to the different affordance perceptions of the users based on their different emotional characteristics.

Affordances are more than an addition to functional reasoning in building design [3]. And affordances for a conference room in space were proposed [5]. However, affordance has been recognized as a cognitive concept in the field of architectural design, and thus there has been very little research into developing a systematic methodology for designing architectural space to reflect the concept of affordance.

In the field of engineering design, there have been considerable research efforts into developing a design theory and methodology to reflect the concept of affordance. Maier and Fadel proposed Affordance-Based Design (ABD) to overcome the weaknesses of Function-Based Design, thus taking the synergy between affordance and function-based approaches to create better design [10, 11]. They also introduced Affordance-Structure-Matrix (ASM) for evaluating and grading the affordances embedded in each component of a product. This matrix can illustrate correlations of affordances and also of components [12].

Figure 10.2 The sitting affordances of a bench and fence [3]

Galvao and Sato proposed the Function–Task Interaction (FTI) Method. This method includes a general product development process and also an affordance method, especially the FTI matrix [14, 15]. In the FTI method, product functions and user tasks were derived from function decomposition and task analysis, and linked to each other in the FTI matrix.

The notion of affordance features in its explicit sense has recently been introduced. Murakami tried a formulation of affordance features for product design by experimenting with some simple shapes [16]. Structural elements tightly related to activities under specific contexts and tasks are identified as affordance features in our own team in two different design domains, in hand-held devices and in interior and space design [5, 6]. Affordance features are used in this research as a way to identify user and activity characteristics.

10.2.2 Personal Creativity Modes

Douglass J. Wilde of Stanford University developed the Personal Creativity Modes Test (PCMT), referred to as Wilde's test, based on the cognitive theory of Jung. Wilde's test has drawn considerable attention since it can be used for promoting the performance of the creative design activities of individuals or teams [17]. Personal creativity modes represent the different creativity modes of individuals, which are intrinsically related to their personal cognitive preferences [18]. At Stanford, Wilde's test has been used in composing design teams so that the personal creativity modes are distributed as evenly as possible throughout the team. As a way to verify the utility of this team composition method, they used the design team performances in a typical design competition as reflected in the quantity and quality of the awards received by Stanford design teams [17]. At the Creative Design Institute, research work is being conducted into various underlying elements of design creativity, including personal creativity modes, with the aim of developing training programs that reflect an individual learner's characteristics in order to enhance their design creativity [19]. Recently, the relationships between personal creativity modes and perceived creativity, and design team interaction have been studied [20].

According to Jung's cognitive theory, there are four aspects of personal cognitive preferences, including perceiving/judging preferences, factual/conceptual perception, thinking/feeling judgment, and introverted/extroverted cognitive motivation. These four aspects can be deployed into eight different modes of creativity [17], as shown in Tables 10.1 and 10.2, and Figure 10.3. The characteristics of each personal creativity mode have also been described in a more recent work [21].

Table 10.1 The eight personal creativity modes [18]

	Perception		Judgment	
	Factual (sensing)	Conceptual (intuitive)	Objective (thinking)	Subjective (feeling)
Introverted	Knowledge-based	Transforming	Analyzing	Evaluating
Extroverted	Experiential	Synthesizing	Organizing	Teamwork

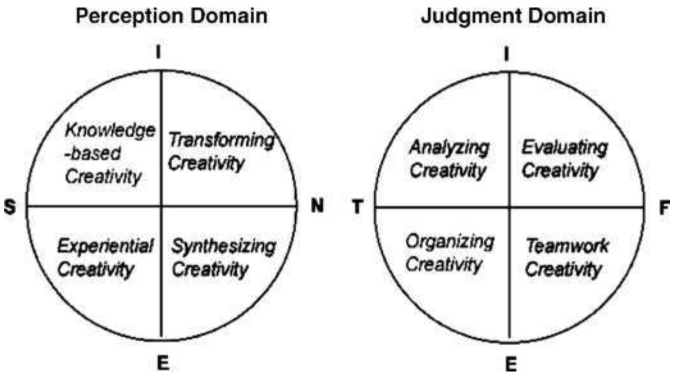
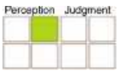
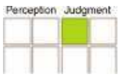
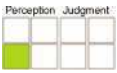
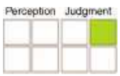


Figure 10.3 The eight personal creative modes

The personal creativity modes of each individual may have significant influence on their activities when they are asked to carry out certain tasks. In a space with a large number of unspecified floating individuals, users show a variety of activities under given tasks. When they are conducting the given tasks, they perceive many affordances and show diverse activities according to their personal characteristics. In other words, users perceive the same affordances in many different ways according to their personal characteristics and carry out various activities to accomplish the given tasks. Therefore, this chapter discusses the relationships between individual personal characteristics and the affordances that they perceive in the space throughout the case study. This research helps in the design of interior space to effectively provide the necessary affordances to its users based on their emotional and personal characteristics.

Table 10.2 Explanations for the eight personal creativity modes [21]

	Perception		Judgement
Synthesizing creativity 	Rearranging various elements into new configurations Seeing external patterns, trends, and relationships Exploring profitable new things and methods	Organizing creativity 	Organizing and managing people and projects to achieve goals Managing resources efficiently and enforcing specifications Setting deadlines, defining procedures, and breaking bottlenecks
Transforming creativity 	Transforming external objects as imaginary things	Analyzing creativity 	Internal reflective reasoning on relations among data and theories Clarifying ideas through analyzing by internal reasoning Comparing results with goals and standards
Experiential creativity 	Discovering new ideas and phenomena by direct experience Providing prompt, practical responses to crises and emergencies Building and testing models and prototypes	Teamwork creativity 	Building environment to support human values Detecting and fixing team interpersonal problems Harmonizing team, client, and consumer
Knowledge-based creativity 	Finding elements of solution in catalogs, handbooks, or class notes Getting or having existing facts and know-how Detecting and correcting mistakes	Evaluating creativity 	Using personal values to distinguish between good and bad Governed by a person's own values – aesthetic, ethical, moral, and spiritual Evaluating human factors and people's needs

10.3 Case Study – Experiments

The case study was conducted in order to investigate the relationships between personal and emotional characteristics and users' perception of affordance. A lobby space was selected for the case study, which had a large number of unspecified floating individuals and required a variety of affordances to perform many user tasks. The personal creativity modes of participating students were identified and their activities under the given tasks were monitored. Then the activities of participating students were analyzed, and their relationships with the personal creativity modes were studied by considering the associated affordances.

10.3.1 Participants and User Tasks

Twenty engineering students from Sungkyunkwan University participated in the case study. It is assumed that these students share similar cultural and societal backgrounds for the given simple tasks. Each student conducted the PCMT and their personal creativity modes were identified.

The lobby space of a commercial building, namely the P-building, which had a large number of unspecified floating individuals, were selected for the case study. We divided this research space into three sub-spaces overall according to their locations, and this was further divided into 11 zones. The floor plan of the research space and associated sub-spaces and zones is given in Figure 10.4. As can be seen in Figure 10.4, Space A was further divided into six zones and Space B consisted of three zones. Finally, Space C was composed of two zones. The details on the spaces and zones are summarized in Table 10.3.

The observation method was used to objectively monitor the participants' activities with video recording and by taking photos. The tests were conducted individually, and each test took 20–30 min with the whole test lasting 4 days. The detailed user tasks assigned to each participant, the associated sub-spaces, and the participants are summarized in Table 10.4. As shown in the table, in Space A, each participant was asked to wait and, at the same time, eat a snack (a cup of beverage and some boiled eggs), and read the magazine. Afterwards, participants moved into Space B's lobby, and were asked to sketch the impressive scene on the paper given. Finally, they were asked to go downstairs, and to lace their shoes and shake some sand out.

In this case study, the four tasks given in Table 10.4 were considered. However, all 20 students did not participate in all four tasks. In the cases of tasks T-1 and T-2, a total of 15 students participated, and these participants are given in Table 10.4. Similarly, 15 students participated in the task T-3 in space B, and these participants were not same as those for T-1 and T-2, as can be seen in Table 10.4.

Finally, in the case of T-4, five participants carried out the task in space A and nine ones in space C, and as a result a total of 14 students performed task T-4.

Five participants carried out the task in space A and nine ones did in space C, and as a result a total of 14 students performed task T-4.

Note that these tasks are simple enough so that the participants' emotional aspects beyond those related to their individual perception of the building lobby structures could be regarded as eliminated. Also note that these tasks are all individually performed so that no direct social interaction issues occur other than their indirect interaction with other unknown people in the lobby. In this way, the personal cognitive characteristics of each user as well as their affordance perception could play significant roles in their activities in performing the tasks.

Table 10.3 Spaces and zones of the lobby space of the P-building in the case study

Space A (on the second basement)	• SA-ZA: Artificial pond zone • SA-ZB: Revolving door zone • SA-ZC: Red stool zone A • SA-ZD: Red stool zone B • SA-ZE: LCD TV zone • SA-ZF: Bench zone
Space B (on the second floor)	• SB-ZA: Escalator sign bar zone • SB-ZB: Lobby zone • SB-ZC: Stair hall zone
Space C (on the first floor)	• SC-ZA: Stair zone • SC-ZB: Bollard zone (outside)

Table 10.4 Tasks and spaces for each student

Task No.	Task	Space	Participants	Sum
T-1	Waiting	Space A	P-01, P-03, P-05, P-07, P-09, P-11, P-12, P-13, P-14, P-15, P-16, P-17, P-18, P-19, P-20	15
T-2	Eating a snack and reading the magazine			
T-3	Sketching on paper	Space B	P-02, P-04, P-06, P-08, P-10, P-11, P-12, P-13, P-14, P-15, P-16, P-17, P-18, P-19, P-20	15
T-4	Lacing shoes and shaking sand	Space A	P-01, P-03, P-05, P-07, P-09	14
		Space C	P-11, P-12, P-14, P-15, P-16, P-17, P-18, P-19, P-20	

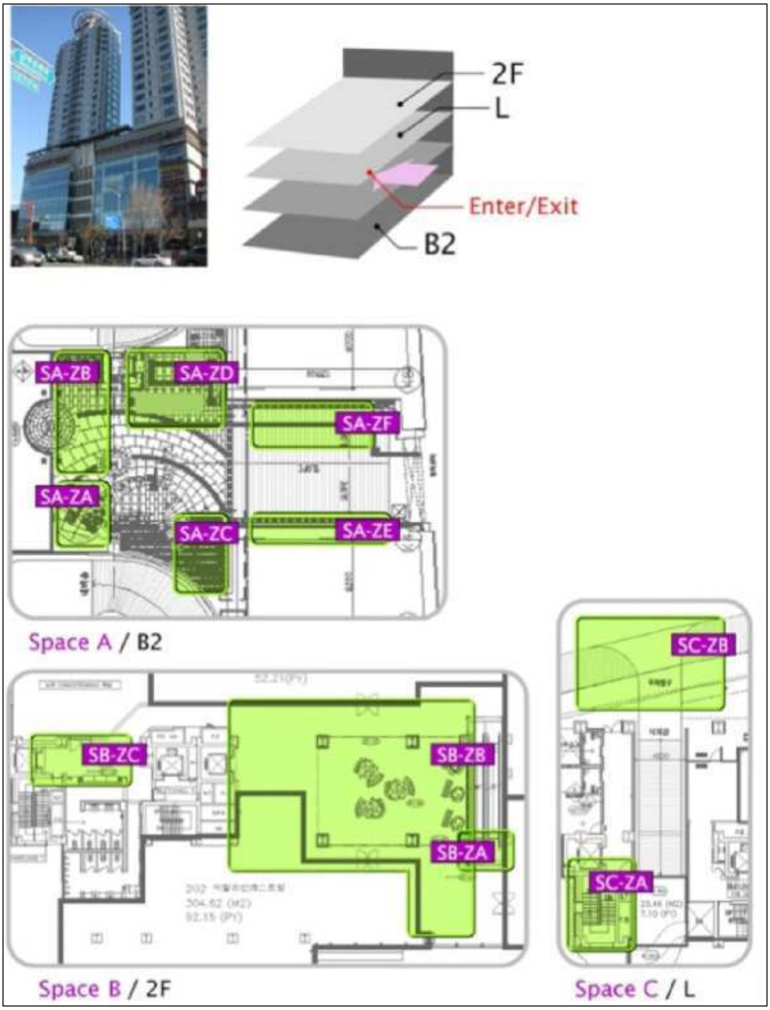


Figure 10.4 Floor plan of the research space for the case study

10.3.2 Activities and Related Affordances

We extracted the task-oriented affordances by observing the participants' activities under given tasks. From the results of the four tasks in the experiment, a total of 35 activities of all 20 participants were observed and listed in a sequential manner, and these are shown in Table 10.5. As can be seen in the table, one or more affordances were mapped to each participant's activity, and a total of 17 affordances were extracted from the activities of the participants.

Table 10.5 Activities and related affordances

Task No.	Activity No.	Activities of Studentes	Related Affordance			
			Physical Feature		Human Feature	
T-1	A-01	Looking for sitting place.	Look	Walk		
	A-02	Sitting.	Sit			
	A-03	Walking.	Walk			
	A-04	Stepping up and down.	Step			
	A-05	Leaning.	Lean			
	A-06	Placing the bag on somewhere.	Place			
	A-07	Supporting the bag on somewhere.	Support			
	A-08	Hanging the bag.			Hang	
	A-09	Looking something.	Look	Information Access		
	A-10	Touching something.	Touch			
T-2	A-11	Placing a cup of beverage and eggs.	Place			
	A-12	Holding a cup of baverage.			Hold	
	A-13	Drinking a baverage.	Drink			
	A-14	Tapping eggs.	Tap			
	A-15	Eating eggs.	Eat			
	A-16	Placing the magazine.	Place		Place	
	A-17	Supporting the magazine.			Support	
	A-18	Hold the magazine.			Hold	
	A-19	Reading the magazine.	Read			
T-3	A-20	Looking for place to draw painting.	Look	Walk		
	A-21	Touching tree.	Touch			
	A-22	Looking at framed picture.	Look			
	A-23	Looking at the warning message.	Information Access			
	A-24	Placing bags and belongings.	Place			
	A-25	Holding and Biting a cup.			Hold	
	A-26	Sitting and Leaning to draw painting.	Sit	Lean		
	A-27	Supporting paper.	Support		Hold	
	A-28	Biting a pen cap with teeth.			Hold	Support
	A-29	Drawing.	Draw			
T-4	A-30	Looking for place to lacing.	Look	Walk		
	A-31	Placing bags and belongings.	Place		Place	
	A-32	Placing foot.	Place			
	A-33	Holding shoe.			Hold	
	A-34	Lacing shoe.	Lace			
	A-35	Tapping shoe.	Tap		Tap	

There are two affordance drivers in the research space. One is a physical feature of the building itself, and the other is the human body or belongings. For instance, as noted in Table 10.5, the activity A-16 is mapped to two different place-abilities, represented by white shading and gray shading respectively. The "place-ability" denoted by the white shading is attributed to the physical feature of the space, and that denoted by the gray shading is due to the human body or a belonging such as a laptop or palmtop computer. The 35 activities given in Table 10.5 cannot be regarded as generic since they were monitored by observing the participants' activities in the particular research space. The relationships between the extracted affordances and activities can be rewritten as shown below. The activities with an apostrophe represent the one associated with the human body or belongings.

- Walk-ability: A-01, A-03, A-20, A-30
- Step-ability: A-04
- Sit-ability: A-02, A-26
- Lean-ability: A-05, A-26
- Place-ability: A-06, A-11, A-16, A-16', A-24, A-31, A-31', A-32
- Support-ability: A-07, A-17', A-27, A-27'
- Look-ability: A-01, A-09, A-20, A-22, A-30
- Information Access-ability: A-09, A-24
- Tap-ability: A-14, A-35, A-35'
- Touch-ability: A-10, A-21
- Hold-ability: A-12', A-18', A-25', A-27', A-28', A-33'
- Eat-ability: A-15
- Drink-ability: A-13
- Read-ability: A-19
- Draw-ability: A-29
- Lace-ability: A-34
- Hang-ability: A-08'

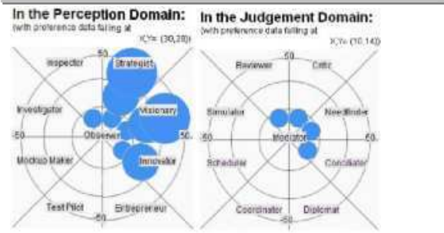
10.3.3 Personal Creativity Modes

The personal creativity modes of all 20 participants are summarized in Table 10.6. Diverse personal creativity modes were observed according to the participants' personal cognitive characteristics. For each student, the perception mode and the judgment mode are shown in separate diagrams. The stronger their preferences are, the bigger the bubbles that are shown, while their numeric values on intuitive/ sensing inclination, feeling/thinking orientation, and intro/extroverted motivation determine the cognitive characteristics.

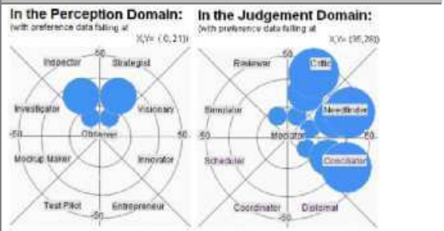
For example, in Table 10.6 student P-04 has very strong preference in extroverted intuition in perception mode as well as very strong feeling orientation in judgment. That is, he has strong synthesizing creativity as well as a mixture of both evaluating and teamwork creativities. On the other hand, student P-10 in Table 10.6 is very weak in both perception and judgment preferences.

Table 10.6 Personal creativity modes of participants

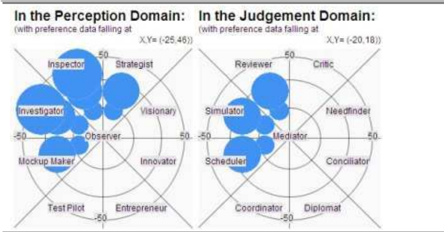
P-01



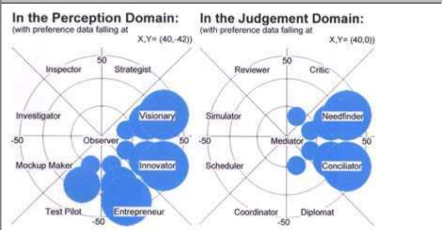
P-02



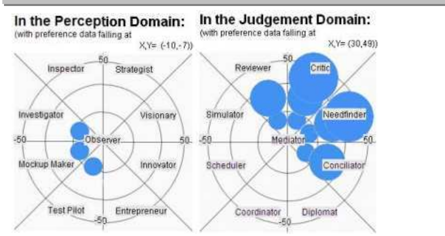
P-03



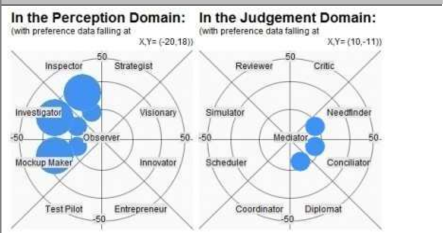
P-04



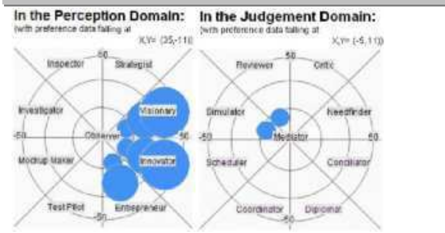
P-05



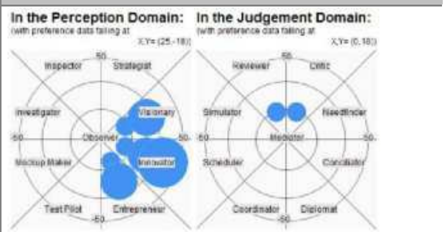
P-06



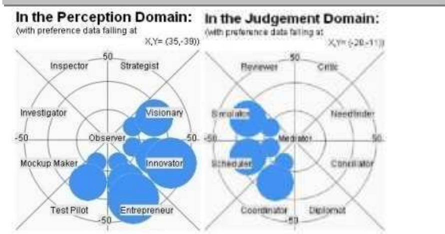
P-07



P-08



P-09



P-10

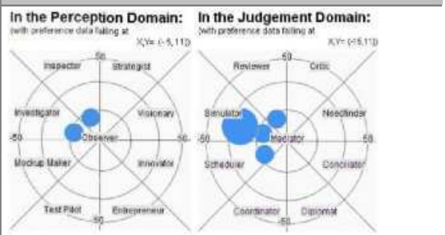
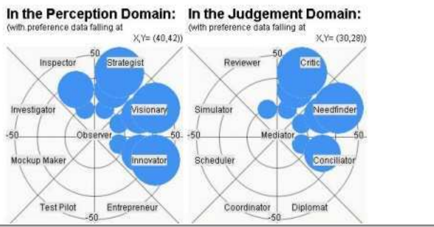
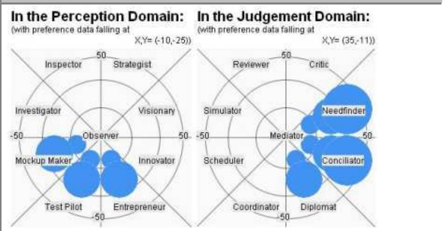
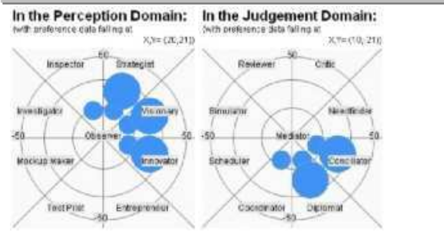
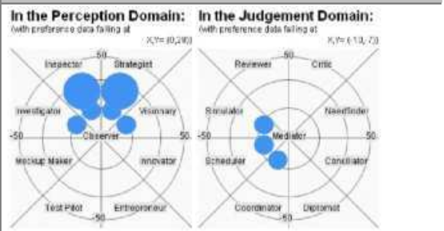
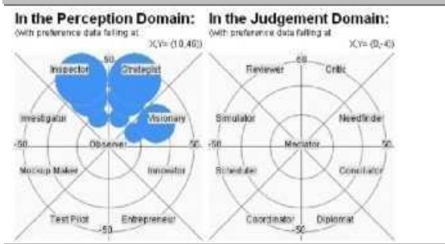
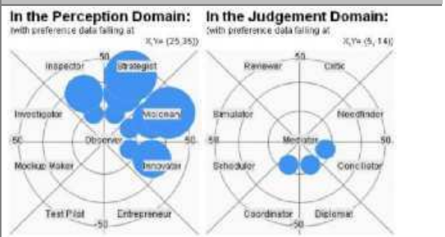
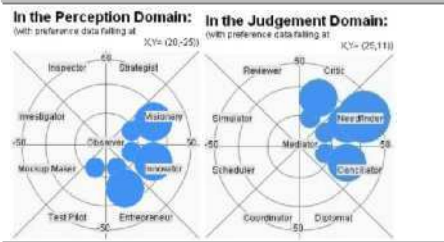
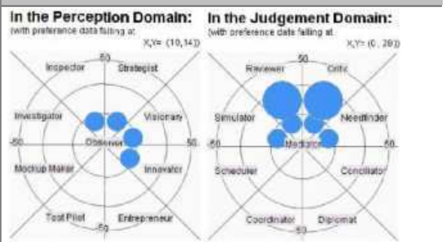
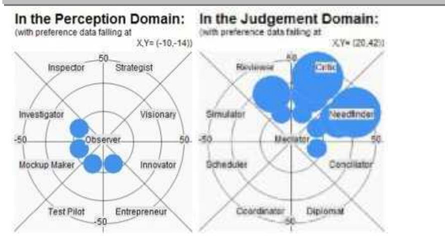
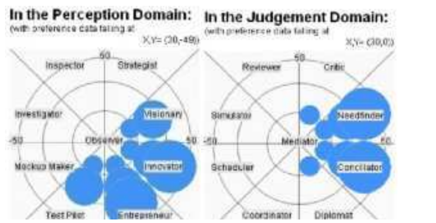


Table 10.6 Continued

<table>
<tr><td>P-11</td><td>P-12</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>P-13</td><td>P-14</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>P-15</td><td>P-16</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>P-17</td><td>P-18</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>P-19</td><td>P-20</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

10.4 Case Study – Analysis on Personal Creativity Modes and Actives

The participants' activities from four user tasks such as waiting, eating and reading, sketching, and lacing shoes and shaking sand were observed and their relationship with the personal creativity modes were analyzed. The observation method was applied to capture the personal and emotional characteristics of each participant from their various activities under the given tasks.

10.4.1 Task: Waiting (T-1)

The task of waiting, denoted by T-1, was assigned to participants in space A. The detailed composition of space A is shown in Figure 10.5. As can be seen in Figure 10.5, there were six zones in space A, namely, SA-ZA, SA-ZB, SA-ZC, SA-ZD, SA-ZE, and SA-ZF.

Figure 10.5 Detailed composition of space A

Basically, the six zones of space A were open zones. However, only SA-ZA, SA-ZB, SA-ZE, and SA-ZF were completely open. We could feel that SA-ZC and SA-ZD were somewhat separated due to the stairs between them. In particular, a more private atmosphere could be found at SA-ZC since it included a hidden area due to the winding stairs.

Scenes of the minor activities associated with the task of waiting (T-1) are shown in Figure 10.6. As can seen in Figure 10.6, P-05 and P-11 showed the activities of appreciating Gogh's painting and looking at the banner and LCD TVs introducing the floor guide and commercial space of the building. Their activities might be attributed to their curiosity, and were related to the affordances of look-ability, walk-ability, and information access-ability. P-11 and P-12 fooled around in space A looking at the displayed cards on the Christmas stall, and these activities were related to the affordances of look-ability, walk-ability, and step-ability. In particular, P-11 showed the activities of seeing the artificial pond, and touching the uneven materials on which water was flowing and the doll symbolizing the bookstore located in space A. These activities were associated with the affordances of look-ability and touch-ability.

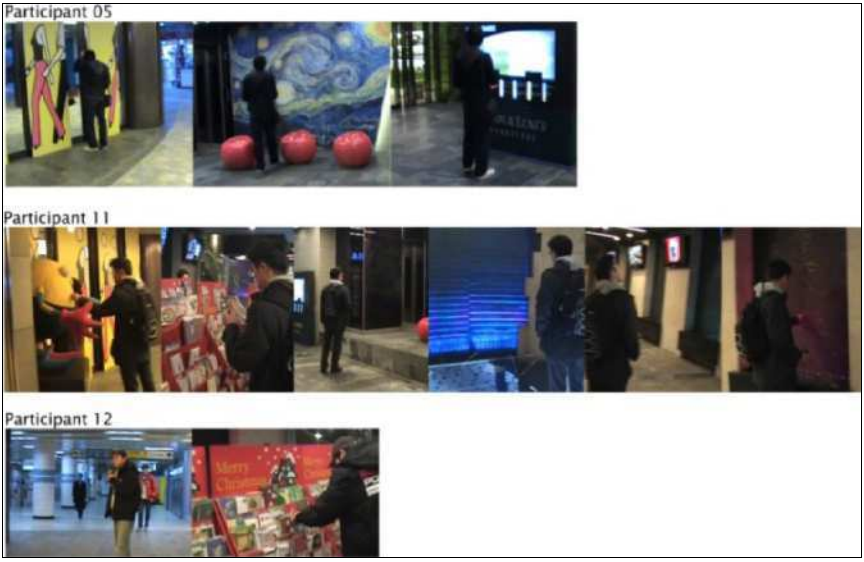


Figure 10.6 Activities associated with the task of waiting (T-1)

The personal creativity modes of P-05, P-11, and P-12 are given in Figure 10.7. As can be seen in Figure 10.7, their personal creativity modes are very similar in the judgment domain and they are feeling oriented. This means that they judged based on their emotion and had a lot of interactions with the features related to look-ability and information access-ability for their pleasure during the task of waiting. In addition, P-11 showed much more touching activities than P-05 and P-12. This means that P-11 perceived more stimuli from the features than others due to a much stronger inclination in the perception domain.

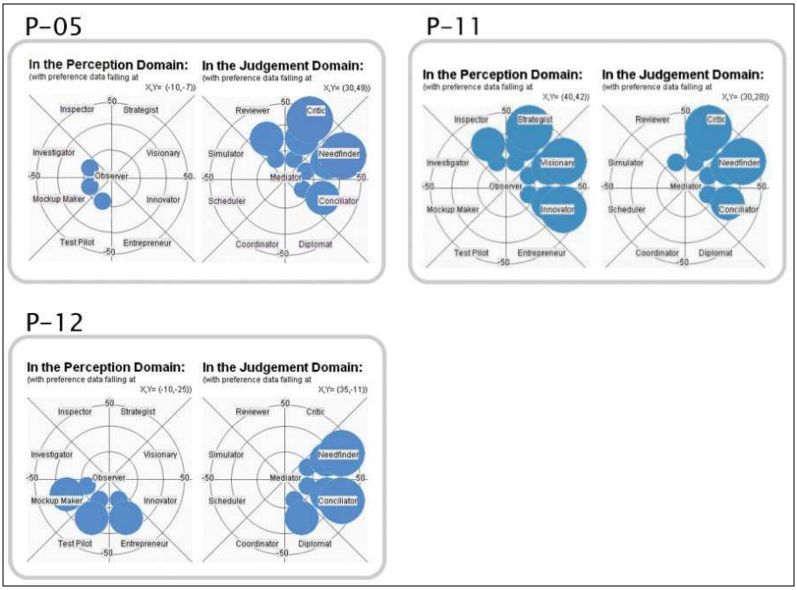


Figure 10.7 Personal creativity modes of P-05, P-11, and P-12

10.4.2 Task: Eating a Snack and Reading the Magazine (T-2)

The tasks of eating a snack and reading the magazine (T-2) were not asked to carry out in the same zone, but most of participants carried out those tasks in the same zone. A total of 15 participants carried out T-2, and ten of those, P-05, P-07, P-09, P-11, P-12, P-13, P-14, P-17, P-19, and P-20, performed T-2 at the bench in zone SA-ZF. This bench can allow several people to sit at the same time. The associated affordances were sit-ability, eat-ability, drink-ability, read-ability, place-ability, support-ability, tap-ability, and hold-ability. The scenes of the various activities of T-2 performed by participants are shown in Figure 10.8. As can be seen in Figure 10.8, most of participants performed T-2 while they were sitting. However, P-20 carried out T-2 while leaning on the feature that afforded lean-ability. This feature also afforded drink-ability and read-ability so that P-20 could eat their snack and read the magazine.

Unlike the above ten participants, P-03 and P-15 performed T-2 at the stone stools located in SA-ZD and SA-ZA respectively. These two zones had a more private atmosphere than SA-ZF. The affordances associated with the stone stool were sit-ability, eat-ability, drink-ability, read-ability, place-ability, support-ability, tap-ability, and hold-ability. The scenes of P-03 and P-15 while they were carrying out T-2 are shown in Figure 10.9. In the case of P-15, he continued to read the magazine while walking after he had performed the reading task in SA-ZA. His activities were related to the affordances of walk-ability, read-ability, and drink-ability.

Figure 10.8 Activities associated with the task of eating a snack and reading the magazine (T-2) – part 1

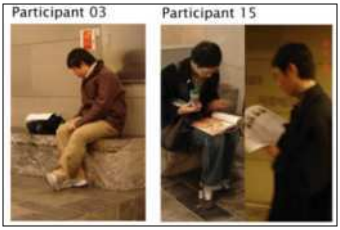


Figure 10.9 Activities associated with the task of eating a snack and reading the magazine (T-2) – part 2

The remaining three participants of the 15, P-01, P-16, and P-18, performed T-2 in the zone SA-ZC, which had the most private atmosphere. There were red stools in SA-ZC which were located at a from distance each other and only allowed one participant at a time to sit. As a result, each participant could have their own private space. These three participants were only ones to put their belongings on the floor. In particular, P-01 also placed his belongings on the feature that protruded from the wall. His activity was associated with the affordance of place-ability. The scenes of P-01, P-16, and P-18 performing T-2 are shown in Figure 10.10.

The personal creativity modes of five participants, P-01, P-03, P-15, P-16, and P-18, who showed minor activities of T-2 are shown in Figures 10.11 and 10.12. As can be seen in Figures 10.11 and 10.12, four of the participants, except P-18, showed a stronger inclination in the perception domain than the judgment domain and possessed a more introverted inclination. In addition, P-01, P-15, and P-16 were more intuitive in the perception domain.

Figure 10.10 Activities associated with the task of eating a snack and reading the magazine (T-2) – part 3

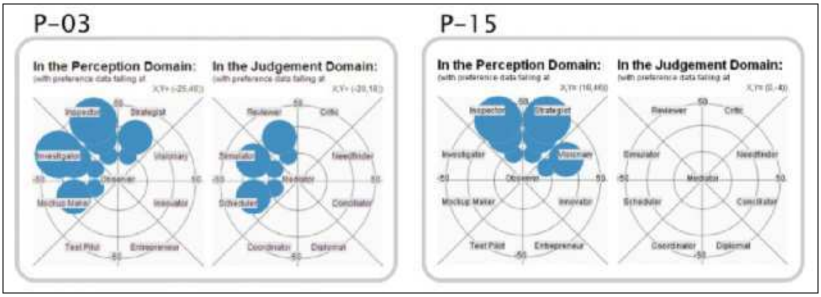


Figure 10.11 Personal creativity modes of P-03 and P-15

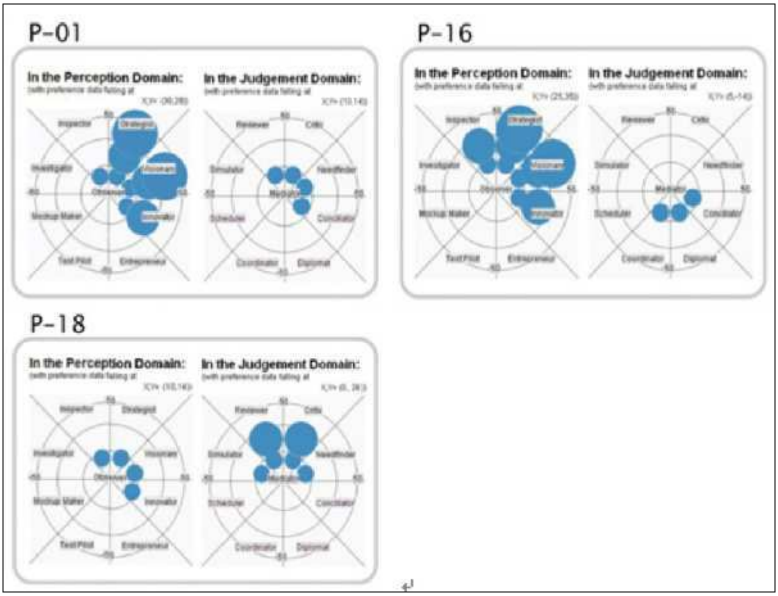


Figure 10.12 Personal creativity modes of P-01, P-16, and P-18

Of the above five participants, P-01 and P-16, who performed T-2 at zone SA-ZC, which was the most private, showed strong transforming creativity modes in the perception domain. This observation could be related to their introverted attitudes and insight with an intuitive sensing capability. In particular, P-01, who recognized the protruding wall and used it as the necessary affordance feature, representatively showed the characteristics of the transforming creativity mode. In the case of P-18, who selected zone SA-ZC for task T-2 but spent too much time on finding the zone to actually perform T-2, showed a stronger inclination in the judgment domain than the perception domain, but the degree of inclination strength in both domains was weaker than in the cases of P-01 and P-16. As a result, he seemed to spend more time to find the features providing the necessary affordances. In addition, his introverted inclination led him to find a quiet place with fewer people where he could sit. Note that none of the other 12 students shared similar personal creativity modes with P-01 and P-16.

It could be noted that P-01, P-15, P-16, and P-18 used their intuition to select spaces to perform their task of eating a snack and reading the magazine rather than judgment based on the fact. The introvertedness of P-01, P-03, P-15, P-16, and P-18 could also have led them to find a more private zone to carry out the task.

10.4.3 Task: Sketching on Paper (T-3)

The task carried out in space B was to sketch the impressive scenes of the lobby space on the sheet of paper provided. This task aimed to discover which features in the space were used to afford the drawing activity on the flexible paper. The detailed view of space B is shown in Figure 10.13.

As can be seen in Figure 10.13, there were three zones in space B, namely, SB-ZA, SB-ZB, and SB-ZC. In particular, there were three sets of artificial trees with heights of about 3 m. These trees each had a cylindrical base with a height of

Figure 10.13 Detailed composition of space B

0.5 m and a diameter of 1.3 m, which could afford the sitting of people. Thus, a message warning “Do not sit” was attached to the tree.

Of the 15 participants who carried out task T-3, five participants, P-08, P-12, P-13, P-14, and P-17, did not sit on the cylindrical base of the artificial trees after looking at the warning message, but used the flat surface of the cylindrical base as the affordance feature to sketch the scene. They placed their belongings and paper on the top surface of the cylindrical base, sat down on their heels, and sketched the scenes on the paper. The associated affordances were draw-ability, support-ability, place-ability, and information access-ability. The activities of the above five participants are shown in Figure 10.14.

Four participants, P-06, P-18, P-19, and P-20, used column and wall features to carry out task T-3. They perceived the affordance of support-ability of column or wall features to hold the paper with their hands in order to sketch the scene. Figure 10.15 shows their activities. The affordances they perceived were draw-ability, support-ability, and hang-ability.

The participants P-04, P-11, and P-15 ignored the warning message and sat on the cylindrical base of the artificial tree while sketching the scene. The scenes of their activities are shown in Figure 10.16. In this case the affordances they perceived were sit-ability, draw-ability, place-ability, support-ability, and information access-ability.

Finally, P-02, P-10, and P-16 used their palms to support the paper and sketch the scene while they were performing task T-3. In particular, P-16 leaned on the column while sketching the scene with the paper supported by his palm. The activities of those three participants are shown in Figure 10.17. The affordances they

Figure 10.14 Activities associated with the task of sketching on paper (T-3) – part 1

Figure 10.15 Activities associated with the task of sketching on paper (T-3) – part 2

Figure 10.16 Activities associated with the task of sketching on paper (T-3) – part 3

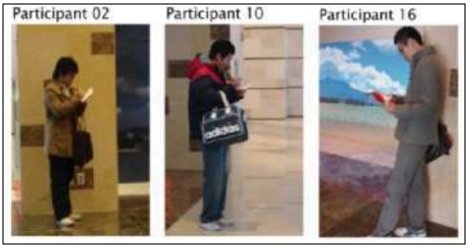


Figure 10.17 Activities associated with the task of sketching on paper (T-3) – part 4

perceived were draw-ability, lean-ability, support-ability, hold-ability, and hang-ability. These three participants did not find any features to place belongings and support the paper, and used their own human body features to perform task T-3.

After analyzing the diverse activities of the participants who carried out task T-3 and their personal creativity modes, those who sat on the cylindrical base in spite of the warning message (P-04, P-11, and P-15) showed very strong cognitive inclinations with the coordinate values larger than 40. Their personal creativity modes are given in Figure 10.18. It is believed that they intuitively responded to the cylindrical base to afford the sit-ability before recognizing the warning message due to their strong inclination in the perception domain.

The personal creativity modes of those who used their palms to support the paper and sketch the scene in task T-3 are given in Figure 10.19. As can be seen in Figure 10.19, P-02 and P-10 showed a stronger inclination in the judgment domain than the perception domain, and were introvert in both domains. They rather passively performed task T-3 without actively searching for the affordance features in the space. They used their own body features to hold bags or a beverage cup and to sketch the scene while performing task T-3. In other words, it seemed that they may not have perceived their surrounding environment and hence made their own judgment in order to perform the task T-3.

On the other hand, P-16's activity in carrying out task T-3 was somewhat different. Although he showed a similar posture when drawing the scene on the paper to the cases of P-02 and P-10, he leaned on the column feature while sketching. P-16's personal creativity mode showed a much stronger inclination in the perception domain than the judgment domain. This characteristic might have led him to find the column feature to lean on while sketching the scene on the paper supported by his palm. Therefore, his different personal creativity mode may have led him to show a different activity from those of P-02 and P-10. It may be the case that the leaning affordance was stronger for his perception preference than any other affordances. Then, P-02 and P-10 may be the only two who did not perceive any affordance features other than their body features, reflecting their very low perception preferences.

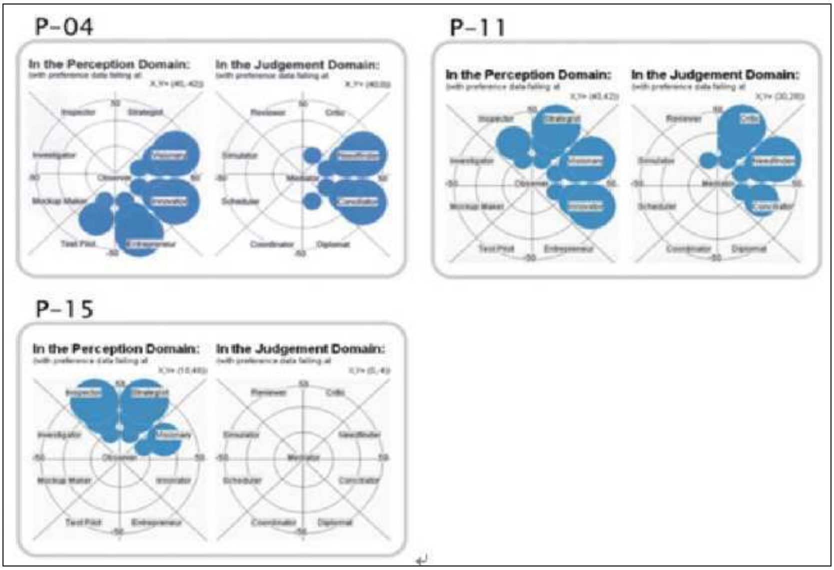


Figure 10.18 Personal creativity modes of P-04, P-11, and P-15

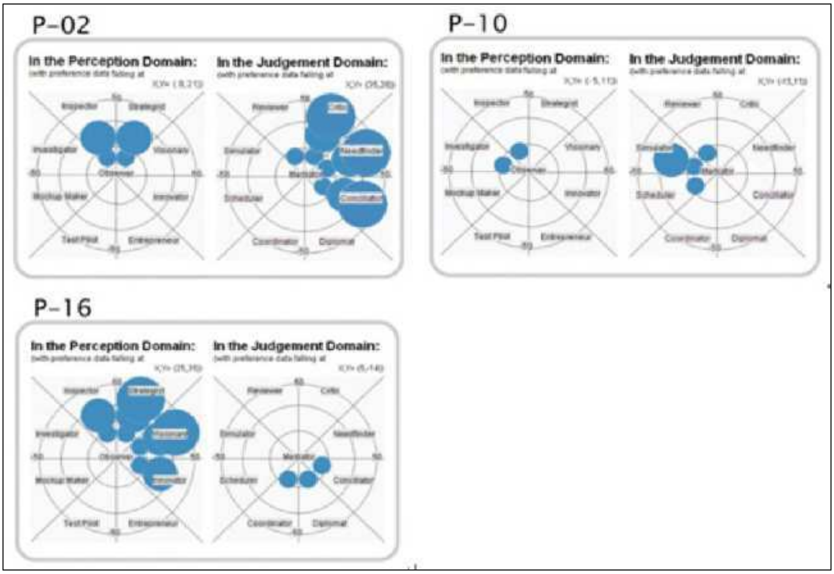


Figure 10.19 Personal creativity modes of P-02, P-10, and P-16

10.4.4 Task: Lacing Shoes and Shaking Sand (T-4)

Two experiments were conducted for the task of lacing shoes and shaking sand (T-4). In the first experiment, the participants carried out task T-4 in space A. In the second experiment, task T-4 was performed in space C. The detailed view of space C is given in Figure 10.20. As can be seen in the figure, there are two zones in space C, namely, SC-ZA and SC-ZB. In particular, zone SC-ZB was the exterior one.

Scenes of the diverse activities of the participants when carrying out task T-4 are shown in Figures 10.21–10.23. As can be seen in Figure 10.21, six out of a total of 14 participants, P-11, P-12, P-16, P-17, P-19, and P-20, performed task T-4 while they were sitting on their heels on the floor. The affordances they perceived were lace-ability and tap-ability.

Figure 10.20 Detailed composition of space C

Figure 10.21 Activities associated with the task of lacing shoes and shaking sand (T-4) – part 1

In addition, five participants were seen lacing shoes and shaking sand while they were sitting on a stool, bench, or bollard. The scenes of their activities are given in Figure 10.22. As can be seen in Figure 10.22, P-01, P-03, P-05, and P-18 laced shoes by leaning forward with their backs bent while they were sitting. On the other hand, P-09 laced his shoes while sitting on the bench by raising one leg and putting it across the other leg without overbending his back. The affordances they perceived were lace-ability, tap-ability, and place-ability.

Finally, P-07, P-14, and P-15 raised their legs and put them on a bench or bollard in order to lace their shoes with their backs bent, overbending their backs. The scenes of their activities are shown in Figure 10.23. As can be seen in Figure 10.23, P-07 used the bench to put her leg on and lace her shoes in zone SA-ZF of space A, and P-14 and P-15 put their legs on the bollards and laced their shoes in zone SC-ZB of space C. The affordances perceived by those three participants were lace-ability, tap-ability, and support-ability.

When investigating the relationship between personal creativity modes and activities of the participants carrying out task T-4, P-14 and P-15, who put their legs on the bollard to lace their shoes, showed similar personal characteristics. The personal creativity modes of P-14 and P-15 are given in Figure 10.24. As can be seen in Figure 10.24, they had a stronger cognitive inclination in the perception domain than the judgment domain. They also showed introvertedness. When they conducted the task of lacing their shoes, it is believed that they perceived the bollard as the affordance feature to support their activities by using both their intuitive sensing and concrete and factual information.

Figure 10.22 Activities associated with the task of lacing shoes and shaking sand (T-4) – part 2

Figure 10.23 Activities associated with the task of lacing shoes and shaking sand (T-4) – part 3

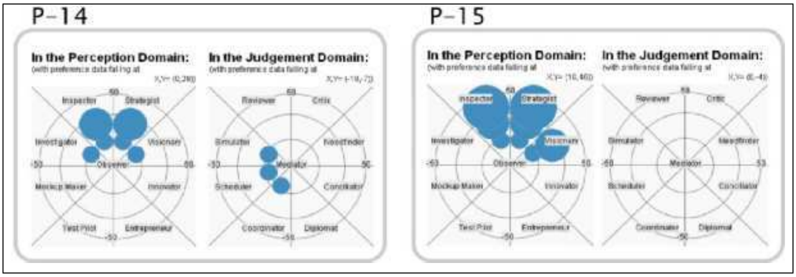


Figure 10.24 Personal creativity modes of P-14 and P-15

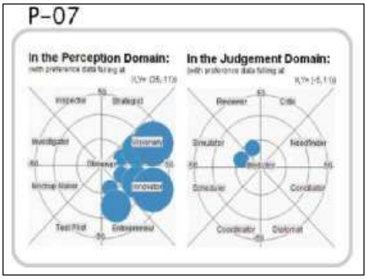


Figure 10.25 Personal creativity mode of P-07

On the other hand, P-07 showed similar activities to P-14 and P-15 when carrying out task T-4. This may be due to her stronger cognitive inclination in the perception domain than the judgment domain, which is similar to the cases of P-14 and P-15. The personal creativity mode of P-07 is shown in Figure 10.25. However, as can be seen in Figure 10.25, she had more extrovertedness in her personal characteristics. This characteristic may have led her to find a more publicly open zone such as SA-ZF while searching for the affordance feature to support her activity, which is different from the cases of P-14 and P-15 who used zone SC-ZB to carry out their task.

10.4.5 Discussions

As discussed in the previous sections, humans perceive diverse features in different ways based on their personal cognitive and emotional characteristics. In other words, humans do not think and act equally although they look at the same features. In addition, humans perceive different affordances based on their personal characteristics, resulting in different activities to carry out the same tasks in the same environment.

In this research, a total of 17 affordances were used as extracted by observing the participants' actual activities. It is observed that those affordances could or could not be perceived according to the participants' personal cognitive characteristics. It was also found that the user of the lobby space could perceive affordances that the designers may not have originally considered. In particular, when participants perceived different affordances, their activities and behaviors varied. If some participants behaved similarly, but in less common manners compared with the majority of the participants, they had similar personal creativity modes. For instance, participants with transforming or synthesizing creativity modes could intuitively perceive affordances that the designers had not originally intended.

Therefore, it may be necessary for designers to consider various users' activities in the course of the design of the lobby space. For example, when designers are to provide the affordance of sit-ability for the lobby space, they may have to provide various features to afford the sit-ability, whilst considering diverse users and contexts. In addition, they also need to expect the users to perceive unintended affordances.

10.5 Conclusions

In this research, the analysis on the user activities and perceived affordances reflecting personal cognitive characteristics was conducted through the case study in a public space – the building lobby. The user activities and behaviors were monitored and analyzed under several tasks. We developed the tasks so that various affordance features could be relevant while eliminating other factors affecting the affordance perception of users. The users' activities were classified and then compared with their personal creativity modes. It was observed that the personal cognitive characteristics could be related to the different and diverse perceptions of affordances. While this study presents a qualitative explanation that personal cognitive characteristics may affect affordance perception, a more systematic user study is desired in order to justify the initial findings, possibly with quantitative support. Similar user studies could also be conducted to understand the relationships between affordance perception and other user characteristics such as physical, cultural, and social aspects. Many such studies when properly compiled and classified would support designers in their affordance-based design processes.

References

1. Norman D.A.: Emotional Design. Basic Books, New York (2004)
2. Norman D.A.: The Design of Everyday Things. Basic Books, New York (2002)
3. Koutamanis A.: Buildings and affordances. In: International Conference on Design, Computing and Cognition, pp. 345–364. Springer, Heidelberg (2006)
4. Visser F.S., Stappers P.J., Derlugt R.V.: Contextmapping: experiences from practice. Int. J. CoCreation in Design and the Arts **1**(2):119–149 (2005)

5. Kim Y.S., Kim M.K., Lee S.W., Lee C.S., Lee C.H., Lim J.S.: Affordances in interior design: a case study of affordances in interior design of conference room using enhanced function and task interaction. In: Proceeding of ASME International Conference on Design Theory and Methodology, Las Vegas, Nevada, DETC2007-35864 (2007)
6. Kim M.K., Jeong J.Y., Kim Y.S., Lee S.W., Lim JS.: Identifying affordance features from user activity study. Paper presented at the International Conference on Design Computing and Cognition, Atlanta, GA (2008)
7. Gibson J.J.: The Theory of Affordances in the Ecological Approach to Visual Perception. Houghton Mifflin (1979)
8. Norman D.A.: The Psychology of Everyday Things. Basic Books, New York (1988)
9. Norman D.A., Ortony A.: Designers and users: two perspectives on emotion and design. In: Symposium on Foundations of Interaction Design, Ivrea, Italy (2003)
10. Maier J.R.A., Fadel GM.: Affordance-based methods for design. In: Proceedings of ASME International Conference on Design Theory and Methodology, Chicago, DETC2003-48673 (2003)
11. Maier J.R.A., Fadel G.M.: A case study contrasting german systematic engineering design with affordance based design. In: Proceedings of ASME International Conference on Design Theory and Methodology, Long Beach, DETC2005-84954 (2005)
12. Maier J.R.A., Ezhilan T., Fadel G.M.: The affordance structure matrix – a concept exploration and attention directing tool for affordance based design. In: Proceedings of ASME International Conference on Design Theory and Methodology, Las Vegas, Nevada, DETC2007-34526 (2007)
13. Brown D.C, Blessing L.: The relationship between function and affordance. In: Proceedings of ASME International Conference on Design Theory and Methodology, Long Beach, DETC2005-85017 (2005)
14. Galvao A.B, Sato K.: Affordances in product architecture: linking technical functions and users' tasks. In: Proceedings of ASME International Conference on Design Theory and Methodology, Long Beach, DETC2005-84525 (2005)
15. Galvao A.B., Sato K.: Incorporating affordances into product architecture: methodology and case study. In: Proceedings of ASME International Conference on Design Theory and Methodology, Philadelphia, DETC2006-99404 (2006)
16. Murakami T., Cheng L.M., Higuchi M., Yanagisawa H.: Trial for formulation of affordance feature for product design. In: Proceedings of the Symposium on Human Interface, Japan, pp. 403–408 (in Japanese) (2006)
17. Wilde D.J.: Design team role. In: Proceedings of ASME International Conference on Design Theory and Methodology, Las Vegas, Nevada (1999)
18. Wilde D.J., Labno D.B.: Personality and the creative impulse. Unpublished manuscript (2001)
19. Kim Y.S., Jin S.T., Lee S.W.: Design activity and personal creativity characteristics: a case study of dual protocol analysis using design information and process. In: Proceedings of ASME International Conference on Design Theory and Methodology, Philadelphia, DETC2006-99654 (2006)
20. Kim M.S., Kim Y.S., Kim T.H.: Analysis of team interaction and team creativity of student design teams based on personal creativity modes. In: Proceedings of ASME International Conference on Design Theory and Methodology, Las Vegas, Nevada, DETC2007-35378 (2007)
21. Levesque L.C.: Breakthrough Creativity: Achieving Top Performance Using the Eight Creative Talents. Davies-Black, Palo Alto, CA (2001)

Chapter 11
Emotion-oriented Human–Computer Interaction

Toshiyuki Yamashita[1], Ahmad Eibo, Takumi Ichimura, and Kazuya Mera

Abstract In the information-oriented society, people spend a lot of time interacting with computer systems. Therefore, we need to develop devices which make it a pleasure to interact with computer systems. In human face-to-face communication, nonverbal messages such as facial expressions, gestures, body movements, and postures can convey a speaker's emotions better than verbal messages. Among verbal messages, chiming in to express a listener's support and repeating the speaker's utterances can convey the listener's emotions, encourage the speaker to continue his/her talk, and make the conversation go smoothly. Therefore, we have developed several human–computer interactive systems using facial expressions as a human interface, and have constructed a human–computer interface which chimed in with the user and repeated the user's responses.

