HOKKAIDO UNIVERSITY

Title	Study of Recognition Methods for Facial Expression and Face Identification
Author(s)	Nagao, Mitsuyoshi
Degree Grantor	北海道大学
Degree Name	博士(工学)
Dissertation Number	甲第5509号
Issue Date	2001-03-23
DOI	https://doi.org/10.11501/3182346
Doc URL	https://hdl.handle.net/2115/51501
Type	doctoral thesis
File Information	000000399928.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Study of Recognition Methods for Facial Expression and Face Identification

Mitsuyoshi NAGAO

Study of Recognition Methods for Facial Expression and Face Identification

System and Information Eng.
Graduate School of Eng.

Mitsuyoshi NAGAO

Hokkaido Univ.
Japan

Submitted in partial fulfillment of the requirements for the Degree of Doctor of Engineering Hokkaido University, December 2000

Contents

1 Face Recognition and Recognition Methods **12**

1.1 Introduction 12
1.2 Face Recognition 13
1.3 Facial Expression Recognition 17
1.3.1 Relation between Facial Expression and Emotion . 17
1.3.2 Facial Action Coding System 19
1.3.3 Facial Expression Recognition Algorithm 22
1.3.3.1 Type 1 22
1.3.3.2 Type 2 24
1.3.3.3 Type 3 25
1.3.3.4 Type 4 26
1.4 Face Identification 27
1.4.1 Face Identification Process 29
1.4.2 Target Facial Image 29
1.4.3 Pre-processing 31
1.4.4 Attribute Extraction 32
1.4.5 Person Identification 35
1.4.6 Issues on Face Identification 35
1.4.7 Related Works and Applications 36
1.5 Face Cognitive Model Acquisition 41
1.5.1 Face Cognitive Model 41
1.5.2 Face Cognitive Model Acquisition based on Facial Image Database Retrieval Task Method 44
1.6 Mahalanobis-Taguchi System 47
1.6.1 Framework of MTS 47
1.6.2 Pattern Recognition based on MTS 49
1.6.3 Optimization of Attribute 54
1.6.4 Mahalanobis-Taguchi-Gram-Schmidt 58
1.7 Interactive Evolutionary Computing 60
1.7.1 Framework and Feature of IEC 61
1.7.2 Applications and Related Works 63

1.7.2.1 Applications in Art 64
1.7.2.2 Applications in Engineering 66
1.7.2.3 Improvement of the IEC 69
1.7.3 Future of IEC 71
1.8 Conclusion 72

2 Facial Expression Recognition based on Mahalanobis-Taguchi System 92
2.1 Introduction 92
2.2 Facial Image Processing 94
2.2.1 Facial Image 94
2.2.2 Edge Extraction 95
2.3 Facial Expression Recognition based on the MTS 98
2.3.1 Recognition Process 98
2.3.2 Extraction of Attribute 99
2.3.3 Construction of Mahalanobis Space 101
2.4 Experimental Method 102
2.5 Experimental Result 104
2.6 Optimization of Attributes 105
2.7 Conclusion 108

3 A Robust Face Identification against Lighting and Face Position Fluctuations 112
3.1 Introduction 112
3.2 Robust Face Identification against Lighting and Face Position Fluctuations 113
3.2.1 Face Identification Process 113
3.2.2 Facial Image 114
3.2.3 Artificial Fluctuations 114
3.2.4 Face Identification based on MTS 115
3.3 Experimental Method 117
3.4 Experimental Result 119
3.4.1 Identification Result 119
3.4.2 Optimization Result 128
3.5 Conclusion 135

4 Face Identification based on Division of Facial Image 142
4.1 Introduction 142
4.2 Face Identification Method based on Division of Facial Image 144
4.2.1 Facial Image 144

4.2.2 Identification Process based on Division of Facial Image . . . 144
4.2.3 Artificial Fluctuation . . . 146
4.2.4 Pattern Recognition based on MTS . . . 147
4.2.5 Identification based on Integration of Mahalanobis distances . . . 149
4.3 Experimental Method . . . 150
4.4 Experimental Result . . . 152
4.5 Discussion . . . 154
4.6 Conclusion . . . 157