11.1 Introduction

Human face-to-face communication consists of both verbal messages and nonverbal messages. Nonverbal messages are composed of facial expressions, hand gestures, body movements, and postures, in addition to paralanguages such as tone, speaking rate, pause, and stumbling. Facial expressions in particular show an informative signature of a person's emotional state and persons can recognize emotions from another's facial expressions with great accuracy and consistency [1]. Therefore, attention to using facial expressions as human interfaces for human–computer interaction (HCI) has grown in recent years [2, 3, 4, 5]. In this chapter, we address our systems that can select and display the facial expression which expresses the user's emotions caused by HCI in order to make HCI pleasant.

[1] T. Yamashita
Tokyo Metropolitan University, 1-1-1, Minami-osawa, Hachioji, 192-0397 Tokyo, Japan
e-mail: yamashita-toshiyuki@center.tmu.ac.jp

Some kinds of verbal messages carry the same functions as nonverbal messages. Imagine a scene where two persons converse pleasantly. If one person suddenly stops replying to the other person, an unpleasant silence will fall upon them. In this human face-to-face communication, chiming in with remarks, repeating the other person's utterances, and so on, can sometimes encourage him/her to continue speaking, give pleasure to the communication, and smooth the conversation. This is also true of HCI. If the computer cannot understand a user's utterances and does not reply, the user will be discouraged and not want to continue the interaction. However, if the computer chimes in, or nods, or repeats the user's utterances, he/she will be encouraged to continue the interaction. For these reasons, we have constructed human interfaces which chime in with the user and repeat the user's utterances. In this chapter, we discuss the effects this has on HCI and describe our systems.

11.2 Facial Expressions as a Human Interface

In order to use faces for a pleasant HCI, we have developed both fuzzy reasoning [6, 7, 8] and neural network models [9, 10] for selecting the facial expression which expresses the emotions caused by several situations and have developed several applications using our models. First, we explain a face selection model by fuzzy reasoning and an application to e-mail.

11.2.1 Face Selection by Fuzzy Reasoning Model

Yamashita *et al.* [6, 7, 8] used nine facial expressions with three levels of brow deflection and three levels of mouth deflection, as shown in Figure 11.1.

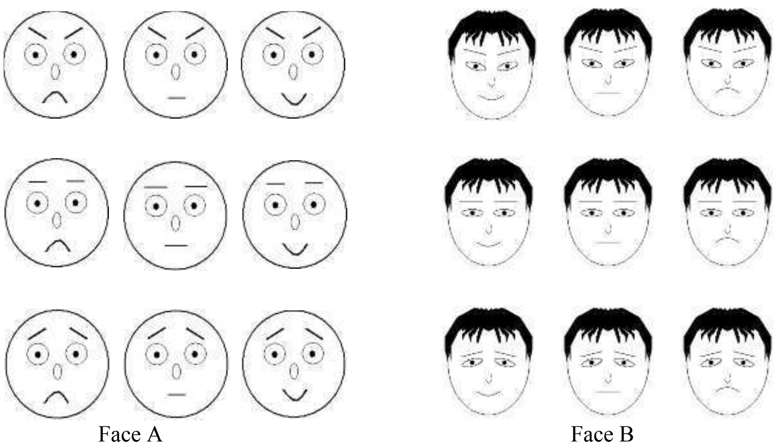


Figure 11.1 Nine facial expressions of Face A and Face B

Following Plutchik [11], we used the following eight emotions: "happiness", "grief", "anger", "disgust", "surprise", "fear", "anticipation", and "resignation".

We asked the subjects to select one or several facial expressions which expressed a given emotion. We assumed that the ratio of the subjects who selected the facial expression for a given emotion was the membership value in the fuzzy set of the facial expressions which expressed the emotion.

As shown in Figure 11.2, Yamashita *et al.* [6, 7, 8] proposed the following fuzzy reasoning model:

$$
\begin{array}{lll}
\text{Rule 1} & : A_1 & \Rightarrow P_1 \\
\text{Rule 2} & : A_2 & \Rightarrow P_2 \\
 & \bullet\ \bullet\ \bullet & \\
\text{Rule 8} & : A_8 & \Rightarrow P_8 \\
\text{Input} & : a_1 \text{ and } a_2 \text{ and} \ldots a_8 & \\
\hline
\text{Conclusion}: & & P',
\end{array}
$$

where $A_1, A_2, \ldots, A_8$ in the antecedent part are fuzzy sets of emotions, and P_1, P_2, …, P_8 in the consequent part are fuzzy sets of facial expressions which express a given emotion.

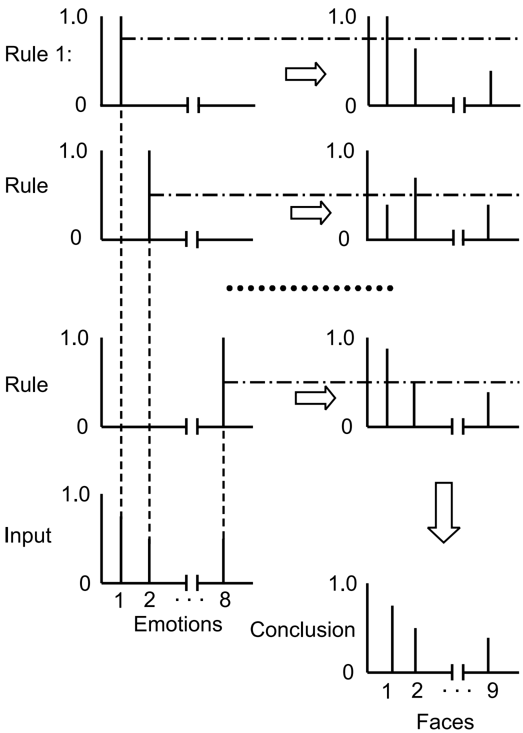


Figure 11.2 Fuzzy reasoning model for selecting a facial expression

If we have a_1, a_2, …, a_8 as the intensity of each emotion, then the fuzzy reasoning result of Rule i ($i = 1, 2, \ldots, 8$) is calculated by:

$$\mu_{P_i'}(z) = a_i \wedge \mu_{P_i}(z),$$

and the combined conclusion of Rule 1, Rule 2, …, Rule 8 is given as:

$$\mu_{P'}(z) = \mu_{P1'}(z) \vee \mu_{P2'}(z) \vee \ldots \vee \mu_{P8'}(z).$$

The facial expression with the highest or the second highest grade of membership should be selected as the facial expression which best expresses the emotions caused by a given situation.

11.2.2 An Application of the Model to e-mail

We applied our fuzzy reasoning model to an e-mail system, which could select and display the facial expression that expressed the emotions caused by the situation written in the sentences on the computer display.

Figure 11.3 shows an illustration of our system. First, the system receives sentences input by the user or from e-mail. Then it decomposes the sentences into words and transforms these words into original forms in terms of inflection [12]. Next, these words are compared to the words in the database for each emotion.

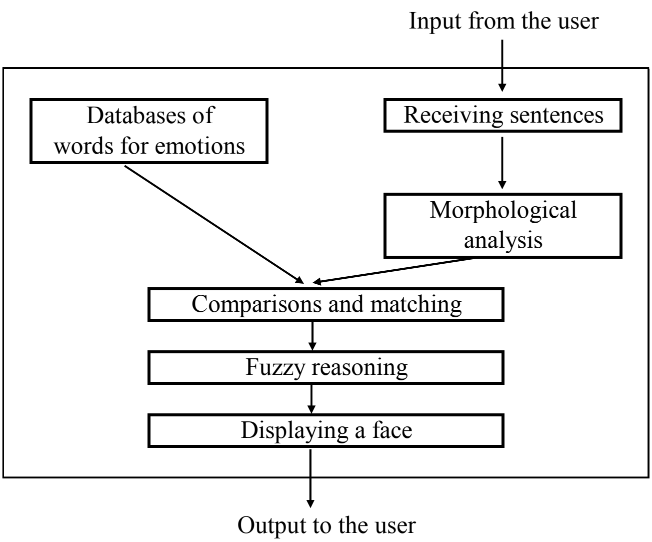


Figure 11.3 Overview of the system

For example, the sentences of e-mail in Figure 11.4 “I am very *glad* to hear that your son entered a junior high school this year. It seems only yesterday that I congratulated you on your son’s entrance to an elementary school. I am *surprised* that children grow very quickly. I can visualize your *joyful* look. Congratulations!”.

In these sentences, the first and third highlighted words “glad” and “joyful” are identified as representing the emotion “happiness”. The second highlighted word “surprise” is identified as “surprise”. Therefore, “surprise” has a membership value of 0.33 (= 1/3) and “happiness” has a membership value of 0.67 (= 2/3). These membership values are used as the input values for fuzzy reasoning. As a result of fuzzy reasoning, the facial expression with the highest membership value is selected and displayed, as shown in Figure 11.4.

Figure 11.5 shows an example of a computer display. Sentences are displayed on the upper side of the computer display, and a face is displayed below the sentences.

Our system can be applied to different interfaces. For example, Figure 11.6 shows our system used in two computer games [6, 8]. The left-hand side shows an example of the display for a computer game similar to the game Breakout. Since the player got a lower score than he expected, a facial expression showing his miserable feelings is displayed. The right-hand side shows an example of the display for a slot machine game. Since the player got the highest score, a facial expression showing happiness is displayed.

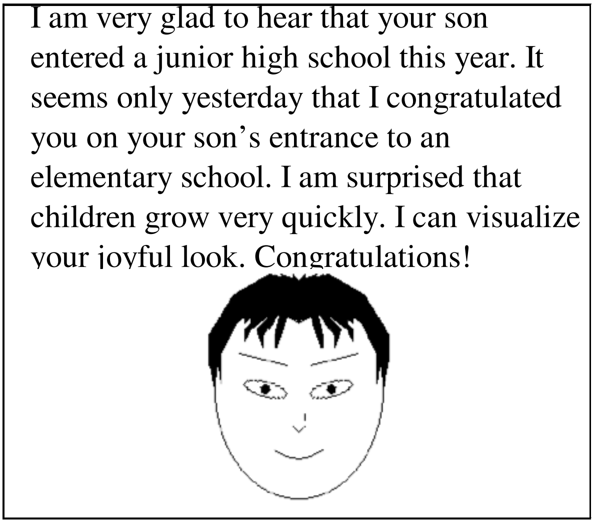


Figure 11.4 Example of e-mail display

Figure 11.5 Example of a computer display

(a) (b)

Figure 11.6 Examples of applications to computer games: (a) the computer game Breakout, and (b) a slot machine game

11.2.3 Comparisons Among Faces

People often insert a so-called "face mark" ("smilies" or "emoticons") immediately after a particular sentence in an e-mail in order to convey the sender's emotions to the reader. As shown in Figure 11.7, the face marks are composed of simple characters. Although the face marks extremely simplify human facial expressions, they strongly convey the sender's emotions and make exchanging e-mails smoother.

Laughing: (^_^) (^o^) :) : -)

Be in a cold sweat : (^_^;) ^_^;; (^^;;

Be in tears: (; _ ;) (>_<) ; -< X -(

Dissatisfied: - (

Put out one's tongue: -P

Figure 11.7 Examples of face marks

We compared the effects of our faces with those of face marks in an e-mail [13]. We used three kinds of faces, that is, Face A and Face B as shown in Figure 11.1, and a face mark, for comparison. Moreover, we adopted two display conditions, that is, the one condition in which a relatively big face was displayed below the sentences, and another condition in which a small face was displayed right after a specific sentence. As shown Figure 11.8, three kinds of faces and two kinds of sizes were combined and produced six display conditions. We added the display condition without a face to the six display conditions. The subjects were asked to view each display and to rate the items such as "cannot obtain information quickly/can obtain information quickly", "unpleasant/pleasant", and "little information/much information" on a five-point scale ranging from 1 ("do not think so") to 5 ("think so"), with a score of 3 being neutral.

A factor analysis extracted three factors, that is, "ease of getting information", "pleasantness and familiarity", and "definition of information". The factor scores of the display conditions suggest the following: (1) facial expressions can add impression of "pleasantness and familiarity" to the information from the computer display. In particular, the small face mark, the big face mark, and small Face A can give the strong impression of "pleasantness and familiarity". (2) A small simple face mark added to a particular sentence can provide the user with the most information from an e-mail. (3) Showing the big face mark, big Face A, and big Face B below the sentences can give the impression that the sentences have "definite information".

I went to see a movie yesterday.
The movie was as amusing as I had expected. 
I recommend you to see this movie.

(a)

I went to see a movie yesterday. The movie
was as amusing as I had expected.
I recommend you to see this movie.

(b)

I went to see a movie yesterday. The movie was
as amusing as I had expected. (^_^)
I recommend you to see this movie.

(c)

I went to see a movie yesterday. The movie was
as amusing as I had expected.
I recommend you to see this movie.
(^_^)

(d)

Figure 11.8 Examples of e-mail sentences with faces: (a) an example of small Face A, (b) an example of big Face B, (c) an example of a small face mark, and (d) an example of a big face mark

11.2.4 User's Faces

We developed the system using user's faces. For generating the facial expression which expresses a given emotion, we used a sand glass type neural network as shown in Figure 11.9. A sand glass type neural network consists of two or more layered neural networks, which can learn some different teaching signals simultaneously and condense the input signals into hidden neurons and restore the original data from the hidden neurons [9, 10]. This section prepares the five-layer neural networks, which have 255 neurons in the first layer and fifth layer, 40 neurons in the second layer and fourth layer, and two neurons in the third layer. The third layers in each network are connected to each other and the neurons in the other layers activate independently in each network. The same teaching signals are set into the input and output neurons in each network to perform an identity mapping. The special feature of faces is compressed into the hidden neurons by Back Propagation Learning. In particular, the two neurons in the third layer can depict the condensed features in the planner lattice after training the network.

The emotional facial expressions of volunteers are used as teaching signals. For each person there are six kinds of emotional facial expressions – “happiness”, “sadness”, “disgust”, “anger”, “fear”, “surprise” – as shown in Figure 11.10. A facial expression model with emotion is constructed by the neural networks so that an emotion space appears in the third layer of the neural network. Figure 11.11 shows an example of the application of our model to e-mail.

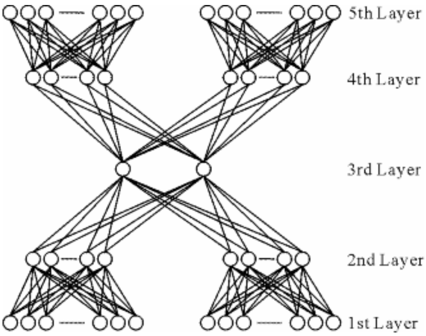


Figure 11.9 An overview of a sand glass neural network

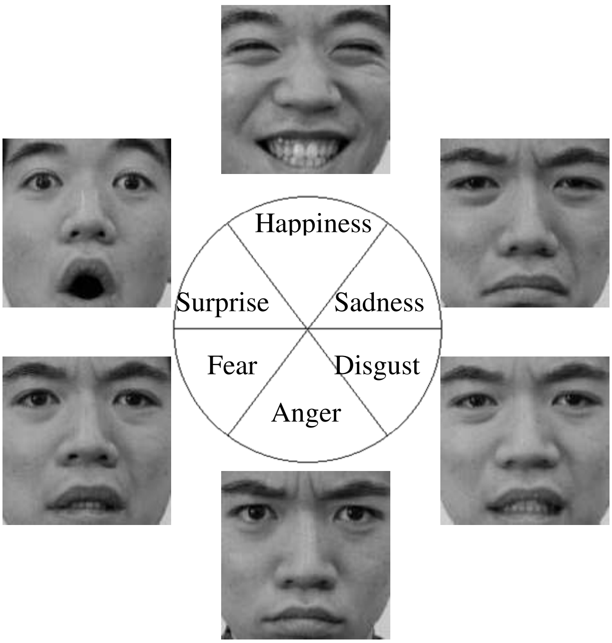


Figure 11.10 Emotion circle of a user’s face

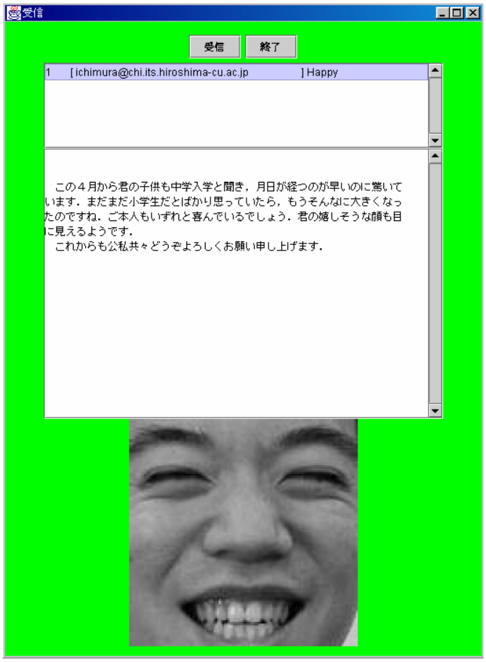


Figure 11.11 Example of an e-mail with a user's face

11.2.5 Noh Mask

Noh masks used in Noh play, which is one of the most popular traditional arts in Japan, are artificial and sometimes ambiguous, and can be interpreted as expressing various emotions [14, 15, 16]. That is, as the Noh mask turns upward and downward or to the left and right when the actor with the Noh mask moves on the stage, the audience can read diverse emotions from the Noh mask from the changing of the angle and the combination of light and shadow. It is usually said that an upward turn of the Noh mask, which is called *terasu* (shining), represents a pleasant and cheerful state of mind, and a downward turn of the Noh mask, which is called *kumorasu* (clouding), represents a gloomy state of mind. The audience enjoys reading delicate emotions from the Noh mask.

People can read various scenarios for Noh play from the Internet without going to Noh theaters, but they cannot enjoy reading the delicate emotions from the Noh mask. We applied the fuzzy reasoning model to selecting the image of the Noh mask which best expresses the emotions caused by some scenario in the Noh play [10]. Moreover, we constructed a computer display system which can allow users to read the scenario of the Noh play and enjoy the delicate emotions that the Noh mask expresses. We used *Koomote* as the Noh mask, because *Koomote* is one of the most popular female masks and is known to have the richest variety of facial expressions. Figure 11.12 shows typical *Koomote* images corresponding to Ekman and Friesen's six basic emotions of "happiness", "sadness", "anger", "disgust", "surprise", and "fear" [17].

We illustrate our system using *Hagoromo* as a Noh play. The story of *Hagoromo* is as follows. A fisherman finds an angel's cloak on a branch of a pine tree on the beach. He steals it and so prevents her from returning to heaven. The angel is very sad because she cannot go back to the sky. After her entreaties, the fisherman decides to give the cloak back to the angel. The angel is very glad and dances to show the glory of heaven.

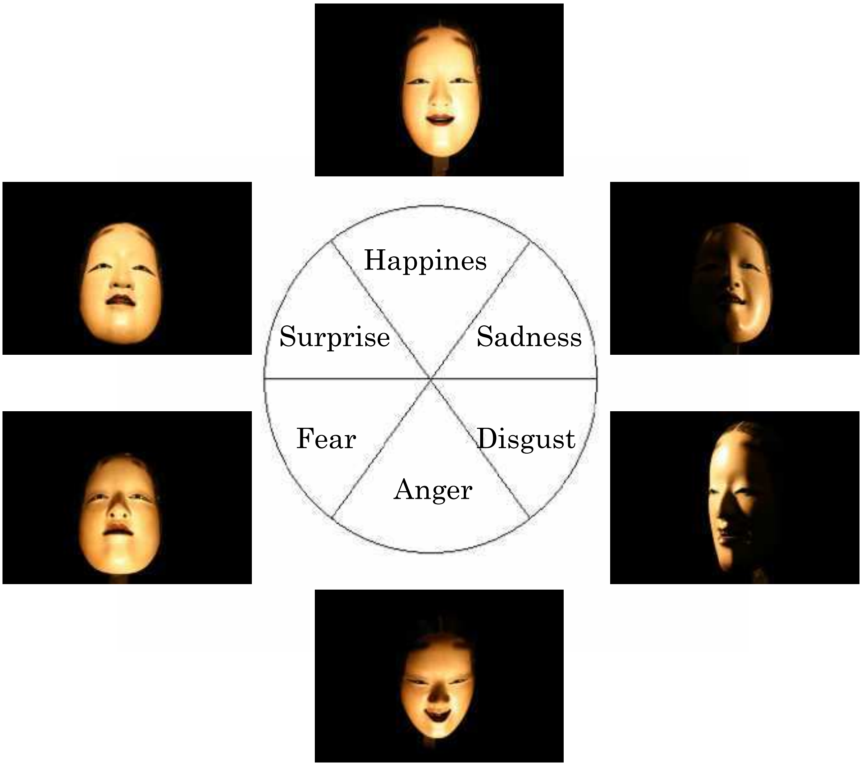


Figure 11.12 Emotion circle for the Noh mask

The examples of the sentences in the scenario [18] and the Noh masks selected by the system are shown in Figures 11.13–11.15. Figure 11.13 shows the scene where the angel is surprised to see a fisherman steal her cloak. Figure 11.14 shows the scene where the angel grieves and entreats the fisherman to give her cloak back. Figure 11.15 shows the scene where the angel is very glad to know that she will be able to go back to the heaven and she dances to show the glory of the heaven.

Angel
Stop!
That cloak is mine.
Why are you going with it?

Figure 11.13 Scene from *Hagoromo* (1)

Angel
Oh pitiful!
How shall I cloakless tread
the wing-ways of the air, how climb
the sky, my home?
Oh, give it back, in charity give it back.

Figure 11.14 Scene from *Hagoromo* (2)

Angel
I am happy, happy.
Now I shall have wings and mount the sky again.
And for thanksgiving I bequeath
a dance for remembrance to the world,
fit for the princes of men:
The dance tune that makes to turn
the towers of the moon,
I will dance it here and as an heirloom leave
it to the sorrowful men of the world.

Figure 11.15 Scene from *Hagoromo* (3)

11.3 Verbal Communication

In verbal communication, we can also make HCI pleasant. In this section, we explain our system which inserts "repeating the user's utterances" and "chiming in with the user" into the HCI [19].

When two persons converse joyfully, an unpleasant silence will fall upon them if one person does not reply to the other. However, chiming in with the speaker's remarks, or nodding, or repeating his/her utterances can encourage him/her to continue speaking, and can give pleasure to the communication, and can smooth the conversation.

This is also true of HCI [20]. If the computer cannot understand the user's responses and does not reply, the user will be discouraged and unwilling to continue the interaction. However, if the computer chimes in, or nods, or repeats his/her utterance, the user will be encouraged to continue the interaction. For these reasons, we have constructed human interfaces that use chiming in and repeating.

Figure 11.16 shows an example of the conversation on a "web-based system of health service" as follows:

System: Can you fill in the forms for your annuity by yourself?
User: I cannot always say that I cannot write it by myself.
System: Cannot you always say that you cannot write it by yourself?
User: I can fill it in a little.
System: Oh, really!

The user's first response to the question from the system is a double negative, and the system cannot understand whether the reply is affirmative or negative. However, in order not to stop the conversation and to encourage the user to continue, the system repeats the user's utterance in an interrogative sentence. Then, from the user's next reply "I can fill it in a little", the system can understand that the user can do it. At once, the system chimes in with him to show the hearer's interest.

If the system again repeats the user's utterance instead of chiming in, the user may go into details. In face-to-face communication, a double negative such as "I cannot always say that I cannot write it by myself" sometimes expresses a desire to say that the accident which prevented him/her from writing happened recently. For example, the system might be able to get the following reply:

System: Can you fill it in a little?
User: I hurt my right hand a few days ago, so I have difficulty in writing.

As shown in this example, repeating and chiming in to show the hearer's interest can make HCI pleasant.

Figure 11.16 Interface for a “web-based system of health service”

11.4 Conclusions

We have tried to construct an emotion-oriented human interface. In this chapter, we have described several examples of our systems which enable human-like communication.

Both verbal and nonverbal messages are needed for smoothing human face-to-face communication. This is also true of HCI. In order to make HCI pleasant, we have adopted facial expressions as nonverbal messages and proposed fuzzy reasoning and neural network models for selecting the facial expression which expressed the emotions caused by several situations. Application of our models to e-mail shows that facial expressions can add impression of “pleasantness and familiarity” to the information on the computer display. Moreover, we applied our model to selecting and displaying a Noh mask in a Noh play scenario.

Chiming in with another’s remarks, repeating their utterances, and so on, have the same functions as nonverbal messages. That is, chiming in and repeating can encourage a speaker to continue his/her speaking, give pleasure to the communication, and smooth the conversation. Since this is true of HCI, we have constructed human interfaces which chimed in with the user and repeated the his/her utterances.

However, many sciences such as physiology, psychology, medical science, philosophy have been struggling to understand human emotions. Our studies aim to achieve smooth and pleasant communications between humans and computers, but for further development of emotion-oriented HCI, a deeper understanding of human emotions is needed.

References

1. Ellison J.W., Massaro DW.: Featural evaluation, integration, and judgment of facial affect. J. Exp. Psychol. Human Perception Perform. **23**(1):213–226(1977)
2. Iwata M., Onisawa T.: Route decision system associating with facial expression. J. Jpn. Soc. Fuzzy Theor. Syst. **8**(3):532–540 (in Japanese) (1996)
3. Onisawa T., Date D.: Personified game playing support system. Jpn. J. Ergonomics **35**(3):157–167 (in Japanese) (1999)
4. Yamashita T. *et al.*: A study on characteristics of faces as a human interface. J. Jpn. Soc. Fuzzy Theor. Syst. **14**(2):248–253 (in Japanese) (2002)
5. Yamashita T.: An application of human mind to engineering. J. Soc. Biomech. **28**(2):72–78 (in Japanese) (2004)
6. Yamashita T., Koseki H.: An application of fuzzy reasoning to facial selection. Jpn. J. Edu. Technol. **22**(3):193–199 (in Japanese) (1998)
7. Yamashita T. *et al.*: An application of facial selection model by fuzzy reasoning to human interface. J. Jpn. Soc. Fuzzy Theor. Syst. **12**(2):313–320 (in Japanese) (2000)
8. Yamashita T., Sakai H.: Face selection model by fuzzy reasoning and its application to computer game. Jpn. J. Ergonomics **36**(3):139–145 (in Japanese) (2000)
9. Ichimura T. *et al.*: Approach to emotion oriented intelligent system by parallel sand glass type neural networks and emotion generating calculations. J. Human Interface Soc. **3**(4):225–238 (in Japanese) (2001)
10. Ichimura T. et al.: Emotion orientated intelligent systems. In: Howlett R.J. *et al.* (eds.) Internet-based Intelligent Information Processing Systems, pp. 183–226. World Scientific Publishing Company, Singapore (2003)
11. Plutchik R.: The multifactor-analytic theory of emotion. J. Psychol. **50**:153–171 (1960)
12. Matsumoto Y. *et al.*: Japanese Morphological Analysis System Chasen, Version 2.0. http://chasen-legacy.sourceforge.jp (in Japanese) (1999)
13. Yamashita T. *et al.*: Facial expressions as a human interface for e-mail. Mem. Tokyo Metropolitan Inst. Technol. **14**:7–11 (2000)
14. Osaka N.: Cross-cultural differences in the perception of facial expressions of ambiguous Noh faces. Bull. Psychonomic Soc. **24**(6):427–430 (1986)
15. Yamashita T. *et al.*: Recognition of affects in the facial expressions of Noh mask by fuzzy reasoning. Jpn. J. Ergonomics **35**(3):193–199 (in Japanese) (1999)
16. Yamashita T., Minoshita S.: Categorical perception in facial expressions on Noh mask “*koomote.*” J. Jpn. Acad. Facial Stud. **5**(1): 49–57 (in Japanese) (2005)
17. Ekman P., Friesen V.W.: Unmasking the Face. Prentice-Hall, New Jersey (1975)
18. Waley A.: Hagoromo. http://etext.virginia.edu/etcbin/toccer-new2?id=WalHago.sgm&images=images/modeng&data=/texts/english/modeng/parsed&tag=public&part=1&division=div1. The University of Virginia Library Electronic Text Center (1997)
19. Sato Y. *et al.*: Proposal of automatic responding system for chat-system detecting affirmative/negative. In: Proceedings of 17th Fuzzy System Symposium, pp. 509–512 (2001)
20. Yamashita T.: Communication in terms of psychology: human computer interaction. Sci. Mach. **59**(1):198–203 (2007)

Chapter 12
BROAFERENCE: A Prototype of an Emotion-based TV Quality Rating System

Terumasa Aoki[1] and Uwe Kowalik[2]

Abstract Whether a television (TV) program is good or not is generally judged by its TV rating, which indicates the audience size. However, it is obvious that a TV rating does not necessarily guarantee the quality of the program. In this chapter, we describe the architecture and experimental results of BROAFERENCE, which is an emotion-based TV quality rating system that makes it possible to judge the quality of a TV program by measuring emotional information from the TV audience. BROAFERENCE utilizes automatic detection of facial expressions and gaze information via a camera observing the TV program's audience in order to collect emotional cues about the impact of a certain program as well as measuring the attention of viewers.

12.1 Current TV Rating System and Its Problems

It is said that television (TV) is the king of the mass media. Actually we can get a lot of information from TV every day.

In general, whether a TV program is good or not is judged by its TV rating, which indicates the audience size. However, as the influence of TV ratings are becoming stronger, the harmful effects of TV ratings are also increasing. One of the most harmful side-effects is that the quality of TV programs is now measured in terms of the viewing-rate instead of in terms of their quality. As a result, many TV programs are becoming vulgar and less meaningful in order to achieve higher ratings. For example, TV programs are usually required to be easy to understand

[1] Terumasa Aoki
Tohoku University, 2-1-1, Katahira, Aoba-ku, Sendai-shi, Miyagi, 9808577, Japan
e-mail: aoki@riec.tohoku.ac.jp
[2] Uwe Kowalik
DC3 Inc., 11-34, Kinomi-cho, Yamagata-shi, Yamagata, 9900044, Japan
e-mail: uwe@dc3.co.jp

for a new audience who tuned into the program a few seconds ago, in order to prevent them from switching channels. Obviously this requirement is correct in terms of increasing the viewing-rate, but it is not always correct in terms of TV program quality. This example shows that TV programs are finding it difficult to cope with both being meaningful and achieving a high TV rating.

Furthermore, there are the following criticisms about the current TV rating system, apart from TV program quality:

1. *Insufficient number of samples*: The number of samples used to provide a TV rating is too small. For example, in Japan, only 600 samples are used in three big cities (Tokyo, Osaka, and Nagoya) and their suburbs, and only 200 in other cities. This means that TV rating includes a 3.3% error from a statistical point of view (*i. e.*, when the TV rating is 20%, the number of samples is 600 and the confidence interval is 95%). In other words, when the published TV rating is 20%, the actual TV rating could be between 16.7 and 23.3%. This error is huge considering that 1% on the TV rating corresponds to one million viewers and it has a value of tens of thousands of US dollars. One may think that increasing the number of samples is easy to do, but this is difficult for political and commercial reasons. Most TV-related companies (TV stations, TV program productions, advertising agencies, *etc.*) – an exception is TV sponsors –neither expect accurate TV ratings nor want accurate values to be publicly available. In fact, it is said that they strongly oppose the disclosure of this information.
2. *Difference between TV "program" rating and TV "advertisement" rating*: As is well-known, TV content generally consists of TV programs and advertisements. Therefore we can define two types of TV ratings: TV program rating and TV advertisement rating. In many countries only the TV program rating is used because the audience is not interested in the TV advertisement rating.
   Indeed the TV advertisement rating is totally different from the TV program rating. In general, the TV advertisement rating is said to be 60 or 70% of the TV program rating. And the ratio of the TV advertisement rating to the TV program rating varies according to the type of TV program. This kind of information is very important for sponsors but is never disclosed in many countries.
3. *Other problems*: Besides the points mentioned above, there exist other problems:
   - TV rating systems in many countries cannot distinguish differences between watching by audience and recording by DVD or Blu-ray recorder
   - In many countries only one company measures the TV ratings, *i. e.*, the objectiveness of the acquired data is not always guaranteed.

In light of this and as a possible extension to existing TV rating systems we have developed an emotion-based TV quality rating system.

The emotion-based TV quality rating system utilizes the automatic detection of facial expressions and gaze information via a camera observing the TV audience while watching the program in order to collect emotional cues about the impact of a particular program as well as measuring the attention of viewers. These data can be sent to a TV rating company or NPO (non-profit organization), providing not only information about the audience size (the current TV rating) but also allowing the measurement of the TV program quality (Figure 12.1).

In this chapter, we will describe the details of this system and its prototyping system, called BROAFERENCE.

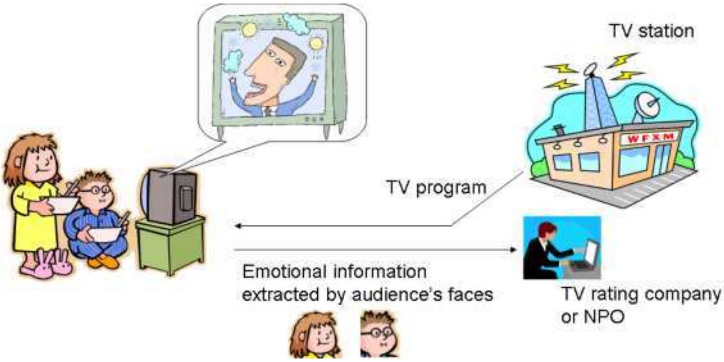


Figure 12.1 BROAFERENCE system

12.2 System Overview: the BROAFERENCE System

BROAFERENCE was developed as a prototype of an emotion-based TV quality rating system allowing for the recording of emotional and interest data acquired from users (TV audience) during TV programs. Figure 12.2 shows the BROAFERENCE terminal. Here, we assume this terminal is connected to the Internet.

BROAFERENCE consists of a display, a receiver unit connected to an IP network, a camera for recording the user in front of the terminal, and two units for measuring the emotional and interest data using vision-based face analysis techniques for non-intrusive measurement (Figure 12.3). Recorded feedback data are synchronized with the media time of the presented TV program and stored on the terminals' hard disk.

Figure 12.2 BROAFERENCE terminal for non-intrusive TV program quality measurement

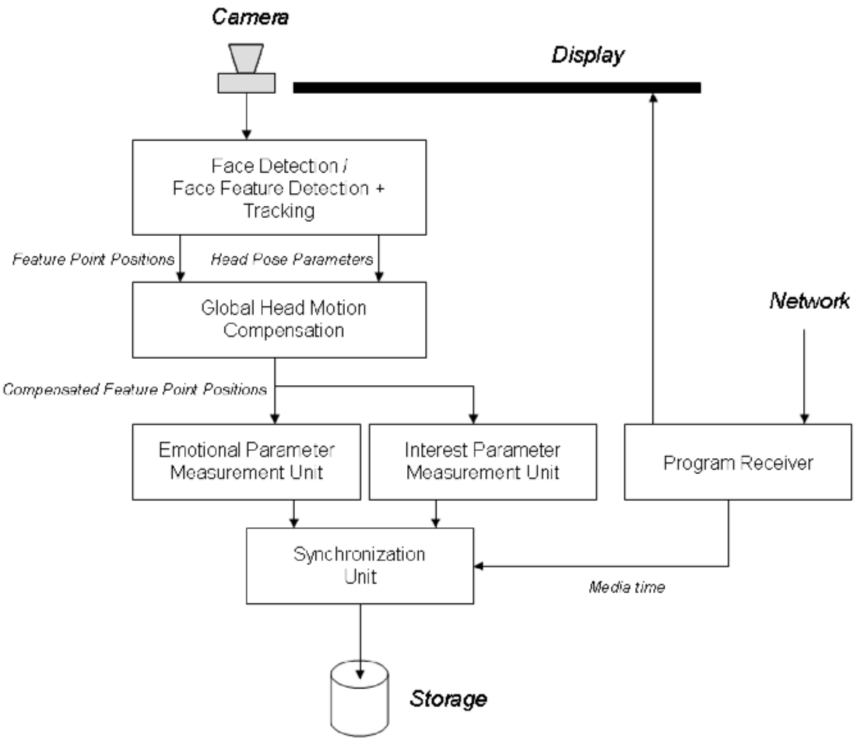


Figure 12.3 System overview

12.3 Face Detection, Face Feature Point Detection, and Tracking

12.3.1 Abstract of Face Detection, Face Feature Point Detection, and Tracking

The BROAFERENCE system acquires user feedback non-intrusively by employing face detection and face parameter analysis. A commercial Face Feature SDK (FFT-SDK) is used for feature detection and tracking. The FFT-SDK provides 22 face feature point positions and additional head pose parameters for each detected face in an image. The layout of the face feature points is shown in Figure 12.4. The FFT-SDK is fully realized in software and has stable performance under changing lighting conditions.

Face features are tracked at full frame rate, *i. e.*, 30 fps in video sequences of size 640×480. Table 12.1 lists the provided head pose parameters.

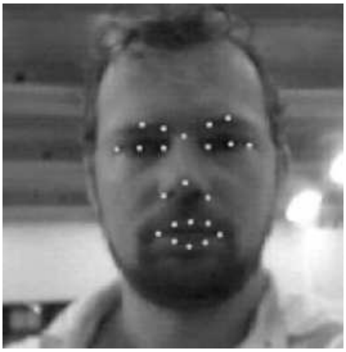


Figure 12.4 FFT-SDK provides 22 feature positions

Table 12.1 Head pose parameters provided by the FFT-SDK

Parameter	Remark
Scale	Apparent face size
Center of gravity	x, y coordinates [pixels]
Rotation around x-axis	[rad] max. 25 degree
Rotation around y-axis	[rad] max 25 degree
Rotation around i-axis	[rad] max 15 degree

12.3.2 Emotional Parameters

The BROAFERENCE system allows the real-time recording of the observed emotions of a TV audience via a description scheme for facial expressions. Facial expressions, spontaneously or intentionally, deliver signals of the current emotional state of a person. Using cross-cultural research, the scientist Paul Ekman devised a list of six basic emotions related to facial expressions that are assumed to be universal to humans [1]. The six basic emotions proposed by Ekman and widely accepted by behavioral science researchers are Anger, Disgust, Fear, Happiness, Sadness, and Surprise. Although it has been found that these basic emotions can be identified by human observers based on the displayed facial expression independent of gender and cultural background, they rarely occur alone but rather as blends of emotions depending, *e. g.*, on individual differences, social context, and the level of self-control. Therefore a one-to-one mapping of facial expressions to emotions is not feasible and the interpretation of facial expressions in terms of emotions is still a subject of psychological research. Ekman and colleagues Wally Friesen and Joseph Hager realized the necessity of separating the appearance of facial expressions from the psychological interpretation and developed the Facial Action Coding System (FACS) in the 1970s [2]. FACS is a systematic description scheme for facial expressions in terms of visible cues and provides a tool and a common syntax for studying the psychological interpretation of facial expressions in a unified framework. FACS defines a set of so-called facial action units (AU) based on visible structural changes on the surface of the face skin caused by contraction and relaxation of the underlying muscles. Any possible facial expression can be described by a combination of multiple AUs. The FACS system defines 46 atomic action units. Each AU may occur with 1 out of 5 different intensity levels, which leads to a large number of theoretically possible combinations. However, not all of those combinations are practically relevant. So far, around 7000 valid combinations have been observed [3]. Table 12.2 shows the AU codes and their description.

The BROAFERENCE system utilizes FACS for recording facial expression occurrence synchronized with the media time of the presented TV program. This allows the emotional experience of the audience throughout certain content to be analyzed by an expert or expert system and thus gain information about the individual impact of the TV program. The current implementation records only a restricted number of action units, namely AU2 (outer brow raiser) and AU12 (lip corner puller). This choice is motivated by the fact that AU12 activity is usually observed when people smile or laugh and is therefore a good indicator for the presence of the emotion *Joy* (or *Happiness*). The emotion *Surprise* is often expressed by pulling the eyebrows upwards, thus AU2 activity provides a cue for the presence of *Surprise*. BROAFERENCE's action unit detection is based on the feature point motion and does not use texture information.

Table 12.2 Description of AU codes defined by the FACS system

AU code	Description	AU code	Description
1	Inner brow raiser	23	Lip tightener
2	Outer brow raiser	24	Lip presser
4	Brow lowerer	25	Lips part
5	Upper lid raiser	26	Jaw drop
6	Cheek raiser	27	Mouth stretch
7	Lid tightener	28	Lip suck
9	Nose wrinkler	29	Jaw thrust
10	Upper lip raiser	30	Jaw sideways
11	Nasolabial deepener	31	Jaw clencher
12	Lip corner puller	35	Suck
13	Cheek puffer	36	Bulge
14	Dimpler	37	Lip wipe
15	Lip corner depressor	38	Nostril dilator
16	Lower lip depressor	39	Nostril compressor
17	Chin raiser	41	Lid droop
18	Lip puckerer	42	Eye to slit
19	Tongue show	43	Eyes closed
20	Lip stretcher	44	Squint
21	Neck tightener	45	Blink
22	Lip funneler	46	Wink

12.3.3 Interest Parameters

In order to estimate the impact of a particular TV program it is important to gain information about the audience's interest. The BROAFERENCE system employs a gaze tracking technique for this purpose. Gaze tracking allows for retrieving detailed information about the user's visual focus on the screen. This naturally also implies information about whether a user is looking at the screen at all. In the same way as for the emotion-related parameters, the gaze information is also recorded synchronously with the present TV program and thus allows for, *e. g.*, detecting parts of the program that were less interesting to the audience (indicated by a loss of screen focus). The gaze information is acquired estimating a "region of interest" (ROI) from the user's eye movement. In comparison to high-end gaze tracking solutions, such as, *e. g.*, [4], a low-cost gaze tracking method was developed for the BROAFERENCE system, which, despite lower accuracy, benefits from the usability of common consumer camera hardware which is available at reasonable costs.

12.4 Global Head Motion Compensation

12.4.1 How to Compensate Global Head Motion

Since the audience in front of a TV may move freely under practical conditions, global head motion compensation is a crucial preprocessing step in order to achieve stable AU detection and gaze tracking results when using only motion information. More specifically, in the context of motion-based facial expressions, detection to distinguish between feature displacements caused by rigid head movement from those caused by actual expression changes is an essential requirement before performing the classification step. The FFT-SDK provides the global head motion parameters with respect to the camera origin. The global translation of the head, the change in size of the shape defined by the facial feature points and the in-plane rotation around the z-axis can be compensated relatively easily by applying an inverse transform based on the provided parameters. However, the feature displacements caused by different depths in the facial structure cannot be compensated for without knowledge of the 3D structure of the face. Since the 3D-head rotation parameters are known, a simple solution for this problem is to recover the depth information from point correspondences between two images. Instead of estimating the actual depth for each feature point, the goal here is estimating an approximate value Z for each face feature point which allows compensation of global out-of-plane head rotation under the assumption of a rigid face model. The Z-value estimation is done once in a self-calibration step before using the system. Figure 12.5 illustrates the influence of a depth value $Z > 0$ on the feature point displacement in the x-direction on the image plane.

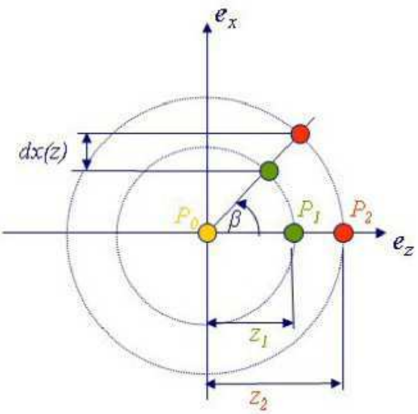


Figure 12.5 Effect of y-rotation on the x-displacement for feature points with different depth values of Z

Assuming only out-of-plane rotations, a 3D-point transformation is given as follows:
Rotation around x-axis with angle α:

$$x' = x,$$

$$y' = y \cdot \cos\alpha - z \cdot \sin\alpha,$$

$$z' = y \cdot \sin\alpha + z \cdot \cos\alpha.$$

Rotation around y-axis with angle β:

$$x' = x \cdot \cos\beta + z \cdot \sin\beta,$$

$$y' = y,$$

$$z' = -x \cdot \sin\beta + z \cdot \cos\beta.$$

Given a rigid reference shape of feature points S_0 captured once for $\alpha, \beta = 0$ in a reference frame I_0 and assuming a z-value of $z = 0$ for each feature point (planar face model), than a compensation value Z can be calculated for each feature point P_k in a shape S_n captured in a frame I_n by referring to its corresponding point in S_0. The expected feature point displacement if $z = 0$ would be

$$dx^k(\beta) = x_0^k - x^k \cdot \cos\beta,$$

$$dy^k(\alpha) = y_0^k - y^k \cdot \cos\alpha.$$

If the actual depth is $z \neq 0$, a displacement error e occurs with $e_x^k = z \cdot \sin\beta$ and $e_y^k = -z \cdot \sin\alpha$. Based on this observation a compensation term $\boldsymbol{Z}$ can be computed with $Z_x^k = e_x^k / (\sin\beta)$ and $Z_y^k = -e_y^k / (\sin\alpha)$ that corrects the displacement error.

The estimation of the compensation term Z is performed once at the initialization phase of the BROAFERENCE system. Figure 12.6 shows an example of the feature point model with estimated depth compensation.

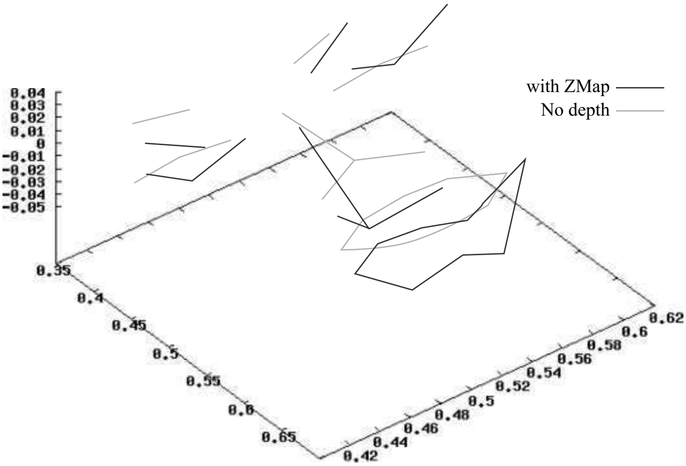


Figure 12.6 Example of a face feature point model with estimated depth compensation

12.4.2 Evaluation

In order to evaluate the influence of the depth compensation an experiment was conducted. First a test video was recorded, showing slow head rotation around the y-axis. Figure 12.7 shows three screen shots showing center, left, and right poses.

Figure 12.7 Screen shots from the head rotation test video

The reference shape was automatically captured from the frontal view. The angle threshold for triggering the Z-value estimation was set to 0.3333 radian deviation from the frontal view.

The feature points were compensated for immediately using the current estimated Z-value and the residual displacement error was recorded. Figure 12.8 shows the feature point displacement in the x-direction during continuous head rotations around the y-axis for the nose tip feature point. "No comp" refers to the case where only the scale and position of the feature points were compensated. "Comp" is the graph for the case of a planar face model and "zmap comp" refers to the global head motion compensation including the compensation with the estimated Z-value. The first negative peak refers to the trigger point of the Z-value computation.

Table 12.3 shows the average displacement error calculated for feature points with actual different depth (normalized feature coordinates). As can be seen, the compensation considering the estimated Z-value reduces the displacement error effectively. The result suggests that the Z-value base compensation performs well in the task of compensating displacement errors caused by out-of-plane head rotations on feature points with a depth $z \neq 0$, and therefore a better performance in the following feature classification can be expected.

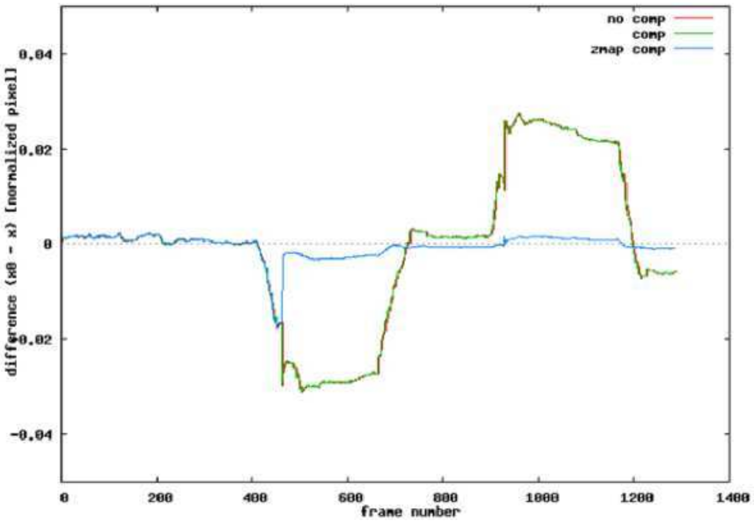


Figure 12.8 Displacement in the x-direction with and without Z-value compensation (nose tip)

Table 12.3 Average displacement error during OOP-rotation around the y-axis

	Average displacement error		
	No comp	Comp (planar model)	Comp (Z-value)
Left nostril	0.005 786	0.001 602 27	0.000 656
Right nostril	0.009 058	0.018 900 664	0.000 982
Nose tip	0.147 976	0.147 902 231	0.000 548

12.5 Action Unit Detection with Artificial Neural Networks

12.5.1 How to Detect Action Units

Action unit detection is based on the geometric properties of the face features only. More specifically the normalized locations of the compensated feature points are used. The BROAFERENCE system uses a set of two *feed-forward artificial neural networks* (ANN) for AU detection. Each ANN was trained separately on different subsets of the compensated face feature positions. The feature selection is necessary due to the fact that not all facial feature points are involved into a certain action unit activity, *e. g.*, AU12 should be not affected by features from the eyebrows and vice versa. Moreover action units may occur in combination depending on the present facial expression thus separate classifiers for each action unit is feasible. Table 12.4 shows the feature locations used for each AU-detector.

The nose root and nose tip are considered to be rigid face features since their relative location inside a face is not very much affected by facial muscle movement. Thus these two features serve as anchor points for estimating the relative movement of left/right eyebrow centers and left/right mouth corners respectively. Figure 12.9 illustrates the simplified model assumptions for both action units.

Table 12.4 Face features used for AU-classifier training

AU2	AU12
LEFT_EYEBROW_CENTER	NOSE_TIP
RIGHT_EYEBROW_CENTER	LEFT_MOUTH_CORNER
NOSE_ROOT	RIGHT_MOUTH_CORNER

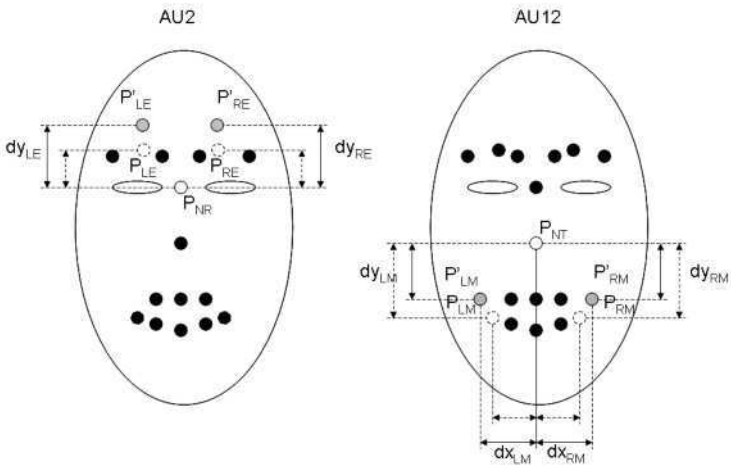


Figure 12.9 Feature point displacement for AU2 (left) and AU12 (right)

In Figure 12.9 P_{NR} denotes the nose root and P_{LE} and P_{RE} refer to the neutral positions of left and right eyebrow center respectively. The model assumes that AU2 is present, if the distances in the y-direction between nose root and eyebrow center points dy_{LE} and dy_{RE} increases by a certain offset when P_{LE} and P_{RE} move to the new locations P'_{RE} and P'_{LE}. This is the case, *e. g.*, when a person shows a surprised facial expression by raising the eyebrows. A similar model assumption is made for the presence of AU12 depicted on the right-hand side in Figure 12.9. Here P_{NT} refers to the nose tip feature position and P_{LM} and P_{RM} denote left and right mouth corner locations respectively. When a person smiles or laughs the mouth corners will usually move up and outwards in a circular trajectory. This leads to a variation of the relative distances in the x- and y-directions with respect to the nose tip, which serves as the anchor point.

Due to individual differences in facial activity and to noise introduced by the camera, as well as due to the limited precision of the feature point detection, it is usually difficult to determine a fixed threshold for each feature point variation. Therefore, the BROAFERENCE system employs a supervised machine learning approach for action unit classification by neural network classifiers, as mentioned earlier.

The input layer of each ANN consists of six neurons accordingly to the input vector size derived from the three facial feature point locations, *i. e.*, the x- and y-components of the features are used. The output layer of each ANN consists of one neuron referring to the classification objective whether the action unit is present or not. As for the number of hidden layers, two configurations were first considered – with one and two hidden layers respectively. The decision as to which architecture should be used was based on the comparison of training errors. The configuration with two hidden layers achieved better adaption to the training set

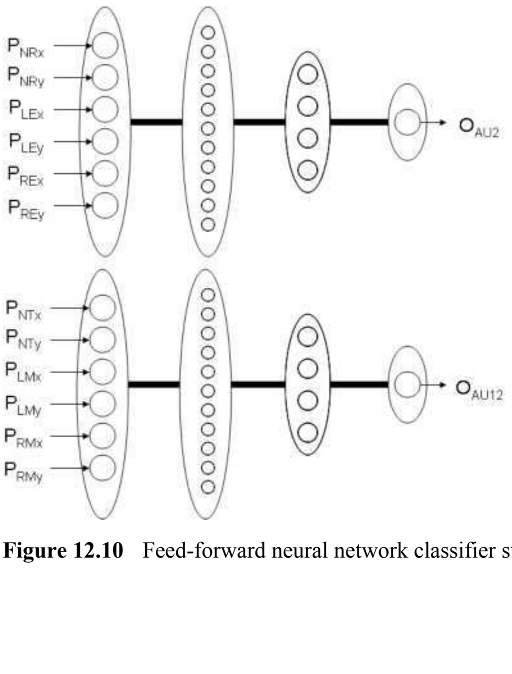


Figure 12.10 Feed-forward neural network classifier structure

and also showed a higher performance on the test data, and thus was used for the system implementation. The final ANN configuration consisting of two hidden layers with 11 and 4 neurons respectively is shown in Figure 12.10. All layers are fully connected to the following layer. The neurons use a Logistic transfer function.

12.5.2 Classifier Training

For creating the training data and evaluation of the classification result, video data was created containing imagery of ten people with different ethnic backgrounds (80% Asian, 10% Latin, and 10% Caucasian) aged between 22 and 39 years. Thirty percent of the subjects were female. The Imagery consists of short video clips recorded with a DV consumer camera with a resolution of 640×480 pixels and a frame rate of 29.97 fps. Some examples are shown in Figure 12.11.

The subjects were recorded while they posed AU2 and AU12 under various head rotation angles. The data for training the classifiers have been extracted from frontal view sequence samples showing the neutral and maximum activity of AU2 and AU12 respectively, including a preceding out-of-plane rotation compensation step. Each training sequence consists of three main parts separated by blue screen frames serving as markers for the automatic distinction between positive and negative image examples. Figure 12.12 shows an example of a training sequence.

Figure 12.11 Video database examples for AU2 and AU12

Figure 12.12 Training sequence containing head rotation, neutral, and maximum AU activity training samples (AU12)

The first part of each training sequence contains a head rotation sequence with the neutral facial expression and slow head rotation around the y-axis necessary to trigger the automatic depth map generation. The second and third part of a training sequence provides neutral and maximum action unit activity samples from an almost frontal view. Although it is not necessary for frontal view imagery, the extracted feature point geometries were compensated with prior training data generation using the estimated Z-compensation term in order to simulate the complete system function.

12.5.3 Evaluation

The goal of the evaluation was to verify how the AU classifiers trained on frontal view examples will perform on sequences showing out-of-plane head rotation. Test sequences have been created that contain the initial head rotation part as described above, followed by a sequence showing minimum and maximum intensity for action units in random order and with varying out-of-plane head rotation. The classifier evaluation was conducted twice with different adaptation angle threshold TH_{depth}. The *true positive rate* (TPR, also referred to as *recall*) was calculated as an evaluation measure. Table 12.5 shows the result for the detection of AU12 with a varying threshold TH_{ann} applied to the ANN classifier output.

A higher TPR is achieved over a wider range when the depth map initialization is performed at higher values of out-of-plane rotations. This suggests that the estimated Z-compensation term is more precise when calculated at higher rotation angles and thus the model fits better to the actual face geometry.

Table 12.5 TPR of the AU12 classifier with varying thresholds

TH_{ann}	$TH_{\text{depth}} = 0.2$	$TH_{\text{depth}} = 0.3$
0	1	1
0.1	0.9	1
0.2	0.9	0.9
0.3	0.9	0.887 324
0.4	0.863 013 7	0.828 947
0.5	0.759 036 14	0.828 947
0.6	0.75	0.828 947
0.7	0.707 865 17	0.768 293
0.8	0.670 212 77	0.707 865
0.9	0.567 567 57	0.623 762
1	0.5	0.5

12.6 Gaze Tracking Approach

12.6.1 How to Estimate Gaze Information

Gaze estimation in the BROAFERENCE system is based on the eye center feature output of the face feature tracker, and incorporates an eye model which is used to calculate the eye rotation from the difference between eye center feature position and a predefined center default position. Figure 12.13 depicts the eye model used.

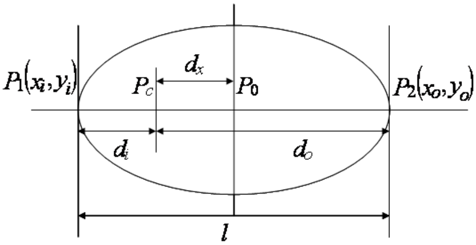


Figure 12.13 Simplified eye model

$P_1(x_i, y_i)$ refers to the inner eye corner and $P_2(x_o, y_o)$ is the outer eye corner position. P_0 denotes the center point of the eye model and $P_c(x_c, y_c)$ is given by the current eye center feature position. First the eye length l is computed by

$$l = |x_i - x_o| \, . \tag{12.1}$$

The deviation from the straight gaze direction towards the camera is given by d_x and can be calculated as

$$d_x = \frac{d_i - d_o}{l} \, . \tag{12.2}$$

If the eye ball is assumed to be an ideal sphere with radius r, then $l = \pi \cdot r$ defines the equatorial line of the visible hemisphere, and

$$d_c = \frac{\pi \cdot d_x}{l} \tag{12.3}$$

gives the deviation in radians.

12.6.2 Mapping to 3D World Coordinates

The gaze tracker uses a 3D head model placed at the origin of a left-handed world coordinate system (Figure 12.14). A virtual camera is placed at the point

$P_{\text{cam}}(0,0,d)$ in the screen center, looking down the negative *z*-axis of the world's coordinate system. The parameter *d* defines the distance to the face model and is calculated by

$$d = f/\text{scale}, \tag{12.4}$$

where *f* is the focal length of the virtual camera and "scale" is the scale factor provided by the feature tracker.

Since *f* is unknown, a calibration step is required. Currently this is done manually by the following steps:

1. Move the face towards the camera until scale = 1.0.
2. Adjust *f* until *d* is equal to the distance between face and camera.

Figure 12.15 shows the relationship between eye model and camera image plane. P_{ol} and P_{or} are the center points of the left and right eye model and P'_{ol}, P'_{or} are their projections in the image plane respectively. The gaze vectors are modeled by two cone geometries scaled to the distance *d*, whereas the projected cone tops refer to P'_{ol} and P'_{or}. The *x*-component of the final gaze position is calculated as

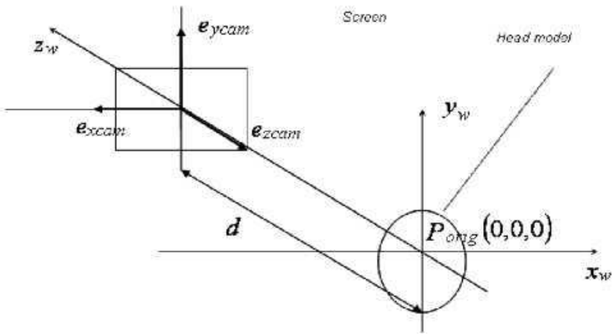


Figure 12.14 Relationship between virtual camera, screen, and 3D head model

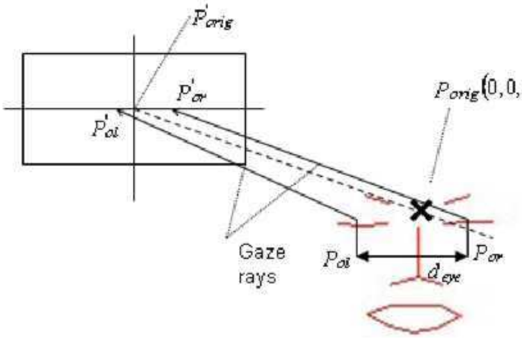


Figure 12.15 Relationship between eye model and camera image

the average $x_{gaze} = \frac{\left(x'_{or} - x'_{ol}\right)}{2}$. In the current implementation the *y*-component is currently not computed from eye movement but only the head rotation is taken into account.

In order to compensate head rotation, the head model builds a transformation tree of rotations (Figure 12.16). The top node (*HeadRot*) of the transformation tree determines the head rotation as delivered by the face tracker module. The two leaf nodes *LeftEyeRot* and *RightEyeRot* are keeping the rotational transform for the left and right eye respectively, calculated as described above.

The visualization of the gaze rays can be seen in screenshots provided in Figure 12.17.

In the left image of Figure 12.17 the cone bottoms rotate independently around the eye center located in the image plane, as determined by the calculated rotation angle. The final virtual intersection points between the gaze rays and the screen are visualized using two colored box geometries. Due to the mirror effect, the intersection points seem to be opposite to the head/eye rotation. However, the virtual intersection points coincide with the intersections of a user's gaze ray and the screen. The rectangular area on the right of Figure 12.17 refers to the region of interest (ROI) that the user is looking at. The screen area is tiled into nine ROIs organized into a 3 × 3 grid, as depicted in Figure 12.18.

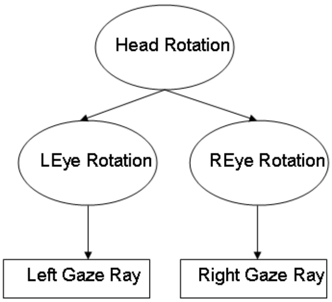


Figure 12.16 Transformation tree for connecting head and eye rotations

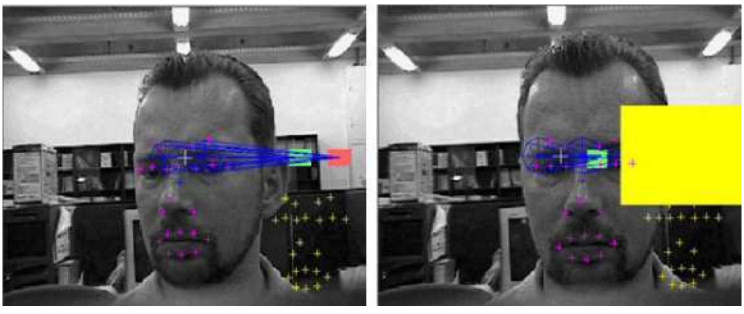


Figure 12.17 Visualization of gaze tracking

2	1	8
3	0	7
4	5	6

Figure 12.18 Nine regions of interest are defined on the screen

In the following a motivation for the number of 3 × 3 ROIs is given based on experimental performance evaluation of the realized gaze tracker. The question during the following experiment was how precise the gaze tracking in the x-direction would perform.

12.6.3 Evaluation

A static face model has been created for evaluating the influence of camera noise onto the estimated gaze vector (Figure 12.19). The face model was placed 1 m away from the camera and, after adjusting the estimated gaze vector pointing to the screen center, the deviation from the center point caused by camera noise was recorded over a period of around 16 s. Figure 12.20 shows the fluctuation of the estimated gaze point recorded from the static face model over time. The display size was 42 inch and distance between user and display approximately 1.50 m. The screen resolution was set to 1024 × 768 pixels for the experiment. A maximum gaze point deviation from the screen center of 153 pixels in the x-direction was measured.

This result suggests introducing a quantization of the screen area into regions rather than relying on the estimated gaze point directly. Taking into account the pixel resolution in the x-direction of 1024 pixels, a number of 1024/153 = 6.67 regions (*i. e.*, six regions) seems to be sufficient in order to cope with the improper estimation caused by camera noise. However, a second experiment was performed in order to evaluate the gaze estimation performance under reality conditions. This

Figure 12.19 Static face model for measuring the influence of camera noise

time five people were asked to focus on the screen center trying to not moving the eyes away from the fixation point. In Figure 12.21 the deviation of estimated gaze points over time is shown.

With the same screen resolution setting of 1024×768 pixels for the experiment, a maximum gaze point deviation from the screen center of 337 pixels in the *x*-direction was measured. The result suggests a reduction of detectable ROIs in the *x*-direction to $1024/337 = 3.04$ regions (*i. e.*, three regions).

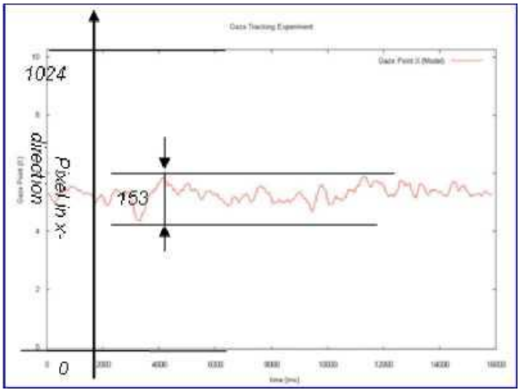


Figure 12.20 Fluctuation in the *x*-direction of estimated gaze point for the static face model

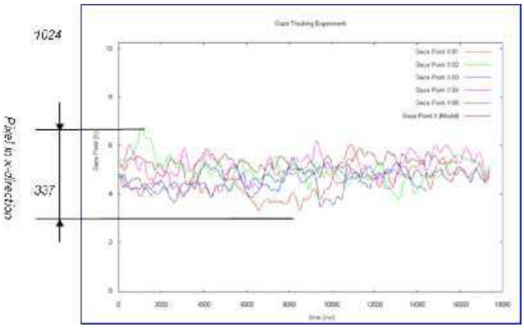


Figure 12.21 Fluctuation of estimated gaze point for five people

12.7 Other Applications

12.7.1 Content Indexing Based on the Emotional Component

Over the last decade the development of new technologies in the multimedia sector has advanced with stunning pace. Due to the availability of high-capacity mass storage devices at low-cost private multimedia libraries containing digital video and audio have recently gained popularity. Although attached meta-data like title, actor's name, and creation time eases the task of finding preferred content, it is still difficult to find a specific section within a previously viewed movie by remembering the time code. The huge number of movies that can be stored, *e. g.*, in a family's computer, requires a search engine in order to locate a certain movie. Current solutions provide indexing mechanisms based on content describing features, such as, *e. g.*, color schemes, shot length, *etc.*, to provide a means of finding a specific section inside a movie. In 0 news video archives are indexed with a lexicon of 100 semantic concepts, *e. g.*, sports, people, *etc.* Users may query by single concepts or combinations of them and access the resulting video at a semantic level. In 0 video documents are automatically indexed with 17 specific concept detectors for speech and noise analysis 0. In contrast BROAFERENCE can facilitate emotion data derived from facial expressions for the task of video indexing. Introducing an indexing approach based on the *emotional cues* could improve the usability of media search engines, since it will enable people to find a section inside a movie related to the specific emotional arousal they had when watching it before. On the other hand, it will be possible to exploit the *emotion-based* metadata for giving recommendations to buyers about which movies have been enjoyed by other people. This idea may have an impact on how digital content is advertised and distributed in the future.

12.7.2 Remote Health Care Prevention and Life Support

The birthrate of children in highly developed industrial countries is decreasing more and more. For example, in Japan it is expected that the ratio of elderly people to total population will reach 26% by 2015 and 35.7% by 2050. Due to the rising costs of care insurance programs there is a reorientation towards prevention in health care rather than treatment. Taking the demographic development into account a huge number of counselors will be needed to monitor and advise people for preventative health care 0, which will benefit from the introduction of support and monitoring systems that exploit broadband and video-communication technology. The remote health care prevention system proposed in 0 provides a computerized health checkup procedure, *i. e.*, a questionnaire and remotely guided, customized sports exercises delivered by means of broadband video to elderly people that are deemed to be at high risk for health conditions that require nursing care.

In the context of such a system the monitoring of individual emotional condition on a daily basis could be an extension to improve the effectiveness of remote counseling. In addition to the questionnaire-based assessment of health condition, remote counseling is proposed, which might benefit from an automatic facial expression classification as presented in this chapter. A counselor may access the facial expressions recorded remotely on the patient's terminal in order to gain information into the emotional state changes of a client during the day, provided the client is continuously monitored by video cameras installed in his/her home environment. An automated expert system can issue alerts if facial expressions indicate bad health conditions caused, *e. g.*, by pain. In particular, for blind people it might be useful to have feedback on the facial expressions of the people they are talking to. A certain facial expression may be related to a specific sound pattern, which also can convey emotions.

12.7.3 Advanced Human–Computer Interface

An interesting area of applications is in game development. The user may interact with the game's characters in a way that enables them to understand his/her emotional state during the game. Depending on the user's attitude the game flow may develop in different directions. Characters may, for instance, refuse to give information if the user does not behave in a friendly manner. Another option could be to control the level of difficulty based on the user's facial expressions. If the user is stressed due to a difficult situation in the game, the difficulty could be decreased, a help screen could be shown, or the game character changes its behavior (see Figure 12.22). This will lead to a higher satisfaction level at the user's side which might be an advantage on the market.

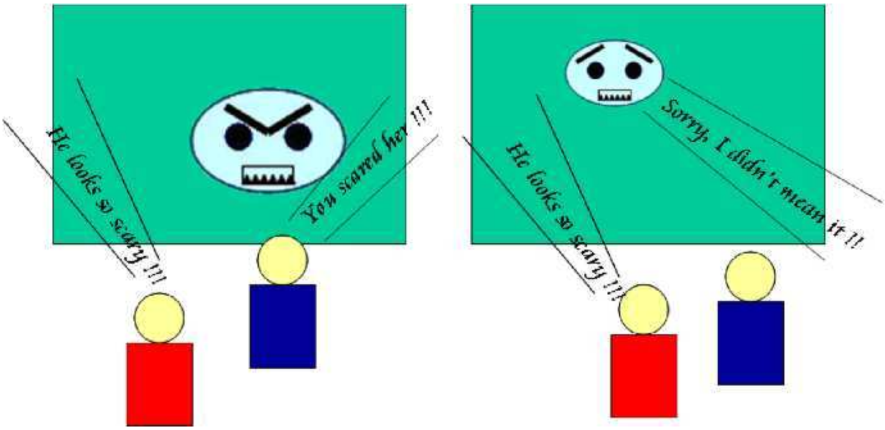


Figure 12.22 Extended user interface for computer games

12.8 Conclusion

The television plays an important role in entertainment, education, and information services. Since a TV can be found in almost every household, the television business has the largest number of customers among all media services. A huge amount of money is invested annually in programs, commercials, equipment, and people. Therefore marketing research institutions have been investigating for some time the watching behavior of the television audience in order to provide statistically proven data for increasing the effectiveness of program making. A common approach of doing this is to place a microcomputer controlled measuring device in a small number of sample households that are members of a carefully selected panel. Those panels statistically reflect the demographic structure of the urban area to be investigated. Examples are the *Telecontrol* system used in Germany, France, and the USA, and the *Radiocontrol* system tested in Norway and Japan in 2003/2004. The *Telecontrol* device, for example, records the consumption of media content by logging user ID, time, channel switch events, and device switch events, *etc.*, over 24 h a day. It also offers a program-rating function based on a flexible rating index scheme. In "interview" mode the device can be used for conducting written interviews where questions are displayed on the TV screen. However, these systems require a certain level of cooperation for user registration and rating which may influence the preciseness of the result, and moreover it will negatively affect the user's media experience. We believe that new systems for TV should not be active but passive (non-intrusive) because TV is a passive media by nature.

Against this background, BROAFERENCE provides a solution for non-intrusive emotion-based TV-quality rating. Automatic face detection and tracking paired with the automatic facial expression classification modules provide a solution for the problem mentioned above. Moreover, incorporating a gaze tracker can provide means to detect loss-of-focus during the program consumption, *i. e.*, when people are leaving the room or talking to each other while the TV is still running. The information provided by a gaze tracker therefore could help in monitoring how many customers have really watched a TV program and what part of the program has been more or less attractive to the audience. Automatic facial expression recognition technology will not replace current techniques, but may improve it, which finally will lead to better customer satisfaction.

References

1. Ekman, P., Friesen, W. V., Ellsworth, P.: Emotion in the Human Face: Guidelines for Research and an Integration of Findings. Pergamon Press, New York (1972)
2. Ekman P., Friesen, W.: Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, CA (1978)
3. Fasel, B., Monay, F., Gatica-Perez, D.: Latent semantic analysis of facial action codes for automatic facial expression recognition. In: Proceedings of the ACM International Conference on Multimedia, Workshop on Multimedia Information Retrieval (ACM MM-MIR), New York, October 2004
4. http://www.interactive-minds.de

5. Snoek, C., Worring, M., Van Gemert, J. *et al.*: MediaMill: exploring news video archives based on learned semantics. In: Proceedings of the 13th ACM International Conference on Multimedia, Singapore, November 2005
6. Worring, M., Hollink, L.,Van Gemert, J., Koelma, D., Nguyen, G.P.: Accessing video archives using interactive search. In: Proceedings of the IEEE International Conference on Multimedia and Expo, IEEE, Taiwan, June 2004
7. Hauptman, A., Baron, R.V., Chen, M.-Y.: Informedia at TRECVID 2003: Analyzing and Searching. Broadcast News Video (2003)
8. http://www.telecontrol.ch
9. Fujiwara, S.: Development and Promotion of System for Reducing the Need for Nursing Care. NTT Technical Review, **4**(2), February (2006)

Chapter 13
Shape Evaluation Properties in Real Space and Virtual Space

Masashi Okubo[1]

Abstract These days, not only designers but also general people have the chance to evaluate the design of products, such as cars, clothes, and electrical appliances, on computer displays or in virtual space. However, there is a high possibility of disagreement between the evaluation of shape in virtual space and that in real space. In this chapter, a 3D shape evaluation support system is introduced which integrates the visual information in virtual space and the tactile and gazing line-action information in real space. The proposed system can control information, for example, visual or tactile, and the linkage between motion and gazing line, which are used to evaluate the product shape, and investigate the role this information plays in evaluation. The preference for 3D shapes in the proposed virtual space is compared with that of real photoformed products made from the same data by the sensory evaluation of paired comparison and questionnaire analysis. It was found that the preference for shapes in both spaces was consistent with the relation of preferences based on the Bradley–Terry model for sensory evaluation. This indicates that the proposed system provides almost the same environment for shape evaluation as in real space. The results of questionnaires also indicate that the proposed system is enough to evaluate 3D shapes in virtual space. Moreover, the influence of differences in the point of view on the evaluation of shapes in virtual space is investigated using the proposed system by the sensory evaluation of paired comparison and questionnaire analysis. As a result, it is found that, for the evaluation of shapes in virtual space, subjects prefer a point of view from which they can see only their forearms' motion. The system enables the investigation of human's sense for shape evaluation by handling visual, tactile, and gazing line-action information.

[1] M. Okubo
Department of Information, Systems Design, Doshisha University
1-3 Miyakodani, Tatara, Kyotanabe, Kyoto, 610-0321, Japan
e-mail: mokubo@mail.doshisha.ac.jp

13.1 Introduction

We evaluate an object on a computer display in place of the real object, for example, in 3D CAD, Internet shopping, and so on. The validity of evaluating the 3D shape in virtual space through the display or a head mounted display (HMD) may be based on the assumption that the evaluation of the shape image in virtual space almost agrees with that of the real object in real space (Figure 13.1). However, it is indicated that there is a high possibility of disagreement between the evaluation of the shape in virtual space and that in real space [1].

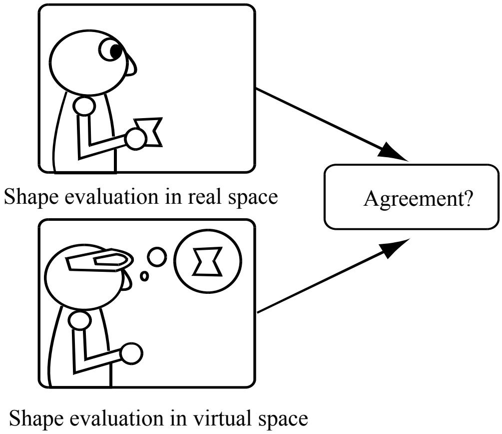


Figure 13.1 Shape evaluation in real and virtual space. The subject handles the two objects in his/her hands in real space and he/she is able to watch the objects' image in virtual space through the HMD. The objects' images that the subject watches are produced in virtual space on the PC on the basis of the angle and position measured by the magnetic sensors (Polhemus FASTRAK) attached to the top of the subject's head and to the objects

In order to clarify the difference between them, a shape evaluation support system has been proposed which integrates the visual information in virtual space and the tactile and gazing line-motion information in real space. The proposed system consists of a PC, an HMD, and magnetic sensors, and enables the handling of tactile and gazing line-action information in virtual space. The subjects obtain visual information about the models of the 3D objects from the HMD, and tactile information from real photoformed products that they hold in their hands. The gazing line-action information is produced on the PC using the output of the magnetic sensors attached to the subject's head and to the objects. The proposed system enables the handling of information on the PC, and analyzes how the subject evaluates the 3D shapes using this information. The system configuration is shown in Figure 13.2.

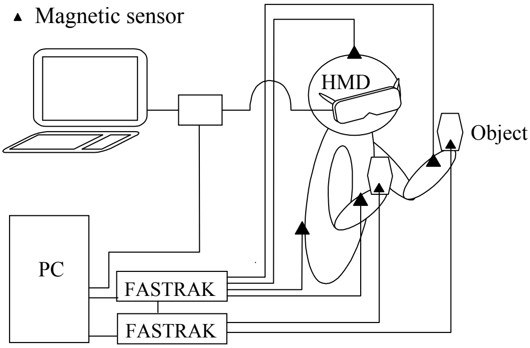


Figure 13.2 3D shape evaluation support system

13.2 System Evaluation

The system was evaluated to confirm the usability of the proposed system. To compare the evaluation of shape in real space and in virtual space, experiments were performed in both spaces for each subject.

13.2.1 3D Models and Objects Used in the Experiment

The 3D models on display were made using 3D CAD and the corresponding real objects were made using a photoforming system (NTT DATA CMET SOUP400GH-SP). To make the 3D models for the shape evaluation experiments, lip motion data was used, because the 3D structures made from lip motion aren't familiar to most people. These 3D models and objects have discrete shapes, so subjects aren't able to visualize an actual use of them. The lip motion images were captured and then sliced into each frame. The 3D models were made from a set of lip contour data on the 3D CAD. Figure 13.3 shows an example of a photoformed product made using this system. The product was made from the lip motion data produced when a person pronounced five concatenated Japanese vowels (/a/, /i/, /u/, /e/, /o/) in 2.3 s, and used approximately 70 frames.

The size of each object was 65 mm in mouth width, 50 mm in mouth height, and 70 mm (1 mm per second) in length. The product characterized the speaker's pronunciation where the speaker's under lip was distorted to the left when he pronounced the vowel /a/. Figure 13.4 shows the photoformed products and 3D models used in the sensory evaluation, which are based on the lip motion data shown in Figure 13.3: Figure 13.4 (a) is an elliptical column and the diameter of each ellipse is the same as each lip contour, Figure 13.4 (b) is also an elliptical column and all ellipses are centered in the column, Figure 13.4 (c) is a column and the area of each circle is the same as each lip area, Figure 13.4 (d) is a column and each section is a dodecagon, Figure 13.4 (e) is a column and each section is a hexagon.

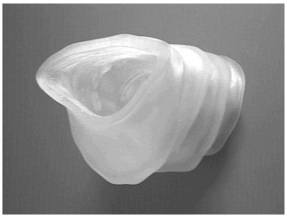


Figure 13.3 Photoformed product made from lip motion data

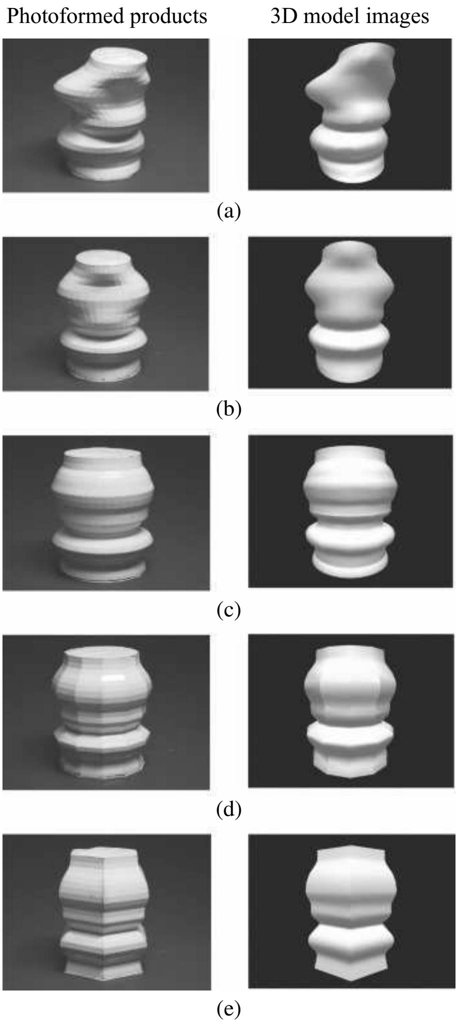


Figure 13.4 Photoformed products and 3D model images used in the experiment: (a) ellipse based on Figure 13.3 (El.), (b) ellipse centered (Ec.), (c) circle centered (Cc.), (d) dodecagon circumscribed (Dc.), and (e) hexagon circumscribed (Hc.)

13.2.2 Sensory Evaluation

The 3D shapes shown in Figure 13.4 (a–c) were used in the sensory evaluation of paired comparison. First, the subject put on the HMD and the magnetic sensors were attached to the subject's head and objects. Two of the three objects were simultaneously presented to each subject in virtual spaces (Figure 13.5), and the subject watched the objects' images from various viewpoints by moving his/her head and/or handling the real objects, as shown in Figure 13.6 (a). Then the subject was asked to say which object he or she preferred. After that, the subject took off the HMD and the same experiment was carried out in real space, as shown in Figure 13.6 (b).

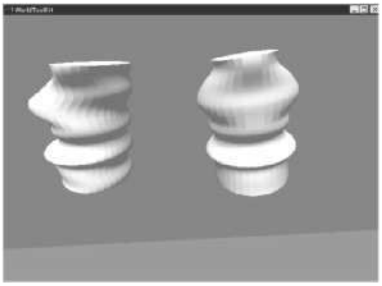


Figure 13.5 Two objects' images in virtual space

(a) (b)

Figure 13.6 Experimental scenes: (a) in virtual space, and (b) in real space

13.2.3 Experimental Results

Table 13.1 (a) shows the result of paired comparison for photoformed products in real space, and Table 13.1 (b) shows that for images viewed on the HMD in virtual space. Each table shows the number of subjects who preferred the object in the column over the object in the row.

Table 13.1 Result of paired comparison: (a) in real space, and (b) in virtual space

(a)

Product	El.	Ec.	Cc.	Total
El.	–	6	6	12
Ec.	22	–	6	28
Cc.	22	22	–	44

(b)

Product	El.	Ec.	Cc.	Total
El.	–	7	6	13
Ec.	21	–	5	26
Cc.	22	23	–	45

The Bradley–Terry model is assumed to evaluate the preference of 3D shape quantitatively, defined as follows [2]:

$$P_{ij} = \frac{\pi_i}{\pi_i + \pi_j}, \tag{13.1}$$

$$\sum_i \pi_i = \text{const. } (= 30), \tag{13.2}$$

where π_i is the intensity of i and P_{ij} is the probability of judgment that i is better than j. π_i shows the intensity of preference of the object.

The model enables the preference to be determined based on the paired comparison. The maximum likelihood method is used to solve π_i. We obtain $\hat{\pi}_i$ by the following formula, where $\hat{\pi}_i^0$ ($i = 1, 2, 3$) were used as initial values:

$$\hat{\pi}_i = \frac{T_i}{\sum_{j(\neq i)} \frac{N}{\hat{\pi}_i^0 + \hat{\pi}_j^0}}, \tag{13.3}$$

where N is the number of objects, and T_i is the total number of winning is. Then $\hat{\pi}_i$ is scaled up or down to satisfy the next formula:

$$\sum_i \hat{\pi}_i = K, \tag{13.4}$$

where K is 30. Therefore,

$$\hat{\pi}_i^1 = \frac{K\hat{\pi}_i}{\sum_i \hat{\pi}_i}. \tag{13.5}$$

We iterated the series of calculation until π_i was settled. The result of the above process is shown in Figure 13.7. As a result, the same shape (Cc.) tends to be preferred on every experiment. It is also obvious that the experimental result using the proposed system is closer to the result in real space than by using 3D CAD.

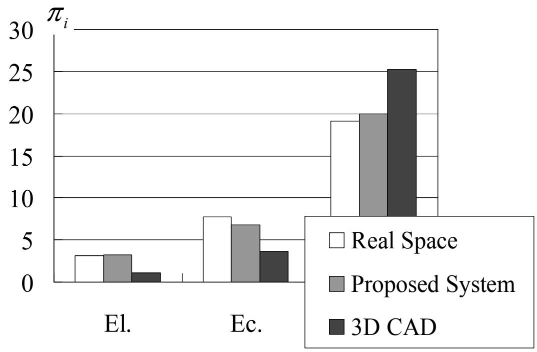


Figure 13.7 Preference of 3D shape

To approve the matching of the models, we apply the goodness-of-fit test and the likelihood ratio test. First, we apply these tests to the Bradley–Terry model of shape evaluation in real space, shown in Figure 13.6.

The goodness-of-fit test for the Bradley–Terry model is as follows:

$$\chi_0^2 = \sum\sum \frac{(X_{ij} - X_{1ij})^2}{X_{1ij}} (= 2.74), \tag{13.6}$$

where X_{ij} is the number of is winning over j, and

$$\chi^2(1, 0.05) = 3.84 > \chi_0^2.$$

The likelihood ratio test is as follows:

$$r = 2\sum\sum X_{ij} \times \log\left(\frac{X_{ij}}{X_{1ij}}\right) (= 2.69), \tag{13.7}$$

$$\chi^2(1, 0.05) = 3.84 > r.$$

These results show that the matching of the model in real space is consistent.

On the other hand, for the tests with the model shape evaluation in virtual space, the goodness-of-fit test is as follows:

$$\chi_0^2 = 2.82 (< \chi^2(1, 0.05) = 3.84).$$

The likelihood ratio test is as follows:

$$r = 2.76 (< \chi^2(1, 0.05) = 3.84).$$

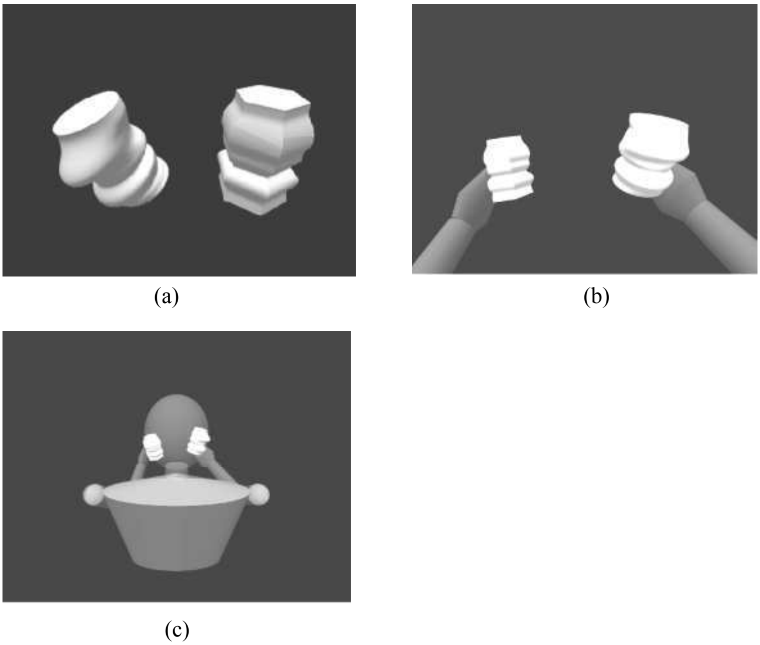
These results show that the matching of the model in virtual space using the proposed system is also consistent. In previous research, however, the matching of the model in virtual space using 3D CAD was inconsistent. This indicates that the tactile, gazing line-action information plays an important role in 3D shape evaluation. This proposed system also enables us to analyze the information which the subjects use in shape evaluation by handling the information.

13.3 Influence of Viewpoint on Shape Evaluation in Virtual Space

In virtual space, we can fix the viewpoint anywhere – although some of these give impossible scenes in real space. However, these viewpoints sometimes encourage the user to do something, *e.g.*, playing a driving game, communicating with a remote partner using the avatars, and so on. In this section, the influence of viewpoint on shape evaluation is described.

13.3.1 Three Types of Viewpoint

We have developed an embodied virtual communication system for human interaction analysis by synthesis, in which remote talkers can share the embodied interaction by observing the interaction of their VirtualActors in the same virtual space [3, 4]. From the user's viewpoint, not only the partner's motion images but also his/her own motion images in virtual space should be supplied in the collaboration system. However, motion images that include themselves would discourage them from performing the individual tasks. Then, the influence of avatar's images in virtual space on the 3D shape evaluation is investigated using the proposed system.



(a) (b) (c)

Figure 13.8 Three types of viewpoint: (a) only objects, (b) objects and forearms of user's avatar, and (c) objects and translucent upper body of user's avatar

Three types of view were prepared for the 3D shape evaluation in virtual space by the sensory evaluation of paired comparison. Figure 13.8 shows images from the HMD that show the two 3D models in virtual space. The colors of the objects are almost the same as that of the real objects. The subjects can change the view-point by moving their heads, and/or the objects in the case of Figure 13.8 (a, b).

13.3.2 Influence of Avatars' Forearms on 3D Shape Evaluation

To investigate the influence of the user's avatar's forearms on the 3D shape evaluation in virtual space, the subjects were first put on the HMD and magnetic sensors were attached to the subject and to the objects. Two of the five 3D models shown in Figure 13.4 were simultaneously presented to each subject in virtual space, and the subject watched the objects' images from various points of view by moving his/her head and handling the real objects in his/her hands in Experiment I-1 (Figure 13.8 (a)). Then the subject was asked to say which object he or she preferred. Next, the same experiment was performed in virtual space with the forearms of the user's avatar visible (Figure 13.8 (b)) in Experiment I-2. Then, the subject took off the HMD and was handed two of the five photoformed products shown in Figure 13.4 to compare them in real space in Experiment I-3. After each experiment, the subjects answered questionnaires about the shape evaluation environment and the system under each experimental condition.

In the next step, the influence of the avatar's motion images were investigated by comparing the two types of viewpoint in virtual space. One is fixed on the rear space and user can see the objects and his avatar's translucent upper body (Figure 13. 8(c)) in Experiment II-1, the other is put on the avatar's head and the user can see the objects and the avatar's forearms (Figure 13.8(b)) in Experiment II-2. To compare the result of sensory evaluation in these virtual spaces with that in real space, Experiment II-3 was carried out in the same way as Experiment I-3.

13.3.3 Experimental Results

Table 13.2 shows the results of paired comparison in Experiment I. Each table shows the number of subjects who preferred the object in the column to the object in the row. The Bradley–Terry model is assumed to evaluate the preference of 3D shape quantitatively.

The Bradley–Terry models for each virtual and real spaces in Experiment I are shown in Figure 13.9. As a result, the same shape (Cc.) tends to be preferred in each experiment. To check that the models match, we applied the goodness-of-fit test and likelihood rate test. We applied these tests to the Bradley–Terry model shown in Figure 13.9.

Table 13.2 Result of paired comparison (Experiment I): (a) in virtual space (Exp. I-1), (b) in virtual space (Exp. I-2), and (c) in real space (Exp. I-3)

(a)

Product	El.	Ec.	Cc.	Dc.	Hc.	Total
El.	–	4	2	4	4	14
Ec.	14	–	2	6	9	31
Cc.	16	16	–	13	12	57
Dc.	14	12	5	–	13	44
Hc.	14	9	6	5	–	34

(b)

Product	El.	Ec.	Cc.	Dc.	Hc.	Total
El.	–	5	4	5	7	21
Ec.	13	–	4	7	9	33
Cc.	14	14	–	13	13	54
Dc.	13	11	5	–	11	40
Hc.	11	9	5	7	–	32

(c)

Product	El.	Ec.	Cc.	Dc.	Hc.	Total
El.	–	4	3	4	5	16
Ec.	14	–	3	6	8	31
Cc.	15	15	–	11	15	56
Dc.	14	12	7	–	16	49
Hc.	13	10	3	2	–	28

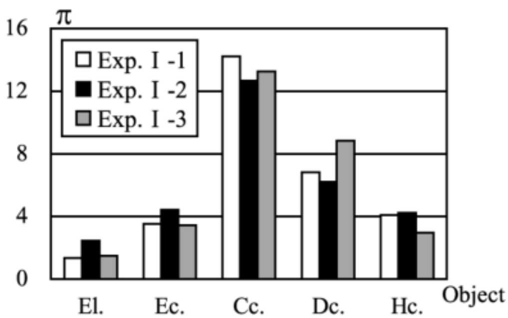


Figure 13.9 Preference of 3D shape in virtual and real space

As a result, the matching of the model in each virtual space using the proposed system and in real space is also consistent. These demonstrate that the forearms of the user's avatar in virtual space do not appear to affect the 3D shape evaluation in virtual space.

To analyze the individual evaluation environment in detail, the number of subjects who changed the selection between paired comparisons in virtual space and

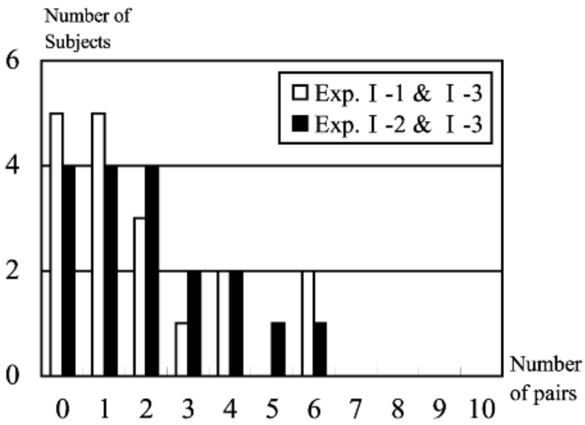


Figure 13.10 Frequency distribution of subjects for the number of changes of selection between paired comparisons in virtual and real space

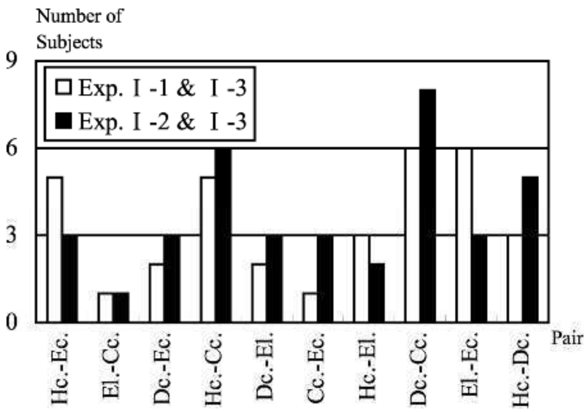


Figure 13.11 Number of subjects who switch to an alternative at paired comparisons in virtual and real space

in real space is investigated. Figure 13.10 shows the frequency distribution of subjects for the number of selection changes between paired comparisons in virtual space without the avatar's forearms (Exp. I-1) and in real space (Exp. I-3), shown as white bars, and between them in virtual space with the avatar's forearms (Exp. I-2) and in real space (Exp. I-3), shown as black bars. It can be seen from this figure that there is no significant difference between them.

Figure 13.11 shows the number of subjects who changed the selection for the each paired comparison between Exp. I-1 and Exp. I-3 (white bars), and between Exp. I-2 and Exp. I-3 (black bars). The differences between each pair are not significant. These results indicate that displaying the avatar's forearms has no influence on the shape evaluation in virtual space.

13.3.4 Influence of Avatar’s Upper Body Image

Table 13.3 shows the results of paired comparison in Experiment II. The Bradley–Terry models for each virtual and real space in Experiment II are also assumed and the results are shown in Figure 13.12. As a result, the same shape (Cc.) tends to be preferred in each experiment, as was found for Experiment I.

To approve the matching of the models, the goodness-of-fit test and likelihood rate test were also applied to the Bradley–Terry model. The results indicate that the matching of the model in each virtual space using the proposed system and in real space are both consistent. These demonstrate that the avatar’s upper body in virtual space doesn’t appear to affect the 3D shape evaluation in virtual space.

To analyze the individual evaluation in detail, the number of subjects who changed their selection between paired comparisons in virtual space and in real space is investigated.

Figure 13.13 shows the frequency distribution of subjects for the number of selection changes between paired comparisons in virtual space with the avatar’s upper half of the body (Exp. II-1) and in real space (Exp. II-3), shown as white bars, and between them in virtual space with avatar’s forearms (Exp. II-2) and in

Table 13.3 Result of paired comparison (Experiment II): (a) in virtual space with avatar’s upper body (Exp. II-1), (b) in virtual space with avatar’s forearms (Exp. II-2), and (c) in real space (Exp. II-3)

(a)

Product	El.	Ec.	Cc.	Dc.	Hc.	Total
El.	–	8	6	7	7	28
Ec.	21	–	11	10	9	51
Cc.	23	18	–	16	21	78
Dc.	22	19	13	–	22	76
Hc.	22	20	8	7	–	57

(b)

Product	El.	Ec.	Cc.	Dc.	Hc.	Total
El.	–	8	4	7	10	29
Ec.	21	–	8	10	6	45
Cc.	25	21	–	17	20	83
Dc.	22	19	12	–	18	71
Hc.	19	23	9	11	–	62

(c)

Product	El.	Ec.	Cc.	Dc.	Hc.	Total
El.	–	5	3	4	9	21
Ec.	24	–	4	7	13	48
Cc.	26	25	–	20	21	92
Dc.	25	22	9	–	23	79
Hc.	20	16	8	6	–	50

real space (Exp. II-3), shown as black bars. There is no significant difference between them. Also, the differences between the number of subjects who changed their selection for each paired comparison between Exp. II-1 and Exp. II-3, and between Exp. II-2 and Exp. II-3, are not significant. These results indicate that displaying the avatar's upper body has no influence on the shape evaluation in virtual space.

After each experiment, the subjects are asked some questions about each virtual environment for 3D shape evaluation and the difference between Exp. II-1 and Exp. II-2. The results of this questionnaire are shown in Figure 13.14. The results for all questionnaires about the difference between Exp. II-1 and Exp. II-2 are significant at a significance level of 1%. These results indicate that the subjects prefer a viewpoint from their own avatar's eyes to one from the rear space of their avatar when performing 3D shape evaluation in virtual space.

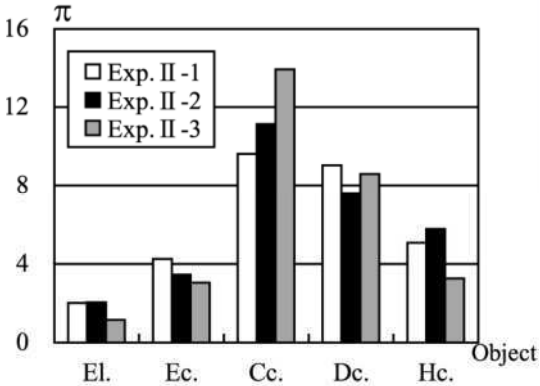


Figure 13.12 Preference of 3D shape using the Bradley–Terry model

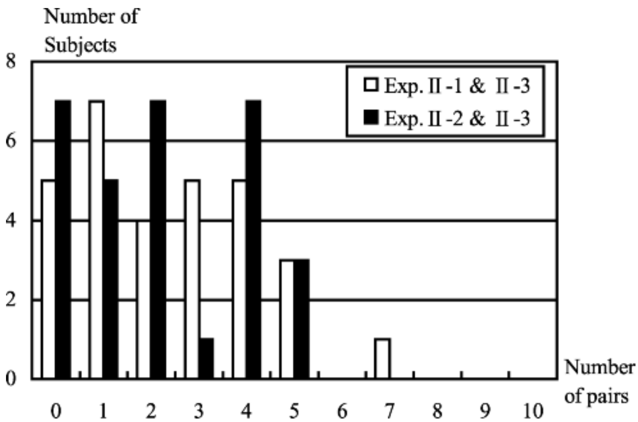


Figure 13.13 Frequency distribution of subjects for the number of changes of selection between paired comparisons in virtual and real space

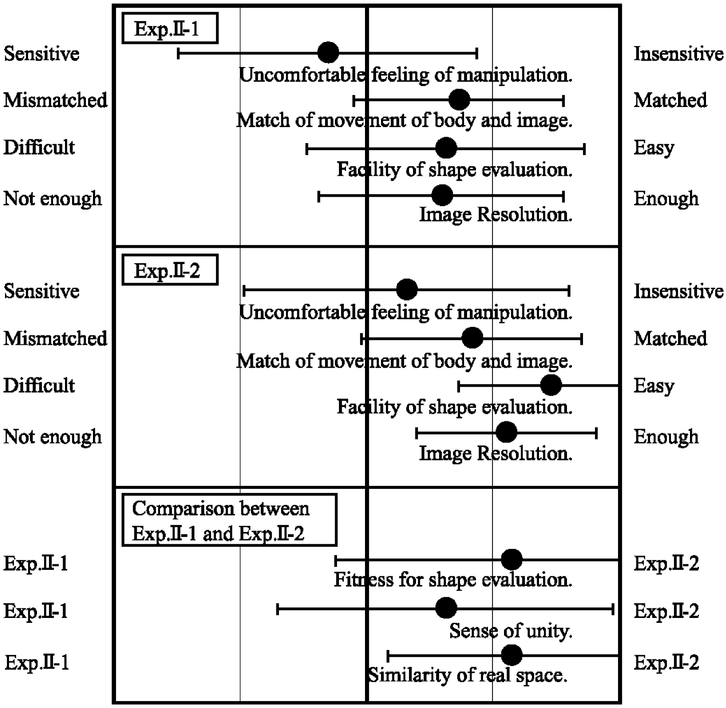


Figure 13.14 Result of questionnaires

13.4 Conclusions

We have proposed a virtual environment for 3D shape evaluation in order to clarify the differences in 3D shape evaluation in virtual and real spaces. Using this system, users are able to obtain visual information in virtual space and tactile and gazing line-action information in real space. The preference of the 3D shape images on a HMD was compared with that of real photoformed products made from the same data by the sensory evaluation of paired comparison. As a result, the same shape tended to be preferred in both spaces, and there was no significant difference in the relations of preference among shapes based on the Bradley–Terry model for the sensory evaluation. This indicates that tactile, gazing line-action information plays an important role in 3D shape evaluation. This proposed system enables us to analyze the information which the subjects use in shape evaluation by handling these information. Using this system, the influence of difference of the viewpoint on shape evaluation in virtual space is investigated by the sensory evaluation of paired comparison and questionnaire analysis. The subjects can see their actions performed by translucent avatars or see only the motion of their forearms. As a result, it is found that the subjects prefer a point of view which shows only the motion of their avatar's forearms in the case of shape evaluation in virtual space.

References

1. Okubo M., Watanabe T.: Sensory evaluation of the preference of 3D shape in virtual and real environments. In: IEEE International Workshop on Robot and Human Communication (RO-MAN '97), pp. 460–464 (1997)
2. Hori T.: Brain and Emotion. Kyoritsu Press (in Japanese) (1991)
3. Watanabe T.: Evaluation on the naturalness of human walking motion by animation. Trans. JSME **54**(502):1291–1296 (in Japanese) (1998)
4. Watanabe T., Masashi O., Nakashige M., Danbara R.: InterActor: speech-driven embodied interactive actor. Int. J. Human-Comp. Interaction **17**(1):43–60 (2004)

Chapter 14
Affective Driving

Yingzi Lin[1]

Abstract In this chapter, we discuss how emotion may play important role in driving especially in terms of driving safety. Vehicle driving is a life-critical process and is pervasive in our daily life. Emotion is, however, often considered to be not particularly relevant to vehicle driving, with the arguments: (1) safety takes precedence over any emotional needs, so any driver assistance systems (DAS) should only look at the driver's performance and *not* emotion, and (2) emotion does not significantly change driving performance. However, several studies conducted by us reveal that emotion can be as important as fatigue in driving applications, and research on how DAS may help to regulate drivers' emotions is highly needed. This chapter gives an overview of our research, leading to the view that future DAS need to consider emotion. At the end, there is an outline of the existing issues and future research directions on incorporating emotion in the design and management of vehicle and transportation systems.

14.1 Introduction

Emotion is a well-known term but lacks a universal agreement on how it works. There appear to be two views of emotion in literature: the first is that emotion is the experience of involuntary physiological changes [1] (*e.g.*, anger accompanied by increased heart rate), while the second is that emotion is the outcome of cognitive evaluation (*e.g.*, whether one's goals are met in the interaction with the environment) [2, 3]. In our view, these two views are inter-related, and involve four processes (see Figure 14.1). First, a human must perform a cognitive task or

[1] Y. Lin
Dept. of Mechanical and Industrial Engineering, Northeastern University,
Boston, 02115 MA, USA
e-mail: yilin@coe.neu.edu

a mixed cognitive and physical task. Second, the performance on the task is evaluated, which is a cognitive process. Third, the result of the evaluation induces changes in psychophysiology. Fourth, the changes in psychophysiology further affect the performance of the task alongside the experiences of these changes, which are transformative to different categories of semantics of the experiences, namely the different labels of these experiences (anger, happiness, *etc.*). These four processes happen in a cycle as shown in Figure 14.1. For the convenience of later discussions, let's put the four cyclic processes together as a model and call it "TCPE" (T: task; C: cognitive evaluation; P: psychophysiological; E: emotion) hereafter.

The discussion thus far suggests that all such notions as changes in psychophysiology, emotion, cognition, and bodily action are inter-related, as further depicted in Figure 14.2, where we also show that the brain serves as an ultimate center to manage the body, cognition, emotion, and their interaction. In addition, emotion is a semantic variable that is dependent on the psychophysiological states and their changes. The usefulness of emotion is such that particular emotional states represent mental states that are both related to human health and task performance. Though task performance depends on the psychophysiological states – which makes it sound that as far as task performance is concerned emotion is a redundant concept – the relationship between emotion and task performance is still useful. This is because (1) we do not know how many types of psychophysiological signals are needed to uniquely determine a particular type of emotion and task performance and (2) the use of psychophysiological states as independent variables to determine task performance can never be achieved because of (1). Therefore, if a subjective or a combined subjective and objective manner can tell one's emotion, this helps. These two reasons suggest that studies to establish the relationship between emotion and task performance will be useful.

Emotional state is a short-term state that usually lasts minutes and hours rather than days [4], compared with other human attributes such as mood, trait, and temperament [5]. Therefore, the interactions between emotion and cognition and between emotion and body are also short-term. Emotion changes can be further classified into transient and steady states; the former being defined as ahead of task performance and the latter being defined as along with task performance. In our study, we consider the steady state of emotion, which was ensured by a proper design of the experiments.

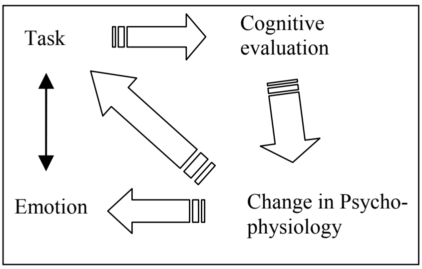


Figure 14.1 A unified view of emotion, task, and change in psychophysiology

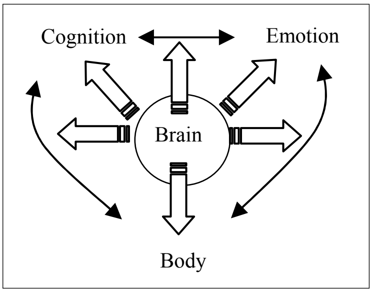


Figure 14.2 Brain, cognition, body, emotion and their conceptual relationship

Incorporating emotion into designing artifacts, especially dynamic systems or machines, is the last task for artificial intelligence research. This view appears to be similar to a view widely agreed among AI researchers in the discipline of computer science and engineering [6], *i.e.*, affective computing or incorporating emotion into computer systems is the last task for AI research. However, these two views are not exactly the same. The major difference is the difference of the target system: machine and computer. For a machine such as robot, it has its physical and chemical state that determines the machine task performance, while for a computer, it is the software running on it that interests us, so the state of the computer is not a determinant factor in the performance of the software. We will use the term "*affective machine*" in this chapter. Our study is on affective machines rather than affective computing.

There are two general questions that need to be answered in designing affective machines. The first question is how the machine knows human emotion (or emotional state) and emotional behavior. Here, emotional behavior means (1) the causes of emotions and (2) the emotional effects on cognition and action. The second question is how the machine responds to humans' emotional cognition and action. The first question can be divided into two sub-questions: how to get cues from humans and how to infer humans' emotion and emotional behavior. The second question can be further divided into two sub-questions: what should be the best response from the machine and how is the best response best expressed or acted in effect on the human? Nevertheless, the ultimate goal of the design is to improve the performance of such human–machine collaborative systems; in the case of vehicle driving it is to make DAS both a safeguard and of high usability.

In this chapter, we provide an overview of our research on "*affective vehicles*", with a focus on our studies towards an answer to the first question above. Therefore, Section 14.2 has an overview of our work on the development of sensors that acquire cues from the driver non-intrusively and in real-time. Section 14.3 has an overview of our work on the development of an algorithm that infers the driver's emotion. Section 14.4 looks at our work on understanding drivers' emotions versus cognition and action (or task performance). Section 14.5 discusses the existing issues and future research directions. Section 14.6 gives a conclusion.

14.2 Natural-contact Biosensor Method

There are generally two methods for sensors to acquire cues from humans: non-contact sensors and wearable sensors. The disadvantage of non-contact sensors is the limited types of cues that they can acquire. The shortcoming of wearable sensors is that they are still intrusive to the human, as they require the human to "wear" the sensor, which is not necessary to an activity per se. For instance, when a human driver drives, the driver has to wear something that is in fact a transducer, according to the wearable sensor method, in order to measure his or her physiological signal (otherwise, the wearing is not a part of human and machine interactive activities).

We proposed a new biosensor method called a natural-contact (NC) biosensor [7]. The idea of the NC biosensor is based on the observation that in any human–machine system there must be some contacts between humans and machines. Sensors can then be designed to be placed on the contact surface of the machine. This is very much like the machine wearing something. Therefore, there is no intrusiveness in philosophy to the human. Based on this idea, we decided on the fundamental problems of the NC biosensor method [7]. In particular, there are four fundamental problems: (1) the transducer elements must be designed to be sufficiently small, being a part of the machine system; (2) the transducer elements must be designed and constructed to cover the contact surface as much as possible to cope with uncertainty in the instant contact location on the contact surface; (3) the signal-to-noise (SNR) ratio must be sufficiently high to cope with noises owing to the varying nature of the human; and (4) multi-signals need to be decoupled as the promise of the NC biosensor is to have all signals measured at one site.

The first proof of the NC biosensor was completed at our research laboratory in 2005 [8, 9], which was to measure the gripping force (GF) and blood volume pulse (BVP) of a driver holding a steering wheel. The second proof was to have more signals measured from the steering wheel, including skin conductance (SC) and skin temperature (ST) [10]. In this design, the second fundamental problem of the NC biosensor was dealt with by two steps. The first step is to analyze the human palm and the contact behavior of the driver, and the second step is to determine the distribution of the NC biosensors over the wheel surface.

In summary, the NC biosensor has more distinct features than the wearable sensor and is more promising than the wearable sensor in terms of non-intrusiveness, acquisition of multi-signals, and acquisition of more types of physiological signals. Since human skin contains a rich set of information with a pathological connection to the human body and mind, the NC biosensor is a promising method for sensors to acquire physiological signals.

14.3 Inference of Human Emotion

The inference of human emotion is to determine human emotional state given a set of pscychophysiological cues and contexts. Contextual information includes information about the task, pre-conditioning of the human operator, and the task-performing environment itself (see again Figure 14.1). This understanding is slightly different from that of others, such as Lisetti and Nasoz [11] and Picard [6], who did not pay attention to contextual information. According to our research contextual information is quite important [12].

Machine learning techniques seem to be most applicable to the problem of inference. These techniques can be further classified into generative learning and discriminative learning. Most approaches are generative learning Bayesian or Markov chain network techniques. Another popular category of techniques is the artificial neural network, which are typically based on the discriminative learning technique [11]. On top of these machine learning techniques, fuzzy logic can be employed to represent a type of imprecise information – vagueness [12, 13]. There are two ideas of how to apply fuzzy logic for inference of emotion (or mental state in a more general sense). The first is that the word becomes the final outcome of the inference, with a typical outcome statement being "the level of anger is very high"; see the work of Mandryk and Atkins [13]. The second is that a number is used for the outcome of the inference, with a typical outcome statement being "the level of anger is 0.7" (out of 1, where 1 = very angry and 0 = not angry at all); see our work [12, 14].

In parallel with the research into cognitive state inference – in particular, a finding described in [15] that no approach is powerful enough to infer cognitive state for all situations and therefore an integrated approach is needed – we stated that this finding can be equally applied to the inference of emotion. Following this line of thinking, we proposed the architecture of an integrated algorithm as shown in Figure 14.3. This architecture has three layers. The *first layer* is a grouping and clustering process (Figure 14.3 (a)). The *Grouping of cues* is based on the "proximity" in their relevance to an inferred target. The *Clustering of cues* is based on the number of instances of the cues in a group. One can observe that grouping is semantic-oriented and clustering is value-oriented. The *second layer* is a classification layer in which algorithms based on various machine learning formalisms (*e.g.*, artificial neural network, Bayesian network, *etc.*) and principle-based knowledge for inference (PB in Figure 14.3 (a)) are integrated. The *third layer* of the architecture is a probability distribution aggregation process. Each classifier (*i.e.*, the second layer) comes up with a crisp value for the inferred cognitive state (CS for short), CS (1), CS (2), *etc.* The CS takes on a value ranging from 0 to 1, which represents the degree of confidence in an inferred target (*e.g.*, fatigue of an operator, tank in a terrain field, *etc.*) with 1 representing the highest and 0 the lowest degree of confidence. The aggregation process integrates all CS (i) ($i = 1, 2, \ldots, n$, where n is the total number of classifiers or inference mappings) to an "agreed" inferred state or target. The aggregation process is built upon a probabilistic uncertainty (the deterministic case is viewed as a special case of the non-deterministic case). Methodologies in the fields of decision-making [16, 17] and expert opinion elicitation [18] are employed for this aggregation process. It is noted that in expert opinion elicitation [18], the problem can often be defined as a weighted average problem, in which each expert is associated with a weight that represents the expertise level of the expert with respect to a decision target. This can be applied to the cognitive inference problem here so that a group of cues corresponds to an expert, and a weight associated with the group represents the degree of inference power of the group with respect to an inferred CS. The notion of inference power makes sense; for example, heart rate variability is more sensitive than blood pressure to a driver's fatigue, and therefore the weight of the heart rate variability cue or its group should be higher than that of the blood pressure cue or its group in the aggregation process for the inferred driver's fatigue. The notion of the weight of the cue in the context of cognitive inference has not been discussed in the literature.

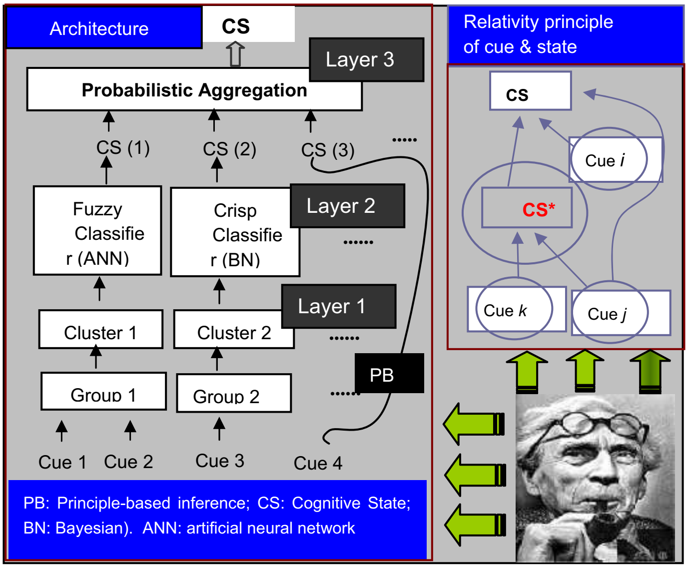


Figure 14.3 Integrated multi-modality cognitive state inference: (a) overall architecture, and (b) cue–state network [16]

Finally, the architecture can simulate an important human inference behavior – that is, a particular state in one context may be viewed as a cue in another context with its relevant inference process (Figure 14.3 (b)). In Figure 14.3 (b), CS* serves as a cue together with other cues (*i*, *j*), some of which may infer CS* in one context, to infer CS. The structure as shown in Figure 14.3 (b) is in its nature a network structure, which further means that the general architecture as proposed here views the cue-state relationship as a network, which is the nature of the human cognitive system. It is noted that the notion of the network relationship among cues and states appears not to be discussed in the current literature.

We applied a simplified version of this architecture to the inference of the fatigue state of the driver. This takes a machine learning model which integrates the first two layers, and the integrated model is simply the TSK (Takagi–Sugeno–Kang) model developed in the ANN literature for control. The approach model of the probability distribution aggregation, *i.e.*, layer 3, is the OWA. In the fatigue state inference, we used the following cues: eye movement (EM); driving hour (DH); sleeping quality (SQ); EEG; and ECG. The EEG and ECG fall into the category of physiological cue. The EM falls into the category of non-contact cue, and the DH and SQ fall into the contextual category of cue. Tables 14.1 and 14.2 show the result of the fatigue state inference for two drivers [19]. The accuracy to infer the fatigue state achieved about 90%.

Table 14.1 Features and simulation result for the first driver

Input features	SQ	DH	ECG	EEG		EM
	0.25	0.5	0.83	0.81	0.85	0.82
TSK output	y_1	y_2	y_3	y_4		y_5
	0.9046	0.5907	0.8925	0.849		0.89
	Overall average: 0.8258					

Table 14.2 Features and simulation result for the second driver

Input features	SQ	DH	ECG	EEG		EM
	0.875	0. 167	0.33	0.38	0.41	0.43
TSK output	y_1	y_2	y_3	y_4		y_5
	0.075	0.21	0.29	0.33		0.41
	Overall average: 0.8258					

We further designed a simplified version of the architecture in the sense that we only include a fuzzy knowledge-based inference engine from the cue to emotion state [20]. In this case, we only considered the following psychophysiological signals: heart rate, skin conductance, skin temperature, and respiration rate, due to their highly relevant to emotional states. Table 14.3 shows an example of a driver's steady emotional state over a relatively long time period, and Figure 14.4 shows an example of driver's transient emotion state over a relatively short time period. The accuracy of inference achieved is about 86% (for anger) and 87% (for disgust).

From our study, emotion is a factor that cannot be ignored in vehicle driving. The inference of the driver's emotional states as well as other mental states such as fatigue, drowsiness, and attention can reach a level of accuracy of just over 85% with the physiological cues that are non-intrusively obtained. The next section will show the state of knowledge about how the emotional states affect the driver's driving performance.

Table 14.3 Emotion rating by self-reporting, expert rating, and fuzzy emotion analyzer

Emotion	Self-reporting	Expert-rating	Emotion analyzer (CF)	Meaning
Calm			–0.61	Unlikely
Joy				
Pleasure			0.19	Unknown
Anger	0.7	0.7	0.61	Probably
Sadness			0.4	Maybe
Excitement			–	–
Surprise			–	–
Fear	0.3		0.40	Maybe
Anxiety			–	–
Frustration			0.48	Probably
Disgust		0.3	0.34	Maybe
Nervousness			–	–

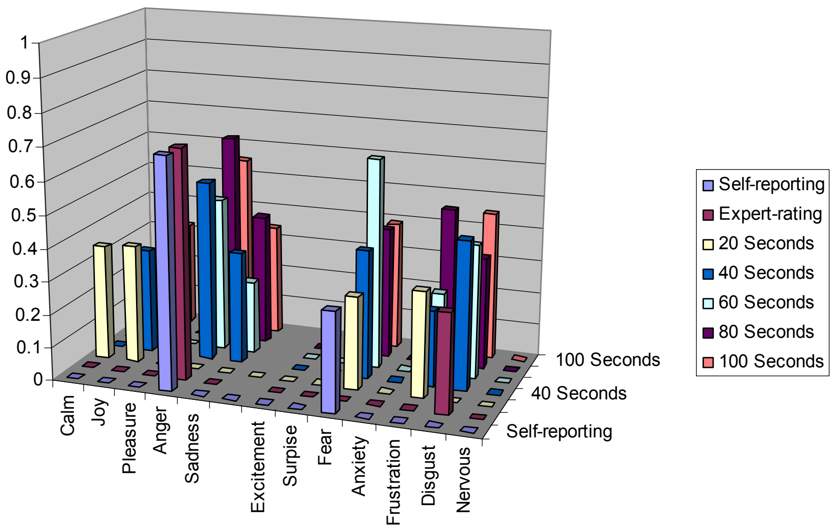


Figure 14.4 An example of emotion transition estimated by the emotion analyzer

14.4 Emotion Versus Performance in Driving

There are two methods of studying the relationship between emotion and task performance: experimental and theoretical. The experimental approach demands a sufficient number of samples and proper generalization. The theoretical approach demands accurate modeling of cues, emotions, tasks, and task performance according to the TCPE model as mentioned before. In this chapter, we give an overview of a study we conducted in an attempt to find a model of the general relationship between a driver's emotions and his/her performance, following the experimental approach.

In the literature, Russell [21] proposed an emotion model based on the two dimensions of arousal and valence (see Figure 14.5). Further, there is a well-known model of the relationship between performance and arousal, called the inverted U-shape model [22] (see Figure 14.6). Further, Brookhuis and De Waard [23] investigated the relationship between drivers' mental workload and driving performance in different situations. They concluded that mental workload has an inverted U-shape relationship with driving performance. This finding is, however, not a surprise, as mental workload has a close association with arousal [24].

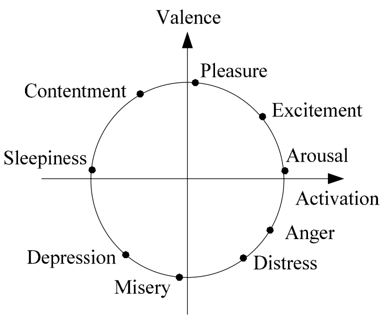


Figure 14.5 2D emotion model [21]

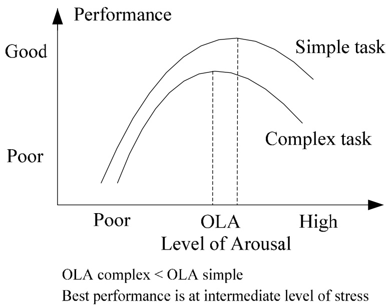


Figure 14.6 Yerkes–Dodson law [22]

It should be noted that the Yerkes–Dodson model is descriptive, so it cannot be used for any quantitative use. Mental workload is viewed as a single variable or attribute, while emotion is considered as a function of two variables according to the two-dimensional model (valence, arousal). Therefore, we investigated the quantitative relationship between emotion and task performance in the context of vehicle driving, which is a pioneer study. In particular, we proposed the following relationship between emotion and task performance [25]:

$$P_{\mathrm{A}} = a\mathrm{Ar}^2 + b\mathrm{Ar} + c\text{, where } a < 0,\ -1 < \mathrm{Ar} < 1, \tag{14.1}$$

$$P_{\mathrm{V}} = d\mathrm{Va} + e\text{, where } d > 0,\ -1 < \mathrm{Va} < 1, \tag{14.2}$$

$$P_{\mathrm{AV}} = P_{\mathrm{A}} \times P_{\mathrm{V}}, \tag{14.3}$$

where Ar and Va are variables that indicate the level of valance and arousal; P_{A} is the task performance associated with the change in arousal and P_{V} is the task performance associated with the change in valence. Specifically, $a < 0$ indicates that the curve P_{A} opens downward, $d > 0$ indicates that the curve P_{V} increases with Va; and b, c, e are parameters relevant to a particular task.

The experiment was designed as follows. Fifty-three samples were taken from 15 participants. The participants saw a movie to stimulate their emotions, and then used a driving simulator. They did multi-tasks: primary driving task and secondary visual search task.

The result confirms that the model is adequate to represent the emotion and task performance relation. In the case of vehicle driving, the parameters in the model were determined through the experiment. This experimental study further concluded: (1) there is a downward U-shape relationship between arousal and task performance; (2) the optimal task performance occurred at the medium arousal range from $0.0 < \mathrm{Ar} < 0.4$; (3) there is a linear increase relationship between task performance and valence when $\mathrm{Va} < 0.5$, but this relationship stops at a positive valence ($\mathrm{Va} < 0.5$) or at high arousal ($\mathrm{Ar} \geq 0.6$); (4) arousal and valence were not perfectly independent in the whole 2D emotion plane; (5) the spaces of the emotion plane such as the planes $\mathrm{Va} > 0.5$ and $\mathrm{Ar} < 0$, and $\mathrm{Va} < -0.5$ and $\mathrm{Ar} < -0.5$ are not occupied at all, which may be due to the slight correlation between arousal and valence; (6) the effect of arousal and valence on the worst driving performance was significant ($p < 0.05$); and (6) the average driving performances at four emotion zones were also significantly different ($p < 0.05$).

In summary, the steady state period of emotion can have a significant effect on task performance in driving. The usefulness of the emotion–performance relationship model can help to predict drivers' task performance degradation and therefore assist drivers in managing emotion to avoid any safety-threatening performance.

14.5 Issues and Future Research Directions

The first issue is to improve the accuracy, robustness, and resilience of the biosensor system with the real-time and non-intrusive measurement capability. This requires an approach that integrates design, fabrication, and signal processing to look into finding a global optimum. The second issue is how to make "soft" or "flexible" biosensors in the context of the NC biosensor method; in particular a large array of transducer elements grown on a soft material. This will enable the biosensor to be more easily worn on the body of the machine that interacts with humans and will therefore be able to cope with any contact uncertainty. The third issue is the relationship between the internal state of the machine and its emotion (emotion engine and emotion expression). This issue is important when DAS is deciding how to express the machine's state to the driver in order to influence the driver's emotion and task performance, which in turn will affect the overall performance of the driver–vehicle system. The fourth issue is the decision-making of the DAS in order to assist the driver in achieving the primary goal of driving safely, along with the secondary goal of being of high usability to the driver.

14.6 Conclusions

This chapter gives an overview of the studies conducted by us on the subject of affective vehicles or affective machines in general. The overview concentrates on how machines can understand humans' emotion and the effect of emotion on their task performance. The following conclusions are made: first, drivers' emotional states have significant effects on their driving performance; second, drivers' emotional states can be inferred through cues that include both contextual and psychophysiological elements. The results demonstrated in this chapter support the above two conclusions.

Finally, research towards such affective or emotional machines is highly important, especially with the development of more and more intelligent machines and/or human–machine interactions in modern society.

Acknowledgements The author gratefully acknowledges the financial support of the National Science Foundation (NSF). The writing of this chapter benefited from the research efforts of her Intelligent Human–Machine Systems Laboratory.

References

1. Ekman P., Davidson R.J., Friese, WV.: Emotional expression and brain physiology II: The Duchenne smile. J. Personality Social Psychol. **58**:342–353 (1990)
2. Ortony A., Clore G.L., Collins A.: The Cognitive Structure of Emotions. Cambridge University Press, New York (1988)
3. Cron W.L., Slocum J.W. Jr, Vandewalle D., Fu Q.: The role of goal orientation on negative emotions and goal setting when initial performance falls short of one's performance goal. Human Performance **18**(1):55–80 (2005)
4. Cowie R., Douglas-Cowie E., Tsapatsoulis N. *et al.*: Emotion recognition in human-computer interaction. IEEE Signal Proc. Mag. **18**(1):32–80 (2001)
5. Blignaut P., Naude A.: The influence of temperament style on a student's choice of and performance in a computer programming course. Comp. Human Behavior **24**(3):1010–1020 (2008)
6. Picard R.W.: Affective Computing. MIT Press (1997)
7. Lin Y.: NC biosensor paradigm. Interacting with Computers (submitted) (in press)
8. Leng H.: Design and application of a novel driver condition monitoring device – smart wheel. Unpublished MS thesis, Concordia University, Montreal (2005)
9. Lin Y., Leng H., Yang G., Cai H.: An intelligent noninvasive sensor for driver pulse wave measurement. IEEE Sensors J. **7**(5):790–799 (2007)

10. Cai H., Lin Y.: An experiment to non-intrusively collect driver physiological parameters towards cognitive/emotional state recognition. SAE Paper No. 2007-01-0401, SAE 2007 World Congress, April 16–19, Cobo Center, Detroit, MI (2007)
11. Lisetti C.L., Nasoz F.: Using noninvasive wearable computers to recognize human emotions from physiological signals. EURASIP J. Appl. Signal Proc. **11**:1672–1687 (2004)
12. Yang G., Lin Y.: Proposal of a neuro-fuzzy TSK OWA network model for multimodality information fusion. In: Dainoff, M.J. (ed.) Lecture Notes in Computer Science, pp. 633–637. Springer, Berlin (2007)
13. Mandryk R.L., Atkins MS.: A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. Int. J. Human Comp. Studies **65**(4):329–347 (2007)
14. Yang G., Lin Y., Bhattacharya P.: A method for multi-modality inferring of human cognitive states based on the integration of neuro-fuzzy network and information fusion techniques, 1. EURASIP J. Adv. Signal Proc. (2008). doi:10.1155/2008/371621
15. Gray W.D.: Integrated Models of Cognitive Systems. Oxford University Press, New York (2007)
16. Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Trans. Systems Man Cybernetics **18**(1):183–190 (1988)
17. Yager R.R.: OWA aggregation over a continuous interval argument with applications to decision making. IEEE Trans. Systems Man Cybernetics Part B **34**(5):1952–1963 (2004)
18. Clemen R.T., Winkler R.L.: Combining probability distribution from experts in risk analysis. Risk Anal. **19**:187–203 (1999)
19. Yang G., Lin Y., Bhattacharya P.: A driver fatigue recognition model based on information fusion and dynamic Bayesian network. Inf. Sci. **180**:1942–1954. Available online at: http://dx.doi.org/10.1016/j.ins.2010.01.011 (2010)
20. Lin Y.: Emotion analyzer. Intelligent human-machine systems laboratory. Northeastern University, Technical Report (2008)
21. Russell J.A.: A circumplex model of affect. J. Personality Social Psychol. **39**(6):1161–1178 (1980)
22. Yerkes R.M., Dodson J.D.: The relation of strength of stimulus to rapidity of habit-formation. J. Compar. Neurol. Psychol. **18**:459–482 (1908)
23. Brookhuis K.A., De Waard D.: Assessment of drivers' workload: performance, subjective and physiological indices. In: Hancock, P., Desmond, P. (eds.) Stress, Workload and Fatigue: Theory, Research and Practice, pp. 321–333. Lawrence Erlbaum, New Jersey (2001)
24. Kroemer K., Kroemer H., Kroemer-Elbert E.: Ergonomics: How to Design For Ease & Efficiency, 2nd edn. Prentice-Hall Inc., Englewood Cliffs, NJ (2001)
25. Cai H., Lin Y., Mourant R.R.: Study on driver emotion in driver-vehicle-environment systems using multiple networked driving simulators. In: Driving Simulation Conference 2007 – North America, September 12–14, Iowa City (2007)

Chapter 15
Approximate Multi-objective Optimization of Medical Foot Support

Masao Arakawa[1]

Abstract Although splayfoot does not seem to be a serious disease, it can cause fatigue in daily life. Therefore, the solution of this problem can make daily life much more comfortable, particularly for elder people. There are some commercial products to treat splayfoot, but they just add a small amount of support and are not personalized for each patient. In addition, if the support height is not correct for the patient, it can make the condition worse. Physical therapists are able to create foot supports for each patient, but use their experience to make them. There have been many studies made on the structure of the bones of the foot, and a summary of the desired positions has been reported by medical doctors, one of which is called the Mizuno standard. In this chapter, we will describe the design of a medical foot support using approximate multi-objective optimization, in order to position the height of the bones to the Mizuno standard, and show its effectiveness through myoelectric potential tests.

15.1 Introduction

These days, a number of people have problems with their feet, such as splayfoot, flatfoot, bowlegs, and so on. Splayfoot is a phenomenon characterized by a lack of cushioning in the vertical arch, which therefore causes a lot of pressure on the bones in the foot. When too much pressure is placed on the bones, it causes fatigue and sometimes sufferers cannot walk because of the pain. This is not only bad for the bones, but also for the muscles of the leg. One reason for leg fatigue is that the muscles are being used in an unusual way.

[1] M. Arakawa
Department of Reliability-based Information Systems Engineering, Kagawa University
2217-20, Hayashicho, Takamatsu, Kagawa, 761-0396, Japan
e-mail: arakawa@eng.kagawa-u.ac.jp

These days, it is possible to buy commercial arch supports (Figure 15.1) through the Internet. However, as these are not designed to heal the patient and sometimes do not fit them properly, they can make the situation worse. In a previous study [1], we examined the effectiveness of support for three different levels of patients through myoelectric potential measurements and showed the different effectiveness for each patient. From these results, we have designed a vertical 2D shape for arch support by comparing the positions of the bones to the Mizuno standard (Figure 15.2 and Table 15.1) using approximate multi-objective optimization [2, 3]. In the previous study, we have shown the effectiveness of the approximate optimization, and the results showed that we can reduce the divergence from the Mizuno standard, and at the same time also reduce the myoelectric potential. In the study, we used experimental data that we had obtained from X-ray results, with various height patterns made from plaster. However, we only designed the vertical shape, and also did not examine the effect of different support materials and plaster on the experiment. In the study described in this chapter, we will try to make a 3D shape using an application of the spline function with control points, and we will use the same material when checking the bone heights. We use the same approximate multi-objective optimization method. As a result, we have had several patterns of the good results, and have validated them with myoelectric potential measurements to show the effectiveness of the results.

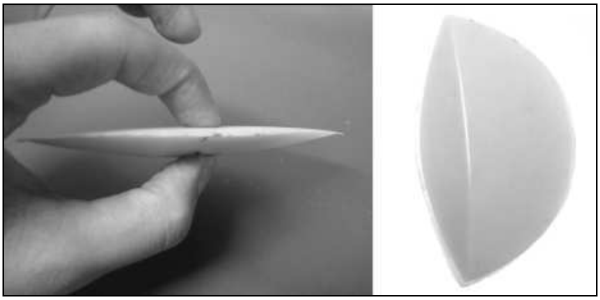


Figure 15.1 Arch support (commercial product)

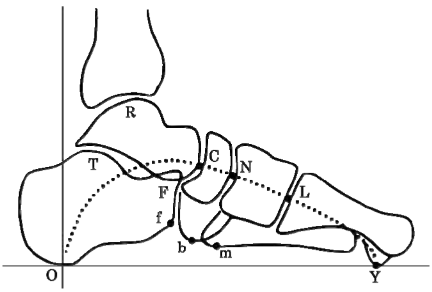


Figure 15.2 Measurement of bone positions in the Mizuno standard [2]

Table 15.1 Mizuno standard foot data

LY	21.66	±0.12%mm
NY	28.36	±0.16%mm
CY	32.71	±0.16%mm
MY	3.02	±0.11%mm
BY	8.11	±0.12%mm
fY	12.08	±0.13%mm

15.2 Approximate Multi-objective Optimization Using Convolute RBF

15.2.1 Convolute RBF

A radial basis function (RBF) network is a multi-layered neural network that gives an output that is a weighted summation of the middle layers. When we use a basis function as a Gaussian distribution with oval form with a different radius for each variable, it becomes:

$$\mathrm{RBF}(\mathbf{x}) = \sum_{i=1}^{m} w_i e^{-\sum_{j}^{n} \frac{(x_j - c_{ij})^2}{r_j^2}} \tag{15.1}$$

where m is the number of the basis function, n is the number of dimensions of variables, r_j is the radius for variable j, and c_{ij} is the ith center of the basis function of variable j. A Gaussian distribution has a character in that its response gets closer to 0 the further it is from its center. Using this property, we propose a convolute RBF as:

$$f_{\mathrm{app}}(\mathbf{x}) = Y(\mathbf{x}) + \sum_{k=1}^{K} \mathrm{RBF}_k(\mathbf{x}), \tag{15.2}$$

where $Y(\mathbf{x})$ is an arbitrary function. With this function the error between the original data set and $Y(\mathbf{x})$ becomes 0 on average. Now, we are going to approximate error between f_{app} and the teaching data. K is the number of convolutions. When there is strong nonlinearity in the original function and/or data, it is impossible to have enough accuracy in a single approximation of RBF. Thus, we need to carry out these approximations several times.

15.2.2 Satisficing Method

A common way of achieving multi-objective optimization is the weighted sum method, which is easy to understand and implement in software. However, even for simple cases, the changes of weights do not coincide with changes in the values of functions, especially when there are more than three objective functions. Therefore, trade-off analysis becomes very difficult to match to the requirements of designers [4].

The satisficing method is an interactive multi-objective optimization method. We set an ideal solution and aspiration level as the desired values of each objective function and obtain a Pareto optimal solution close to the aspiration level. The solution is obtained by minimizing the extended Thebychev scalarization function, as follows:

$$\min_{x\in\Omega}\left\{\max_{i\in N} w_i \frac{f_i(x)-f_i^{\text{asp}}}{f_i^{\text{asp}}-f_i^{\text{ideal}}}+\alpha\sum_{i=1}^{N} w_i \frac{f_i(x)-f_i^{\text{asp}}}{f_i^{\text{asp}}-f_i^{\text{ideal}}}\right\}, \tag{15.3}$$

where Ω is the design space and α is a small value to correct for correct a weak Pareto solution. The superscript "asp" stands for aspiration level, and "ideal" stands for the ideal value that each function can take. w_i is a weight factor that may play a role in distinguishing between objective function and constraints.

15.2.3 Data Addition

In approximate optimization, it is important to indicate where we need to add data for global optimization. If we can indicate appropriately, we can reduce the number of function calls and/or experiments. There are two main reasons for data addition: (1) to obtain a better accuracy around the optimum solution, and (2) to obtain a better approximation for the overall response. The second reason is not only for a better response surface, but also for global optimization to overcome the approximation error in areas with a lower density data distribution. In the multi-objective optimization case, we need to find an approximate scalar function of Equation 15.3 that is strongly nonlinear. When we approximate this function as $\mathrm{RBF}_\Phi(\mathbf{x})$ then, to have good response surface, it is important to know the maximum and minimum values of the original $\Phi(x)$. Thus, we would like to find a maximum absolute value of $\mathrm{RBF}_\Phi(\mathbf{x})$. In order to estimate the data density, we use +1 for the data that follow the constraints and –1 for those which don't, and we make RBF the function of $\mathrm{RBF}_{\text{ex}}(\mathbf{x})$. We would like to minimize the absolute value of $\mathrm{RBF}_{\text{ex}}(\mathbf{x})$. To make both purposes easier, we introduce the following recommendation function, following the Nash solution as:

$$\mathrm{Rmd}(\mathbf{x}) = (|\mathrm{RBF}_\Phi(\mathbf{x})|+\alpha)\times(\beta-|\mathrm{RBF}_{\text{ex}}(\mathbf{x})|), \tag{15.4}$$

where α is included to avoid the approximation error of $\mathrm{RBF}_\Phi(\mathbf{x})$, and β is slightly greater than 1. In this recommendation function, the output of $\mathrm{RBF}_{\text{ex}}(\mathbf{x})$ near the active constraints becomes close to 0, so that we can add data both around the area of lower density and surrounding the active constraints. In this paper, we will add data by maximizing the recommendation function using the method that we proposed earlier [5].

15.3 Arch Support

Medical arch support is equipment used for splayfoot patients to raise the arch of the foot and should be designed to return the arch to its normal shape. There are many different kinds of commercial arch support, and one particular example is shown in Figure 15.1.

In our previous study, we examined the effectiveness of commercial arch support for three different subjects – one patient with severe problems, one patient with mild problems, and one with a normal height arch. We asked each subject to walk on a treadmill for a while, and measured the myoelectric potential. We integrated absolute values for them and averaged them over one step and for five trials. Table 15.2 shows the results of the validation. The absolute values differ between each other, so that it does not have any meanings, but the reduction rates show the effectiveness of the support. It is obvious that the support has a postive effect on the severe patient, but a negative effect on the normal subject. There is little affect on the mild patient. From this result, we can see the importance of designing a suitable arch support for each patient.

Table 15.2 Results of effectiveness of foot support

Patient	With support		Without support	
	Big toe	Little toe	Big toe	Little toe
Severe	80.66	81.37	92.38	99.42
Not severe	81.85	55.95	82.79	68.66
None	57.43	44.63	53.90	41.56

15.4 Model of an Arch Support

15.4.1 Spline Curve

A spline curve goes through every control point, and has second-order continuity. In this study, we will use a third-order spline curve. For each interval, it follows the following equation for the interval variable $\tau = x_{i+1} - x_i \;\; (1 \le i \le 3)$:

$$S_j(\tau) = a(\tau)^3 + b(\tau)^2 + c(\tau) + d \;\; (1 \le j \le 3), \tag{15.4}$$

where Equation 15.4 has the following conditions:

- it has the same gradient for both sides of control points;
- it has the same changing ratio of gradient for both sides of the control points;
- gradient = 0 for start and end points.

15.4.2 How to Make a Foot Support Model

- Take a picture of a foot as in Figure 15.3 and fit the spline curve using control points, which are denoted by ○ in Figure 15.4. Draw line C. Choose six additional points denoted ●.
- Calculate the spline curve in the vertical direction using four points on lines A and B in Figure 15.4.
- With three spline curves A, B, and C, calculate each start and end point with line C, using two random control points from A and B. With these four points, create a vertical spline curve as shown in Figure 15.5. When four control points are not available, reduce the order of the spline curve and replace those areas.
- From the spline curves shown in Figure 15.5, calculate arbitrary control points to convert the spline curve from vertical to horizontal, as shown in Figure 15.6. As we would like to make an experimental foot support from layers of hard sponge, the shape needs to be in the horizontal plane. Four points (○) in Figure 15.7 become fixed points for the patient's arch, and the height of the other points become the design variables.

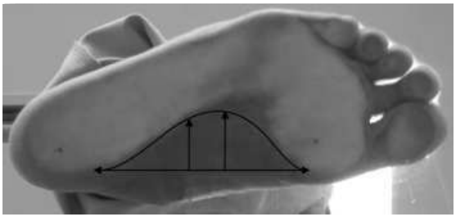


Figure 15.3 Position of the spline curve

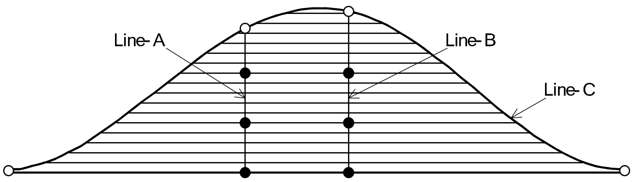


Figure 15.4 Overhead view

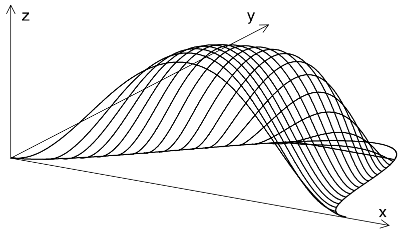


Figure 15.5 Vertical model

15.5 Design of Personal Medical Foot Support

In our previous study, we used plaster to change the height of bones of foot, and we used hard sponge to make the final foot support. Therefore, there may be some difference between experimental result and the actual response of the foot because of the difference in materials. In order to get rid of these differences, we use the same material (hard sponge) in the experiment.

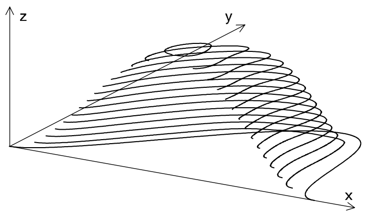


Figure 15.6 Horizontal model

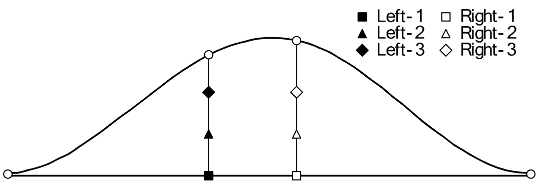


Figure 15.7 Design variable

Figure 15.8 Plaster footprint

15.5.1 Initial Pattern

We use plaster to calculate the shape of the patient's foot arch as shown in Figure 15.8. Working from the plaster model, we measure four fixed points as in Figure 15.7.

15.5.2 Ideal Position of Arch

In this study, we assumed that the Mizuno standard from Figure 15.2 and Table 15.1 shows the ideal position of the arch. Table 15.1 shows values for the vertical length from point y to OY, divided by OY in Figure 15.2. We call it % mm.

15.5.3 Detection from X-ray Digital Data

We developed an automatic calculator of each length in Table 15.2, when we indicated the points O, Y, L, N, C, m, b, f on the X-ray digital data as shown in Figure 15.9.

Figure 15.9 X-ray data analysis software

15.5.4 Procedure of the Proposed Method

1. For initial sets, give each design variable randomly.
2. Make arch supports such as shown Figure 15.10. Fit them to the patient, and take X-ray digital data. Measure OY, Ly, Ny, Cy, my, by, and fy.
3. If you are happy with the results, then end. Otherwise go to step 4.
4. Add new data to the database. Make an RBF approximation using the scalar function with:

$$\Phi = \max_{i\in\{L,N,C,m,b,f\}} w_i \frac{{}^{100iy}\!/_{\mathrm{OY}} - \mathrm{asp}_i}{1+\mathrm{asp}_i}, \tag{15.5}$$

   where asp_i is the square of the error in Table 15.2. Use the average as an arbitrary function. We will use this RBF approximation for data addition. Next, make an RBF approximation for each of six data points (Ly, Ny, Cy, my, by, and fy). Then, using the following scalar function, we will add one data point:

$$\Phi = \max_{i\in\{L,N,C,m,b,f\}} w_i \frac{{}^{100\times \mathrm{RBF}_{iy}(\mathbf{x})}\!/_{\mathrm{OY}} - \mathrm{asp}_i}{1+\mathrm{asp}_i}. \tag{15.6}$$

In this case, we had 27 initial data points, and for the initial iteration of the approximate optimization, we gained ten additional data points. We carried out the iteration three times. Table 15.3 shows the complete list of measured data. Figure 15.11 shows the best pattern of arch support.

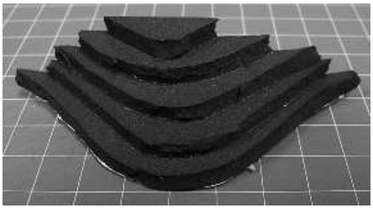


Figure 15.10 Sample pattern of arch support

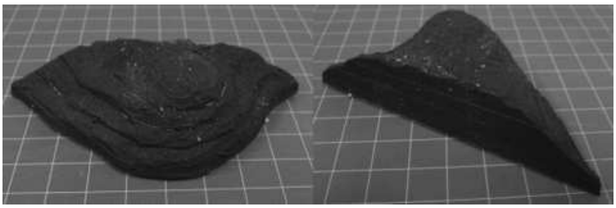


Figure 15.11 Best pattern of arch support

Table 15.3 Results of measurements of all 57 patterns

Pattern	left1	left2	left3	right1	right2	right3	F	F_Ly	F_Ny	F_Cy	F_my	F_by	F_fy
1	21.9	14.6	7.31	20.1	13.4	6.69	40.4	0.255	8.07	6.94	14.9	3.34	4.16
2	25.9	17.3	8.64	23.7	15.8	7.91	96.5	1.76	12.3	18.8	9.80	1.18	1.96
3	20.7	13.8	5.80	18.9	12.6	6.30	55.4	1.08	2.94	7.24	33.3	10.3	10.1
4	25.4	17.4	8.70	18.1	11.2	4.13	25.6	0.434	2.77	3.80	25.3	7.91	8.53
5	25.4	17.4	8.70	26.8	14.5	7.25	25.9	0.512	1.54	3.44	26.2	7.97	9.56
6	20.7	13.8	6.88	26.8	14.5	7.25	38.5	0.720	5.17	7.88	19.2	3.93	4.39
7	19.3	12.8	6.42	20.3	10.8	4.73	28.9	0.666	3.11	2.02	29.2	9.35	11.20
8	16.9	10.1	3.38	20.3	10.8	4.73	26.2	1.67	3.56	3.01	18.5	5.77	6.65
9	16.9	10.1	3.38	16.9	8.8	3.38	37.2	0.677	5.70	5.26	10.6	2.98	4.58
10	13.5	8.1	2.03	18.9	12.2	6.08	42.9	1.42	7.29	8.82	8.69	1.84	2.24
11	20.3	12.2	4.73	13.5	7.43	2.70	42.0	1.08	8.39	4.39	12.5	2.54	3.79
12	13.5	6.1	2.03	15.5	7.43	2.70	66.9	1.21	11.2	10.9	10.7	1.51	3.17
13	19.3	12.8	6.42	17.6	11.8	5.88	45.1	0.510	8.26	7.08	9.17	2.04	2.62
14	23.6	16.2	5.41	17.6	11.8	5.88	23.4	1.25	4.83	3.67	7.56	1.62	3.31
15	23.6	18.9	5.41	24.3	20.3	4.73	35.7	0.494	4.10	7.34	12.1	3.42	2.86
16	17.4	14.5	3.62	23.2	19.6	5.07	35.1	0.534	7.23	5.80	14.3	2.88	4.25
17	17.4	14.5	3.62	16.7	12.3	2.90	42.9	0.526	7.69	8.44	5.29	0.761	1.95
18	14.5	11.6	2.17	18.9	14.5	4.35	33.2	0.738	5.37	5.81	6.33	1.08	2.10
19	21.7	15.9	7.25	14.5	10.1	3.62	39.8	2.31	8.19	5.68	5.26	0.636	2.07
20	20.7	15.9	6.52	18.9	9.42	5.07	26.2	1.64	4.16	4.95	10.9	2.75	3.26
21	26.9	23.1	7.69	28.5	23.1	9.23	38.0	1.08	6.70	7.83	5.79	1.21	1.14
22	31.8	23.6	12.7	30.9	20.9	8.18	40.1	1.31	5.66	8.04	4.61	0.91	1.55
23	14.9	11.9	2.38	14.3	12.5	2.98	34.0	0.152	6.56	6.31	14.3	3.66	5.67
24	17.0	13.7	2.98	15.5	12.5	1.79	49.9	0.503	6.26	10.1	6.70	1.33	2.80
25	22.0	20.2	1.79	20.8	17.3	3.57	21.8	1.73	2.69	3.19	12.6	3.10	3.93
26	22.0	20.2	1.79	15.5	10.7	1.79	41.0	1.45	8.07	6.37	8.80	1.47	3.35
27	17.0	16.1	2.98	22.6	19.6	4.17	29.7	1.95	3.40	5.01	11.1	2.95	5.17
28	20.0	19.7	2.49	26.3	19.8	5.68	42.4	1.23	3.32	8.73	6.50	1.35	0.789
29	35.8	12.8	1.50	26.4	17.0	9.74	43.2	0.288	5.14	8.00	6.99	1.10	0.907
30	34.9	25.0	11.8	26.4	17.0	8.57	11.4	1.14	0.694	1.49	8.48	2.63	1.22
31	37.2	11.6	1.50	34.2	28.5	12.8	11.7	0.909	1.64	1.38	7.20	2.51	1.77
32	25.3	37.9	19.9	29.6	16.3	13.6	18.0	1.10	2.66	3.43	9.42	1.96	2.25
33	37.0	24.2	12.1	35.0	28.7	3.38	40.3	0.628	3.51	8.30	16.20	4.05	3.00
34	27.9	15.0	4.72	26.1	12.4	1.50	33.9	0.477	2.27	6.98	9.04	2.32	1.79
35	33.8	43.1	4.60	36.9	19.3	3.71	16.2	1.15	0.531	2.49	4.28	0.819	0.840
36	35.1	18.1	7.75	27.3	17.8	13.9	23.1	3.78	0.305	1.90	10.3	2.60	2.45
37	36.9	18.4	8.71	34.9	29.0	9.24	46.3	0.691	6.10	8.40	9.16	1.04	1.25
38(C)	28.3	16.9	8.33	17.1	11.0	5.85	85.0	0.852	11.9	16.0	11.4	1.42	1.26
39	15.6	25.9	7.53	27.0	16.9	8.94	50.6	0.339	5.28	10.1	10.0	2.10	1.35
40	39.8	24.9	18.3	26.3	16.2	8.55	28.8	1.36	1.44	5.38	9.37	2.75	1.90
41	15.4	11.2	3.07	34.9	28.7	13.0	44.0	1.08	6.31	6.50	10.3	2.57	1.84
42	31.7	12.7	2.93	35.3	27.8	13.7	33.2	0.356	2.70	6.84	3.33	0.690	0.414
43	37.5	43.0	10.8	36.4	19.3	3.86	24.0	2.18	1.12	3.78	5.19	1.60	0.96
44	23.7	43.0	3.38	36.7	19.4	3.77	23.5	2.68	0.28	3.69	3.81	1.18	1.40
45	33.6	12.0	7.36	33.9	28.8	12.5	30.4	1.93	3.42	6.27	5.65	1.47	1.20
46	30.2	42.7	16.5	37.5	19.1	3.70	20.7	0.993	0.670	4.28	0.950	0.261	0.063
47	10.8	13.0	11.8	35.0	28.8	13.0	80.3	0.921	8.90	16.50	6.81	1.23	0.770
48	22.1	16.3	5.63	17.1	11.3	5.68	28.5	0.282	2.78	5.65	7.54	1.18	0.950
49	21.9	17.8	14.4	19.6	11.0	4.26	9.05	0.417	1.12	1.63	4.04	1.05	0.252
50	22.5	17.8	18.9	18.3	11.0	4.42	10.1	1.50	0.340	0.877	4.52	1.24	1.52
51	21.6	16.8	23.9	19.8	11.4	4.49	13.1	0.977	1.11	1.20	13.00	4.02	2.95
52	32.8	43.4	17.1	34.8	18.4	4.05	8.94	1.83	0.0582	0.378	0.909	0.249	0.293
53(best)	22.4	20.4	23.9	20.3	16.7	3.03	4.61	0.462	0.344	0.529	4.68	0.821	1.54
54(A)	27.3	17.3	8.97	29.1	10.7	4.19	10.7	0.143	1.38	2.22	1.73	0.116	0.084
55	25.8	17.4	6.03	36.9	11.2	4.13	28.4	0.659	2.03	5.67	4.16	0.810	1.05
56(B)	34.6	24.2	23.2	25.7	17.5	8.46	38.7	0.033	5.49	6.07	4.59	0.646	0.582
57	22.0	21.5	15.2	18.8	17.6	3.93	57.4	0.417	5.18	11.10	4.74	0.540	0.661

15.6 Validation Test of Personal Medical Foot Support

To validate the model, we made three patterns of arch supports (marked A, B, and C in Table 15.3).

15.6.1 Myoelectric Potential Measurement

The myoelectric potential measurement records the electricity caused by moving a muscle. In the experiments, we used three pairs of electrodes to measure the movement of the calf, big toe, and little toe. In this study, we asked patients to walk on running machines with a speed of 5 km/h, and a slope of 4%.

15.6.2 Experiment

In this validation, a splayfoot patient and a normal person are selected. We prepared a commercial foot support and four kinds of foot supports (A, B, C and Best), as shown in Figures 15.11–15.14. Their data are shown in Table 15.3. To analyze the myoelectric potential measurements, we followed the steps below:

1. Use a low-pass filter to get rid of the landing noise.
2. Calculate a moving average of about four steps.
3. Calculate the absolute moving average difference for each value.
4. Average each value in step 3.
5. Carry out steps 1 to 4 three times and average the results.

The results of the experiment are shown in Table 15.4. Each figure is the amount of movement over 100 μs. It differs from Table 15.2. Therefore, this time we use the moving average as standard following doctors' advice.

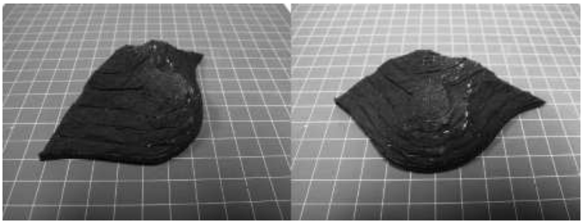


Figure 15.12 Foot support – pattern A

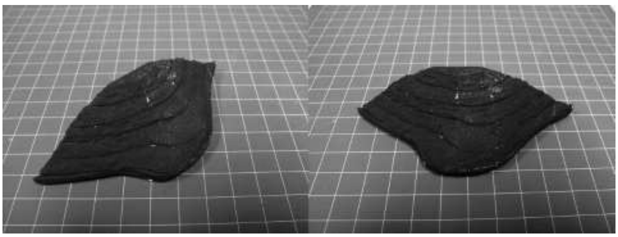


Figure 15.13 Foot support – pattern B

Figure 15.14 Foot support – pattern C

Table 15.4 Results of validation

Pattern	Splay foot			Normal		
	Calf	Big toe	Small toe	Calf	Big toe	Small toe
Bare foot	27.4	96.1	86.3	69.4	83.1	89.0
Comercial	26.9	98.5	84.4	70.9	82.7	80.5
Best	25.0	93.3	80.8	71.6	87.2	80.4
A	25.1	104.0	89.4	71.0	86.7	78.0
B	25.1	99.4	79.4	75.6	86.5	76.5
C	25.3	98.8	86.7	73.6	84.1	81.0

15.6.3 Results

Looking at Table 15.4, we can see that all four foot supports cause a reduction in calf muscle use for the splayfoot patient. However, support patterns A, B, and C have a negative effect, particularly on the big toe. The optimum support is the only support that can simultaneously reduce all three areas of muscle use. Looking at Table 15.3, support pattern A shows better error results for small toes; however, the response on that side is not good at all. It seems that error for the big toe will play a more important role in reducing muscle use. Looking at the results for the normal person, even the commercial support had a negative effect on him, and all the patterns for splayfoot patients show an even worse effect. This result is quite rational, as a foot support should be made personally for each patient if we want to reduce the pressure of muscle use.

15.7 Conclusions

In this chapter, we have used approximate multi-objective optimization using convolute RBF networks in designing made-to-order medical foot supports for splayfoot patients. If we can design the support in the correct way, we can reduce the pressure on bones and muscles for the patients. Of course, it is impossible to make any general models of human beings because we need to identify too many parameters for each individual, and also the measurement must be done in non-destructive way. In this study, we used the Mizuno standard position of bones, and tried to minimize the distance of six positions of bone from them, which gives us six objective functions. Under this condition, we need to carry out approximate multi-objective optimization. In addition, we would like to reduce the number of X-ray inspections, both because of the cost and because of the negative effect on patients. In a previous study, we succeeded in designing a vertical shape using two-dimensional shape optimization. In this study, we extended the model to three-dimensional shape optimization and introduced a 3D spline expression with six control points, and formed a six-variable, six-objective-function optimization problem. Starting from 27 initial data points, we added 30 additional data points to obtain satisfactory results. For validation, we used myoelectronic potential measurement, to show the amount of muscle use. We prepared four different patterns of foot support, with one commercial support. From these validation tests, we have shown that only the best solution can simultaneously reduce all three areas of muscle use – and all four supports have a negative effect on a normal person. These results show the capability of approximate multi-objective optimization: a convolute RBF approximation can calculate an approximate function correctly even in a human model which might have a lot of nonlinearity, and the data addition rule is effective so we only need 57 experiments. Moreover, we have optimized according to the Mizuno standard, but results show that it has a positive effect on the reduction of the pressure on muscles. This means that the effectiveness of the Mizuno standard has been proven from a medical point of view.

References

1. Miyake S., Arakawa M.: Approximate multi-objective optimization of medical foot support. In: Proceedings of IDETC, ASME, New York (2008)
2. Mizuno S.: Study on human foot: extension from flatfoot. Ishiyakushuppan (in Japanese) (1973)
3. Arakawa M.: Development of convolute approximation of radial basis network for sequential approximation multi-objective optimization. Trans. JSME **75**(765C):2141–2147 (in Japanese) (2009)
4. Nakayama H., Tanino T.: Theory and application of multi-objective programming. Keisoku-jidouseigyogakkai (in Japanese) (1994)
5. Arakawa M., Miyashita T., Ishikawa H.: Development of genetic range genetic algorithms (3rd report, proposal of fitness function to search multiple local optima). Trans. JSME **70**(695C):2062–2069 (2004)

Chapter 16
Kansei Quality in Product Design

Hideyoshi Yanagisawa[1]

Abstract Kansei quality is a product quality that evokes the customer's specific impressions, feelings, and emotions – such as comfort, luxury, delight, *etc.* It is important to consider kansei quality especially when designing mature consumer products. One of the most important issues in the design of kansei quality is to construct the evaluation criteria of the target kansei qualities. Without such criteria, the designer has to rely only on his/her subjective criteria for designing and evaluating kansei quality. We often express such a quality using adjectives that are subjective and often ambiguous. However, the meaning and sensitivity of adjectives may differ from person to person. Moreover, we possess latent evaluation criteria that are often evoked by new experiences, such as seeing a new product that provides a new kansei quality. In this chapter, the author presents a method to construct evaluation criteria for kansei quality, taking into consideration the diverse and latent kansei of customers. As a case study, the author applies the method to the design of a product's sound quality.

16.1 What Is Kansei Quality?

Customers' needs towards consumer products become diverse in a mature market where many products have similar functionalities and performance. With such diversification, the customer has come to focus more on the emotional and sensuous quality of a product.

A *kansei quality* is a product's quality that evokes the customer's specific impressions, feelings, or emotions towards a product (*e.g.*, comfort, luxury, delight) [1]. *Kansei* is originally a Japanese word that refers to the sensitivity of a human

[1] H. Yanagisawa
The University of Tokyo, 7-3-1, Hongo, Bunkyo, Tokyo, Japan
e-mail: hide@design.t.u-tokyo.ac.jp

sensory organ at which sensations or perceptions take place in response to stimuli (*e. g.*, a product) from the external world. Kansei includes evoked senses, feelings, emotions, and impressions. The word kansei has begun to be used internationally because there is no suitable translation in English.

According to the psychologist A. Maslow, human needs shift from common physiological needs to personal psychological needs [2]. This view suggests that sensuous needs are likely to increase further in the future. To respond to such continuously diversifying sensuous needs, it has become important to determine the customer's effective needs and to design according to those needs. It is, however, difficult to grasp such sensuous needs as compared to other needs because sensuous qualities are subjective, not easy to externalize, and often latent.

The most important issue in the design of kansei quality is extracting quantitative criteria for evaluating such a quality. Without such quantitative criteria, the designer cannot set a clear goal for design because he/she is not sure what kind of evaluation criteria the customers have, or how to design a product to increase a target kansei quality. The designer has to rely on his/her assumptions based on sensitivity and tacit knowledge. If there is a gap between the designer and customers in terms of their sensitivities, the designer may misinterpret how a customer will evaluate his/her design.

16.2 Approaches Towards Kansei Quality in Product Design

Kansei qualities are often represented by subjective words, such as adjectives. Several methods employing sensory tests have been used to evaluate a product's kansei quality represented using words. The *semantic differential method* (SD method) [3], which is widely used, uses pairs of opposite words to evaluate the kansei quality of evaluation samples. The subjects score the degree of their impression towards product samples according to five to seven ranks between the word pairs. To measure the emotional score more precisely, pair-wise comparison [4] is often used. Two samples are randomly selected and the subject scores one sample by comparing it with the other sample in terms of a specific kansei quality. Although this method enables the precise measurement of kansei qualities, the number of trials increases exponentially with the number of samples because of the number of possible combinations.

Most approaches to sensitivity quantification aim to make generalizations of sensitivity by averaging the subjects' evaluations. A mapping between the averaged kansei score and measurable design attributes is created. Several mapping methods have been developed and used, such as multi regression analysis, fuzzy reasoning, and neural networks [5]. *Reduct* in rough set theory [6, 7] has recently been noticed as a method of knowledge acquisition that is useful for design [8].

The constructed mapping serves as a metric of general kansei qualities. However, by nature, human sensitivity differs from person to person. Highly subjective perceptions that are directly related to product value, such as pleasantness and preference, are particularly highly individualistic. Few researches deal with the individuality of sensitivity. Nakamori applies fuzzy set theory to represent sensitivity individuality [9]. In his method, the degree of individuality is represented by the fuzziness of a fuzzy set whose center is an average value. This approach regards individuality as errors from the average value. Causes of personal differences cannot be explained using this method.

Yanagisawa and Fukuda found semantic differences in emotional words between designer and customer using the SD method and *principal component analysis* (PCA) [10]. An emotional word contains multiple scales of semantics that vary from person to person. There is little point in averaging between scales that have different meanings. An averaged kansei quality cannot be used to represent diverse sensitivities.

There is another approach in which the individual customer is assisted to evolve design samples to meet his/her psychological satisfaction through the interaction of analysis and synthesis. A customer's evaluations of provided design samples generate new design samples, allowing recursive refinement. However, from the standpoint of the customer, it is desirable to keep questionnaires to a minimum – *interactive evolutional computation* (IEC) [11] is one solution to this problem. Human evaluation is regarded as the fitting function of an evolutional computation (EC) such as a genetic algorithm (GA), where the design parameters are coded as chromosomes. Users evaluate the design samples generated by the EC, and then the EC generates the next generation's design samples. Users only need to evaluate the design samples until they arrive at a satisfying result. There are some applications of IEC for design support systems [12, 13]. In this approach, analysis and synthesis interact with each other for short periods. Because human preferences and sensitivities change as they are influenced by the design samples, this scheme is suitable for supporting personal design.

16.3 Towards Diverse and Latent Kansei

In this section, we focus on two characteristics of human sensitivity that are important to consider when designing kansei quality. The first characteristic is the *diversity* of human sensitivity. Human sensitivity towards kansei quality is different from person to person. In other words, individual differences of sensitivity exist. Although we share common senses for some basic kansei qualities, they do not always cover all of the kansei qualities that relate to a product's value.

When we say "individual differences" of sensitivity, we must be careful to note that there are two kinds of individual differences: *variation* and *diversification*. *Variation* is an individual difference that can be measured by a unique scale (see Figure 16.1 (a)). For example, the sense of heaviness in a sound may be slightly different from person to person but the meaning of heaviness should be universal. We can apply statistical operations such as averaging to such data. We can use the standard deviation as an indicator of the individual difference. Most conventional approaches to individual differences of sensitivity deal with this type of difference. In other words, individuality is regarded as an error from the average value.

On the other hand, *diversification* represents individual differences of the scales themselves (see Figure 16.1 (b)). For example, the sense of beauty may have multiple different scales depending on personal viewpoints. We cannot use statistical operations between different scales, such as averaging between length and weight. Higher and more complex kansei qualities tend to have this diversification. These kansei qualities tend to directly relate to product value.

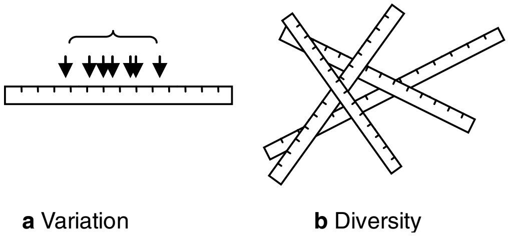


Figure 16.1 Difference between (a) variation and (b) diversity

The second characteristic of human sensitivity is the *latent* sensitivities that we potentially possess – although we are not consciously aware of them – which are evoked by new experiences. Most conventional approaches construct the evaluation criteria for kansei qualities using existing products on the market as evaluation samples, and their sensory evaluations are scored by human subjects. The constructed evaluation criteria are then used to evaluate new designs. However, existing products are not exhaustive enough in the design space to construct general evaluation criteria that can be used for future design. People may evoke different latent sensitivities towards an unknown design sample. In other words, the constructed evaluation criteria may not be applicable for evaluating and designing new designs.

16.4 A Method for Extraction of Diverse and Latent Evaluation Criteria of Kansei Quality

This section presents a new method on how to extract and formalize *diverse* and *latent* evaluation criteria toward kansei quality [14, 15]. The method consists of two sensory tests. The first sensory test uses evaluation samples of existing products; the second uses both composite samples and existing samples in order to extract latent evaluation criteria. Composite samples are created in the design feature areas untouched by existing products by modifying the features toward those directions that increase the target kansei quality based on multiple criteria obtained from an analysis of the results of the first sensory test. The method used is as follows:

1. Design samples were prepared from existing products. Using the samples, we conducted a sensory test based on the SD method. In the test, multiple subjects gave their impressions of the samples using pairs of opposing adjectives, called SD scales.
2. Next, we extracted the design feature values from each sample.
3. From the results of the first sensory test, we analyzed the multiplicity of each SD scale, which are different from person to person, and extracted patterns of subjective scales considering the diversity of personal sensitivities. The patterns were extracted using cluster analysis, based on the correlation coefficients between subjects for each SD scale that represent similarities of sensitivity. We formulated each extracted scale using the design features and interpreted the semantics of each subjective scale. (Details of this process are given in Section 16.4.1.)
4. Based on the formulated scales, we set feature values that are used to synthesize composite design samples. We selected an SD scale as the target kansei quality and set feature values so that they are dispersed on the scale. To extract evaluation criteria that can be used for future design, the design feature values are required to cover areas of the feature space untouched by existing products. We synthesized composite design samples to fit the set feature values by modifying the original samples of existing products.
5. A second sensory test was conducted using both the created samples and existing ones in the same manner as the first sensory test. For the SD scales, we added new SD scales or deleted old ones from the first sensory test based on their contributions.
6. We extracted and formulated multiple scales from the result.
7. To check the repeatability of the scale, the results of the first and second sensory tests were compared in terms of the SD scales commonly used in both experiments. We then analyzed the changes in kansei quality due to the addition of the new composite design samples. Finally, we extracted potential factors of the kansei evaluation criteria for designing new samples and applied factor analysis using the multiple scales obtained from the results of the second sensory test.

16.4.1 Extraction of Multiple Kansei Scales Considering Diverse Kansei Qualities

As discussed in Section 16.3, we assume that the semantics and sensitivities for SD scales (*i. e.*, pairs of adjectives), which are used as evaluation scales in the SD method-based sensory test, are diverse. In other words, subjects have *different potential scales* for each SD scale. We extracted patterns of potential scales for each adjective based on the similarity of scores obtained from the first sensory test. Figure 16.2 shows how to extract them. We used correlation coefficients between scores given by different subjects as an indicator of their similarity. It is assumed that two subjects have similar psychological scales if their respective scores for the samples are similar. To break down a scale, we apply *cluster analysis* using the correlation coefficient-based distance.

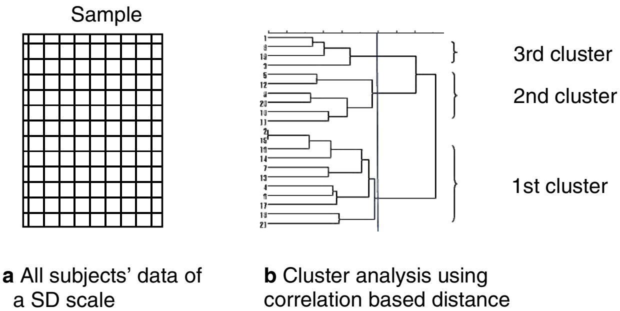


Figure 16.2 How to extract patters of personal evaluation criteria from an SD scale: (a) all subjects' data from an SD scale, and (b) cluster analysis using a correlation-based distance

First, select N_p products S_i ($i = 1, 2, \ldots, N_p$) as evaluation samples of the sensory test. N_s subjects T_j ($j = 1, 2, \ldots, N_s$) evaluate their impressions of the samples S_i using the SD method with N_w pairs of adjectives (SD scales) I_k ($k = 1, 2, \ldots, N_w$). Let $\boldsymbol{E}_{jk} = < e_{ijk} >$ be a vector of scores given by the jth subject for all evaluation samples in terms of I_k. If the pth subject has a similar sensitivity to the qth subject in the SD scale I_k, the correlation coefficient of E_{pk} and E_{qk} should be close to 1.0. If the pth subject has the opposite sensitivity to the qth subject for I_k, the correlation coefficient should be close to –1.0. We define the distance between the pth and qth subjects in terms of the sensitivity of an SD word I_k as follows:

$$d_{kpq} = 1 - r(\boldsymbol{E}_{pk}, \boldsymbol{E}_{qk}) \quad (p \neq q), \tag{16.1}$$

where d_{kpq} is the distance and $r(\boldsymbol{a}, \boldsymbol{b})$ denotes the correlation coefficient between vectors $\boldsymbol{a}$ and $\boldsymbol{b}$.

For each SD scale, we next classify all subjects into clusters using the distance and cluster analysis; members of each cluster have similar sensitivities for that SD scale. The obtained clusters reflect a division (*i. e.*, breakdown) of the SD scale, and represent the multiple viewpoints or sensitivities of that SD scale. We derive the threshold value of the distance for cluster formation, where members of each cluster are not statistically different from each other, in terms of the sensitivity of the SD scale with significance level α.

Kansei qualities comprise a hierarchical structure as shown in Figure 16.3 [16]. The lower level of the hierarchy more directly reflects human perceptions of external stimuli. The higher level consists of more subjective kansei, such as subjective impressions (*e. g.*, beauty) and preferences. It is assumed that individuality increases as one goes higher up the hierarchical level, as shown in Figure 16.3.

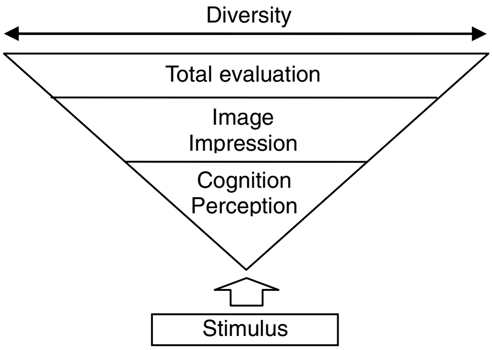


Figure 16.3 Hierarchy and diversity of emotional response

If a SD scale cannot be broken down into smaller groups by this process, which means that all subjects give statistically similar values for all evaluation samples, then that kansei is considered to be low in the kansei hierarchy. If a kansei scale can be broken down into many clusters and each cluster consists of a small number of subjects, that kansei is high in the hierarchy. We use the relative proportion of subjects that belong to a divided (*i. e.*, broken down) kansei cluster as an indicator of its *commonality*. This commonality can be used to decide the hierarchical order of each kansei word.

We define the commonality of an emotional word for each cluster as follows:

$$\mathrm{Com}_l = \#T_c \,/\, \#T \times 100, \tag{16.2}$$

where $\#T$ denotes the number of subjects and $\#T_c$ denotes the number of subjects included in the lth cluster C_l. The commonality is equal to 100 if all subjects have a similar (*i. e.*, statistically the same) SD scale sensitivity.

We use the average value of each cluster as a central value of the cluster. The central value of the lth cluster for the ith SD scale and kth evaluation sample is calculated as follows:

$$y = \textstyle\sum_{j \in Cl} e_{ijk} \,/\, \#T_c. \tag{16.3}$$

To interpret the semantics for each cluster scale, we use correlation coefficients with the cluster scale value of other words whose commonality is higher (*i. e.*, low-level kansei quality).

16.4.2 Strategies of Setting Design Feature Values

There are two strategies for creating composite design samples, as follows: (1) cover areas of the design feature space *untouched by existing product*, and (2) disperse features on a target kansei quality considering its *diversity*. These are described in this section.

1. Finding a Design Feature Area Unexplored by Existing Products

The first strategy is to create design samples in a feature area where no existing sample appears. If the number of features is two, we can visualize the mapping of existing products to find such an area. However, if the number of dimensions of features is more than three, it is difficult to find such areas visually. To reduce the number of dimensions, we apply *principal component analysis* (PCA). PCA reduces a multidimensional space to a lower-dimensional space (2D or 3D) while retaining as much information as possible.

We extract N_f design features $P=<p_1, p_2, \ldots, p_{Nf}>$ for the evaluation samples of the sensory test. The ith principal component F_i is obtained as follows:

$$F_i = W_i'P\,, \tag{16.4}$$

where $W_i = < w_1, w_2, \ldots, w_{Np} >$ is a principal component loading. The obtained principal components are orthogonal to each other. The variance of a principal component denotes the degree to which the principal component explains the original data P. We use the two top principal components in terms of their variances to visualize the mapping of the evaluation samples. The 2D scatter graph using the selected components allows us to visually find areas unexplored by existing designs in the feature space.

2. Setting Feature Values Dispersed on the Target Kansei Quality

We set a *target kansei quality* that the designer aims to increase for future design. We create composite samples by dispersing this quality in unexplored feature areas. By analyzing the relation between such samples and their emotional responses, we can extract latent factors for such a quality.

To set such design features, we formalize a target kansei quality using the results of the first sensory test. Assume we obtain N_c clusters from the target kansei quality. We can apply *multiple regression analysis* (MRA) to formalize the target kansei quality for each cluster, using extracted design features as explanatory variables. The central value of the lth cluster $Y_l = < y_{l1}, y_{l2}, \ldots, y_{lNs} >$ is estimated as follows:

$$Y_l = A_l'F + \beta\,, \tag{16.5}$$

where $A_l = < a_{l1}, a_{l2} >$ denotes the weight vector, F is the principal component vector, and β is the error vector. A_l represents a direction in feature space for creating design samples. This direction has the potential that composite samples will be dispersed in terms of the target kansei quality.

16.5 Case Study: Product Sound Quality

The sound made by a product is an important factor that affects the product's kansei quality. For quite a long time, sound engineering dealt mainly with the reduction of the overall sound pressure level (SPL) emitted by a product. Within the last decade, however, the focus has started to switch more towards aspects related to the sound quality. The biggest change is that the design goal switched from objective values, such as “decibel” levels that can be physically measured, to subjective ones, such as kansei qualities. To design for kansei qualities of product sound, it is necessary to develop metrics to quantitatively evaluate such subjective qualities. Zwicker *et al.* [17] developed sound quality metrics (SQM) as an evaluation metric of the product sound quality. SQM provides values for simple perceptions of sound, such as loudness and sharpness. However, the kansei qualities of a product’s sound include more complex affective perceptions, such as pleasant, annoying, luxurious, *etc.* To deal with such complex sensitivities in sound design, most conventional approaches conduct sensory tests using affect-laden words to score target kansei qualities. Statistical methods are used to compose a map between SQM and complex kansei qualities [18]. Several applications have been studied based on this approach [19–22]. Most research so far, however, has not considered the diversity and potentiality of human sensitivity.

16.5.1 Sound Quality Metric as Design Parameters

As measurable design parameters of product sounds, we use four basic SQM [17]: loudness, sharpness, roughness, and fluctuation strength. These are widely used and well defined. Recent studies have demonstrated that these metrics are independent of the meaning of a sound [21, 22]. We extracted these four SQM from the stationary sounds of ten products.

Loudness

Loudness is a perceptual measure of the effect of the energy content of sound in the ear. It is related to the decibel level and also depends on the frequency content of a sound. For example, a very low-frequency sound such as a 20 Hz tone at 40 dB would be perceived to be quieter than a 1 kHz tone at 40 dB. The loudness level of a sound is defined as the sound pressure level of a 1 kHz tone in a plane wave and frontal incident that is as loud as the sound; its unit is the “phon”. ISO 226 constructs equal loudness contours using data from 12 references [23].

Third-octave bands can be used as an approximation to critical bandwidths, which is a measure of the frequency resolution of the ear [24]. A specific loudness can be calculated from the decibel level for each third-octave band. The value of loudness (N) is calculated as the integral of the value of the specific loudness (N'). The unit of loudness is the “sone”. One sone equals 40 phons, which is the loudness of a 1 kHz tone at 40 dB in a plane wave.

Sharpness

Sharpness is a measure of the high-frequency content of a sound; the greater the proportion of high frequencies, the “sharper” the sound. Using Zwicker and Fastl’s approach [17], sharpness is calculated as the weighted first moment of the specific loudness. The unit of sharpness is the “acum”.

Roughness
People perceive a rapid amplitude modulation around 70 Hz of a sound as “rough”. The unit of roughness is the “asper”. One asper is defined as the roughness produced by a 1000 Hz tone of 60 dB, which is 100% amplitude modulated at 70 Hz.

Fluctuation strength
Fluctuation strength is similar in principle to roughness except it quantifies subjective perception of slower (up to 20 Hz) amplitude modulation of a sound. Maximal values are found to occur at a modulation frequency of 4 Hz. The unit of fluctuation strength is the “vacil”. One vacil is defined as the fluctuation strength produced by a 1000 Hz tone of 60 dB, which is 100% amplitude modulated at 4 Hz.

16.5.2 Sensory Test Using Existing Samples (First Experiment)

We first carried out an impression evaluation experiment based on the SD method with 21 subjects. We recorded the stationary sounds from ten selected products of different makers in an anechoic chamber and used them as evaluation samples.

Table 16.1 SD words used with SD method-based sensory test

No.	Pair of SD words
1	hard–soft
2	dull–clear
3	silent–noisy
4	square–round
5	opaque–limpid
6	weak–strong
7	discomposed–composed
8	ugly–beautiful
9	static–dynamic
10	cheerful–gloomy
11	poor–rich
12	small–big
13	high–low
14	dislike–like
15	untypical–typical (sounds like the machine)
16	unentertaining–entertaining
17	cheap–expensive
18	effective–not effective (How good a job the machine is doing)
19	elegant–inelegant
20	agreeable–annoying
21	western–Japanese
22	unpleasant–pleasant

We selected 22 pairs of adjectives related to the target product sounds, as shown in Table 16.1. These words contain different levels of the kansei hierarchy. For example, “loud–silent” is a perceptional level kansei (low level) and “like–dislike” is a preference level kansei (high level).

We divided the subjects into four groups of five people each. The five subjects sat in front of a speaker. Each sound was played for 5 s and the subjects gave their impressions of the sounds by filling out a questionnaire consisting of word pairs. Two trials of the same experiment were conducted in order to test the reliability of the data. To avoid the influence of the learning curve, the subjects practiced responding before conducting the experiment.

16.5.3 Extracting Patterns of Personal Kansei Scales

We divided each SD scale using cluster analysis with the correlation-based distance discussed in Section 16.4.1.

Figure 16.4 shows examples of the clusters demonstrated in 2D space using *multi-dimensional scaling* (MDS) [25]. MDS is often used to compose a 2D space using only distances among samples. Each point represents a subject. The coordinate system does not have any meaning but the distances between points correspond to the correlation-based distance.

The scale of “big–small” comprises one cluster (Figure 16.4 (a)). This means that all subjects evaluated all evaluation samples with similar scores on the “big–small” scale. In other words, we can say that the “big–small” scale of a product sound is a common scale where all subjects perceive in a similar manner. For such a SD scale, the conventional approach can be used where the average value is the representative value.

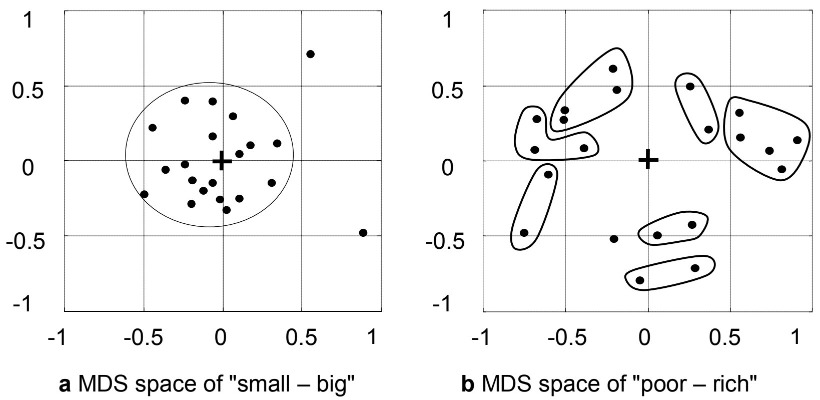


Figure 16.4 Examples of SD scales divided by correlation-based distance and cluster analysis. Each point represents a subject. The 2D space is composed by MDS: (a) MDS space of “small–big”, and (b) MDS space of “poor–rich”

The scale of "poor–rich" contains six clusters, indicating six different sensitivities (Figure 16.4 (b)). For such subjective SD scales containing different sensitivities, the average value should not be taken as a representative value because the average value might not be chosen by anyone. In fact, there are no data around the average value, which is the zero point of the space in Figure 16.4 (b). Thus, the average value of "poor–rich" represents nobody's kansei.

Figure 16.5 shows the proportions of subjects included in each cluster for each SD scale. These proportions denote the commonalities of the divided SD scales. We discard clusters having only one subject because they represent extremely personal evaluation criteria.

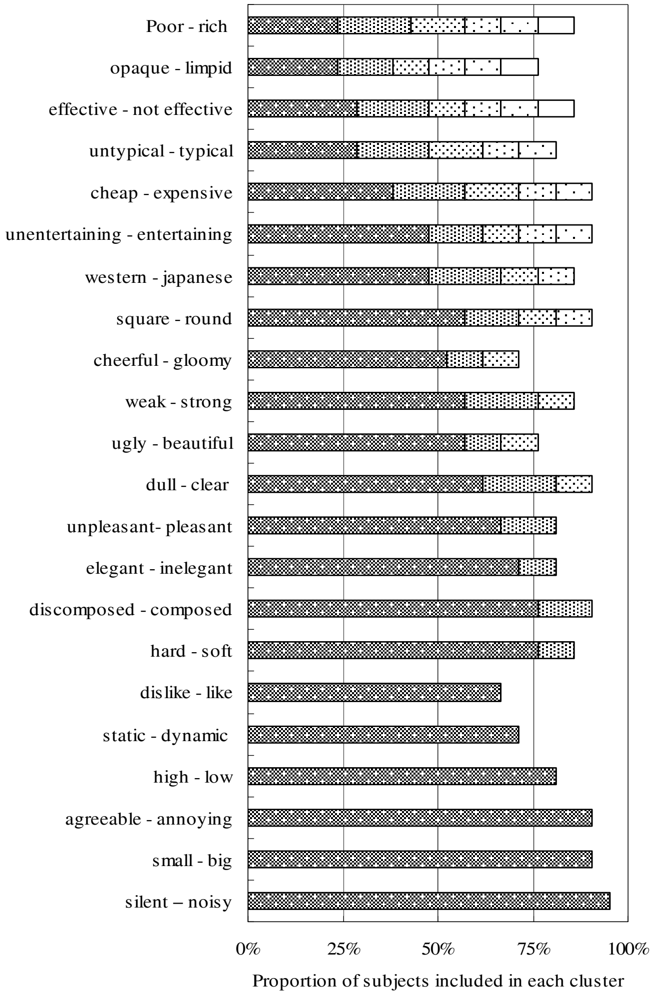


Figure 16.5 Proportion of subjects included in each cluster for all SD scales

The scales in Figure 16.5 are sorted by the number of divisions and commonality. For low-level perceptual scales, such as “silent–noisy”, “small–big”, and “low–high”, all subjects are grouped into one cluster. In other words, all subjects graded each sound sample in a similar manner and displayed similar sensitivity.

On the other hand, high-level impression scales, such as “cheap–expensive”, “effective–not effective”, and “poor–rich”, are divided into many clusters. Such highly subjective words contain multiple viewpoints of its cognition. These results suggest that commonality can be used as an indicator of the hierarchical order.

Some highly subjective scales, such as “like–dislike” and “pleasant–unpleasant”, even though they are assumed to contain different viewpoints that vary with individual subjects, are exceptions to the above rule, and are divided into only a few clusters. This result suggests that the commonality of highly subjective SD scales varies depending on the target design. For example, “like–dislike” is statistically similar among the subjects for the machine sound used in the experiment, although we can assume that the preference differs from person to person. Using the clustering method and commonality, the designer can extract such particular instances of divergent subjective scale characteristics.

16.5.4 Multiple Scales of a Target Kansei Quality

We use the SD scale “expensive–cheap” as a *target* kansei quality because we, as designers, believe that a luxurious sound increases a product’s emotional value. The target kansei quality is an SD scale that the designer temporarily sets as the design concept of the machine sound.

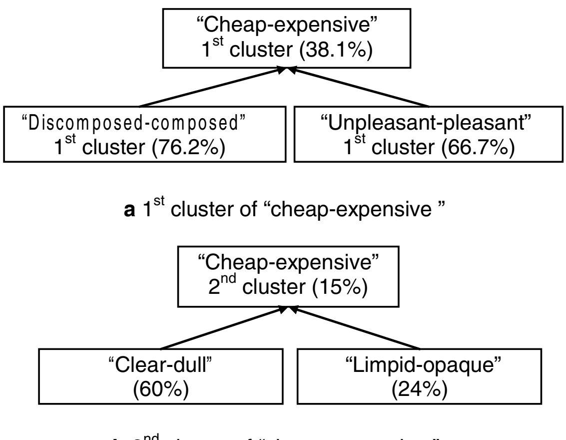


Figure 16.6 Semantic differences among multiple viewpoints of the target kansei quality “cheap–expensive” scale explained by correlations with other words whose commonality is higher: (a) first cluster of “cheap–expensive”, and (b) second cluster of “cheap–expensive”

The target kansei quality “expensive–cheap” contains five clusters, *i. e.*, five types of evaluation scales. Figure 16.6 shows an example of interpreting two different viewpoints of “cheap–expensive” using cluster scales of other words whose commonality is higher. The first cluster scale is related to the first clusters of “pleasant–unpleasant” and “composed–discomposed”, *i. e.*, the subjects who adopt the first scale perceive the “expensiveness” of the machine sound from the viewpoints of pleasantness and composedness. The second scale is related to the first clusters of “clear–dull” and “limpid–opaque”. The second scale is different from the first scale in terms of these semantics.

To consider the multiplicity of the definition of the target SD scale, we used the above two scales to set the design feature values of composite sounds.

16.5.5 Unexplored Design Area in Feature Space

We used SQM as the design features. To find untouched areas in the SQM space, we constructed a two-dimensional space using PCA. Figure 16.7 shows the result of PCA. The areas where no data appear are the untouched areas.

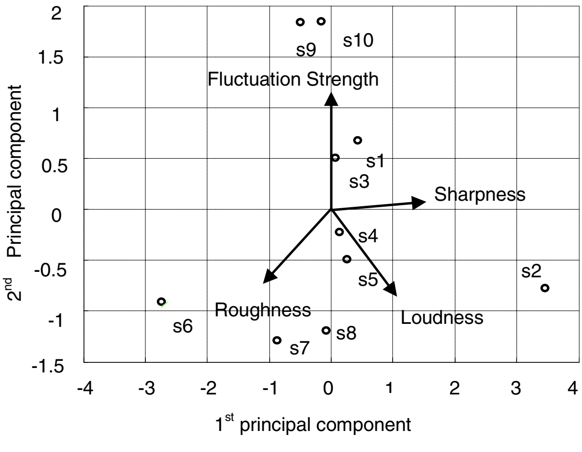


Figure 16.7 SQM space in low dimension constructed using PCA

16.5.6 Gradient of a Target Kansei Quality in Feature Space

To set design feature values in areas untouched by existing products, we need to establish the local gradients towards the target kansei quality along boundary areas between touched and untouched areas. If the target kansei quality has a nonlinear relationship with SQM, the regression plane derived from MRA using all data of existing products will not correspond to the gradient in the vicinity of the boundary areas. Furthermore, to cover multiple untouched areas, we should obtain multiple directions for setting features for creating composite sounds.

For the above reasons, we split the SQM space into several subspaces, each with the same number of sounds for existing products, and conducted an MRA for each split space.

Figure 16.8 shows regression planes of the first scale for "expensive–cheap" for each divided local area. The regression planes are represented in contour. The numbers on the contour lines denote the estimated value of the SD scale. According to the result, the gradients of all areas face in the same direction. The direction towards the upper left area is estimated as high in terms of the scale. Loudness and sharpness negatively relate to the scale. Meanwhile, we found three directions to increase the expensiveness of the sound in the second scale, as shown in Figure 16.9.

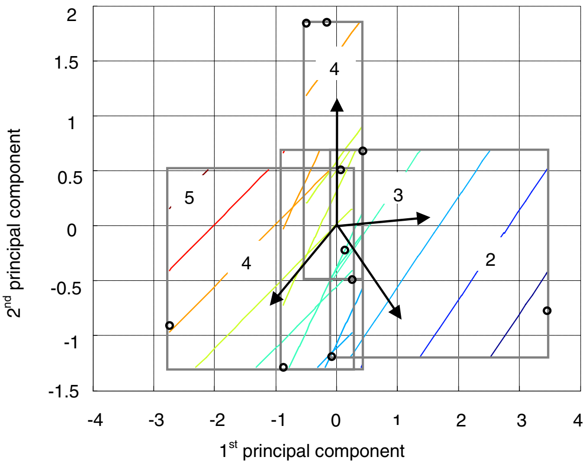


Figure 16.8 Example of local-regression surfaces of first cluster scale of "expensive–cheap"

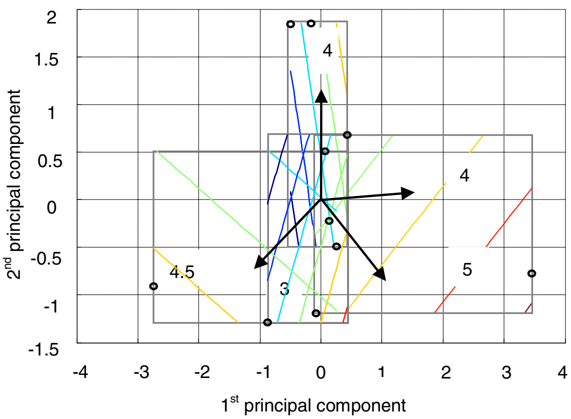


Figure 16.9 Example of local-regression surfaces of second cluster scale of "expensive–cheap"

16.5.7 Setting Feature Values and Creating New Sounds

By considering the obtained directions that increase the target kansei quality from the two major points of view and the untouched areas, we set the SQM values for creating sounds. We selected six original sounds from existing products and synthesized them so that they satisfied the above two conditions. The strategy of synthesizing sounds is based on increasing or decreasing the SQM values of the original sounds. We created 18 sounds as shown in Figure 16.10.

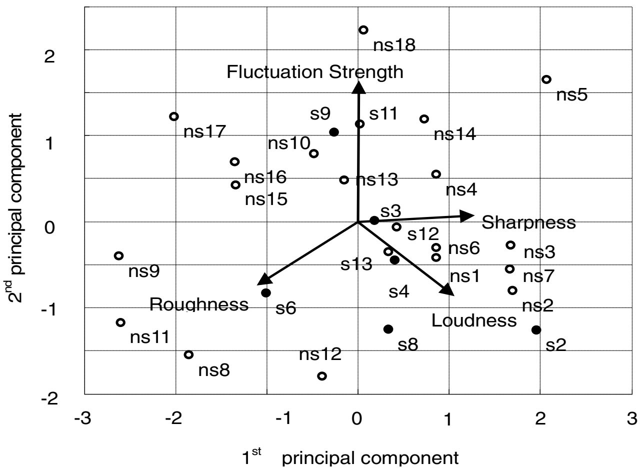


Figure 16.10 Original and composite sounds in SQM space

16.5.8 Sensory Test Using Composite Samples (Second Sensory Test)

We conducted the second sensory test using the created sounds and the original sounds. The purpose of the test is to find a potential evaluation factor that is effective for designing new sounds of a product. We used 18 created sounds and eight sounds of existing products as evaluation samples, all of which were stationary sounds. Thirty subjects, who were different to the subjects used in the first sensory test, evaluated the kansei qualities of each sound sample based on the SD method. We selected 11 SD scales (pairs of adjectives). We selected six SD scales – "cheap–expensive", "dislike–like", "agreeable–annoying", "silent–noisy", and "powerful–weak" – from the first sensory test and confirmed that they are independently effective SD scales. The remaining SD scales are newly introduced. The subjects were divided into three groups with ten people in each. The ten subjects listened to each sample using a headphone. Each sound was played for five seconds and the subjects evaluated the sounds, based on seven levels, by filling out a questionnaire consisting of word pairs.

16.5.9 Comparison of SD Scales Obtained from First and Second Experiment Data

First, we extracted patterns of personal SD scales from each SD scale using the proposed statistical method. In this method, we calculated correlation coefficients of the score vectors for each SD scale between subjects and conduct cluster analysis using the correlation coefficients to classify the subjects into groups of similar sensitivities. We used the average value of the SD scale in each cluster (group) as the representative score of the personal SD scale.

Figure 16.11 shows a comparison between the first and second sensory tests in terms of the proportion of subjects in each cluster for SD scales which are used in both tests. The proportions of the largest cluster for each scale (the black portion) are greater than 50% for both test results, except for the SD scale "expensive–cheap". The cluster that contains the largest proportion of subjects in a SD scale is called its "major cluster", and the scale that is composed of only the major cluster is called a "major scale". A major cluster represents a set of majority subjects who have a similar sensitivity for a SD scale.

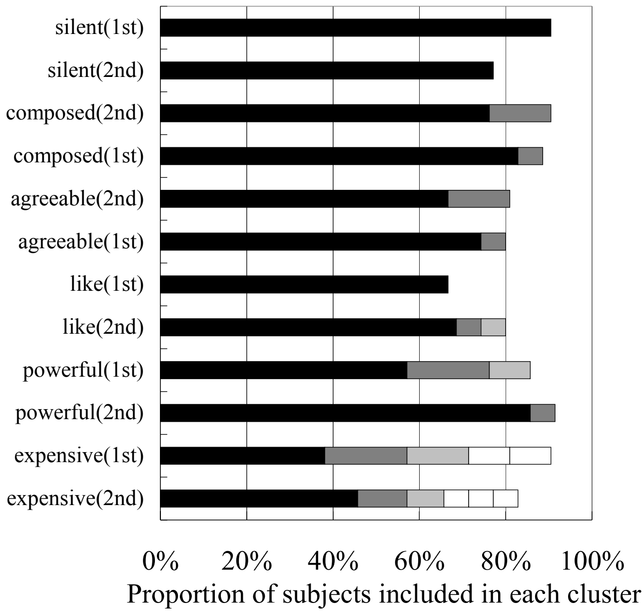


Figure 16.11 Comparison of proportion of subjects included in each cluster for SD scales that are used in first and second experiments

The target kansei quality "expensive–cheap" contains multiple SD clusters. The commonality of the major cluster is less than 50%. In other words, people have different sensitivities. To see the difference of semantics between SD cluster scales, we calculated the correlation coefficients between the top three SD cluster scales for “expensive–cheap” and the major scales of other SD scales. Table 16.2 shows the results. The three SD cluster scales of “expensive–cheap” all relate to the major scales of “composed–discomposed” and “silent–noisy”. We found that only the SD scale “reliable–unreliable”, which was newly introduced in the second test, discriminates the major scale from the other scales, as shown in Table 16.2. The SD scale “reliable–unreliable” contains two major scales.

Table 16.3 shows the correlation coefficients between the two cluster scales of “reliable–unreliable” and the major scales of other SD scales. The major scale of “reliable–unreliable”, whose proportion is 31.4%, relates to “composed–discomposed” and “silent–noisy”. The second cluster scale relates to “powerful–weak” and “unobstructed–obstructed”. Those two feelings related to reliability have totally different contexts. The major cluster scale of “expensive–cheap” relates to both cluster scales of “reliable–unreliable”. Only the third cluster scale relate to one of the scales of reliable. Thus, the major cluster scale of “expensive–cheap” is a complex scale that contains two different feelings of reliable sound.

Table 16.2 Correlation coefficients between SD cluster scales of “expensive–cheap” and related major scales

Expensive	Composed (82.9%)	Silent (77.1%)	Reliable (31.4%)	Reliable (28.6%)
Major scale (45.7%)	0.88	0.74	0.86	0.72
Second cluster scale (11.4%)	0.60	0.62	0.53	0.14
Third cluster scale (8.6%)	0.75	0.83	0.69	–0.06

Table 16.3 Correlation coefficients between SD cluster scales of “reliable–unreliable” and related major scales

Reliable	Composed (82.9%)	Silent (77.1%)	Powerful (85.7%)	Unobstructed (11.4%)
Major scale (31.4%)	0.95	0.90	–0.14	0.44
Second cluster scale (28.6%)	0.39	0.22	0.76	0.76

16.5.10 Finding Kansei Factors

The target SD scale “expensive–cheap” is a complex scale. To extract independent factors that are used to evaluate the sound quality of a product, we conducted a factor analysis using the SD cluster scales obtained from the results of the second sensory test.

Figure 16.12 shows the factor loadings of the first factor. The top three cluster scales of “expensive–cheap” all positively relate to the first factor. Only the major scale of “reliable–unreliable (31.4%)” relates to the first factor. Major scales related to it, such as “silent–noisy”, “composed–discomposed”, and “agreeable–annoying”, positively relate to the first factor.

Meanwhile, the second factor positively relates to the second cluster scale of “reliable–unreliable (28.6%)” and its related major scales such as “powerful–weak” and “unobstructed–obstructed”, as shown in Figure 16.13. The major scale of “expensive–cheap (45.7%)” relates to both factors, so that the factors are individual scales that can explain the complex feelings related to a sound’s expensiveness.

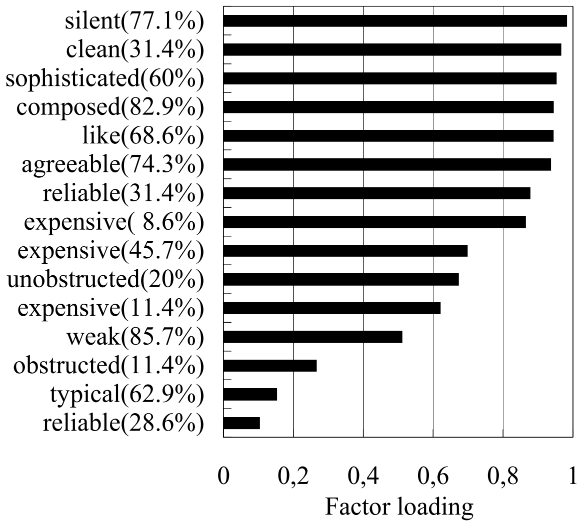


Figure 16.12 Factor loading of first factor (contribution ratio = 57.8%)

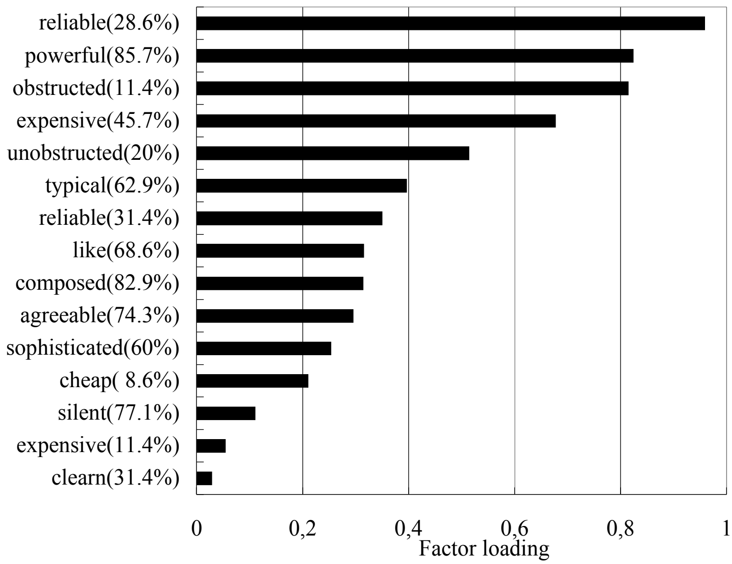


Figure 16.13 Factor loading of second factor (contribution ratio = 24.5%)

To formalize the factors, we conducted multi-regression analysis using the SQMs as explanatory variables. Table 16.4 shows the results of the analysis for the first factor. Both the regression coefficient and correlation coefficient of loudness are dominant. Loudness negatively relates to the factor. "Silent–noisy" has the highest factor loading, so that the value of the first factor increases when the sound is silent.

The fact that the first factor is related to loudness is adequate for conventional works that aim to reduce the loudness of the product sound. This factor represents a simple and clear goal when designing sound.

There is, however, a technical and cost limitation to reducing loudness. We focused on the second factor to design a sound without reducing loudness. The result of MRA using the second factor and SQMs shows that sharpness negatively relates to the factor, so that high-frequency sounds do not get higher scores for the factor. Loudness positively related to the factor, which means that reducing loudness reduces the evaluation score of the second factor. The first and second factors have a trade-off in terms of loudness. A measurable indicator, "tone-to-noise ratio" (TNR) [26], is newly introduced to explain the feelings of reliable and power-

Table 16.4 Result of multiple regression analysis using the first factor and SQM ($R = 0.87^{**}$)

SQM	Standardized regression coefficient	*p*	Partial correlation coefficient	Correlation coefficient
Loudness	–0.88	0.00	–0.87	–0.89
Sharpness	–0.15	0.27	–0.24	–0.66
Roughness	–0.07	0.40	–0.18	0.23
F.S.	–0.20	0.09	–0.35	–0.07

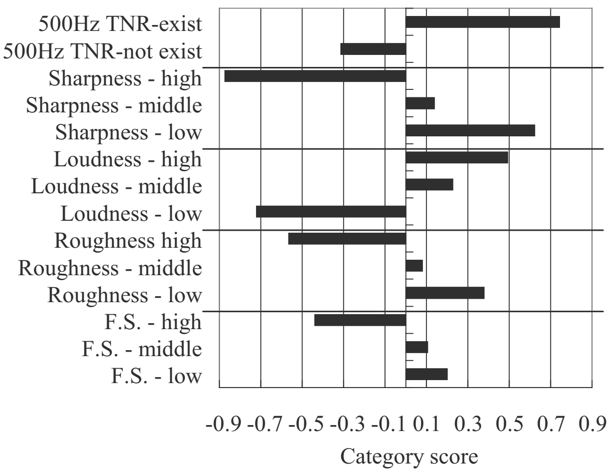


Figure 16.14 Results of quantification theory 1 using the second factor as an objective variable ($R = 0.9^{**}$)

ful, which highly relate to the second factor. The TNR is the difference between the tone and the sound pressure level of the noise in a critical band centered around the tone. Most vacuum cleaners have a peak tone around 500 Hz because of the frequency of the motor. We applied quantification theory using the TNR and SQMs as explanatory variables. Figure 16.14 shows the category scores, which represent the weights of each category of a feature. From the result, a sound having a perceivable TNR around 500 Hz gets high scores for the second factor. Thus we found that a perceivable motor sound is important to increase the second factor. This is a new criterion because conventional approaches have aimed at reducing TNR and loudness.

References

1. Yanagisawa H., Fukuda S.: Interactive reduct evolutionary computation for aesthetic design. Trans. AMSE J. Comp. Inf. Sci. Eng. **5**(1):1–7 (2005)
2. Maslow A.H.: A theory of human motivation. Psychol. Rev. **50**:370–396 (1943)
3. Osgood C.G., Suci G.J., Tannenbaum P.: The Measurement of Meaning. University of Illinois Press, Champaign, IL (1957)
4. Thurstone L.L.: The Measurement of Values. The University of Chicago Press, Chicago, IL (1959)
5. Hsiao S.W., Huang H.C.: A neural network based approach for product form design. Design Studies **23**(1):67–84 (2002)
6. Pawlak Z.: Rough sets. J. Comp. Inf. Sci. **11**:341–356 (1982)
7. Pawlak Z.: Rough classification. J Man-Machine Studies **20**:469–483 (1984)
8. Mori N., Takanashi R.: Knowledge acquisition from the data consisting of categories added with degrees of conformity – proposal of extended reduct calculation in rough set theory. Kansei Eng. Int. **1**(4):19–24 (2000)
9. Nakamori Y., Ryoke M.: Modeling of fuzziness in multivariate analysis. In: Proceedings of IEEE SMC Conference 2004, vol. 1, pp. 294–298 (2004)
10. Yanagisawa H., Fukuda S.: Development of industrial design support system considering customers' subjective evaluation. Trans. Jpn. Soc. Mech. Eng. **68**(675):3431–3438 (2001)
11. Takagi H.: Interactive evolutionary computation. Proc. IEEE **89**(9):1275–1296 (2001)
12. Yanagisawa H., Fukuda S.: Development of interactive industrial design support system considering customer's evaluation. JMSE Int. J. Series C **47**(2):762–769 (2004)
13. Smyth S.N., Wallace D.R.: Toward the synthesis of aesthetic product form. In: Proceedings of ASME DETEC'00 Computers and Information in Engineering Conference (2000)
14. Yanagisawa H., Murakami T., Noguchi S. *et al.*: Quantification method of diverse kansei quality for emotional design – application of product sound design. In: Proceedings of DETC'07/ASME International 27th Computers and Information in Engineering (2007)
15. Yanagisawa H., Kataoka A., Murakami T. *et al.*: Extraction of latent emotional factors by analyzing human sensitivity towards unexplored design: application to product sound design. In: Proceedings of International Conference on Engineering Design 2009 (2009)
16. Yanagisawa H., Murakami T., Ohtomi K. *et al.*: A method for quantifying product emotional quality considering diversity of personal sensitivities. In: Proceedings of Design Engineering Workshop, pp. 102–107 (2008)
17. Zwicker E., Fastl H.: Psychoacoustics: Facts and Models, 3rd edn. Springer, New York (2006)
18. Otto N., Amman S., Eaton *et al.*: Guidelines for jury evaluations of automotive sounds. In: Proceedings of the 1999 SAE Noise and Vibration Conference (1999)

19. Chatterley J.J., Sommerfeldt S., Blotter J. *et al.*: Sound quality testing and analysis of multiple brands of sewing machines. In: Proceedings of ASME IDETC/CIE 2005 (2005)
20. Vastfjall D., Gulbpl M., Kleiner M. *et al.*: Affective evaluation of and reactions to exterior and interior vehicle auditory quality. J. Sound Vibration **255**(3):501–518 (2002)
21. Fastl H.: Neutralizing the meaning of sound for sound quality evaluations. Internoise (2001)
22. Ellermeier W., Zeitler A., Fastl H.: Predicting annoyance judgments from psychoacoustic metrics: identifiable versus neutralized sounds. Internoise (2004)
23. ISO 226: Acoustics – normal equal loudness level contours (2003)
24. ISO 532/R: Acoustics – method for calculating loudness level (1975)
25. Schiffman S.S., Reynolds M.L., Young F.W.: Introduction to Multidimensional Scaling. Academic Press, New York (1981)
26. ISO 7779 ANSI S1.13: Tone to noise ratio (1995)

Chapter 17
Emotion in Engineering Design Teams

Neeraj Sonalkar[1], Malte Jung[2], Ade Mabogunje[3]

17.1 Introduction

Knowledge that is relevant to the practice of engineering can be categorized into three domains. First is the knowledge of the natural world that we fashion into engineering artifacts. This includes knowledge domains such as physics, chemistry, biology, and thermodynamics. Second is the knowledge of processes that we may use to transform the natural world into engineered artifacts. These include various engineering design methods, production processes, and mathematical methods. The third is the knowledge of the humans creating and using the engineering artifacts. This involves understanding and improving how engineers perceive, think, and act individually or collectively, such as in teams or organizations, when they are engaged in the daily practice of engineering; and also understanding how the users of these artifacts perceive and interact with them in the course of their life cycle. This domain uses and synthesizes knowledge from other fields such as psychology, group work, cognitive science, sociology, and anthropology that focus on the human as a subject of study. However, it differs in one key respect from these fields in that its focus on the human is rooted in an engineering value system that seeks to understand in order to re-create artifacts and situations for the better. The study of emotion is an important part of the domain of humans creating and using engineering artifacts.

[1] N. Sonalkar
Center for Design Research, Stanford University, Stanford, CA, USA
e-mail: sonalkar@stanford.edu
[2] M. Jung
Center for Design Research, Stanford University, Stanford, CA, USA
e-mail: mjung@stanford.edu
[3] A. Mabogunje
Center for Design Research, Stanford University, Stanford, CA, USA
e-mail: ade@stanford.edu

At the Center for Design Research at Stanford University, we are predominantly engaged in understanding how engineers design new artifacts and how we can help them in designing more effectively. As such, we deal with the domain of humans creating engineering artifacts. In this chapter, we will explore the role of emotions in the activity of designing new engineering artifacts. We begin with a discussion on emotion in the context of engineering design teams, followed by a brief overview of perspectives in emotion research. We then present two research studies which examine the influence of emotion in two situations, namely group ideation and group conflict, which are common to engineering design teams.

17.2 Why Study Emotion in Engineering Design?

The context of our work is creative collaboration amongst engineering designers leading to new product development and eventually product and service innovation in the market place. While engineering has often been defined as the application of mathematics and science to the needs of the society, a definition that nominally has nothing to do with emotions, research on the process of design, in other words – how design happens – reveals quite the opposite.

Engineering designers increasingly work in multidisciplinary *teams* on *complex* and *functional* tasks whose outcome while *uncertain* in terms of the reception in the market place is expected to be *creative* (Figure 17.1).

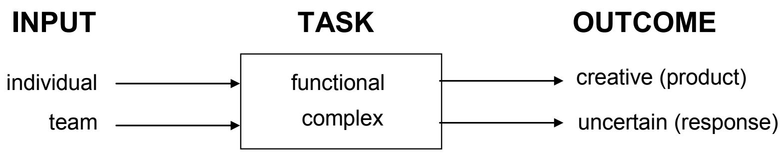


Figure 17.1 Engineers work on complex and functional tasks that require abstract and conceptual ideas to be translated into concrete and physical products, effects, and services for humans. This work is highly contextual and dependent on emotional cues. A systematic understanding of engineering cannot be done without a basic understanding of the emotional dynamics of their work situation

Within this six-element context of their work the only element that is devoid of emotion is the functional task, which essentially relies on mathematical and scientific knowledge. All other elements involve the implicit or explicit use of emotions. This could be, amongst other emotions, the look and feel of the tangible product or service experience (aesthetic), the interpersonal conflict resulting from an unresolved disagreement or perceived humiliation, the excitement over a particular product concept, the fear associated with the danger of product failure, or the shame associated with the utterance of a wild idea. Some of these effects would have had less of an impact if engineering designers did not have to work in teams, if the fraction of time during which they are engaged in purely functional tasks was large, or if they were not responsible for coming up with imaginative ideas. Indeed, according to Crispin Hales, engineering designers working in industry spend almost 50% of their time and effort attending to tasks that are more social in nature, and more associated with the other five elements where emotions predominate [1]. Needless to say, any attempt to improve the performance of engineering designers through systematic study must pay attention to their emotions.

Given our interest in creativity for which human imagination is the source, we see a direct relation to emotion as expressed by the following simultaneous relations:

- Imagination → Emotion → Action (*e.g.*, verbal expression of an idea).
- Outcome ← Emotion ← Reaction (*e.g.*, non-verbal expression of disinterest by team).

The outcome will include the possibility that the idea is not written in the meeting minutes or public whiteboard used for brainstorming. Also the speaker may feel embarrassed and may no longer wish to contribute to the discussion. In this way, a designer's emotional state allows us to account not only for his or her subjective appraisal of a situation but also for their perception and motivation. When we study engineering in this way, we are able create in a more systematic manner an alignment between people, tasks, and situations that increases the chances for creative breakthroughs.

17.3 Perspectives in Emotion Research

Imagine three engineers grouped around a table. They are all part of a team and are engaged in a discussion about what they should have accomplished as a final project deliverable (see Table 17.1).

If we are attuned only to speech, this is all we can observe. However, once we start attending to emotional cues, a very different quality of the interaction emerges (see Table 17.2).

Table 17.1 Interaction segment with verbal content only

Speaker	Content
Sarah:	So… what should we have done for the final presentation?
Becky:	So I think this was this conversation that we started while we were walking here, right?
Sarah:	Yeah.
Tom:	Well, do we want to start with what the ideal thing is and then start taking away from there?
Becky:	Ok
Sarah:	We all know what the ideal thing is.
Becky:	We all know what the ideal thing is, it works, it's beautiful, like…

Table 17.2 Interaction segment with verbal content and emotion expression

Speaker	Content	Emotion
Sarah:	So … what should we have done for the final presentation?	[Neutral]
Becky:	So I think this was this conversation that we started while we were walking here, right?	[Neutral]
Sarah:	Yeah.	[Neutral]
Tom:	Well, do we want to start with what the ideal thing is and then start taking away from there?	[there is a dragging quality in his voice and gestures that expresses disinterest in what is being said and disengagement]
Becky:	Ok	[Ok; sounds uninterested and there is an expression of contempt on her face]
Sarah:	We all know what the ideal thing is.	[Neutral]
Becky:	We all know what the ideal thing is, it works, it's beautiful, like …	[laughs nervously and there is a frustrated tone in her voice]
Tom:		[rolls his eyes]

The few seconds of dialog pictured here provide a context for our research. Several past studies have studied internal emotion-related mechanisms, such as the effects of emotional states on individual creativity or decision-making [2, 3, 4]. Little has been done to explore the moment-to-moment role of emotions in creative interactions. For the research described in this chapter we therefore focus on how emotions shape the flow of an interaction and how the quality of that interaction influences the emergence of outcomes. For example, how does Becky's expression of contempt in the example above influence the engagement of her teammates? How will it influence whether Peter and Sarah will listen and build on the ideas she might propose next? As these questions are in the service of developing understanding our research, at the same time, is undertaken in the service of an interest in change. We want to imagine and develop better ways of interacting in an engineering team. For example, can we imagine different reactions by Becky to Tom's proposal that might have led to a more constructive discussion?

Interactions such as the one between Becky, Peter, and Tom are subject to a wide range of emotion-related phenomena. Peter's bad mood over the last week as well as Becky's stress due to an upcoming project deadline both can be assumed to have impacted the interaction above. If we want to study the emotion-related phenomena that are particularly relevant to the moment-to-moment dynamics of an interaction it makes sense to distinguish emotions from a broader range of emotion-related phenomena. One way to achieve this is to define emotion as a specific event in time. Ekman defines emotions to be less than 4 s in duration [5]. This distinguishes it from other phenomena like mood or stress that sustain over longer durations of time. However, this still does not help us answer the basic question "what exactly do we mean by emotion?".

17.3.1 Assigning the Label "Emotion"

Reviewing emotion research over the past 40 years, we realized that there are different perspectives regarding which phenomena the labels emotion and affect can be assigned to. A useful way to anchor ourselves is through the signs and signals that we can detect and experience. These can be broken down into three major categories:

1. internal physiology or biology;
2. subjective experience or feelings;
3. externally displayed behaviors.

The debate then centers on which phenomena to include within the label emotion and how then, what we mean by emotion, comes about.

In terms of assigning labels, we are maintaining two different perspectives in our research.

The first perspective by Tomkins as described by Nathanson [6] assigns the label "affect" to biological signals (Figure 17.2). Feeling is the consciousness of these signals. It is a sensation that can be checked against previous experiences and labeled. It is personal and biographical because every person has a distinct set of previous sensations from which to draw when interpreting and labeling his or her feelings. Emotion is the triggering of memories by feelings and the projection or display of a feeling. For example, Tom's eye roll response may have been a display of a feeling of contempt triggered by memories. Unlike feelings, the display of emotion can be either genuine or feigned [6].

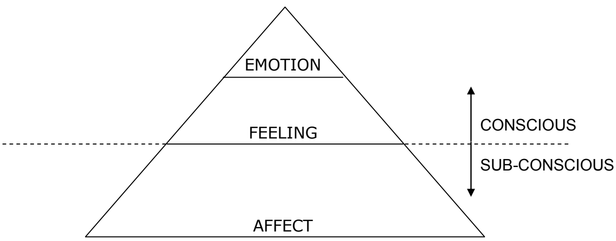


Figure 17.2 Emotion as an emergent phenomena distinguished by its level of awareness and display

The other perspective we maintain is influenced by the work of Scherer and Gross. According to this viewpoint the term "affect" is assigned to a super category referring to all kinds of motivational phenomena [7, 8]. Emotion is a subcategory of affect referring to clearly delineated, intensive patterns of affective processes such as the frustration expressed and probably experienced by Becky in the scenario above. Emotion then is a label assigned to certain combinations of physiology, feeling and behavior [9]. Among other sub-categories of affect we can then make distinctions between stress responses, moods such as depression or euphoria, and other motivational impulses such as those related to eating, sex, aggression, or pain [7] (see Figure 17.3).

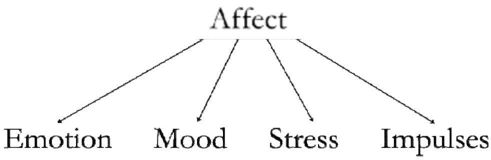


Figure 17.3 Distinction of emotion from other affective phenomena as laid out by Gross [7]

17.3.2 How Does an Emotion Come About?

From a process viewpoint emotion can decomposed into major functional elements and their causal interrelations. The decomposition of an emotion from a process perspective is particularly helpful in thinking about when and how to change emotion dynamics in interactions. To date many different emotion process models exist. LeDoux has written an excellent overview about different models and how they emerged throughout history [10].

Corresponding to the two perspectives in labeling, we can describe two different perspectives on how emotions come about. The perspective on emotion as a result of feelings triggered by memories [6] considers the physiological signal or affect as the beginning of the process that leads to an emotional response. The consciousness of affect leads to feelings which when situated in the context of past memories lead to an emotional response (Figure 17.4). As mentioned earlier, this perspective stresses the biological and biographical elements in the emergence of emotions.

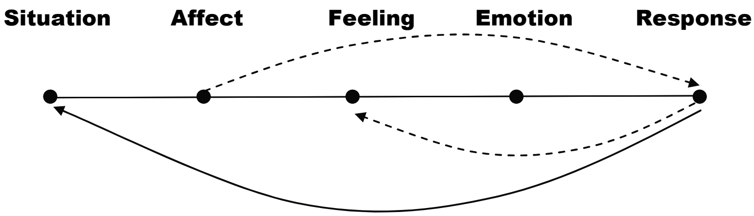


Figure 17.4 An adapted version of Tomkin's stratified model. The dotted lines indicate that the process of emotional responding need not follow a linear sequence

The other perspective corresponds to the modal model of emotion proposed by Gross [7]. The modal model of emotion defines emotion as a person-situation interaction that compels attention, has particular meaning to an individual, and gives rise to a coordinated yet flexible multi-system response to the ongoing person-situation interaction [7]. It decomposes the process out of which an emotion emerges into situation, attention, appraisal, and response and differentiates between different intervention strategies dependent on the point where the intervention is made (Figure 17.5). The modal model thus assigns meaning in the formation of emotion to the individual.

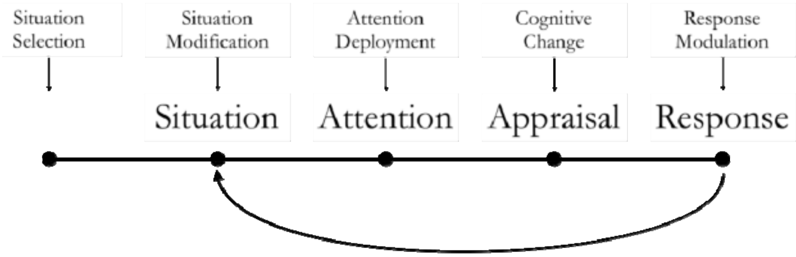


Figure 17.5 The modal model of emotion with specific emotion regulation strategies. Redrawn from [7]

17.3.3 Distinguishing Between Different Emotions

Even within the category of emotion there are a number of labels that vary by culture and context to distinguish between different emotional patterns. How these distinctions should be drawn is a matter of longstanding debate between proponents of what can be called natural kinds perspectives and proponents of core affect perspectives. The natural kinds perspective asserts that there are a limited number of emotion categories that are inherent in nature, and they are discovered rather than socially constructed. They are given by nature, and rely on distinct neuro-biological mechanisms. Under a natural-kinds perspectives as proposed by Tomkins or Izard, emotions are typically categorized into six [11, 12] to nine [13] "basic emotions" and a number of blends or schemas [12] that refer to the situated occurrence and experience of these basic emotions. Examples of typically listed "basic emotions" are interest, joy/happiness, anger, disgust, and fear. The core affect perspective, on the other hand, asserts that there are no emotion-specific mechanisms but rather a core affect system as a mechanism capable of generating all kinds of emotion-related phenomena [14]. Under a "core affect" perspective, emotions are often categorized dimensionally along characteristics such as valence and arousal [15].

17.3.4 Emotion in the Context of Current Studies

Table 17.3 shows that it is important to be aware that many different conceptualizations of emotion exist. Rather than pointing out a particular conceptualization as the right one we think that each perspective adds particular value dependent on the specific research questions.

In the context of the research studies described here, we define emotion as a person-situation interaction that compels the attention of and is meaningful to a trained observer within a particular cultural context. We draw distinctions between emotions from a social functionalist perspective [16]. From this perspective it has been useful for us to categorize emotion-patterns based on the effects they have in social interactions. This perspective particularly puts behavioral aspects of emotions such as gestures, body posture, facial muscle movement, voice tone, and speech in the foreground. The definition used here is particularly helpful in studying not only what happens from an emotion perspective when engineers interact but is also helpful in imagining and developing more effective interaction patterns.

Table 17.3 Simplified overview of emotion perspectives

Question	Alternative 1	Alternative 2	Alternative 3
What phenomena do we assign the label affect?	Physiology only. Tomkins as described in Nathanson [6]	An aggregation of different motivational phenomena including emotion, stress, mood, and impulse (Gross [7])	
What phenomena do we assign the label emotion?	Tomkins argues that emotion is the triggering of memories by feelings and the projection or display of a feeling (Nathanson [6])	A combination of physiology, feelings, and displayed behavior. This combination is true for other phenomena besides emotion [7]. Emotion is that which lasts for less than 4 s [5]	
How do emotions come about?	Biologically and biographically determined [6]	The process involves a cognitive mechanism of attention and appraisal. Modal model of emotion [7]	
How do we distinguish between different emotions?	There exist six to nine basic emotions that are biologically hardwired [11, 12, 13] (natural kinds perspective)	Emotions are categorized based on the effect they have in social interactions (socio-functionalist perspective [16])	Emotions are categorized dimensionally along characteristics such as valence and arousal (core affects perspective)

17.3.5 Measuring Emotion

A wide array of instruments are available that are useful in investigating particular emotion-related phenomena in the dimensions of physiology, subjective experience, and behavior. Subjective experience, for example, can be assessed in discrete intervals using the Positive and Negative Affect Schedule (PANAS) [17], or continuously using the affect rating dial method [18]. With the Facial Affect Coding System (FACS) [19] and the Specific Affect Coding System (SPAFF) [20, 21] powerful assessment tools have been developed to codify facial muscle movement relevant to display behavior (FACS) or complex behavior patterns including, voice tone, gestures, speech, and facial movement (SPAFF). Physiology is often assessed through heart rate, blood pressure, or skin conductivity [22]. In the studies described below we are measuring behavior in interpersonal interactions in teams using a version of SPAFF adapted for design team situations.

17.4 Research Study 1 – The Role of Emotion in Group Ideation in Engineering Design Teams

Any conversation or interaction that involves the discussion of ideas between two or more individuals in a design team can be considered an instance of group ideation. This can range from informal "water-cooler" discussions to more formal design review and ideation meetings. These interpersonal interactions are not devoid of emotions. The emotions expressed and felt during ideation interactions may influence the ideas generated by the design team.

In order to study the effect of emotions during interpersonal interaction on the ideas developed, we video-taped nine design teams discussing ideas for a product concept in an engineering design project course at Stanford University. These tapes were analyzed to study the quality of interpersonal interactions and how they affect the quality ideas that are being discussed. We used a modified version of SPAFF developed by Gottman [21] to manually code the emotions expressed during the ideation interactions in the team. From a preliminary analysis of the video tapes, we can infer the following:

1. Emotional quality of interaction varies across time within a team.
   The emotional quality of an interaction is determined by the emotions expressed by individuals participating in the interaction during the course of the interaction. We identified seven different emotion codes based on the modified version of SPAFF coding scheme. These emotion codes were pushy, frustration, tension, neutral, validation, play/humor, and excitement. The emotion codes are described briefly in Table 17.4.
   Arranging these emotions on a negative to positive scale over time, we could plot the emotional variation within a design team. As an example, Figure 17.6 show an emotional variation graph for Team A over a duration of 40 min.
   We can see from Figure 17.6 that Team A had moments of positive and negative emotions. The next question is whether the emotions are related to the quality of ideas discussed in these moments.

Table 17.4 Emotion code descriptions used in the study

Emotion code	Description	Examples
Pushy	This emotion code describes behavior that pushes a particular agenda onto others. It is generally recognized through use of non-conditional terms in language like "obviously", "absolutely", *etc.* It is also recognized though the tone of voice that is forceful and even at times aggressive. Body posture associated with the code is erect and rigid	"For sure, we are lacking post-its"
Frustration	Frustration is a code that denotes constrained anger. The anger will appear constrained or out of the obvious awareness of the speaker. Voice cues include the lowering of the voice and speaking in an even, staccato rhythm, as if to communicate to the partner that the speaker is at the end of her rope	"I always wanted to buy the chair, but you guys were against it" (spoken in a staccato rhythm)
Tension	Tension communicates anxiety, worry, and nervousness. Indicators include nervous gestures, fidgeting, stuttering speech, and incomplete or unfinished statements	"Well, I uh... it's just that whenever... I mean, umm, when I wa... want to uh... want to go out I feel that I... it's like I always have to ask" (speech disturbance)
Neutral	Neutral code represents the dividing line between positive and negative codes. Indicators include relaxed quality with even pitch and volume of voice	"Shall we begin?" (asked in an even volume)
Validation	Validation communicates sincere understanding and acceptance. Indicators include head nods, verbal sounds of acceptance like yes, yeah, ok, paraphrasing, and finishing each other's sentences	"Yes! It is projector screen floss"
Play/humor	Play and humor have been grouped together in this code. Indicators include shared laughter, and playful actions accompanied by smile or laughter	Playful interaction over a whiteboard marker
Excitement	Expression of passionate interest in person, idea or activity. Indicators include expression of joy, delight, high volume, pitch exclamations, and shouts	"When your boss is angry the room shakes!" (spoken in a high-pitched tone with laughter)

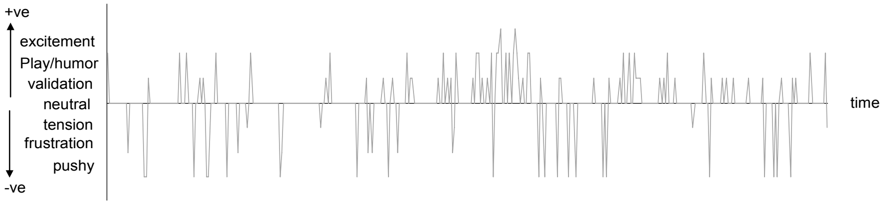


Figure 17.6 Emotional variations over 40-minute duration in Team A. Excitement, play/humor, and validation codes indicate positive emotions; tension, frustration, and pushy codes indicate negative emotions

2. Positive emotions more than negative emotions are associated with interactions where ideational responses are emergent and spontaneous.

   The video segments of the nine ideation session were decomposed into interaction segments based on topical continuity of individual responses. If an individual initiates a new topic through a verbal or non-verbal action, then that indicates the start of a distinct interaction segment. If others respond to it and their responses are linked to each other topically, then their interaction could be categorized as one interaction segment. When we analyzed the interaction segments in a team ideation session, we realized that:

   1. not all of them were related to product ideas;
   2. they had different qualities depending on how a response was linked to previous actions. For example, in question answer interactions, the responses were directly constrained or even demanded by the previous action, the question. In some idea proposal interactions, the responses to an action were more emergent and spontaneous. By emergent response, we mean that the response was not fully determined by the previous actions. Emergence is an event in complex systems that cannot be predicted in advance, but can be found to be emerging from an interaction of several dynamic elements [23, 24]. By spontaneous responses, we mean the responses that were observed to be qualitatively without premeditation.

   The ideation interactions with emergent and spontaneous responses were associated with positive emotions like validation, play/humor, and at times even excitement. Table 17.5 contains a brief section illustrating the emotional cues in an emergent, spontaneous interaction segment.

Positive emotions, thus, seem to be closely linked to the spontaneous emergence of ideas in an engineering design team. While further studies are necessary to understand and model the role emotions play in creative interactions, we can propose a few hypotheses on how emotions factor into group creativity. Emotions could play a two-fold role in group ideation:

1. Positive emotions could act as indicators of personal resonance with an idea. We have observed in our video data that the occurrence of positive emotions, especially excitement, is associated with remarks on personal significance of the idea being discussed. In other instances, positive emotions are associated with the shared imagination of a scenario being discussed. In the latter case, positive emotions tend to be shared while in the first case, they may be limited to an individual.
2. Positive emotional expression about an idea may facilitate the development of the idea. Prior research on group creativity that focuses on the process of creative collaboration [25] shows that group creativity depends on accepting what is presented and extending it further into imagination. The key is to not let the present action constrain the future possibilities, and yet the future emergent possibilities should build on what is accepted in the present. In this delicate balance, emotions could play the role of creating resonance with a presented idea while motivating further engagement with the idea.

Table 17.5 Emotional cues in a spontaneous ideation segment

Speaker	Content	Emotional expression	Gesture	Tone of voice	Facial expression
A	What about like, people need lighting that inspires them, but how do you know when they need to be inspired? What if we had watch, right, that senses your vital signs and it's like… ok, Brendan's falling asleep and is hungry, something like that…	Neutral	Waving hands	Neutral	Smiling
B	It senses your heart rate…	Neutral	None	Neutral	Not visible on camera
A	And it like…	Neutral	None	Neutral	Smiling
B	Shocks you!	Play/humor	Waves his hands	Playful	Not visible on camera
A	Shocks you	Play/humor	Mimics B's gesture	Playful	Smiling
C	(laughing)	Humor	Not visible on camera	Genuine laughter	Not visible on camera
D	Oh, yeah, you know like...	Validation	Standing up and moving towards A and B	Interested	Not visible on camera
A	It all of a sudden like, plays some pumped up music like	Play/humor	Rocking on his feet	Neutral	Smiling
B	(makes a guitar jamming sound)	Play/humor	Bobbing his head	Playful	Not visible on camera
A	Jam, truck jam, or something	Neutral	Moving to face D	Neutral	Smiling

17.5 Research Study 2 – Team Conflict in Engineering Design Teams

Conflict in an engineering team can have adverse consequences for the product being developed like inhibition of ideas, skewed decision-making priorities, delayed product completion and even failure in the marketplace. Team conflict is an important topic of research that has been studied in the past [26, 27]. Jehn [27] proposed that intragroup conflict could be categorized into task conflict and relationship conflict. Task conflict is defined as "*disagreements among group members about the content of the tasks being performed, including differences in viewpoints, ideas, and opinions*" [27]. Relationship conflict is defined as "*interpersonal incompatibilities among group members, which typically includes tension, animosity, and annoyance among members within a group*" [27]. It is currently held that task conflict is beneficial and relationship conflict is detrimental to team performance [27]. However most of the studies pertain to the content of conflict and not to the process of how it unfolds over time. When we study what is happening moment-to-moment in a team conflict situation, we need to pay attention to the role of emotions expressed by the team members.

We can examine a sample of team conflict that we recorded on video as part of a study on disagreement in engineering design teams in order to illustrate how emotions form an important dimension of conflict that can be separated from the content of a disagreement.

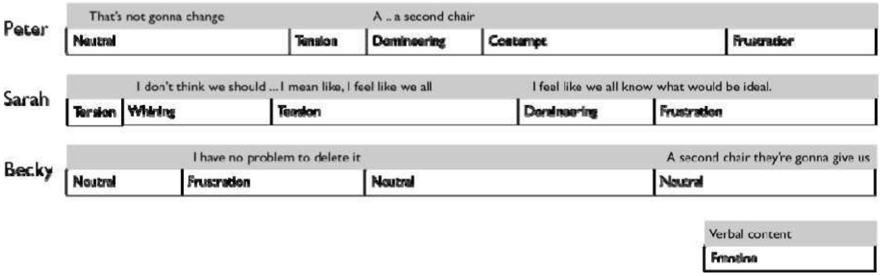


Figure 17.7 A disagreement conversation with verbal content and emotion expression

In the example in Figure 17.7 the quality of the interaction is characterized by negative emotions such as frustration, domineering, and contempt. In a social context these emotions often have the function of alienating people from each other. The very same content however could be delivered with a different set of emotions. Peter could deliver his statement “a second chair” without the lecturing and somewhat condescending tone in his voice and instead say the same words expressing interest. Following the flow of emotions during conflict situations turns the spotlight away from what a team is disagreeing about (whether it be relationship-related issues or task-related issues) and towards *how* a team is engaging in a particular disagreement.

Furthermore, building on the model of marital conflict proposed by Gottman [28], we can propose a model of conflict for engineering design teams that incorporates the role of emotion as a mediator of good or bad team conflict.

Gottman [28] proposed a relationship between the flow of emotional expressions that are exchanged during an interaction, the way this flow of emotions is perceived, and the physiological responses. The emotion flow can consist of positive emotions like interest, validation, and humor, or negative emotions like contempt, belligerence, and domineering. Based on the ratio of positive to negative emotions expressed in a flow over time, the perception of the partner changes from a state of well-being to a state of distress. If more positive emotions are expressed, the state remains in well-being. If the ratio of positive to negative emotions falls below 1, then the state shifts to a state of distress. Once a shift has occurred in perception, further expression of negative emotions by the other partner is attributed to the person rather than to the situation thus precipitating a downward spiral in the relationship. While being in a state of well-being negative emotions are generally attributed to the situation and not the person. If there are negative emotions one might think, “Maybe he or she is just having a bad day.”. However, while being in a state of distress, those same negative emotions are attributed to the person. One might then think, “He or she is just that kind of a person who does this.”.

We propose that in a team conflict situation as well, the amount of positive to negative emotions expressed during a disagreement discussion influences how team members perceive each other. We call a disagreement discussion over a task related to the design project a task conflict. If in a task conflict, greater amount of positive emotions are expressed, then it is a positive task conflict which enables the team to remain in a state of well-being. If in a task conflict, greater amount of negative emotions are expressed, then it is a negative task conflict. With repeated occurrences, negative task conflict can lead to state of distress, in which team members make personal attributions that are stable over time. This state of conflict where team members attribute negative emotions to each other’s personalities is similar to the relationship conflict defined in literature [27] as being detrimental to the team. Thus, emotional expression could be the mediator of good team conflict (task conflict with participants in a state of well-being) or bad team conflict (relationship conflict with participants in a state of distress).

Understanding the role of emotion in team conflict enables us to recognize when to intervene in team disagreement situations. Currently, we are engaged in developing this model of team conflict further and designing facilitation tools that can be used by engineering design teams to resolve negative team conflict.

17.6 The Cultural Context of Emotion

The emotional categories we mention in these research studies are labels that we assign to behavior. The coding of observable behavior into meaningful categories is with reference to the culture in which the situation of study occurs. The two studies described here occurred in a European-American cultural context and were analyzed in the same cultural context. Hence we should note that the categories are salient from a European-American cultural perspective. If we were to conduct studies in a different cultural context say East-Asian or African, the behaviors and their interpretation by both the participants in the study and the researchers could be different.

17.7 Looking Forward

Now that we can detect and categorize behavior into different emotions, what can we do with it? One alternative is to train engineers to recognize and understand the different emotional categories and how emotions play into their activity of designing. This may enable them to be more aware of emotions as they occur in practice and respond to them appropriately. Another alternative is to give real-time feedback to engineering design teams about the condition of their team interaction based on emotional expression as a variable of measurement. This could be achieved through technological tools that could detect emotions and display feedback, or through human observation and coaching. The paradigm of coaching in engineering design is useful for providing an informal knowledge channel to the design team [29]. Emotion as an indicator for team performance could be beneficial to a coach in guiding a team effectively.

17.8 Conclusion

Schön [30] describes the world of the practitioner as dealing with messy realities that the technical rationality of academic researchers too often abstracts away from their studies. Emotion often resides in the part that is abstracted away. However, we have seen in this chapter that we do have tools and methods to study emotion and to gain a deeper understanding of engineering practice. We therefore foresee a future in which the study of emotion is an integral part of engineering knowledge and practice much like thermodynamics, mathematics, and manufacturing.

Acknowledgments The research presented here is the result of several years of research funding and intellectual collaboration. We would like to acknowledge the Kempe Foundation, National Science Foundation under grant no. 0230450, the Kozmetsky Global Collaboratory's Real-time Venture Design Laboratory, and our colleagues at the Center for Design Research at Stanford University.

References

1. Hales C.: Analysis of the Engineering Design Process in an Industrial Context. Unpublished PhD dissertation, University of Cambridge (1987)
2. Fredrickson B.L.: The broaden-and-build theory of positive emotions. Philos. Trans. R. Soc. Lond. B Biol. Sci. **359**(1449):1367–1377 (2004)
3. Isen A.M., Daubman K.A., Nowicki G.P.: Positive affect facilitates creative problem solving. In: Jenkins J., Oatley K., Stein N. (eds.) Human Emotions: A Reader. Blackwell, Massachusetts (1998)
4. Isen A.M., Rosenzweig A.S., Young M.J.: The influence of positive affect on clinical problem solving. Med. Decis. Making **11**(3):221 (1991)
5. Ekman P., Friesen W.: Felt, false, and miserable smiles. J. Nonverbal Behav. **6**(4):238–252 (1982)
6. Nathanson D.L.: Shame and Pride: Affect, Sex, and the Birth of the Self. Norton, New York (1992)
7. Gross J., Thompson R.: Emotion regulation: conceptual foundations. In: Gross J. (ed.) Handbook of Emotion Regulation, pp. 3–24. Guilford Press, New York (2007)

8. Scherer K.: On the nature and function of emotion: a component process approach. In: Scherer, K. (ed.) Approaches to Emotion, pp. 293–317. Lawrence Erlbaum, New Jersey (1984)
9. Mauss I., Levenson R., McCarter L., Wilhelm F., Gross J.: The tie that binds? Coherence among emotion experience, behavior, and physiology. Emotion **5**(2):175–190 (2005)
10. LeDoux J.: The Emotional Brain: The Mysterious Underpinnings of Emotional Life. Simon & Schuster, New York (1996)
11. Ekman P.: An argument for basic emotions. Cogn. Emot. **6**(3–4):169–200 (1992)
12. Izard C.: Basic emotions, natural kinds, emotion schemas, and a new paradigm. Perspect. Psychol. Sci. **2**(3):260–280 (2007)
13. Tomkins S.: The quest for primary motives: Biography and autobiography of an idea. J. Pers. Soc. Psychol. **41**(2):306–329 (1981)
14. Barrett L.: Are emotions natural kinds? Perspect. Psychol. Sci. **1**(1):28–58 (2006)
15. Watson D., Wiese D., Vaidya J., Tellegen A.: The two general activation systems of affect: Structural findings, evolutionary considerations and psychobiological evidence. J. Pers. Soc. Psychol. **76**(5):820–K38 (1999)
16. Morris M., Keltner D.: How emotions work: The social functions of emotional expression in negotiations. Res. Organ Behav. **22**:1–50 (2000)
17. Watson D., Clark L., Tellegen A.: Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Pers. Soc. Psychol. **54**(6):1063–1070 (1988)
18. Ruef A.M., Levenson R.W.: Continuous measurement of emotion. In: Coan J.A., Allen J.B. (eds.) Handbook of Emotion Elicitation and Assessment. Oxford University Press, New York (2007)
19. Ekman P., Friesen W.V., Hager J.C.: Facial Action Coding System: Manual: Consulting Psychologists Press, Palo Alto, CA (1978)
20. Coan J.A., Gottman J.M.: The Specific Affect Coding System (SPAFF). In: Coan J.A., Allen J.B. (eds.) Handbook of Emotion Elicitation and Assessment. Oxford University Press, New York (2007)
21. Gottman J.M.: What predicts divorce: the measures. Unpublished coding manuals. Lawrence Erlbaum Associates, New Jersey (1996)
22. Levenson R.: Emotion and the autonomic nervous system: a prospectus for research on autonomic specificity. In: Wagner H. (ed) Social Psychophysiology and Emotion: Theory and Clinical Applications, pp. 17–42. Wiley, London (1988)
23. Mead GH.: The Philosophy of the Present. University of Chicago Press, Chicago, IL (1932)
24. Sawyer R.K.: The emergence of creativity. Philos. Psychol. **12**(4), 447–469 (1999)
25. Sawyer RK.: Group Creativity: Music, Theater, Collaboration. Lawrence Erlbaum, New Jersey (2003)
26. De Dreu C.K.W., Weingart L.R.: Task versus relationship conflict: a meta-analysis. J. Appl. Psychol. **88**:741–749 (2003)
27. Jehn K.A.: A multimethod examination of the benefits and detriments of intragroup conflict. Adm. Sci. Q. **40**(2) (1995)
28. Gottman J.M.: What Predicts Divorce? The Relationship Between Marital Processes and Marital Outcomes. Lawrence Erlbaum, New Jersey(1994)
29. Eris O., Leifer L.: Faciliating product development knowledge acquisition: interaction between the expert and the team. Int. J. Eng. Educ. **19**(1):142–152 (2003)
30. Schön D.: The Reflective Practitioner: How Professionals Think in Action. Basic Books, New York (1983)

Chapter 18
Measurement of Wakuwaku Feeling of Interactive Systems Using Biological Signals

Michiko Ohkura[1], Masahide Hamano, Hiroyuki Watanabe, and Tetsuro Aoto

Abstract To evaluate the kansei values of interactive systems, subjective evaluation methods such as questionnaires are commonly used, even though they have some drawbacks such as linguistic ambiguity and interfusion of experimenter and/or participant intention to the results. We began our research to objectively evaluate interactive systems by quantifying sensations using biological signals to redeem the above questionnaire drawbacks. We utilize biological signals to estimate participants' feelings of relaxation, comfort, and excitement, which are considered positive sensations. However, relaxation and comfort are considered static compared with dynamic feelings such as excitement. We focus on a positive and dynamic feeling called "wakuwaku" in this chapter, and construct various systems to evaluate the kansei values used to derive wakuwaku feelings using biological signals, in order to clarify the relationship between the wakuwaku feeling and biological signals. In addition, we derive a kansei model of interactive systems using biological signals to objectively evaluate their wakuwaku degree.

18.1 Introduction

Recently, the kansei value has become very important in manufacturing in Japan. Following function, reliability, and cost, the kansei has been determined as the fourth value axis of industrial products by the Japanese Ministry of Economy, Trade and Industry (METI). According to METI, it is important not only to offer new functions and competitive prices but also to create a new value to strengthen Japan's industrial competitiveness. Focusing on kansei as a new value axis, METI launched

[1] M. Ohkura
Shibaura Institute of Technology, 3-7-5, Toyosu, Koto-ku, 135-8548 Tokyo, Japan
e-mail: ohkura@sic.shibaura-it.ac.jp

the “Kansei Value Creation Initiative” in 2007 [1, 2] and held a kansei value creation fair called the “Kansei-Japan Design Exhibition” at Les Arts Decoratifs (Museum of Decorative Arts) at the Palais du Louvre, Paris, in December 2008. Launched as an event of the “Kansei Value Creation Year”, the exhibition had more than 10,000 visitors during its ten-day run and was received favorably [3].

To evaluate the kansei values of interactive systems, subjective evaluation methods such as questionnaires are commonly used, even though they suffer from the following drawbacks:

- linguistic ambiguity;
- interfusion of experimenter and/or participant intention into the results;
- interruption of the system’s stream of information input/output.

Solving these problems is crucial in order to evaluate the degree of interest in and/or excitement of a constructed interactive system, and to identify the moment of excitement. Evaluating the kansei value of an interactive system by using only subjective evaluation methods such as questionnaires is almost impossible.

We began our research into objectively evaluating interactive systems by quantifying sensations using biological signals that offer the following merits, and can be used to redeem the above questionnaire drawbacks:

- can be measured using physical quantities;
- avoids influence from the intentions of experimenter and participants;
- can be measured continuously.

Many previous researches have measured mental sensations using biological signals. Some researchers have used biological signals to measure mental stress or simulator sickness [4–7], which are considered negative sensations. On the other hand, some research have looked at relaxation or comfort [8, 9]. We have previously utilized EEG alpha waves to estimate participant feelings of relaxation [10]. Relaxation and comfort are considered to be positive sensations.

In this chapter, we focus on a feeling called “wakuwaku”, which is a Japanese word for the positive sensation caused when someone feels something exciting or captivating. The word means “thrilling” or “exhilarating” in English. A wakuwaku feeling is also considered a positive sensation, as are relaxation and comfort. However, a big difference exists between those sensations; a wakuwaku feeling is considered dynamic, especially compared to the static sensations of relaxation and comfort. Little previous research exists on positive and dynamic sensations such as the wakuwaku feeling.

The purposes of this chapter include the following:

- to clarify the relationship between the dynamic, positive feeling of the wakuwaku sensation and biological signals;
- to derive a kansei model of interactive systems using biological signals to objectively evaluate their wakuwaku degree.

Based on Russel's circumplex model [11], we propose a two-dimensional model called the wakuwaku model by the following procedures:

1. evaluation of the wakuwaku sensation by studying indexes;
2. construction of examples of interactive systems;
3. evaluation of the constructed systems;
4. derivation of the model.

18.2 Indexes of the Wakuwaku Model

Our survey for wakuwaku experiences included the following items:

1. How would you describe your wakuwaku experience?
2. Why do you think you felt wakuwaku?
3. How else would you describe your feeling except wakuwaku?

From the questionnaire answers of 176 student participants in their twenties, we extracted keywords related to the wakuwaku sensation, such as excitement, pleasure, happiness, enjoyment, astonishment, and fright. Since the most commonly related keyword to wakuwaku was excitement, we employed "exciting" as the horizontal axis to construct a two-dimensional model called the wakuwaku model. Although other candidates such as happiness and fright were also considered, we chose "enjoyable" as the vertical axis (Figure 18.1).

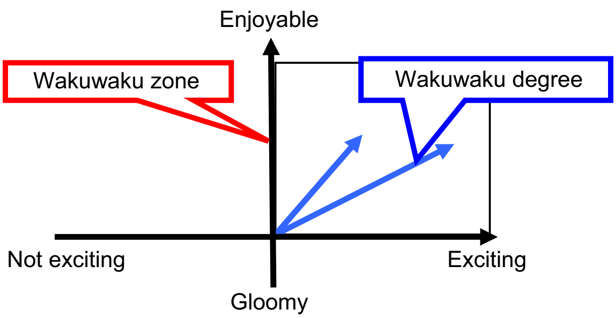


Figure 18.1 Wakuwaku model

18.3 Evaluation Experiment

18.3.1 Construction of Evaluation Systems

We constructed various systems based on a whack-a-mole game to evaluate their degrees of wakuwaku. After setting parameters and their factors (Table 18.1), eight types of systems were constructed (Table 18.2).

Table 18.1 Parameters and their factors

Parameters	Factor 1	Factor 2
Input device	Hammer	Keypad
Sound (BGM and effects)	With	Without
Target	Mole	Ball

Table 18.2 Constructed systems

System	Input device	Sound	Target
A	Keypad	With	Mole
B	Keypad	With	Ball
C	Keypad	Without	Mole
D	Keypad	Without	Ball
E	Hammer	With	Mole
F	Hammer	With	Ball
G	Hammer	Without	Mole
H	Hammer	Without	Ball

Figure 18.2 shows the system diagram for the interactive systems. For a wide viewing angle and high resolution, a 100-inch screen with rear projection was employed as the system's visual output device. The auditory output device was a pair of speakers. The input device was a hammer device or a keypad. Figure 18.3 shows the hammer, and Figure 18.4 shows a mole and a ball for the target. A PC was employed for the system display. The game took place as follows:

1. The experimenter explains the game (Figure 18.5 (a)).
2. After the game begins, players try to hit the targets with the input device (hammer or keypad) to increase their scores (Figure 18.5 (b)).
3. After a particular number of targets appear, the speed with which the target appears increases.
4. After another particular number of targets appear, the game is finished, and the results are displayed (Figure 18.5 (c)).

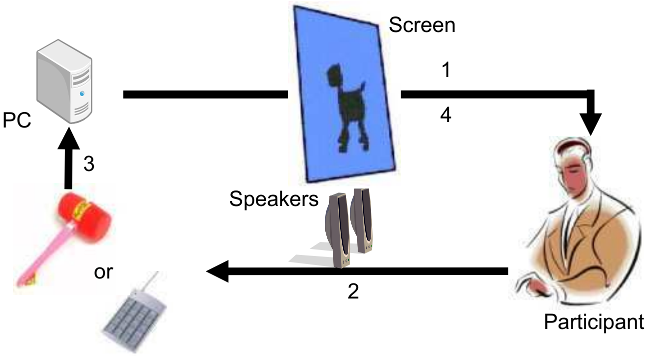


Figure 18.2 System diagram

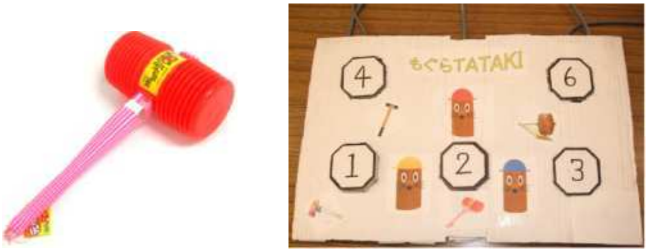


Figure 18.3 Hammer (left: hammer part, right: input part)

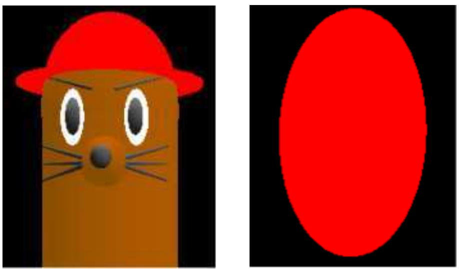
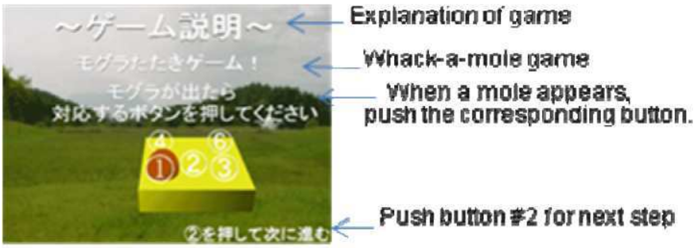
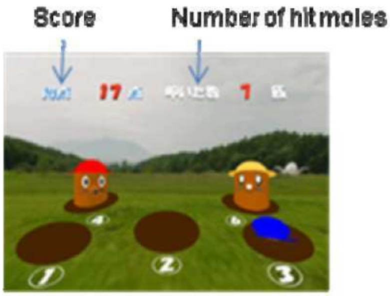


Figure 18.4 Targets (left: mole, right: ball)



(a)



(b)

(c)

Figure 18.5 Screen shots: (a) explanation, (b) start of the game, and (c) results

18.3.2 Experiment to Evaluate the Systems

A questionnaire and biological signals were employed to evaluate the wakuwaku degrees of the constructed systems. We employed two types of questionnaire: five-scale evaluation and free description. For the former, 30 items were employed such as “enjoyable”, “exciting”, and “interesting”. The free description questionnaire contained the following questions:

- When did you feel excited during the game?
- When did you feel enjoyment during the game?
- What do you consider the game’s exciting aspects?

- What are its good points?
- What are its bad points?

The following biological signals were measured: galvanic skin reflex (GSR), heart rate, and breathing rate. GSR, which is affected by states of emotion, was used as a physiological index to detect emotions such as anxiety and mental stress. Heart rate is changed not only by physical exercise but also by mental factors such as anxiety and stress. In addition, breathing rates and patterns can also be used as indexes of stress and anxiety. Specific sensors for each biological signal described above, BIOPAC measurement equipment (BIOPAC Systems, Inc.), and a PC were used to measure these biological signals.

The experimental procedure was as follows:

1. Participants sat on chairs.
2. The experimenter explained the experiment.
3. Participants wore electrodes.
4. Participants remained quiet for 30 s.
5. Participants played one of eight possible games.
6. Participants answered questionnaires.
7. Steps 5 and 6 were repeated four times.

The participants played four out of eight games per day and completed all eight games over two days. The biological signals were constantly measured during the experiments.

18.3.3 Experimental Results

Experiments were performed with eight male students in their twenties who served as volunteers. Figure 18.6 shows a scene of the experiment. The orders of the eight games were counter-balanced.

Figure 18.6 Experiment in progress

From the questionnaire results, averages scores for all 30 items were calculated. Figure 18.7 shows the result of "enjoyable" as an example of the averaged scores. These questionnaire results show that the existence of sound and the mole images of the target effectively increased the "enjoyable" scores. On the other hand, differences between the hammer and the keypad as input devices were not obvious.

Next, we did a three-factor analysis of the variance, with input device, sound, and image of target as factors. Table 18.3 shows the results for "enjoyable". Sound and target have significant main effects at a 1% level. Three-factor analysis results for the other items closely resembled this result.

The following results were confirmed:

- the degree of wakuwaku depends on both sound and target;
- the degree of wakuwaku is independent of the input device.

From the results of the free descriptions, the most common moments of excitement were as follows:

- when the mole's speed increased;
- when its movement got faster;
- when participants heard the sound effects.

The first two statements show the effectiveness of the speed change of the mole movement for increasing participant excitement. In addition, the third shows the effectiveness of sound for increasing excitement.

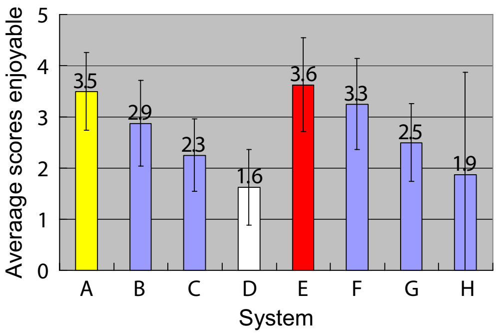


Figure 18.7 Questionnaire results (for "enjoyable")

Table 18.3 Analysis of variance (for "enjoyable")

Factor	Sum of squared deviation	DOF	Mean square	F-value	P-value
Input device	1.00	1	1.00	1.10	0.30**
Sound	27.56	1	27.56	30.41	0.00**
Target	9.00	1	9.00	9.93	0.00**
Error	54.38	60	0.91		
Total	91.94	63			

Like the free description results for the moment of enjoyment during the game, some participants cited the moment of hitting the moles, which shows that the hammer effectively promotes enjoyment. In addition, participants listed the following as the game's most popular good points:

- sound effectively increases the feeling of enjoyment;
- hitting the moles with a hammer is fun.

These answers suggest that the hammer as an input device, the existence of BGM and the sound effects, and the mole as a target all effectively increase the game's wakuwaku feeling – as we intended.

The effectiveness of the sound and the mole as a target were confirmed both by the five-scale evaluation and the free description. However, the effectiveness of the hammer as an input device was confirmed only by free description and not by the five-scale evaluation, as mentioned above. To solve this contradiction, the answers for the final free description question were very helpful because they provided a list of some of the bad points of the game.

These answers suggest the difficulty of using the hammer, which explains the contradiction between the five-scale evaluation and the free description, where the feeling of the enjoyment at the moment of using the hammer may be reduced by the difficulties of its continuous operation. This assumption might indicate a limitation of the five-scale evaluation.

18.3.4 Derivation of Wakuwaku Model

Based on our previous experiments [11–14], we selected the following physiological indexes for the model:

- average GSR;
- average heart rate;
- variance of heart rate;
- average R-R interval;
- variance of R-R interval;
- number of breaths;
- average breath magnitude;
- variance of breath magnitude.

Because the temporal change of GSR moves in one direction, the averaged GSR value reflects the magnitude and repetition of its temporal change. The averaged heart rate and its variance are defined as the average number of heart beats per minute and its variance, respectively. The average R-R interval and its variance are defined as the average interval time between the R-waves of the ECG and its variance, respectively. These indexes are known as the indexes for stress, uneasiness, or relaxation, and are used to measure dynamic feelings [11–14].

The number of breaths and the variance of breath are defined as breaths per minute and its variance, respectively. The breath magnitude is defined as the difference of the maximum and minimum values in one indrawn breath. As already mentioned, our wakuwaku model consists of “exciting” and “enjoyable” axes. We derived equations to explain both axes using a neural network as follows:

- *input*: all the physiological indexes listed above were calculated from when the speed of the target appearance increased at the end;
- *output*: averaged score of “exciting” for each system;
- *method*: clipping;
- *hidden layer*: one;
- *sample data for learning*: 60%.

Table 18.4 shows the neural network results for “exciting”. We expressed the “exciting” axis as follows:

$$\text{exciting} = 0.55 * \text{average breath} + 0.45 * \text{average heart rate}. \qquad (18.1)$$

In the same manner (see Table 18.5), the “enjoyable” axis was expressed as follows:

$$\text{enjoyable} = 0.71 * \text{average heart rate} + 0.45 * \text{variance of breath}. \qquad (18.2)$$

Table 18.4 Results of neural network (for “exciting”)

Psychological indexes	Relative importance
Average breaths	0.546 233
Average heart rate	0.445 581
Variance of heart rate	0.382 679
Average GSR	0.305 76
Number of breaths	0.288 106
Variance of breath	0.186 697

Table 18.5 Results of neural network (for “enjoyable”)

Psychological indexes	Relative importance
Average heart rate	0.709 154
Variance of breath	0.450 969
Average breaths	0.350 584
Average R-R interval	0.333 264
Number of breaths	0.266 689
Variance of heart rate	0.252 472

18.3.5 Confirmation of Wakuwaku Model

Using the two derived equations, our wakuwaku model was applied to the eight systems. Figure 18.8 shows the results for each system, where the values of the horizontal axis were calculated from Equation 18.1, and the values of the vertical axis were calculated from Equation 18.2.

System E, where the input device is a hammer, sound exists, and the targets are moles, has the highest scores both for "exciting" and "enjoyable". On the other hand, system D has the lowest score, where the input device is a keypad, sound doesn't exist, and the targets are balls:

$$E > G > F > H > A > B > C > D, \tag{18.3}$$

where the first four systems use a hammer as an input device.

From this result, the system with the most wakuwaku has the following combination of parameters:

- *input device*: hammer;
- *sound*: yes;
- *target*: mole.

The system with the least wakuwaku has the following combination of parameters:

- *input device*: keypad;
- *sound*: no;
- *target*: ball.

These results are the same as the five-scale evaluation results shown in Figure 18.7.

However, Figure 18.8 shows that the effect of a different input device is stronger than the effects of sound and different target images, which is completely different from the five-scale evaluation results (Figure 18.7). On the other hand, these results agree with the free description that shows that using the hammer effectively increased participant enjoyment. This fact suggests the appropriateness of our proposed model.

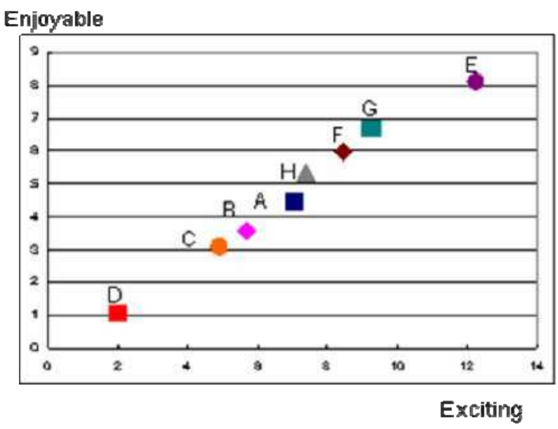


Figure 18.8 Results of wakuwaku model

Deriving a two-dimensional wakuwaku model based on biological signals using neural networks, we can reasonably explain the varieties of the wakuwaku degree of eight constructed systems (Figure 18.8). Moreover, the limitation of the five-scale evaluation as a questionnaire was overcome by employing this model, suggesting its effectiveness based on biological signals to qualify the degree of the wakuwaku feeling of the interactive systems.

18.4 Experiment 2

This section introduces another experiment to measure the wakuwaku feeling using biological signals.

18.4.1 System Construction

We constructed various systems based on a treasure chest game to evaluate the degrees of wakuwaku feeling. Compared with the previous experiment, these constructed systems have various complicated components to promote the wakuwaku feeling, such as the appearances of figures, their combination, and the actions of the combined figures. The parameters of these systems were the design of the boxes and the sound, the BGM and effects, as shown in Table 18.6. The three box designs are shown in Figure 18.9. The constructed systems are shown in Table 18.7.

Table 18.6 System parameters

Parameters	Factor 1	Factor 2
Box design	Designs[1]	White
Sound (BGM and effects)	With sound	Without sound

[1] Shown in Figure 18.9

Figure 18.9 Three types of boxes

Table 18.7 Constructed systems

System	Box design	Sound
A	Decorated	With
B	Decorated	Without
C	White	With
D	White	Without

The procedures of the game were as follows:

1. Confirm the figures in the boxes (Figure 18.10 (a)).
2. Choose one of the boxes (Figure 18.10 (b)).
3. Watch the figure in the chosen box (Figure 18.10 (c)).
4. Repeat the above procedures (Figure 18.10 (d)).
5. Watch the combinations of the two figures (Figure 18.10 (e)).
6. Watch the combined figure (Figure 18.10 (f)).

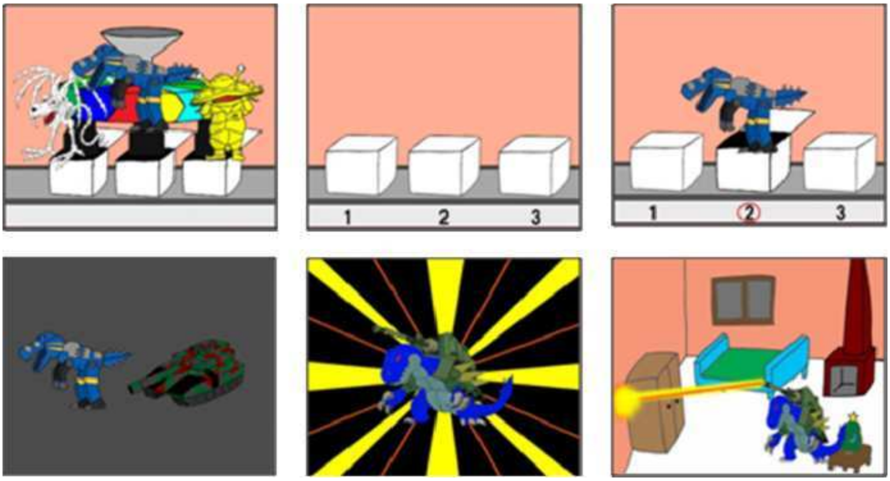


Figure 18.10 System flow

These procedures were designed to promote wakuwaku feelings when expecting a figure's appearance from the chosen box and when combining two figures. Questionnaires and biological signals were employed to evaluate the degree of wakuwaku feeling of each system.

Figure 18.11 shows the system diagram. The input device was a keypad, and the output devices were a 17-inch LCD display and a pair of speakers. Biological signals were measured by sensors and BIOPAC measurement equipment. Two PCs were employed for system display and to measure the biological signals.

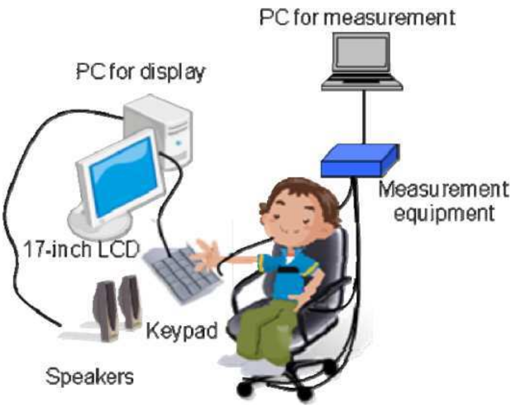


Figure 18.11 System setup

18.4.2 Experiment to Evaluate the Systems

18.4.2.1 Experimental Method

The participants randomly played the four games shown in Table 18.7 and answered questionnaires about each system. The questionnaire about the wakuwaku feeling consisted of 23 items of paired seven-point evaluations such as "fun–boring," and five items of unpaired five-scale evaluation, such as "pounding." For the paired seven-point evaluation items, 4 indicates neutral, 7 the best, and 1 the worst. For the unpaired five-point evaluation items, 5 indicates the best and 1 the worst. In addition, participants were asked some free description questions after playing the final game.

The following biological signals were measured constantly during the experiments to detect the degree of wakuwaku feeling: galvanic skin reflex (GSR), electrocardiogram (ECG), and breathing rate.

18.4.2.2 Experimental Results

Experiments were performed with 12 male students in their twenties who served as volunteers. From the results of the analysis of variance for each questionnaire item with parameters in Table 18.6, the main effect of sound was found to be significant for almost all questionnaire items, including "exciting" and "enjoyable." On the other hand, the main effect of the box design was not significant for almost all items. Table 18.8 shows the result of the analysis of variance for "enjoyable." In the free description answers, some participants pointed out that the BGM and the sound effects were good points of the system, suggesting that sound is effective for creating wakuwaku feelings.

For the biological signals, we selected the same physiological indexes as the previous experiment described above. Since we designed the game flow with various events to promote wakuwaku feelings, we chose the following three moments for analysis:

- *moment I*: when the first box opened;
- *moment II*: when the second box opened;
- *moment III*: just after combining the two figures.

The first and second moments were in the first half of the game, while the third moment was in the second part. Using a paired difference test, the heart rate at each moment of the first and the second choices of boxes was found to be significantly different between the systems with different box designs. On the other hand, the heart rate at moments I and II was not significantly different between the systems with and without sound. However, the averages of GSR at moment III were significantly different only between the systems with and without sound. Table 18.9 summarizes the results of all tests.

Table 18.8 Analysis of variance (enjoyable)

Factor	Sum of squared deviation	DOF	Mean square	F-value	P-value
Sound	20.02	1	20.02	14.47	0.00**
Box design	0.19	1	0.19	0.14	0.71
Error	62.27	45	1.38		
Total	82.48	47			

Table 18.9 Results of difference tests

Physiological index	Parameter	Moment I	Moment II	Moment III
Averages of GSR	Box design	–	–	–
Averages of GSR	Sound	–	–	*
Heart rate	Box design	**	**	–
Heart rate	Sound	–	–	–
R-R Interval	Box design	**	**	–
R-R Interval	Sound	–	–	–

–: Not significant, *: significant at 5% level, **: significant at 1% level

18.4.3 Discussion

The above experimental results suggest that heart rate and GSR averages may show the wakuwaku feeling of the users of interactive systems. Moreover, the heart rate results are related to the system's former part, and the results of the GSR averages are related to its latter part. Since the questionnaire results agreed with the results of the GSR averages and disagreed with the heart rate results, they might reflect the wakuwaku feeling of the latter part of the systems. The questionnaires may reflect the wakuwaku feeling of the system's last part because that is the only part the participants can remember.

18.5 Conclusion

To evaluate the kansei values of interactive systems, we utilized biological signals for objective evaluations. We evaluated wakuwaku, which is a positive and dynamic feeling in Japanese, of interactive systems using biological signals and proposed a two-dimensional model that followed these procedures:

- reviewed indexes to evaluate wakuwaku sensations;
- constructed several systems;
- conducted evaluation experiments of the constructed systems;
- derived a wakuwaku model using a neural network.

We designed and constructed eight different systems based on a whack-a-mole game for various wakuwaku degrees. From the experimental results, we derived two equations with biological indexes as parameters to form the wakuwaku model. The results of the model showed that the derived wakuwaku model based on biological signals reasonably explained the wakuwaku degree of the constructed systems.

We conclude that our proposed model can play a role in the evaluation of the kansei values of interactive systems from the aspects of positive and dynamic sensations by qualifying biological signals.

In addition, we performed another experiment to measure the degree of wakuwaku feeling by constructing systems based on a treasure chest game. From analysis of the experimental results, we obtained the following useful knowledge:

- the degree of wakuwaku feeling may vary depending on such parameters as object design and sound effects;
- the degree of wakuwaku feeling may be measured by biological signals such as GSR and ECG.

The measurement of such positive and dynamic feelings as wakuwaku derived by interactive systems is considered an evaluation of the kansei value of the interactive systems. Thus, these works above are the first step in objectively evaluating the kansei value of interactive systems by measuring biological signals. Future work will include more detailed research.

Acknowledgments This research was partly supported by the Shibaura Institute of Technology Research Promotion Funds. We thank the students of the Shibaura Institute of Technology who served as volunteers.

References

1. Ministry of Economy, Trade and Industry, Japan (METI): “Kansei” initiative – suggestion of the fourth value axis. http://www.meti.go.jp/english/information/downloadfiles/PressRelease/080620KANSEI.pdf (2008). Accessed June 20, 2008
2. Araki J.: Kansei and value creation initiative. J. Jpn. Soc. Kansei Eng. **7**:417–419 (in Japanese) (2007)
3. METI.: Announcement of “Kansei Value Creation Museum”. http://www.meti.go.jp/english/press/data/20090119_02.html (2009). Accessed January 19, 2009
4. Hirao N. *et al.*: Assessment of mental stress during a cooperative work using stress detection method triggered by physiological data. ITE Tech. Rep. **21**:67–72 (in Japanese) (1997)
5. Shimono F. *et al.*: Method for assessment of mental stress during high-tension and monotonous tasks using heart rate, respiration and blood pressure. Jpn. J. Ergon. **34**(3):107–115 (in Japanese) (1998)
6. Hioki K. *et al.*: Physiological evaluation of mental workload in time pressure. Trans. Inst. Electr. Eng. Jpn. C **127**:1000–1006 (in Japanese) (2007)
7. Ishii Y. *et al.*: Study on mental stress using near-infrared spectroscopy, electroencephalography, and peripheral arterial tonometry. In: Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2008, Vancouver, BC, pp. 4992–4995 (2008)
8. Omori M. *et al.*: Relation between psychological and physiological responses on color stimulus. J. Color Sci. Assoc. Jpn. **26**(2):50–63 (in Japanese) (2002)
9. Kawashima T.: Development of active rocking chair with changing swing according to heartbeat fluctuation. JSME Int. J. Ser. C **49**:814–820 (2006)
10. Ohkura M., Oishi M.: An alpha-wave-based motion control system for a mechanical pet. Kansei Eng. Int. **6**(2):29–34 (2006)
11. Aoki Y. *et al.*: Assessment of relief level of living space using immersive space simulator(report 4) – examination of physiological indices to detect “uneasiness”. In: VRSJ 11th Annual Conference (CD-ROM), vol. 11, art. 2B1-4 (in Japanese) (2006)
12. Ohkura M. *et al.*: Evaluation of comfortable spaces for women using virtual environment – objective evaluation by biological signals. Kansei Eng. Int. **8**(1):67–72 (2009)
13. Aoto T., Ohkura M.: Study on usage of biological signal to evaluate kansei of a system. In: Proceedings of the 9th Annual Conference of JSKE 2007 (CD-ROM), art. H-27 (in Japanese) (2007)
14. Aoto T., Ohkura M.: Study on usage of biological signal to evaluate kansei of a system. In: Proceedings of KEER2007 (CD-ROM), art. L-9 (2007)

Chapter 19
Footwear Fitting Design

Tsuyoshi Nishiwaki[1]

Abstract The fitting is one of the most important properties of footwear. The fitting property mainly depends on the upper material and structure. This indicates that an investigation of the deformation of the footwear upper is required in designing the fitting property. In this chapter a new quantitative evaluation method of the fitting property is introduced. This evaluation method is based on the difference of strain distributions in foot and upper. First, the strain distribution on a bare foot is measured under the heel up condition using a non-contact typed three-dimensional strain measuring system. The secondary upper strain distribution is also measured by the same method. Based on the difference between the two, the *fitting parameter* is quantitatively defined. Through the above measurements it was found that the strain distributions on a bare foot are not uniform and are different on the medial and lateral sides. Considering these barefoot results, as an example of practical footwear design, a new upper structure is also introduced.

19.1 Introduction

Footwear is an essential component of sports activities as well as in daily life. According to the purpose for which they are intended, items of footwear must have various structures. Running shoes mainly have EVA(ethylene-vinyl acetate) foam soles and PE(poly-ethylene) mesh uppers. Sprint spikes have hard plastic plate soles with metal pins and artificial leather uppers. These footwear structures depend on the individual properties required. In running shoes, shock attenuation, excessive motion control, and breathability are very important properties for comfortable long-term running and the prevention of running injuries. In sprint shoes,

[1] T. Nishiwaki
ASICS Corporation, 6-2-1, Takatsukadai, Nishi-ku, Kobe, 651-2271, Japan
e-mail: nishiwaki-ts@asics.co.jp

motion control and grip properties are the most important in order to produce an effective propulsion kicking force. Compared with casual shoes, the property requirements in sports shoes have been clarified through study, which means motion analysis in sports scenes and case analysis for the various sports injuries. In sports shoes, the most popular footwear are running shoes, which are predominantly used in long-term running with relatively low speeds.

In this chapter, running shoe construction and the properties required for this are first introduced and the importance of fitting is discussed. Second, the conventional optical measurement technique which is effective for measuring shoe upper deformation is described and a new fitting quantitative evaluation method based on the measurement technique is introduced. Finally as an example of practical design, some running shoes with excellent fitting are introduced.

19.1.1 Running Shoe Construction

Figure 19.1 shows a lateral side view of a standard running shoe with the parts named. The general running shoe can be divided into two major parts, sole and upper. Moreover, the sole can be divided into two portions, midsole and outersole. The midsole and outersole have important functions which produce high shock attenuation and grip on surfaces, respectively. To meet these functional and dura-

Figure 19.1 Running shoe construction with names of parts

bility requirements, the midsole is made from ethylene vinyl acetate foam (EVA foam) [1–3]. In order to improve the shock attenuation, high-damping materials such as gel are often inserted into the midsole. In the midfoot area, plastic reinforcement called a trusstic is inserted. This reinforcement is effective for controlling excessive foot joint deformation [4–7]. In the outer soles, butadiene rubber blended with natural rubber is dominant. To reduce the weight of the shoe, rubber foam is also used. The upper contains a number of separate parts, such as toe reinforcement, eyelet stays, heel counter, and so on. Polyethylene mesh, artificial leather, and local resin reinforcements are also applied. Runners can control the tightness of the shoe by changing the shoelace binding through eyelets. The heel counter is a resin device inserted at the back of the upper, in order to control excessive heel motion during the stance phase in running.

19.1.2 Property Requirements

A running shoe has three main important roles:

- injury prevention;
- enhancement of the runner's performance;
- maintenance of the in-shoe condition.

In order to satisfy these roles, eight important functions for running shoes are defined, and these are listed in Table 19.1. On the whole these properties are very important in the design of high-performance running shoes. However, the priorities of these eight properties depend on running conditions, which means the skill of runners, running speed, surface condition, and so on. As an example, high cushioning, to effectively absorb the ground reaction force during the contact phase in running, is very important for preventing running injuries. High cushioning is caused by large midsole deformation, which is due to large deformation, large strain energy changes in thermal energy. This thermal energy is the factor that

Table 19.1 Eight major properties required in running shoes

Property	Function	Main design target
Flexibility	Curvature matching in kicking phase	Mid and outer sole in forefoot area
Stability	Control of excessive foot motions	Midsole
Cushioning	Shock attenuation in heel contact phase	Midsole
Breathability	Control of in-shoe temperature and humidity	Upper
Light weight	Reduction in shoe weight	Whole
Durability	Increasing usable period	Outer sole and upper toe area
Grip	Prevention from fall accidents	Outer sole
Fitting	Production of foot comfort	Upper

produces shock attenuation. However, the large deformation has a negative impact on high-speed running. In the design process of shoes for high-speed running, such as marathons and road relay races, designers must provide a smooth guidance from heel contact to toe off. However, the large midsole deformation interrupts the smooth guidance. Therefore, marathon shoes have a thinner midsole than running shoes. It is said that cushioning isn't too important for highly skilled runners like marathon runners, but is more important for less skilled runners like joggers. Therefore, the importance of cushioning depends on the runner's skills. Similarly, stability which controls excessive foot motion also depends on runner's skills. On the other hand, durability and fitting are important properties for every runner.

19.1.3 Running Shoe Design Flow

In this section, the general design flow for sports shoes is explained. Figure 19.2 shows the flowchart. At first, motion analysis for the target product is carried out using various instruments, such as motion capture, high-speed video, force platforms (ground reaction force measurement system), accelerometry, electromyog-

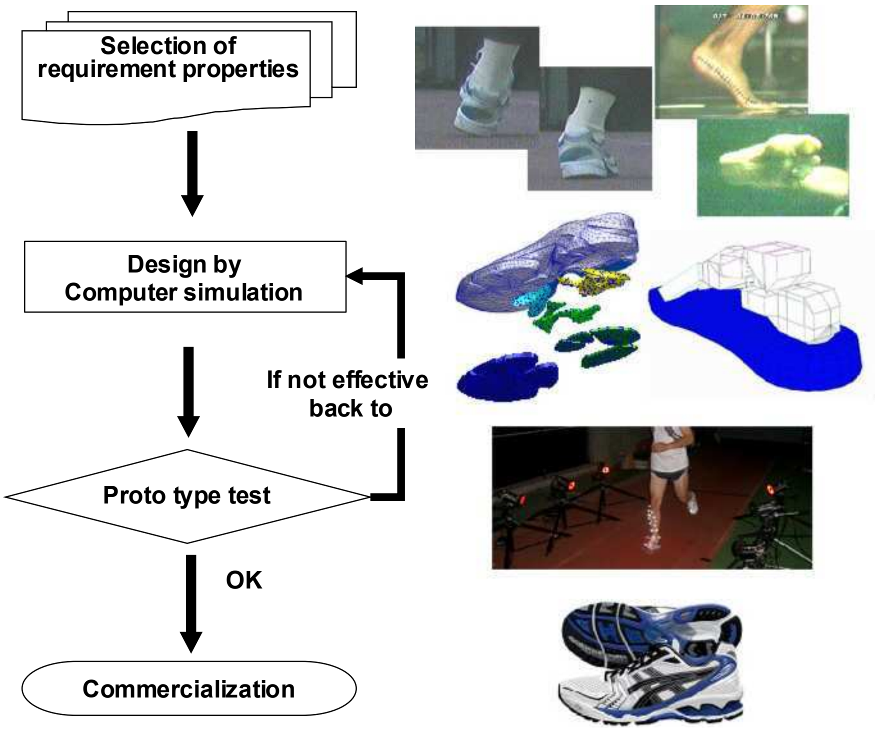


Figure 19.2 Footwear design flowchart

raphy, and so on. Through this motion analysis, runners' responses can be widely investigated and the required properties selected. In the case of running shoes, the most important properties are stability, forefoot flexibility, cushioning, and fitting. Then the sole and upper materials are designed or selected from conventional materials in order to optimally produce the properties required.

Based on the material properties decided, numerical simulations such as the finite-element method are performed, and the shoe materials and structure are decided accordingly. Prototype shoes are practically fabricated in the factory and the required properties are evaluated. Here the evaluation methods should be quantitative, because performing only sensual tests cannot communicate improvement points to the designers. Namely, the establishment of quantitative evaluation method for shoe function is a key part of designing better shoes. Moreover, a validity check of the evaluation method is also important. However, wearing tests are essential, because there is a limit to the quantitative engineered evaluation method in the laboratory – shoes made for the Japanese don't always work in America, Australia, and so on. Once the validities of the new shoes' properties are confirmed, they are shipped to the market.

In these required properties, quantitative evaluation methods for cushioning [1], stability [8, 9], and flexibility [10] have been already reported. In the next section, the meaning and design of the key property of fitting will be discussed.

19.2 Fitting

For anything which touches a human body, *e. g.*, glasses, clothes, socks, hat, chair, bed, shoes, the fitting property is one of the most significant properties in the object's design. In running shoes, systematic research into the fitting property has not been reported. The main reason for this can be seen by asking yourself, "What is the fitting?". If we ask about the fitting of running shoes, we are surprised at the variety of the answers. Some runners focus on foot arch support as an index of fitting, whilst others focus on the holding function around the heel. The indices of shoe fitting evaluations are infinitely variable – it is impossible to consider them all in the shoe fitting design process. In order to improve the fitting property in current running shoes, a new solid fitting index and quantitative evaluation techniques should be established.

It has been considered that the conventional fitting design technique is equivalent with last design. Here, the last is an aluminum and/or plastic foot shape which is used in the shoe fabrication process to keep a good silhouette. Figure 19.3 shows some examples of last geometries. To correspond to various runners with different foot shapes, shoes with different length and width are produced. It is very important for each runner to select the shoe that closely matches his or her foot shape. Recently, the design of personal lasts has been an expanding market, as shown in Figure 19.4. In this design, the runner's foot geometry is first measured by a laser scanning system. Based on the geometry scanned, a personal last is automatically produced. However, this is not a perfect solution. As shown in Figure 19.5, a large deformation in not only the sole but also the upper can be observed, especially under the heel rise condition. Upper deformation, *i. e.*, wrinkling, has a negative influence on the fitting properties because runners feel local pressure due to line contacts between the boot and upper. This means that upper deformation caused by running motion must be considered when designing running shoes to have a better fitting property.

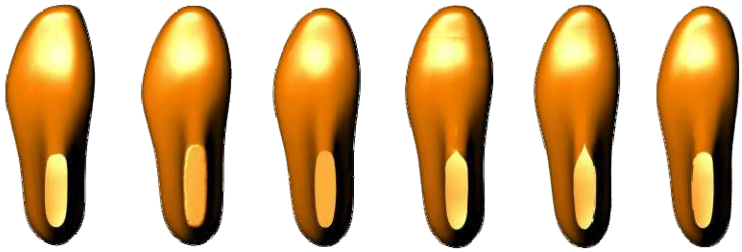


Figure 19.3 Examples of last geometry. To correspond to various runners with different foot shapes, shoe manufactures prepare lasts of various shapes. Starting from the left: oblique type, racing wide, racing regular, super wide, regular, and ladies wide

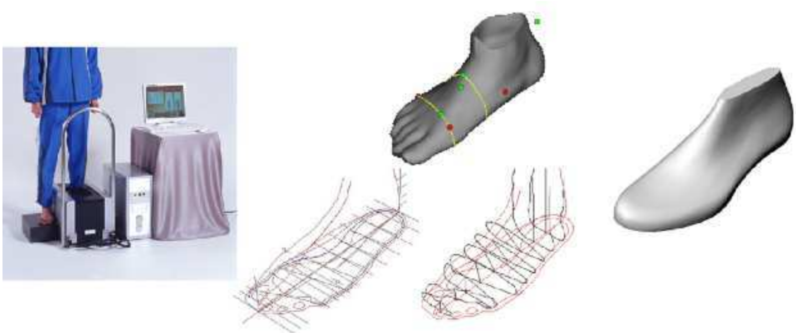


Figure 19.4 Fabrication process of a personal last. Based on the surface data in the CAD system, a personal last can be produced

Figure 19.5 Example of deformation of running shoe upper just before toe off. Some wrinkles can be observed in this magnified photo

Therefore, it is clear that measurement of the upper deformation will be an important factor in running shoe fitting design. As already mentioned, most of the upper is fabricated from a textile such as PE double-russell knitted mesh with high breathability. However, the measurement as well as predictions of the textile structural behavior is difficult to achieve. To be precise, a textile can be considered to be a hybrid material composed of fiber and air. As the friction between the components is very low, textile deformation is very large under small external loading. At the same time, most textiles have severe anisotropic and nonlinear properties. Conventional measurement techniques such as the strain gauge method and computer simulation such as the finite-element method have a poor predictive performance for this measurement. In the next section, a method that uses optical instruments to measure the strain distribution in the whole view field on the knit structure, which is effective for running shoe fitting design, will be described.

19.3 Optical Measuring Instrument

19.3.1 Current Technique

For the measurement of the strain distribution in shoe uppers, a non-contact and full-field typed instrument is required. In other words, point-by-point measurement such as using a strain gauge is not adequate. To satisfy the above conditions. an optical measurement system has been proposed. For example, interferometries using moiré technique, holography, speckle pattern, and photo geometry are well known [11, 12]. Electronic Speckle Pattern Interferometry (ESPI), which has excellent accuracy and resolution, is a powerful tool for the measurement of material surface deformation. However, the strain range that can be measured is limited to only a small percentage, which, when compared with the practical deformation of shoe uppers, is too small to be useful. On the other hand, the grid method is better adapted for the measurement of large deformations. Moreover, this method can be used at high temperatures, especially those associated with the rapidly changing strain field around the crack tip in a structure. It is well known that a very large strain field is caused at the crack tip under various external loadings. The biggest disadvantage of this method is that the grid must be precisely drawn on the measurement target surface. In the case of running shoes uppers, the grid drawing is very hard work.

19.3.2 New Technique

19.3.2.1 Fundamental Principles

Against the background of the issues about current optimal measurement techniques, some new techniques have been developed. One typical example is Aramis® (GOMmbH, Braunschweig, Germany [13]). Aramis® is a non-contact strain distribution measuring system based on image-processing technology, and has the dual advantages of easy handling in the ESPI and the capacity to respond to large deformations in the grid method.

The fundamental principle of Aramis® is based on the fact that the distribution of grayscale values in a rectangular area, a so-called facet, in the non-deformed state corresponds to the distribution of grayscale values of the same area in the deformed state, as shown in Figure 19.6 (a) and (b), respectively. The surface of the measurement target must exhibit a pattern in order to unambiguously match the pixels of the recorded images. A pixel in the reference image can thus be matched to the corresponding pixel in the destination image. This pattern can be created by alternate spraying with black and white ink.

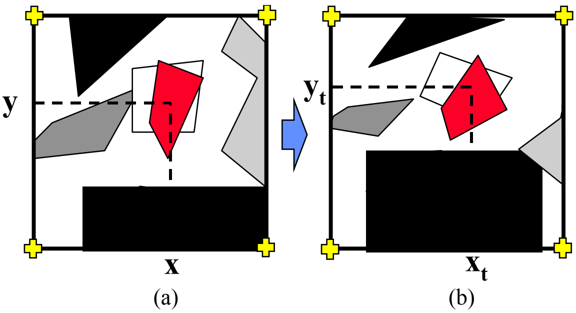


Figure 19.6 Facets in the non-deformed and deformed state: (a) before, and (b) after deformation. From the comparison of grayscale patterns between (a) and (b), the system can recognize that these are in the same position

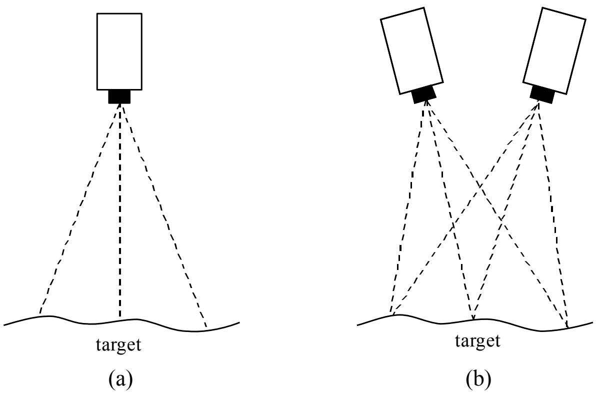


Figure 19.7 CCD camera configuration examples: (a) two-dimensional measurement, and (b) three-dimensional measurement

Figure 19.7 shows the CCD camera configurations. In two-dimensional strain distribution measurement, one CCD camera is used. In three-dimensional measurement, which considers out-of-plane deformation, two CCD cameras are needed.

The facet is a minimum unit in strain calculation. Namely, one facet is equivalent to one element in the finite-element method. The strain is calculated from the displacement between facets on non-deformed and deformed targets, as shown in Figure 19.6 (a) and (b), respectively. The most important point is how to match these two facets with the same position in the global system. The accuracy of this matching in real applications lies between 0.1 and 0.01 pixels. Considering continuity through so-called multi-facet matching, facets in non-deformed and deformed targets can be easily matched [14].

19.3.2.2 Measurement Example

In order to explain the above measurement procedure and check its accuracy, a simple example is introduced. The target specimen shown in Figure 19.8 is a dog bone shaped EVA foam sheet with 15 mm width and 2 mm thickness. Considering the large deformation, the specimen has a central hole with a diameter of 6 mm. Crosses in this figure indicate square facet corners. In this case, one dimension of each facet is 17 pixels. By using an Instron® testing machine with computer control (4204 type Instron, MA, USA), the specimen is stretched until 20% tensile strain is reached. To check the accuracy, finite-element analysis is carried out. Figure 19.9 shows comparisons of y-directional and xy-shear strain distributions obtained from Aramis® and FEM. It can be confirmed that both have almost the same distributions. Higher normal and shear strain components are demonstrated in the vicinity of the center hole in not only Aramis® but also in the FEM results. From these comparisons it can be confirmed that Aramis® can calculate the strain distribution with satisfactory accuracy under high strain conditions.

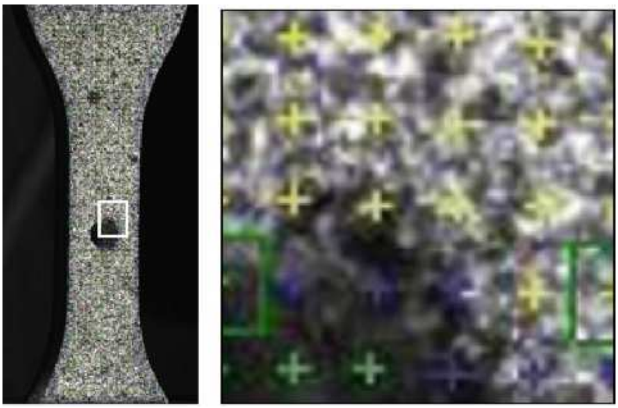


Figure 19.8 Test specimen sprayed and close-up picture at the vicinity of central hole. This is a tiff format file

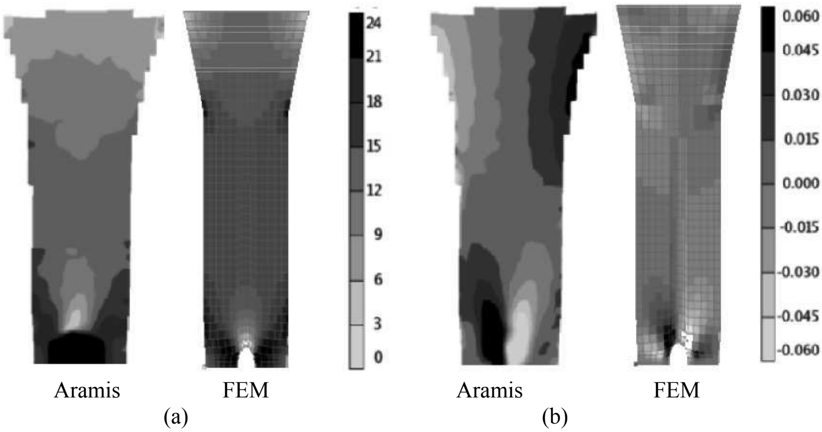


Figure 19.9 Comparisons of longitudinal normal and in-plane shear strain distributions in Aramis and FEM: (a) y-directional normal strain, and (b) xy-shear strain

19.4 Evaluation Method of Shoes Fitting

As already mentioned, conventional design of the fitting property is equivalent to the geometrical design of a shoe last. This design can only correspond to the fitting when standing. Therefore, some wrinkles as shown in Figure 19.5 will be caused when running, and this is considered to be an unavoidable problem. Moreover, the fitting property is dependent on runners' individual feeling. In order to produce running shoes with a better fitting property, a new evaluation index and a method of quantitatively evaluating it must be established. As a new index of the fitting property, we have focused on the reduction of wrinkles. It is clear that the wrinkles have a negative influence on the fitting property, because line contacts on the foot due to the upper wrinkles are associated with a feeling of discomfort. These wrinkles are mainly caused by the difference in foot and upper deformations. In this section, a new quantitative evaluation method for the fitting property is introduced which focuses on the reduction of wrinkles in running [15].

19.4.1 Definition of Measurement Areas

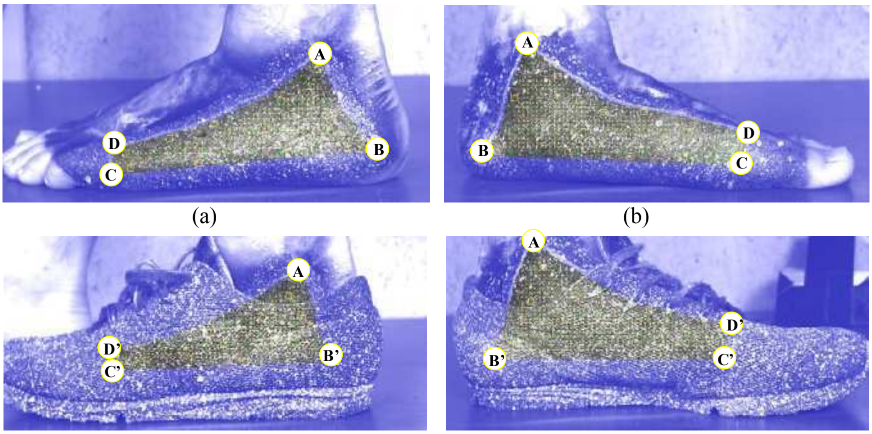
In order to reduce wrinkles, deformation matching between the bare foot and upper must be considered. In the case where both match, the strain distributions should also match. Namely, the strain components of two points on the bare foot and upper with the same coordinate should be compared. First, a volunteer's left foot and a test shoe (Figure 19.10) upper are sprayed with black and white inks.

As already mentioned, this gray pattern is absolutely necessary to the strain measurement using the Aramis® system. As shown in Figure 19.5, the wrinkles tend to occur in both the lateral and medial sides. In other words, the toe and heel areas do not wrinkle in running.

(a) (b)

Figure 19.10 Test shoes used: (a) lateral side view, and (b) medial side view. This is a commercial running shoe

To compare both strain distributions, the quadrilateral measurement area whose corners are lateral malleolus (A in Figure 19.11 (a)), calcaneus (B in Figure 19.11 (a)), fifth metatarsal head (C in Figure 19.11 (a)), and side surface (D in Figure 19.11 (a)) on the subject's foot is defined. Next, the measurement area constructed by four points, A, B′, C′, and D′, is also defined on the shoe upper as shown in Figure 19.11 (c). Here two measurement areas, ABCD and AB′C′D′, have the same geometry and area. By using these area definitions it is possible to compare the strain values at the same coordinates in the measurement area. The same procedures are also carried out on the medial side.



(c) (d)

Figure 19.11 Target area definitions: (a) target area on barefoot lateral side, (b) target area on barefoot medial side, (c) target area on upper lateral side, and (b) target area on upper medial side

19.4.2 Measurement

Figures 19.12 and 13 show the measurement condition and CCD camera configuration. In this system, the barefoot and upper strain distributions under 45 mm heel up conditions are measured on both the lateral and medial sides. In the barefoot measurement, it is defined that the foot flat condition has a strain field of zero. In the upper measurement, it is defined that the foot flat condition without shoelace binding has a strain field of zero. Here, a 45 mm heel up condition corresponds to the heel rise phase when most upper wrinkles are observed.

(a) (b)

Figure 19.12 Photograph of CCD camera configuration: (a) foot flat condition, and (b) heel rise condition

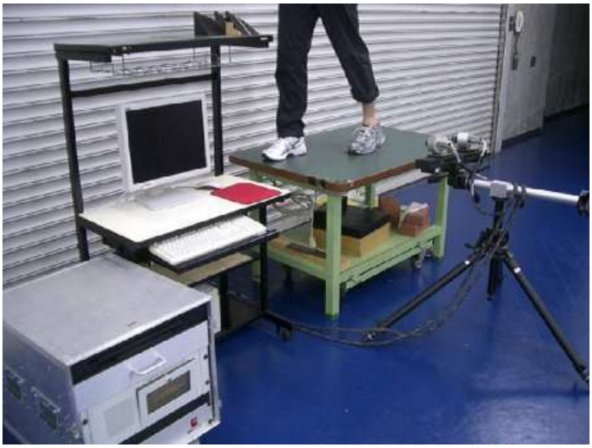


Figure 19.13 Measurement conditions

19.4.3 Measurement Results

The normal strain distributions of the *x* and *y* components in the measurement area on the bare foot are shown in Figure 19.14 (a) and (b), respectively. Here, the *x* and y directions denote the foot longitudinal and height directions, respectively. Figure 19.14 (c) and (d) show the strain distributions on the upper obtained from the conventional running shoes illustrated in Figure 19.10. Comparing the barefoot strain distributions with those on the upper, we can see large differences in both the x and y strain components. It is supposed that these different distributions cause the upper wrinkles.

For production of running shoes with better fitting, these difference should be reduced. At the same time, these difference can be an index for the evaluation of shoes fitting properties.

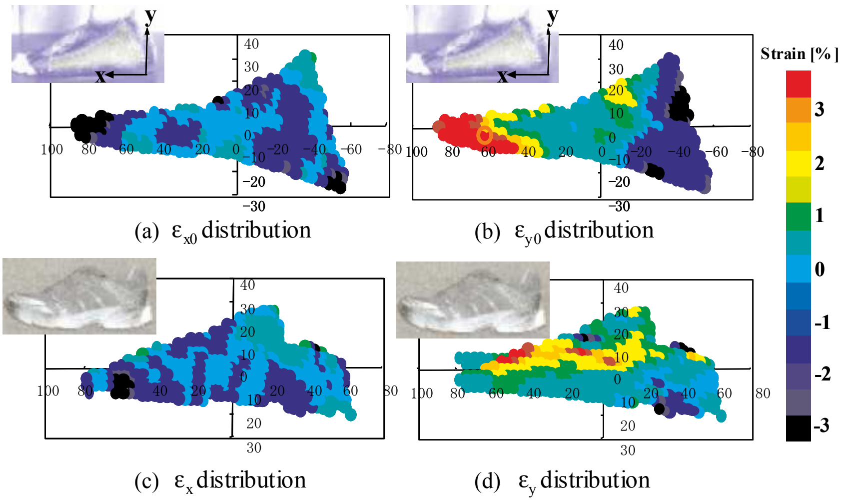


Figure 19.14 Comparison of normal strain distributions in lateral side: (a) normal x-strain on bare foot, (b) normal y-strain on bare foot, (c) normal x-strain on upper, and (d) normal y-strain on upper

19.4.4 Fitting Quantitative Evaluation Method

This section describes a quantitative evaluation method of the differences between the strain components at the same coordinate points on the bare foot and upper. As two identical measurement areas were individually defined, we can easily search the same coordinate points on foot and upper. It was defined that the *fitting parameter* (FP), which can be an index of shoe fitting, could be calculated from the following equation:

$$\mathrm{FP} \equiv \sum_{i=1}^{N} \left[\frac{\left|(\varepsilon_{x0})_i\right| \sqrt{\left\{(\varepsilon_x)_i - (\varepsilon_{x0})_i\right\}^2}}{\sum_{i=1}^{N'} \left|(\varepsilon_{x0})_i\right|} + \frac{\left|(\varepsilon_{y0})_i\right| \sqrt{\left\{(\varepsilon_y)_i - (\varepsilon_{y0})_i\right\}^2}}{\sum_{i=1}^{N'} \left|(\varepsilon_{y0})_i\right|} \right]. \quad (19.1)$$

Here, N is the total number of comparable points, i, in measuring areas ABCD and AB′C′D′, and $(\varepsilon_x)_i$ and $(\varepsilon_{x0})_i$ are the normal strain components at point i on the shoe upper and foot in the x-direction. According to this formula, FP is the non-dimensional value which is the weighted strain difference between the bare foot and shoe upper on the same coordinate. Shoes with large FP, which means there is a large strain difference between the bare foot and upper, produce poor following properties to the bare foot when running. Namely, a large FP indicates a large difference between the barefoot and upper deformation states and a poor fitting property. On the other hand, a small FP indicates a better fitting property. The FPs obtained from the test shoes in Figure 19.10 are shown in Figure 19.15. The total point numbers used in the comparison, N in Equation 19.1, are 423 on the lateral side and 392 on the medial side. From this result, it can be seen that these running shoes have a better fitting property on the medial side than on the lateral side.

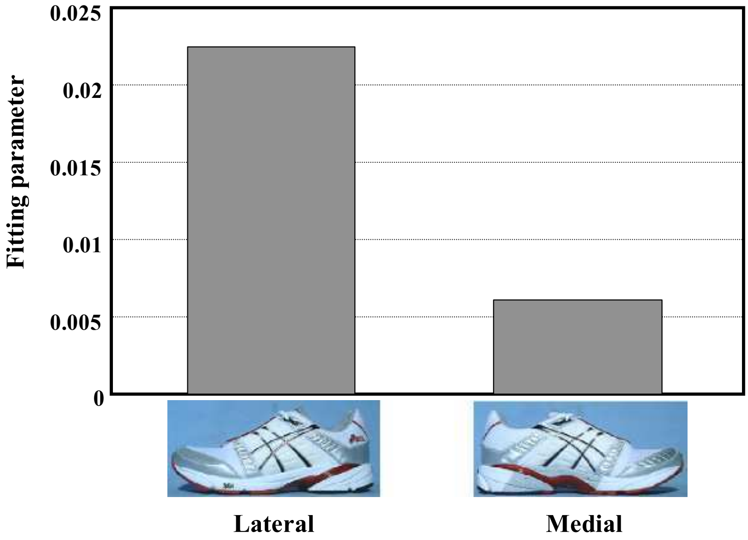


Figure 19.15 FP comparison between medial and lateral side obtained

19.4.5 Validity Checking

In order to check the validity of the FP value, FPs of five kinds of shoes with different upper structure and material (Figure 19.16) are calculated, and compared with the sensual testing results for fitting. Five subjects wore the above shoes and ranked the five shoes in turn. It was found that all subjects indicated the same ranking order, with type-5 shoes having the best fitting and type-2 with the worst. In all shoes, the FP on the lateral side is always larger than that on the medial side. In other words, all shoes have a poorer fitting property on the lateral side. Figure 19.17 shows the comparison results with the FP values calculated on the lateral side. Both results have the same tendency. This indicates that the FP calculated using the above method can be an important index to the quantitative fitting evaluation.

Figure 19.16 Test shoes with different upper material and structures. Shoes are numbered type-1 to 5, beginning at the top

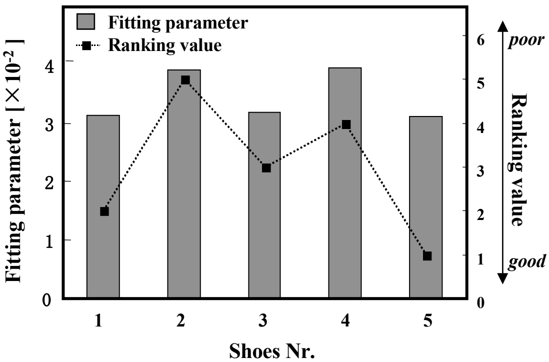


Figure 19.17 Comparison between lateral FP and sensual testing results. Bar chart denotes FP and plot denotes sensual testing results. In both results, smaller values mean better fitting property

19.5 Design Examples

In this section, a practical fitting design is introduced. By using the above measurement method, a number of barefoot strain distributions were measured. It was reconfirmed that the strain distributions on both the lateral and medial sides were quite different and not uniform. As shown in Figure 19.5, wrinkles in the vicinity of the metatarsophalangeal joint are caused on the conventional running shoe upper in running. As already mentioned, the upper material are mainly knitted fabric, which has non-zero tensile and zero compressive moduli. From these phenomena, it was found that the main reason for the wrinkles which reduce the fitting property was derived from the conventional unified one-piece upper. Figure 19.18 shows the major principal barefoot strain distributions on both sides. From these strain distributions, the existence of the borderlines between large and small strain fields on both the medial and lateral sides could be confirmed. In the case where there is an interruption in the strain continuity on the upper in the longitudinal direction, wrinkles can be eliminated and the upper fitting can be improved. In order to interrupt it, stretchy sheet materials are inserted in both sides. It is important to note that the insertion position should correspond to the borderline positions between the small and large strain areas shown in Figure 19.18. Figure 19.19 has an expansion view of a new upper with stretch material. In the figure, dotted bold lines denote the stitching lines. By seaming them each other, this two-dimensional expansion view changes to a three-dimensional upper. Figure 19.20 shows the finished product fabricated from the expansion.

The upper is mainly divided into two pieces and these pieces are connected with stretch material at the borderline positions. Vertical lines in the stretch material are resin reinforcements. As already mentioned, running shoes have many function requirements – not just fitting. Judging from other function evaluations, it is confirmed that only inserting stretch materials has a negative effect on stability. Stretch materials have a role in interrupting the strain continuity in the longitudinal direction. In order to simultaneously satisfy the requirements of high stability and better fitting property, an anisotropic stretch material, with high and low elas-

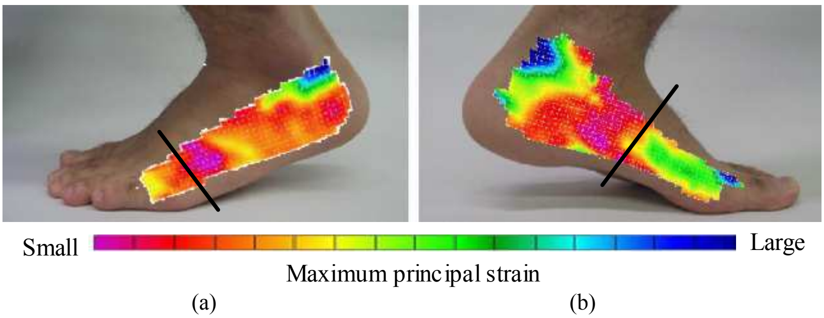


Figure 19.18 The results of the barefoot distribution in running: (a) medial side, and (b) lateral side. Lines denote the borderline positions in both the sides

tic moduli in the vertical and longitudinal directions respectively, is inserted. In the practical wearing test, many runners from all over the world confirm the wrinkle reduction and the better fitting property. From FP comparison results with that of conventional shoes, it was found that the FP calculated from these new running shoes fell by 25 and 50% on the medial and lateral sides, respectively.

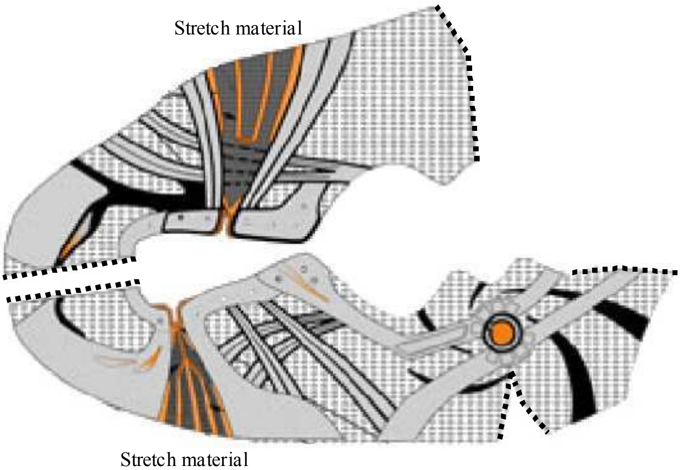


Figure 19.19 Development view of the upper with better fitting

(a) (b)

Figure 19.20 New product with improvement of the fitting property: (a) medial side view, and (b) lateral side view

19.6 Summary

As an example of footwear fitting design, the design process in running shoes was introduced. In the past, the fitting property in footwear has been subjectively evaluated using wearing tests. However, it is impossible to continue to improve the fitting in footwear just using wearing tests, because footwear designers cannot find out the direction the fitting needs to be improved.

In this chapter a quantitative evaluation technique based on the difference between barefoot and upper strain distributions was introduced using a new optimal strain measurement system. After checking the validity of a new index, the *fitting parameter*, a practical design example of a new product with better fitting was also introduced.

Figure 19.21 shows a pair of commercially supplied running shoes from 1965. Since then sports shoes have had a slow but large evolution year by year. Many new materials have been developed and new measurement techniques and evaluation methods for the function requirements have been also developed. You can easily see the big difference between the two running shoes shown in Figures 19.19 and 19.20. However, the biggest problem remaining is the fitting design. The design technique introduced still has problems to be worked out; however, it is a big step in the evolution of footwear fitting design.

Figure 19.21 Running shoes developed in 1965

References

1. Nishiwaki T., Nakabe N.: Numerical design of running shoes based on human response. Eng. Sport **4**:393–399 (2002)
2. Nigg B., Bahlsen A., Luethi M., Stokes S.: The influence of running velocity and midsole hardness on external impact force in heel-toe running. J. Biomech. **20**(10):951–959 (1987)
3. Yingling R., Yack J., White C.: The effect of rearfoot motion on attenuation of the impulse wave at impact during running. J. Appl. Biomech. **12**:313–325 (1996)
4. Wit D., Clercq D., Aerts P.: Biomechanical analysis of the stance phase during barefoot and shod running. J. Biomech **33**:269–278 (2000)
5. Stacoff A. *et al.*: Effects of shoe sole construction on skeletal motion during running. Med. Sci. Sports Exerc. **33**:311–319 (2001)
6. Nigg B. *et al.*: Effects of arch height of the foot on angular motion of the lower extremities in running. J. Biomech., **26**:909–916 (1993)
7. van Woensel W., Cavanagh P.: A perturbation study of lower extremity motion during running. Int. J. Sports Biomech. **8**:30–47 (1992)

8. Nakabe N., Nishiwaki T.: Development of simplified numerical foot model for sole stability design. Eng. Sport **4**:824–830 (2002)
9. Nakabe N., Nishiwaki T.: Evaluation of shoe sole stability based on knee joint kinetics. Impact Tech. Sport **1**:499–504 (2005)
10. Nishiwaki T.: Running shoe sole stiffness evaluation method based on eigenvibration analysis. Sports Technol. **1**(1):76–82 (2008)
11. Kobayashi A.: Handbook on Experimental Mechanics, 2nd edn. Wiley, New York (1993)
12. Rastogi P., Inaudi D.: Trends in Optical Non-Destructive Testing and Inspection, illustrated edn. Elsevier Science, Amsterdam (2000)
13. http://www.gom.com/EN/index.html
14. http://www.gom.com/EN/measuring.systems/aramis/system/workflow.html
15. Nishiwaki T.: New fitting evaluation method based on barefoot deformation. Proc. Footwear Symp.**7**:36–37 (2005)

Chapter 20
Psychological Teamology, Emotional Engineering, and the Myers–Briggs Type Indicator (MBTI)

Douglass J. Wilde[1]

Abstract This chapter summarizes the author's *Teamology: The Construction and Organization of Effective Teams*, and also describes later advances that simplify and clarify the theory. First, it develops the personality theory of psychiatrist C.G. Jung, which theorizes that people solve problems by eight mental processes called "function-attitudes" or "cognitive modes". Then it discusses how to determine these preferred modes for each person in a group to be formed into teams. Next, ways are suggested of using this modal information to construct teams covering as many of the eight modes as possible. Finally, a novel graphical way is displayed that organizes each team to focus on every mode while reducing duplication of effort.

20.1 Introduction

This chapter was originally intended to be exclusively about psychological teamology: the construction and organization of effective teams [1]. The word "psychological" is inserted here because by itself the word "teamology" is often used, especially in England, to describe team-building exercises rather than team construction and organization. After *Teamology* was published, it was soon realized that using data from the Myers–Briggs Type Indicator (MBTI) [2] to form teams was more effective than merely using the MBTI categories (*E*xtraversion, *Introversion*, *S*ensing, *iNtuition, T*hinking, *F*eeling, *J*udgment, and *P*erception) because twice as many Jungian cognitive mode preferences [3] can thereby be identified. Also the occasional attitude error [4] of categorical type dynamics can be avoided. This should be of great importance, not only to the many certified type profession-

[1] D.J. Wilde
Stanford University, Mechanical Engineering, Stanford, CA 94305 USA
e-mail: wilde@stanford.edu

als trained in this incomplete and sometimes incorrect method of analysis, but also to the many more school counselors and educators who use the MBTI for teaching and student career guidance.

Therefore, while a book is in preparation to cover the subject more clearly, this chapter, which was originally on teams, has been supplemented with the results described above. The new findings appear as elegant simplifications of the earlier teamology formulas.

Engineers who would not flinch at the elementary arithmetic used here are unlikely to be interested in the application to personality psychology, except in the context of teamology as a bridge to the new discipline of emotional engineering. They may, however, find themselves in demand with their human resource departments or university counseling centers to help them double the power of the MBTI in these endeavors.

The chapter begins with a brief outline of the personality theory of renowned Swiss psychiatrist Carl Gustav Jung because it underlies teamology's psychological approach. Jung theorized that people solve problems by eight mental processes called function-attitudes [5] or cognitive modes [6]. For brevity they will be called "modes" here. Then how to determine the modes preferred by each person in a group to be formed into teams will be discussed. Next, ways will be suggested for using this modal information to construct teams covering as many of the eight modes as possible. Finally, a novel graphical way to organize each team to focus on every mode while reducing duplication of effort will be discussed.

20.2 Underlying Psychological Theory

Psychological teamology depends on scores for Jung's cognitive modes calculated from each potential team member's response to a forced-choice questionnaire like the MBTI or the abbreviated version given on p. 10 of [1]. The underlying variables were formulated mainly by C.G. Jung, one important pair being proposed later by Myers and Briggs [7]. Here is a brief outline of Jung's three postulates which led to the MBTI instrument. Surprisingly perhaps, the second of these has never before been expressed constructively in a form suitable for quantitative interpretation using MBTI data.

Jung's first contribution was the well-known formulation of the pair *extraversion–introversion* (*E-I* in MBTI notation). Whereas extraverted actions and decisions are determined by objective external conditions [3] (p. 333), introverted consciousness selects the subjective determinants as the decisive ones [3] (pp. 373–374). A sample *E-I* item on the teamology questionnaire is "You are more: (*e*) talkative (or) (*i*) quiet".

His second postulate was that (*E-I*) varies according to whether information is being collected, a situation known as "perception", or a decision is to be made with the information, a situation known as "judgment". This concept has never before been applied quantitatively to the MBTI instrument, but it did generate a second MB variable pair *judgment–perception* (*J-P*) contributed by Myers and Briggs [7]. Both variable pairs are called *attitudes*. A sample *J-P* item on the teamology questionnaire is "You are more: (*j*) methodical (or) (*i*) improvisational".

This second postulate implicitly defines a different attitude pair for each of the two situations – information collection or decision-making. Define then *extraverted* (*introverted*) *judgment* (*E*j-*I*j) and *extraverted* (*introverted*) *perception* (*E*p-*I*p) to be these new attitude pairs. If *e*, *i*, *j*, and *p* are respectively the *E*, *I*, *J*, and *P* scores, these new attitudes turn out to be the following simple averages:

$$Ej = (e + j)/2; \qquad Ij = -Ej = (i + p)/2,$$
$$Ep = (e + p)/2; \qquad Ip = -Ep = (i + j)/2. \tag{20.1}$$

The rank of these four equations is only two because those for introversion are not independent, being the negatives of those for extraversion. The equations define the simplest possible linear approximation matching the MBTI maximum scores to those for the new attitudes. For example, maximum MBTI scores of 30 for both *e* and *j* would produce the same score of 30 for the maximum value of extraverted judgment *E*j. Smaller values are interpolated linearly. These new attitudes deserve names: *control* for *E*j, *evaluation* for *I*j, *exploration* for *E*p, and *focus* for *I*p. Collectively they will be described as "decoupled" because, unlike the original "coupled" variables *E-I* and *J-P*, which simultaneously influence both perception and judgment, the *E*j-*I*j pair applies only to judgment modes; the *E*p-*I*p pair only to perception modes.

Jung's third postulate has two parts, one for information collection and the other for decision-making. Each situation can be carried out in two ways, called *psychological functions* in both cases. The two perception functions (*S-N*) are named *Sensing* and *iNtuition*; the judgment functions (*J-P*) are *Thinking* and *Feeling*. A sample *S-N* item on the teamology questionnaire is "You are more: (*s*) hands-on (or) (*n*) theoretical"; a (*T-F*) sample is "You are more: (*t*) truthful".

20.3 Cognitive Modes

From these three postulates come the four pairs of MBTI variables, of which the coupled attitudes *E-I* and *J-P* will be replaced in the sequel by the decoupled attitudes *E*j-*I*j and *E*p-*I*p. This replacement will be done to simplify the final expressions for the cognitive mode scores used to form and organize teams.

The cognitive modes themselves were formulated by Jung as combinations of a decoupled attitude and a function, both of course applying to the same situation, perception or judgment. This then gives the eight modal combinations shown in Table 20.1. Descriptive keywords have been added to aid understanding, and more detailed descriptions appear in Table 20.2. Each cell of Table 20.1 also has formulas for the associated mode scores computed from the psychological instrument used. As an example, consider the following teamology questionnaire values: *E*1, *P*3, *N*5, T1. The Extraverted Sensing score *S*e is to be calculated for these values. The upper equation in the Extraverted Sensing cell is $S\text{e} = (e+p) + 2s = (1\ \ +3)\ \ +2(-5)\ -\ -6 < 0$. This upper equation, as for all upper equations, is for computational convenience designed to generate only integers, the input scores themselves being entirely integer. The negative result signals that this particular mode is *not* preferred. It does indicate indirectly that the mode Introverted iNtuition Ni diagonally opposite *is* preferred with a positive score of 6. This is true because every variable for Ni is opposite to its corresponding Se variable and must therefore have the opposite sign.

Table 20.1 Cognitive modes with transformation formulas and keywords: (a) information collection (perception) modes, and (b) decision-making (judgment) modes

(a) (b)

Extraverted Sensing $S\text{e} = (e+p) + 2s$ $\%S\text{e}$ $= (\%E\text{p} + \%S)/2$ EXPERIMENT	Extraverted iNtuition $N\text{e} = (e+p) + 2n$ $\%N\text{e}$ $= (\%E\text{p} + \%N)/2$ IDEATION	–	Extraverted Thinking $T\text{e} = (e+j) + 2t$ $\%T\text{e}$ $= (\%E\text{j} + \%T)/2$ ORGANIZATION	Extraverted Feeling $F\text{e} = (e+j) + 2f$ $\%F\text{e}$ $= (\%E\text{j} + \%F)/2$ COMMUNITY
Introverted Sensing $S\text{i} = (i+j) + 2s$ $\%S\text{i}$ $= (\%I\text{p} + \%S)/2$ KNOWLEDGE	Introverted iNtuition $N\text{i} = (i+j) + 2n$ $\%N\text{i}$ $= (\%I\text{p} + \%N)/2$ IMAGINATION	–	Introverted Thinking $T\text{i} = (i+p) + 2t$ $\%T\text{i}$ $= (\%I\text{j} + \%T)/2$ ANALYSIS	Introverted Feeling $F\text{i} = (i+p) + 2f$ $\%F\text{i}$ $= (\%I\text{j} + \%F)/2$ EVALUATION

Table 20.2 Cognitive mode descriptions: (a) information collection (perception) modes, and (b) decision-making (judgment) modes

(a) (b)

Extraverted Sensing Discovers new ideas and phenomena by direct experience EXPERIMENT	Extraverted iNtuition Rearranges known concepts into novel systems IDEATION	–	Extraverted Thinking Efficiently manages resources, decisive, imposes structure ORGANIZATION	Extraverted Feeling Expressive, tactful builder of group morale COMMUNITY
Introverted Sensing Physically self-aware, values practice and known technique KNOWLEDGE	Introverted iNtuition Prophetic, guided by inner fantasies and visions IMAGINATION	–	Introverted Thinking Logically improves rational performance ANALYSIS	Introverted Feeling Uses personal values to distinguish good from bad EVALUATION

It is reasonable to ask if the score of 6 is high enough to reflect a real interest in the mode. To answer this, consider that experience with the MBTI predicts that retesting will rarely change any score more than 20% of its maximum, which for the teamology instrument is a single unit. Notice that the score formula takes on the value 4 when all pertinent questionnaire variables assume this minimum value 1: $Se = e + p + 2(n) = 1 + 1 + 2 = 4$. Thus scores of 4 or more can be considered *significant* in the sense that retaking the instrument would not be expected to change the sign of the mode score. This calculation suggests therefore that the person to whom the scores belong would be comfortable in the *N*i Imagination mode on a team, but definitely not in the *S*e Experiment mode.

The results of computing the three other modes possible are displayed in Table 20.3. Two other modes *N*e and *T*i turn out to have positive scores; the last *T*e (as well as *F*i) cancel out to zero, indicating an unclear preference that could well change upon retesting. Thus on a team the subject of the instrument could be used in any or all of three modes. It is important to note that the type dynamic method of interpreting the MBTI would never predict more than two, in this case *N*e and *T*i. The other mode *N*i would be completely overlooked!

Table 20.3 Mode Map with example integer and relative scores. Modes with slight and/or negative scores are shown in gray. Relative scores are in parentheses

Extraverted Sensing EXPERIMENT	Extraverted iNtuition 14 (70%) IDEATION	–	Extraverted Thinking (0) ORGANIZATION	Extraverted Feeling COMMUNITY
Introverted Sensing KNOWLEDGE	Introverted iNtuition 6 (30%) IMAGINATION	–	Introverted Thinking 4 (20%) ANALYSIS	Introverted Feeling 0 EVALUATION

20.3.1 Relative Scores and Coupled Attitudes

The top line formulas generate numbers that differ from one instrument to another. To have universal formulas applicable to all instruments, which is useful when some in a personnel pool have used different instruments, let x^* be the maximum value of any variable x, and define the *relative* value $\%x$ by $\%x = 100\%(x/x^*)$. Then for a typical top line formula in which all variables have a common maximum for *S*e, the mode maximum will be four times the common maximum: $Se^* = 4e^* = 4p^* = 4s^*$, 4 being the sum of the formula coefficients. It follows that $\%Se^* = 100\%(Se/Se^*) = 100\%\ [(e/e^*\ + p/p^*) + 2s/s^*]/4 = [(\%e\ + \%p)/2\ + \%s]/2$. But by Equation 20.1, $(\%e\ + \%p)/2 = Ep$, extraverted perception. Hence, $\%Se = (\%Ep\ + \%S)/2$ as on the lower line of the Extraverted Sensing cell of Table 1. The other lower line formulas are derivable in the same way. For the example, the typical variable maximum is 5 and the mode maximum is $4(5) = 20$, so the relative mode scores are those shown in parentheses in Table 20.3. Relative scores 20% and higher are considered significant.

Equations 20.1 are more significant than merely changes of variable for abbreviation. They stealthily prove that the expressions indeed partition the attitudes into one pair *E*p-*I*p working only for perception and another *E*j-*I*j applying only to judgment, as in Jung's second postulate. The top line equations are themselves proven in Chap. 2 of [1].

20.3.2 Graphical Interpretation

A principal advantage of using relative scores is that they are easily graphed, as demonstrated in Figure 20.1 [8]. The graph may be considered a superposition of two coordinate systems, a rectangular one for the attitudes and functions, and a diagonal one for the cognitive modes, The linear approximation is exact at the corners and at the origins, interpolating linearly in between. Notice that, unlike the customary Cartesian system in which one variable runs from left to right and the other from bottom to top, this composite system has its coordinate axes running radially from the origins so that all coordinates have positive values, a concession to the psychological community. The variable identifications such as *e* thus behave like algebraic signs. The small squares at 20% surrounding the origins mark the lower limits of significance. As will be illustrated later, this graphical representation is helpful mainly for visualizing the organization of an existing team rather than for constructing teams in the first place, a task better performed by guided individual choice or on a spreadsheet, as will be discussed next.

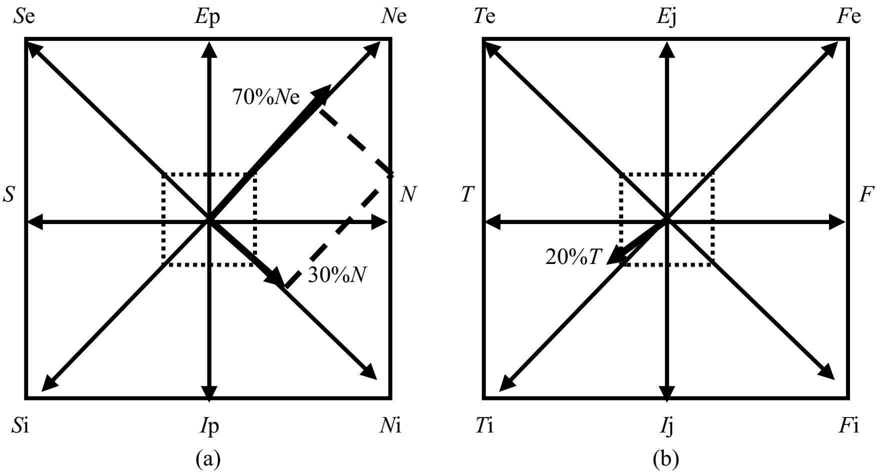


Figure 20.1 Graphical representation of the cognitive modes: (a) perception modes, and (b) judgment modes. Example: 40%*E*p & 100%*N* give 70%*N*e & 30%*N*i. 20%*I*j & 20%*T* give 20%*T*i & 0%*F*i. (Rectangular and diagonal axes run from 0% at origin to 100%)

20.4 Team Construction

Once the cognitive mode scores have been determined for everyone in a personnel pool, teams can be formed, either by a “team-meister” such as a faculty member or superintendent, or more casually by the members themselves. The goal is to have, for every cognitive mode, someone on every team with a significant preference for that mode. This prescription seems to be supported by the ongoing studies on raw MBTI scores [9]. This may not be possible in practice if some modes are “scarce”, in the sense that few prefer them. When this happens, team construction must be content to use the largest scores available, even if they are slight or even negative.

Constructing a set of teams from cognitive mode scores is a little too complicated for summary here; for details see in Chap. 3 of [1]. Instead, a simpler construction strategy will be presented that is useful for beginning team-meisters willing to settle for marked but not maximum improvement in effectiveness of a set of teams. Experience has shown that it is too difficult to cover all eight modes with only three people, so for this discussion it will be assumed that all teams have at least four members. If the total number p in the personnel pool is not an exact multiple of four, then the total number of teams t will be the largest integer less than $p/4$. If, for instance, $p=41$, there would be $t=10$ teams since $p/4=10.25$. This would require $p-4t=41-40=1$ quintet to supplement the remaining $10-1=9$ quartets.

An easy way to ensure improved performance for every team is to have someone with a significant extraverted intuition *N*e “Ideation” score on every team. A team-meister can accomplish this simply by designating the people with the t top *N*e scores as team “seeds”, one per team, around whom each team is formed. The seeds are not to be considered team “leaders” in any way; they are merely expected by example to help the entire team generate a wealth of alternatives “outside the box” for evaluation by all together. This mode *N*e was selected because it is closely related to the Institute for Personality Assessment’s (IPAR) “creativity index”, which is discussed on pp. 214–215 of [7]. This simple seeding of Stanford Design’s teams with Ideation types by itself doubled the fraction of its teams winning Lincoln Foundation design awards during a five-year test period.

A further intervention is to identify the top t extraverted feeling *F*e “Community” people and make sure every team has one. This seems to make each team more unified in the sense that the *F*e communitarian naturally seeks to harmonize the diverse personalities and reduce conflict. This combination of the *F*e and *N*e modes echoes changes resulting from two-week creativity workshops for engineering design faculty conducted by Stanford design professors Bernard Roth, Rolf Faste, and the author [10]. A three-year test at Stanford of teams with both an *F*e communitarian and an *N*e ideator produced a tripling of the fraction of prize-winners from one-quarter to three quarters. Incidentally, whenever a top ideator is also a top communitarian, a second *F*e person is not needed on the team. Such a combined ideator-communitarian can then have three choices instead of the usual two when seeking to complete the quartet.

20.5 Team Organization

When a team is formed, it is wise to convene a half hour meeting to organize the team psychologically, that is, in terms of the Jungian mode information. This is true whether or not the team has been constructed for maximum diversity, as in Chap. 3 of [1]. The fourth chapter indeed discusses such a meeting using the mode scores of an example student team known as “Trio Four”. This team has psychological gaps because it is what remains of a diversified foursome after one of the members has left. Table 20.4 shows a mode map for Trio Four with the mode scores for each member expressed as relative (%) scores instead of the raw scores in the book. This will lend itself to the graphical display and analysis to follow.

Figure 20.2 is a graphical representation of the situation, with the mode arrows not shown. Notice that the ranges of the two maps is different, the usual 100% for perception (Figure 20.2 (a)) and only 20% for judgment (Figure 20.2 (b)), where all modes are slight. The dots labeled ***B***, ***I***, and ***F*** for the three members show visually that none of the three show any interest in either introverted sens-

Table 20.4 Trio four-mode map with relative scores

*S*e *I* 25%	*N*e *F* 35% *I* 5% *B* 0%	–	*T*e *B* 5% *I* 0%	*F*e *I* 10% *B* 5% *F* 5%
*S*i *B* 0%	*N*i *B* 10% *F* 5%	–	*T*i	*F*i *F* 15% *I* 0%

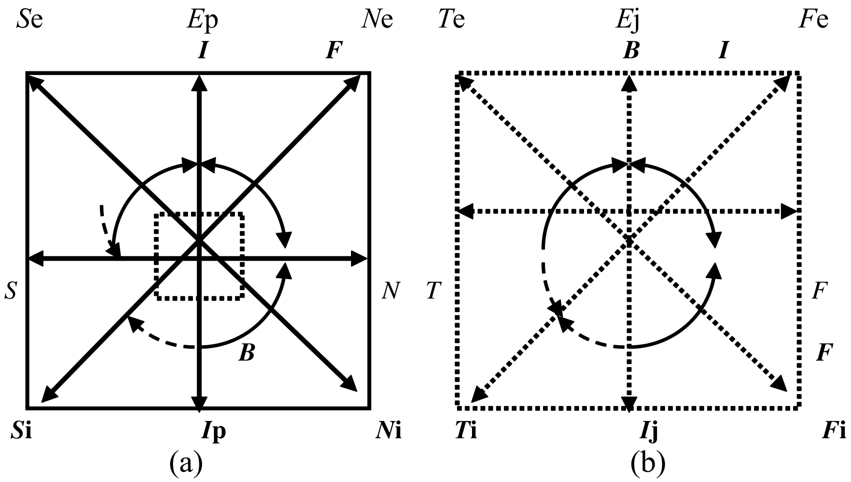


Figure 20.2 Graphical representation of trio four organization: (a) perception modes (range 100%), and (b) judgment modes (range 20%)

ing *S*i (Knowledge) or introverted thinking *T*i (Analysis), so the team's problem is to decide how to cover these important modes.

The solid arcs indicate that ***B*** can take responsibility naturally for *N*i Imagination and *T*e Organization, ***F*** can cover *N*e Ideation and *F*i Evaluation, while ***I*** can do *S*e Experiment and *F*e Community. The dashed arcs shows how the members "reach" into adjacent modes to fill the gaps. Specifically, ***B*** extends herself into halves of modes *S*i and *T*i; ***F*** also into the other half of *T*i; while ***I*** covers the *S*i half not taken by ***B***.

These half modes correspond to actual "roles" – activities to be performed by the team members. Figure 20.3 displays the 16 "team roles", discussed in more detail with short descriptions in Chap. 4 of [1]. Of particular interest in the example here are the roles Inspector and Specialist reached by ***B***. One would suspect that without ***B***'s assumption of these responsibilities, the team could suffer from careless failures to meet specifications. ***I***'s extra role is that of Investigator, the one who looks up the patent literature and catalogs to prevent the "reinvention of the wheel". ***F*** reaches into the Reviewer role, most often employed to edit any reports.

The other 12 roles come more naturally to the members. It is helpful to list them for each member. Thus ***B*** is expected to cover the three perception roles Vision, Strategy, and Inspection, and two judgment roles, Coordination and Methodology. ***F*** looks at perception roles Entrepreneur and Innovator while covering both Evaluation roles Need-finder and Critiquer. ***I*** then fills the three perception roles Tester, Prototyper, and Investigator, as well as the three judgment roles Organizer, Methodologist, and Specialist.

Figure 20.4 displays these role assignments graphically. A similar diagram (Figure 4.1 on p. 50 of [1]) differs slightly from Figure 20.4 here. This merely means that, because some of the interests are so slight, the two organization plans are about equally effective. Anyway, it's up to the team to decide as a group how to distribute the responsibilities.

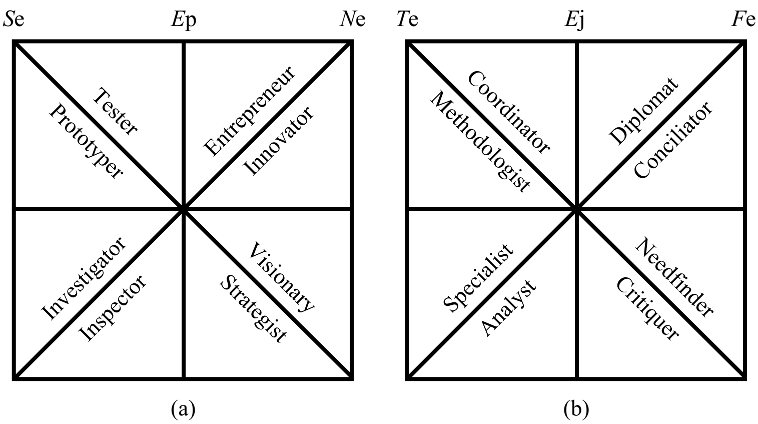


Figure 20.3 Team role keywords: (a) perception modes, and (b) judgment modes

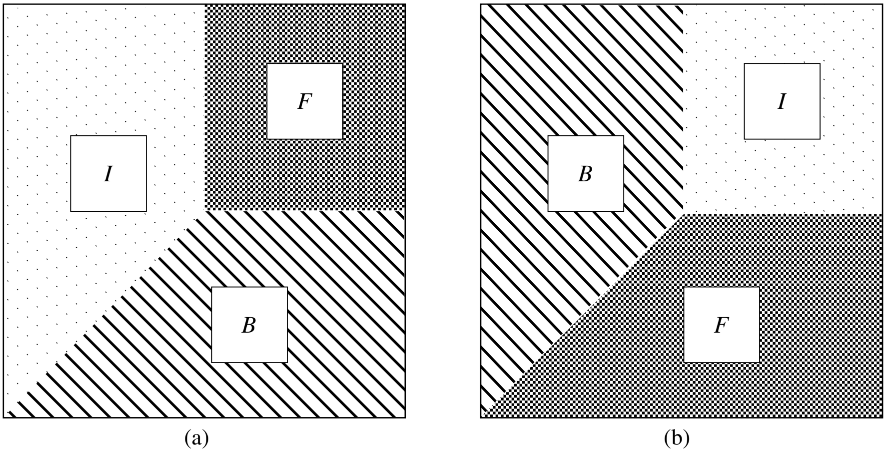


Figure 20.4 Trio four-role pattern: (a) perception roles, and (b) judgment roles

20.6 Conclusions

This chapter has not only reviewed the book *Teamology*, it has updated the underlying theory transforming the questionnaire data into the cognitive mode scores used to construct and organize effective teams from a pool of personnel. This new theory for the first time decouples the two *coupled* Jungian attitude pairs *E-I* and *J-P* into two equivalent *decoupled* attitude pairs *E*p-*I*p for perception (information collection), and *E*j-*I*j for judgment (decision-making). This development simplifies later formulas for the cognitive modes so that all calculations only involve simple averages of sums and differences. These procedures effectively *double* the amount of personality information extracted from the questionnaire compared to the traditional MBTI categorical analysis method known as "type dynamics". A further update of *Teamology* is the introduction and demonstration of graphical procedures. All this is useful not just for engineers and managers constructing teams, it can also be used by human resource departments for doubling the effectiveness of their MBTI personality questionnaires in other applications.

References

1. Wilde D.J.: Teamology: The Construction and Organization of Effective Teams. Springer, London (2009)
2. Myers I.B., McCaulley M.H., Quenk N.L., Hammer A.L.: MBTI Manual: A Guide to the Development and Use of the Myers–Briggs Type Indicator, 3rd edn. Consulting Psychologists' Press, Palo Alto, CA (1998)

3. Jung C.G.: Psychological Types. Princeton University Press, Princeton, NJ (originally published 1923) (1971)
4. Reynierse J.H.: The case against type dynamics. J. Psychol. Type **69**(1) (2009)
5. Thompson H.L.: Jung's Function-attitudes Explained. Wormhole, Watkinsville, GA (1996)
6. Singer J., Loomis M.: Interpretive Guide for the Singer-Loomis Inventory of Personality. Consulting Psychologists Press, Palo Alto, CA (1984)
7. Myers I.B., McCaulley M.H.: MBTI Manual: A Guide to the Development and Use of the Myers–Briggs Type Indicator, 2nd edn. Consulting Psychologists' Press, Palo Alto, CA (1989)
8. Wilde D.J.: Graphical interpretation of the teamology transformation, DETC2009-86036. In: Proceedings of the American Society of Mechanical Engineers, Design Engineering Technical Conference (2009)
9. Forootan M.Z., Freiheit T.: Modeling Team Personalities from Member Meyers (sic) – Briggs Inventory Scores. In: Proceedings of the 2009 Industrial Engineering Research Conference (2009)
10. Wilde D.J.: Changes among ASEE creativity workshop participants. J. Eng. Educ. **82**(3) (1993)

Chapter 21
Designing Emotional and Creative Motion

Toshiharu Taura[1] and Yukari Nagai

Abstract We propose a method for designing emotional and creative motion that extends beyond images currently generated by human imagination and resonate with the deep feelings that reside within us. We discuss design from two perspectives: (1) the motion of natural objects are reflected in the sources of emotional motion, and (2) creative and emotional motion, which resonates more with human feelings than natural objects do, is generated by extending beyond the motion of natural objects. We also focus on another perspective of the method for designing such motion, that is, whether or not the motion can be generated using the traditional method. On the basis of these discussions, we develop a method that involves "analogy to natural objects" and "emphatic blending of motions". We have developed this method by focusing on rhythmic features, and we will illustrate an example of designing an emotional and creative motion using the proposed method.

21.1 Introduction

21.1.1 Emotion and Design

The most significant ambition in design is to create objects that resonate with the deep feelings felt by humans. Needless to say, we (humans) are deeply impressed when we come across natural scenery that is delicate or magnificent or wonderful. We learn such emotional experiences when we "feel the beauty of nature". However, nature is not the only thing that resonates with human feelings. In fact, we gather emotional experiences not only from nature but also from artifacts. Music is a good example – humans are sometimes deeply impressed when listening to mu-

[1] T. Taura
Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe, 657-8501, Japan
e-mail: taura@kobe-u.ac.jp

sic. Thus, we can say that music, which is created by humans, stimulates human feelings. Perhaps it is the musical instruments that effectively trigger human feeling with their music. Musical instruments are man-made tools, and it is difficult to play music without musical instruments. Both the creation and usage of musical instruments facilitates music's capability of resonating with human feelings. There are many other artifacts that resonate with human feelings – paintings, movies, *etc.* – and we collectively identify them as "art". True art has the power to provide special symbols that stimulates emotional experiences within humans [1]. Human feelings (emotions) must be taken into account if we aim to design a truly desired artifact. Moreover, it is necessary to develop a method for designing the truly desired artifact. This method should address the following question: "How can we create an artifact that resonates with human feeling in the same way as musical instruments?". In other words, our goal is to create artifacts that resonate with human feeling and stimulate human emotion.

21.1.2 Designing the Emotional and Creative Motion

In recent years, design has mainly targeted shapes; thus, most designs provide the shapes and forms of objects. Further, as mentioned in the previous subsection, music has been another subject of design. In this study, we challenge further developing the subject of designing the motions. There are numerous types of motion created by humans, for example, the behaviors of vehicles or robots [2], and animation and dance. Interestingly, the motions of animation are constructed from static pictures. Animation looks dynamic, but is actually motion that is stimulated on the basis of visual mechanism [3]. Thus, each animation can be separated into a time-sequential sheet that represents each static image; it is difficult to reproduce a complete image of a motion. Sequential sheets are not adequate tools for designing an animation that portrays an entire motion, since the dynamic features of motion cannot be described easily and must be recognized by reading between the lines. An entire image of the motion is stored only in the human mind. We still do not have a good instrument to describe and design motions, unlike, for example, music which is designed using musical instruments. In addition, we cannot compose the motion, in contrast to music which can be composed using music scores, nor can we play an instrument of motion in the same way that we play musical instruments.

To consider the issue of how we can create or design motions, we can take a hint from dance. A dancer probably imagines the complete motion of the dance and composes the dance by moving his/her body. However, this method limits the imagination of motion to one's individual body. To overcome this limitation, dance directors use choreography, which is based on an expert knowledge of dance. "Choreography" implies "dance-writing" – the art of designing sequences of movements, with each movement formed by various motions. Consider, the famous choreographer, Maurice Béjart, who created an advanced style of ballet [4].

He developed upon traditional ballet by integrating modern dance styles and adapting other dances to ballet. For instance, he studied Japanese Kabuki and mixed blended the motions of Kabuki actors with those of western ballet. By employing such methods, he created a partial combination of many forms or motions that he borrowed from different sources, such as foreign dances, children's actions, and movie scenes. Béjart might not have been able to imagine all the motions nor design novel motions without the use of the body. Nonetheless, the human body strongly constrains the development of the design of motions. If we design a method and instruments that allow us to operate the motions more freely, we can perhaps design novel motions that are more creative and beyond human imagination. In addition, we further expect that these novel motions can more strongly resonate with human feelings.

Nowadays, it is possible to create, record, and replay motions using computers. Thus, we can computationally design new motions that humans cannot easily imagine.

Therefore, paying attention to computational methodology, we aim to develop a method for designing emotional and creative motions that extend beyond the images currently generated by human imagination and that resonate with the deep feelings that reside within us.

21.2 Method of Designing Emotional and Creative Motion

To design emotional and creative motion that resonates with deep feelings, we determined on the following strategies. First, the motion of natural objects are reflected in the sources of emotional motion, since humans have evolved in the natural environment and are thought to have images of nature imprinted in their minds. Second, emotional and creative motion, which resonates with deep feelings more than natural objects do, is generated by extending beyond the motion of natural objects. From these two perspectives, we develop a method that involves an "analogy to natural objects" and "emphatic blending of motions". Furthermore, we focus on another perspective of our method, that is, whether or not such motion can be generated using traditional methods, such as composing conventional time-sequential pictures.

21.2.1 Analogy to Natural Objects

Analogy (property mapping) is known to play a very important role in creative design [5, 6]. Analogy is considered to be a method for concept creation and involves the transfer of certain features from an existing concept (base concept) to another. For example, the "Swan chair", a famous example in the field of design, was imaged using analogy [7]. On the other hand, the importance of learning from nature has long been recognized, and attempts have been made by researchers like Vogel *et al.* [8] to learn from nature for the purpose of product development. Systems in the natural world and their functionality are seen as rich sources of inspiration for the generation of ideas. For learning from nature, in addition to an effective design method, we focus on the feeling of nature that is expected to dwell deep within the human mind. Based on these considerations, we develop "analogy to natural objects" as the source for designing emotional motion.

21.2.2 Emphatic Blending of Motions

In studies on cognitive linguistics, Fauconnier analyzed how conceptual integration creates mental products and how mapping and blending systems are deployed between mental spaces [9].

From the perspective of mental space theory, he showed that conceptual integration operates on two input mental spaces to yield a third space that is called "the blend". The blend inherits partial structures from the input spaces and has emergent structures of its own. Concept blending is considered to develop a more novel concept as compared to analogy because a concept developed by the former process does not belong to the domain of the input concepts (base concepts), while analogy cannot extend beyond the domain of the base concept [10].

Therefore, we develop a method to blend the motions generated by analogy from natural objects so that a more creative motion can be designed.

Furthermore, we try to emphasize the rhythmic feature of motion in the process of blending such that the generated motion can extend even further beyond the simple blended motion of natural objects.

21.2.3 Outline of Proposed Method

An outline of the proposed method to design emotional and creative motion is shown in Figure 21.1.

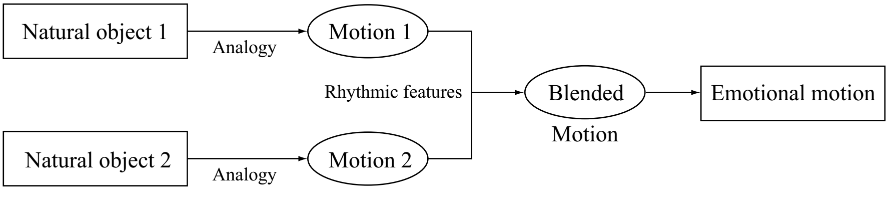


Figure 21.1 Outline of the proposed method

21.3 Detailed Method and Example

The method addressed in the previous subsection was developed by Tsujimoto *et al.* [11]. The outline of their study is shown as follows. In the developed method, it was assumed that the design target has a linked-pole structure. This assumption was made on the basis of the following two points. First, it is possible to represent all natural objects with interesting motions as skeletal structures using computer graphics (CG). Second, this linked-pole structure is the most fundamental structure and can be extended to other structures such as the meatball and polygon structures.

In this method, the motion was captured as a sequence of pictures, and smooth images were obtained from each picture using Hermitian interpolation.

21.3.1 Designing Emotional and Creative Motion by Focusing on Rhythmic Features

Wavelet analysis and genetic algorithms were used to transfer rhythmic motion features to design targets in order to emphasis the rhythmic features of the motion.

21.3.1.1 Wavelet Analysis

The main purpose of using wavelets is to enable the analysis of signals on the basis of their magnitude. Wavelets are functions that satisfy certain mathematical conditions and are used to represent data or other functions. In wavelet analysis, the scale that one uses for viewing data is important. Wavelet algorithms process data at different scales or resolutions. Using wavelet analysis, we can observe the data in detail and analyze discontinuities in the changes. The analysis was performed using a contracted high-frequency version of the prototype wavelet. The wavelet analysis procedure involves the adoption of a wavelet prototype function called an “analyzing wavelet” or “mother wavelet” [12]. In this study, Daubechies wavelets was used as the wavelet-prototype functions. Since the original signal or function can be represented in terms of a wavelet expansion (using a linear combination of the wavelet functions), data operations can be performed using only the corresponding wavelet’s coefficients. Moreover, if the best wavelets that can be adapted to the data are chosen or if the coefficients below a threshold are truncated, the amount of data can be reduced. This “sparse coding” makes wavelets an excellent tool in the field of data compression.

21.3.1.2 Genetic Algorithm

In general, an algorithm starts with a set of solutions called a population [13]. The solutions from one population are used to form a new population. This is done to ensure that the new population will be better than the old one. The selection of solutions to form new offspring is based on their fitness. This process is repeated until a particular condition is satisfied. The genetic algorithm (GA) transforms a population of individual behaviors into a new generation using crossover and mutation. In this method, each individual in the population represents a possible behavior. The individuals in the population are usually fixed-length character strings patterned after chromosome strings. Crossover and mutation are defined as follows:

- *crossover*: the act of swapping gene values between two potential solutions, simulating the "mating" of the two solutions;
- *mutation*: the act of randomly altering the value of a gene in a potential solution.

It is necessary to be able to evaluate how "good" a potential solution is relative to the other potential solutions. The "fitness function" is responsible for performing this evaluation and returning a "fitness value" that indicates whether the solution is an optimal one. The fitness values are then used in a process of natural selection to choose the potential solutions that will continue onto the next generation.

21.3.1.3 Obtaining Rhythmic Features

First, the standardized changes in the angles of joints corresponding to the design target were analyzed using Daubechies wavelets, and the results were stored in a database.

Next, in order to introduce rhythmic features in the designed motion, GA was used. It is necessary to decide on the makeup of the chromosomes, which includes the number of genes and what those genes represent. In this study, a chromosome represents the time series of the angles of a joint. A gene to represent an angle was defined in a frame (Figure 21.2).

In this study, it was assumed that the degree of correlation between the rhythms of natural objects and those of the design target represents a measure of the fitness of the GA. In a GA, the system analyzes the standardized changes in the angles of joints in the design target using Daubechies wavelets and then evaluates the correlation. This method can facilitate the design of motions that contain rhythmic features, even if the motions are not similar. The procedure for obtaining rhythmic features is shown in Figure 21.2.

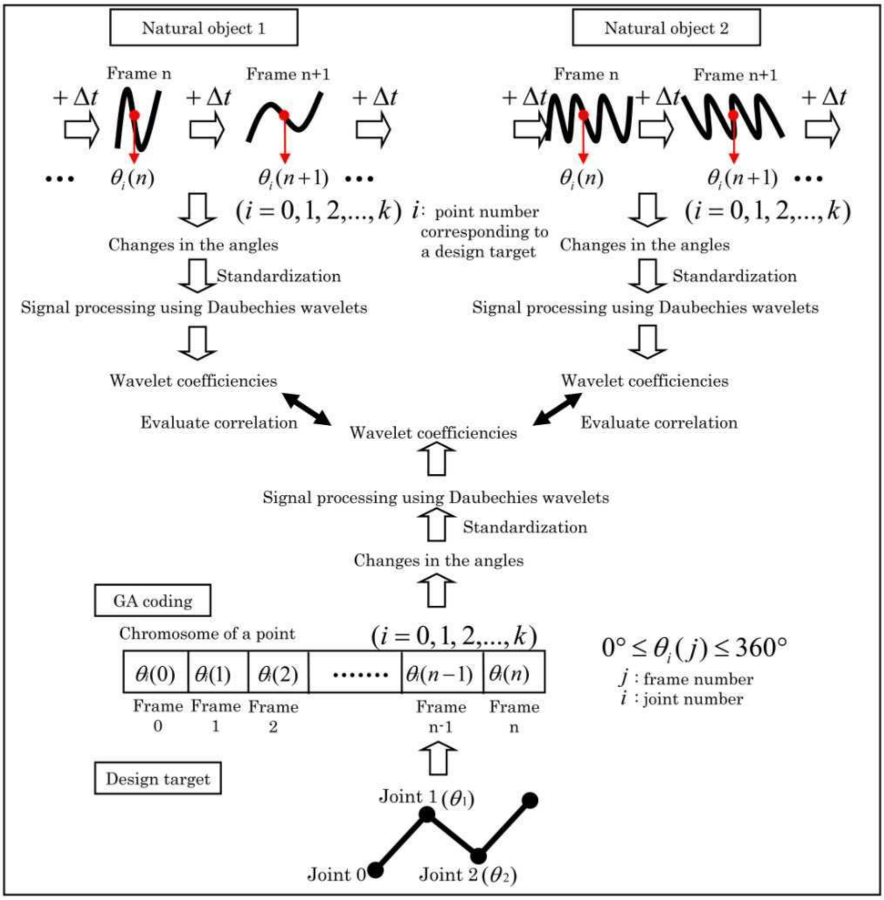


Figure 21.2 Process for obtaining rhythmic features

21.3.2 Example of Designing an Emotional and Creative Motion Using the Proposed Method

A new emotional and creative motion was designed for a robot arm. The structure of the robot arm is shown in Figure 21.3. In this example, the motions of frogs and snakes were used as the base objects for the analogy.

Figure 21.4 shows some frames from the sequence of motions produced by a frog and a snake.

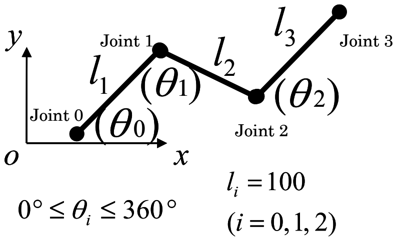


Figure 21.3 Structure of a design target

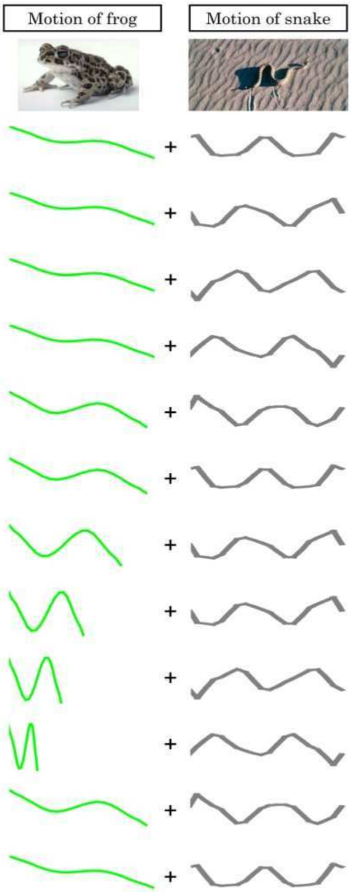


Figure 21.4 Sequential frames of the motions of a frog and a snake

The extracted rhythmic features are shown in Figures 21.5 and 21.6.

The designed motion is shown in Figure 21.7. The features of the frog's motion are recognized from the instantaneous movement of the arm, while those of the snake's motion are recognized in the steady-rhythmic movement of the arm. On the other hand, the generated motion is recognized as extending beyond the motion of a frog and a snake. Furthermore, it is assumed that the generated motion cannot be produced using traditional methods, since it is very rhythmic.

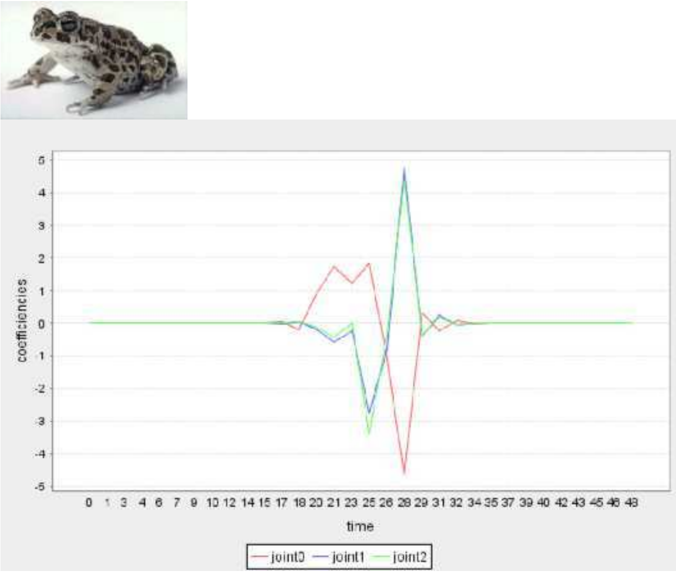


Figure 21.5 Rhythmic features of a frog's motion

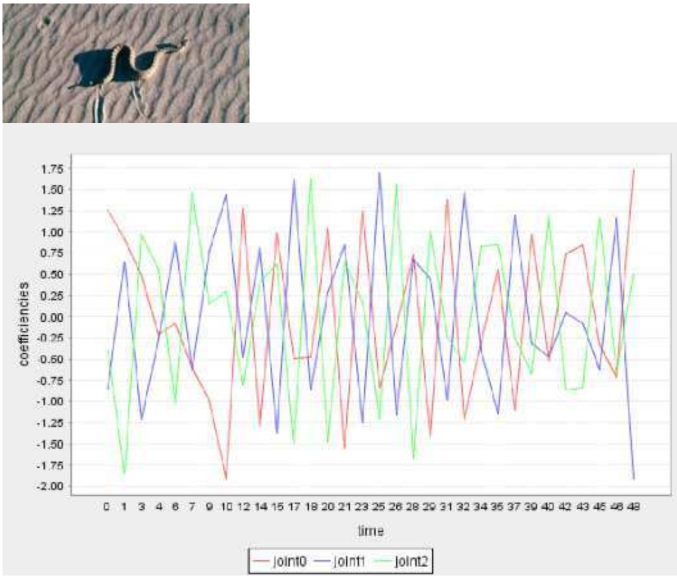


Figure 21.6 Rhythmic features of a snake's motion

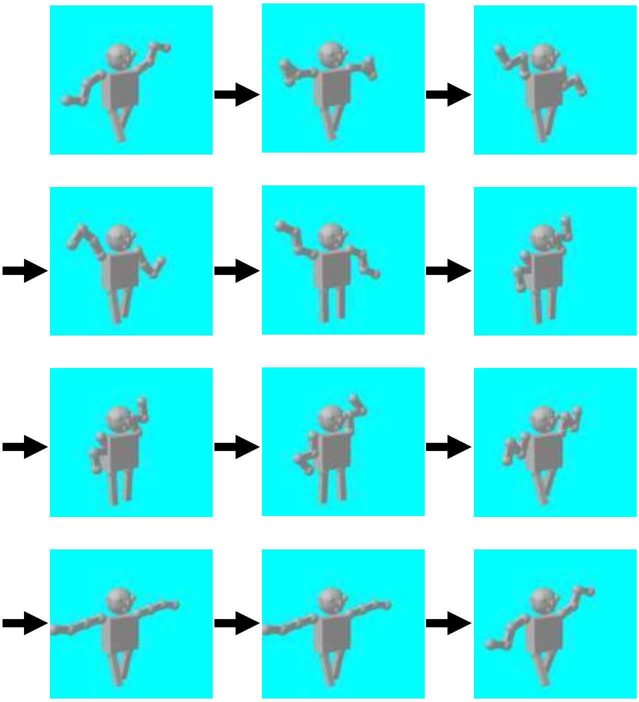


Figure 21.7 Designed Motion

21.4 Summary and Conclusions

In this paper, we discuss a notion for designing emotional and creative motion. We determined that emotional and creative motion is expected to extend beyond the images produced by human imagination and resonates with the deep feelings residing within us. The study also pointed out that emotional experiences arise from not only nature but also artifacts and discussed the issue of designing emotional and creative motion from two perspectives: (1) the motion of natural objects are reflected in the sources of emotional motion, and (2) creative and emotional motion, which resonates more with human feelings than do natural objects, is generated by extending beyond the motion of natural objects. On the basis of these discussions, we developed a method that involves “analogy to natural objects” and “emphatic blending of motions”. We developed this method by focusing on rhythmic features and provided an example of designing an emotional and creative motion using the proposed method.

Acknowledgments We would like to express our gratitude to Mr. Kazuya Tsujimoto and Mr. Shinji Miura for developing the computer system and to Prof. Akira Tsumaya and Prof. Amaresh Chakrabarti for presenting us with useful comments.

References

1. Langer S.: Feeling and Form: A Theory of Art Developed from Philosophy in a New Key. Routledge & Kegan Paul, London (1959)
2. Billard A.: Dreams and imitation in robot. In: Proceedings of AAAI Fall Symposium on Socially Intelligent Agents – the Human in the Loop, Menlo Park, CA, pp. 9–12 (2000)
3. Auzenne V.: The Visualization Quest: A History of Computer Animation. Fairleigh Dickinson University Press, New Jersey (1993)
4. Krasovskaya V.: An Anatomy of Genius. Taylor & Francis, London (1990)
5. Gero J.S., Maher M.L.: Mutation and analogy to support creativity in computer-aided design. In: CAAD Futures '91, pp. 261–270. Bertelsman, New York (1991)
6. Casakin H., Goldschmidt G.: Reasoning by visual analogy in design problem-solving: the role of guidance. Planning Des. **27**:105–119 (2000)
7. Taura T., Nagai Y.: Primitives and principles of synthetic process for creative design. In: Computational and Cognitive Models of Creative Design VI, pp. 177–194 (2005)
8. Vogel C.M., Cagan J., Boatwright P.: The Design of Things to Come: How Ordinary People Create Extraordinary Products. Wharton School Publishing, New Jersey (2005)
9. Fauconnier G., Turner M.: The Way We Think – Conceptual Blending and the Mind's Hidden Complexities. Basic Books, New York (2002)
10. Taura T., Nagai Y., Morita J., Takeuchi T.: A study on design creative process focused on concept combination types in comparison with linguistic interpretation process. In: Proceedings of ICED07, the 16th International Conference on Engineering Design (CD-ROM) (2007)
11. Tsujimoto K., Miura S., Tsumaya A., Nagai Y., Chakrabarti A., Taura T.: A method for creative behavioral design based on analogy and blending from natural things. In: Proceedings of the ASME 2008 International Design Engineering Technical Conference and Computers and Information in Engineering Conference IDETC/CIE2008 (CD-ROM) (2008)
12. Percival D., Walden A.: Wavelet Methods for Time Series Analysis. Cambridge University Press, Cambridge (2000)
13. Miettinen K., Neittaanmaki P., Periaux J., Makela M. (eds.): Evolutionary Algorithms in Engineering and Computer Science: Recent Advances in Genetic Algorithms, Evolution Strategies, Evolutionary Programming, Genetic Programming and Industrial Applications. Wiley, Chichester (1999)

About the Editor

Shuichi Fukuda is currently consulting professor at Stanford University, USA, and visiting professor at the Open University of Japan and Cranfield University, UK. He is also a member of the Science Council of Japan.

He received his bachelor's, master's, and doctor of engineering degrees from the Department of Mechanical Engineering at the University of Tokyo. He worked as research instructor at the Department of Precision Machinery, University of Tokyo and as associate professor at the Welding Research Institute, Osaka University, and later concurrently as associate professor at the Institute of Industrial Science, University of Tokyo.

He joined Tokyo Metropolitan Institute of Technology in 1991 as a full professor and worked there as professor, department chair, dean of engineering, director of the center for university–industry collaboration, and dean of the library and information systems departments until his retirement in 2007. During this time, he was also visiting researcher at Stanford University, and visiting professor at West Virginia University, Stanford University, and Osaka University. He is now emeritus professor at both Tokyo Metropolitan University and Tokyo Metropolitan Institute of Technology.

His research interests are design engineering, emotional engineering, collaborative or concurrent engineering, reliability or trust engineering, and the management of technology.

He is a Fellow of ASME, JSME, ISPE (International Society for Productivity Enhancement, where he is also currently president), and IEICE (Institute of Electronics, Information and Communication Engineers).

Index

1

1/f fluctuation, 144, 150

A

Abduction, 9
Affect, 315
Affordance, 180, 182
Affordance, 3
Affordance features, 183
Affordance perception, 183
Ambient information, 22
Analogy, 379
Analysis by synthesis, 168
Annotation, 52
Approximate multi-objective optimization, 277
Arch support, 276, 279
Attitude, 367
Attitude-pair, 367
Augmented Reality (AR), 57
Augmented reality, 108
Avatar, 163

B

Ba place, 40
Back propagation learning, 214, 215
Basic emotions, 124, 127
Behavior, 122, 129, 130, 131, 133, 136
Behavior selection, 121
Behavior system, 135
Biological signal, 328, 333
  Breathing rate, 333
  Galvanic Skin Reflex (GSR), 333
  Heart rate, 333
Blending, 380
BoP (Bottom of the Pyramid), 16
Bounded rationality, 2
Bradley-Terry model, 252
Brand, 12

C

Carl Gustav Jung, 366
Challenge, 13
Closed world, 7
Cluster analysis, 294
Cluster analysis, 95
Cognitive development, 5
  Cognitive development theory, 5
Cognitive emotion, 14
Cognitive mode, 365, 366
Common scale, 299
Commonality, 295
Communication terminal, 24
Computer game, 211
Confidence, 12
Conflict, 322, 323
Consummatory communication, 22
Continuous prototyping, 15
Control, 367
Cooperative behavior, 139
Coping behavior, 137
Coupled, 367
Coupled attitude, 369
Creative collaboration, 312
Creative customer, 3
Creative design, 379
Creative motions, 379
Creativity, 4
Creativity, 72, 76, 78, 81

Customer, 65
Customization, 73

D

Decision making, 2
Decision making, 367
Decoupled, 367
Deduction, 9
Design, 67
Design and engineering, 64, 70
Designer-centric design, 2
Deterioration curve, 15
Development psychology, 5
Different potential scale, 294
Disagreement, 322
Driver's emotion, 265
Driver's emotion, 88
Drives, 122, 129, 131
Driving performance, 269

E

e-mail, 210
Embodied communication, 162
Embodied interaction, 254
Embodied knowledge, 40
Embodied media, 176
Embodied virtual communication, 254
Embodiment, 161
Emotion, 4, 313, 314, 317, 318, 319, 323, 324
  Cultural context, 324
  Define, 314
  Distinguishing between, 317
  Flow, 323
  Interpersonal interaction, 319
  Measuring, 318
  Moment-to-moment, 314
  Perspectives, 313, 318
Emotion engine, 131, 137
Emotion equation, 6
Emotion model, 124, 125, 126
Emotion recognition, 121
Emotion recognition, 33
Emotion regulation, 121, 123, 126, 137
Emotion representation, 121
Emotional communication, 111
Emotional competence, 120, 121, 124, 126, 135, 138
Emotional design, 115
Emotional engineering, 366
Emotional engineering, 64
Emotional intelligence, 121
Emotional interaction, 53
Emotional motions, 379
Emotional quality, 319
Emotions, 122, 124, 125, 133
Enactive interface, 75, 78
Enactive knowledge, 75
Engineering, 67
Engineering design team, 322, 323
Entrainment, 161
EQ emotional quotient, 12
Evaluation, 367
Evaluation criteria, 292
Evaluation environment, 256
Expectation, 2
Expectations, 12
Explicit knowledge, 42
Exploration, 367
Expression method, 27, 36
Expressive behavior, 137
eXtensible Markup Language (XML), 50
eXtensivle Virtual world description Language (XVL), 50
Extraversion-introversion, 366
Extraverted (introverted) judgment, 367
Extraverted (introverted) perception, 367

F

Face detection, 226
Face feature point detection, 226
Face mark, 213
Face-to-face communication, 207
Facets, 352
Facial Action Coding System (FACS), 227
Facial expression, 207
Factor analysis, 93
Factor loading, 306
Feeling, 367
Feeling of attachment, 15
Feelings, 315
Fitting Parameter (FP), 357
Fluctuations, 5
  Brownian motion, 6
  Chaos, 5
  Complex systems, 5
  Fluctuation equation, 5
  Random theory, 5
Focus, 367
Footwear design flow chart, 349
Force feedback device, 52
Function-attitude, 366
Future design, 293
Fuzzy emotion model, 128
Fuzzy reasoning, 208
Fuzzy set, 209

G

Gaze tracking, 237
Gazing line-action information, 248
Global head motion compensation, 229
Goodness-of-fit test, 253

H

Haptic interface, 79
Haptic sensation, 106
Haptic-sound interface, 80
Head-up display (HUD), 87
Hierarchical structure, 294
Highly subjective scale, 301
Homo faber, 8
Homo ludens, 8
Homo mobens, 13
Human interaction, 161
Human interface, 163
Human interface, 207
Human needs, 11
Human perception, 294
Human–Computer Interaction (HCI), 207

I

Ideation, 319
  Spontaneous, 321
Image processing technology, 352
Immersive projection display, 106, 108
Immersive virtual environmental system, 52
Impression, 294
Individual differences, 291
Induction, 9
Industrial design, 68
Industrial engineering, 68
Inference of human emotion, 266
Influence of avatar’s image, 254
Influence of viewpoint, 254
Information and communication technology (ICT), 40
Information collection, 367
Interactive system, 328
iNtuition, 367

J

Japanese teacup, 144
Judgment-perception, 367

K

Kansei, 289
Kansei engineering, 88
Kansei quality, 289
Kansei value, 327
  Kansei value creation, 328

L

Laser scanning system, 349
Laser-CCD displacement sensing, 146
Last, 350
Latent sensitivities, 292
Learning curve, 15
Learning from failure, 10
Lifetime value, 16
Likelihood ratio test, 253
Long-duration emotion, 14

M

Magnetic sensors, 248
Major scale, 306
Market creation, 12
Maslow’s hierarchy of human needs, 11
Maximum likelihood method, 252
Media communication, 22
Medical arch support, 279
Membership value, 209
Metatarsophalangeal joint, 360
Method of trial and error, 10
Mixed reality, 82
Modal combinations, 367
Mode map, 369
Motivation, 4
Multi-Dimensional Scaling (MDS), 299
Multi-facet matching, 353
Multimedia, 40
Multiple Regression Analysis (MRA), 296
Muscles, 10
Myers-Briggs Type Indicator (MBTI), 365

N

Natural-contact biosensor method, 266
Neural network, 208
Noh mask, 216
Noh play, 216
Non-verbal message, 207

O

Objective, 366
Observation method, 193
One time value, 16
On-the-job training, 41
Open world, 7

P

Paired comparison, 252
Participatory design, 73
PDSA/PDCA cycle, 45
Perceived affordance, 182
Perception-action interaction, 75
Personal characteristics, 182
Personal creativity modes, 182, 183
Personality psychology, 366
Personality theory, 366
Photoforming system, 249
Physiological index, 333
Potential factors, 293
Power spectrum, 147
Pragmatism, 9
Preference, 294
Presentation image, 89
Principal Component Analysis (PCA), 296
Problem formulation, 5
Problem solving, 5
Process value, 18
Product development process, 66, 72
Product evaluation, 82
Product testing, 71
Product value, 18
Products, 65, 82
Prototypes, 69, 71
Prototyping, 82
Psychological functions, 367
Psychological teamology, 365

Q

Quality function deployment (QFD), 5
Quantification theory, 309
Quantification theory type 2, 100

R

Radial basis function (FBF) network, 277
Rapid prototyping, 156
Rationalize actions, 6
Reflective practice, 8, 9
Regression plane, 303
Regulation, 131
Relationship, 322
  Marital, 323
Relative scores, 369
Reliable sound, 306
Repair, 15
Resonate, 377
Rhythmic feature, 380
Risk engineering, 6
Robotics architectures, 120

S

Sand glass type neural network, 214
Satisfaction, 6
Satisficing method, 277
SECI model, 44
Semantic Differential method (SD method), 290
Sensation, 328
  Dynamic, 328
  Mental, 328
  Negative, 328
  Positive, 328
  Static, 328
Sensing, 367
Sensitivity quantification, 290
Sensory evaluation, 255
Sensuous, 290
Shape design, 150
Shape evaluation, 248
Shape evaluation, 79
Shewhart cycle, 9
Shoji semi transparent room divider, 29
SHOJI Symbolic Hosting Online Jog Instrument, 29
Short-duration emotion, 14
Skill, 64, 79
Skill transfer, 42
Social robotics, 120
Sound quality, 296
Spatial frequency, 147
Stimulus-response (reaction) cycle, 8
Stretch material, 360
Subjective, 366
Synchronized Multimedia Integration Language (SMIL), 50

T

Tacit knowledge, 42
Task, 322
Team construction, 365, 371
Team organization, 372
Team-meister, 371
Teamology, 365
Teamology questionnaire, 366
Tele-immersion, 111
Thinking, 367
Tone-to-Noise Ratio (TNR), 308
Tools, 70
Topographical feature, 147
Transformation formulas, 368
Trust, 11
TV quality rating system, 225

TV rating, 223
Type dynamics, 365

U

Unexplored design area, 302
User activity, 182
User-centric design, 2, 74
  Design, 75
  Designers, 75
Utility, 6

V

Variation and diversification, 291
Value, 17
  Value co-creation, 17
  Value engineering, 17
Verbal communication, 219
Verbal message, 207
Video avatar, 112
Virtual communication, 163
Virtual Reality (VR), 40
Virtual reality, 104
Virtual training system, 56
VirtualActor, 254
Visceral design, 8
Visual information, 248

W

Wakuwaku, 328
  Degree of wakuwaku, 329
  Wakuwaku feeling, 328
Wakuwaku model, 329, 335, 336
Weber-Fechner's law, 16
Wrinkling, 350