5 Face Cognitive Model Acquisition using Interactive Genetic Algorithm 160
5.1 Introduction . . . 160
5.2 Facial Image Database Retrieval . . . 162
5.2.1 Facial Image . . . 162
5.2.2 Facial Image Retrieval Process . . . 164
5.3 Face Cognitive Model Acquisition Method using IGA . . . 165
5.3.1 Cording . . . 165
5.3.2 Presumption Method of Face Cognitive Model based on IGA . . . 166
5.3.3 Genetic Operation . . . 167
5.3.3.1 Fitness Function . . . 167
5.3.3.2 Natural Selection . . . 170
5.3.3.3 Crossover . . . 170
5.3.3.4 Mutation . . . 170
5.4 Experimental Method . . . 170
5.5 Experimental Result . . . 172
5.5.1 Experimental Result of Face Impression Retrieval Task . . . 173
5.5.2 Experimental Result of Similar Face Retrieval Task 176
5.6 Discussion . . . 179
5.7 Conclusion . . . 180

6 Prospect for Development of Comprehensive Face Recognition System 183

7 Summary 187

List of Figures

1.1 Relation model between facial expression and emotion . . 18
1.2 Six basic facial expressions based on FACS 21
1.3 Process of face identification 30
1.4 An example of photographic fluctuations (From MIT face database) 37
1.5 An example of appearance fluctuations (From ORL face database and Purdue University face database) 38
1.6 Retrieval process for the ambiguous retrieval target 46
1.7 Image of sensuous recognition ability acquisition process . 48
1.8 Pattern recognition based on the MTS 50
1.9 Extraction method of the attribute based on differential and integral characteristics 52
1.10 Orthogonal array L_{128} and SN ratio 55
1.11 User's operation and internal behavior in the system based on the IEC 62
1.12 Biomorph by R. Dawkins 65
1.13 Genetic Images by K. Sims 65
1.14 Multi-objective Interactive Genetic Algorithm 72

2.1 An example of facial image 95
2.2 Neighboring patterns 97
2.3 An example of the edge extracted facial image 98
2.4 An example of the attributes based on the differential characteristic and the integral characteristic 100
2.5 An example of the facial image on each facial expression . 103
2.6 Mahalanobis and Euclidean distances of unknown facial images 105
2.7 Optimization result 108

3.1 An example of the edge extracted facial image for each value of α 116
3.2 An example of face position fluctuations 116
3.3 An example of training facial images in LP 118

3.4 Correlation matrix on each subject (A) 122
3.5 Correlation matrix on each subject (B) 123
3.6 Mahalanobis distances between each Mahalanobis space and the unknown images (A) 124
3.7 Mahalanobis distances between each Mahalanobis space and the unknown images (B) 125
3.8 Mahalanobis distances between each Mahalanobis space and the unknown images (C) 126
3.9 Mahalanobis distances between each Mahalanobis space and the unknown images (D) 127
3.10 Identification result of s4 127
3.11 Factorial effects on each subject (A) 130
3.12 Factorial effects on each subject (B) 131
3.13 Factorial effects on each subject (C) 132
3.14 Factorial effects on each subject (D) 133
3.15 Factorial effects on each subject (E) 134
3.16 Mahalanobis distances before and after the optimization between abnormal facial images and the optimization target Mahalanobis space 137
3.17 Mahalanobis distances before and after the optimization of unknown facial images which should belong to the optimization target Mahalanobis space 138
3.18 Mahalanobis distances before and after the optimization between the Mahalanobis space on each subject and the unknown facial images which are created from s4's original unknown facial image 139

4.1 An example of facial images 145
4.2 Identification process based on the division of facial image 145
4.3 An example of artificial face size fluctuations 147
4.4 An example of training facial images involving the combination of three artificial photographic fluctuations 148
4.5 An example of original training facial images and original unknown facial images 151
4.6 Frequency of shortest Mahalanobis distance on each region of unknown facial images which were correctly identified . 153
4.7 Frequency of shortest Mahalanobis distance on each region of unknown facial images which were correctly identified (2×2) 155

4.8 Frequency of shortest Mahalanobis distance on each region of unknown facial images which were correctly identified (4×4) 156
4.9 Frequency of shortest Mahalanobis distance on each region of unknown facial images which were correctly identified (6×6) 156

5.1 An example of facial image 163
5.2 Interface for the facial image database retrieve task method 165
5.3 An example of the coding 166
5.4 Calculation method of the fitness value 168
5.5 An example of uniform crossover 171
5.6 Table for pairwise comparison 175
5.7 Quality of the retrieved facial images 175
5.8 Sum of Euclidean distances between the facial images presented on each step and the target face 178
5.9 Variance of the facial images presented on each step . . . 178

List of Tables

1.1 An example of Action Units 20
1.2 Combination of AUs for six basic facial expressions 21
1.3 Typical biometrics and its classification 29
1.4 Strong and weak points of each clue for face detection . . . 32
1.5 List of studies on face identification 39
1.6 List of face cognitive model acquisitions 42

2.1 Recognition result (%) 105
2.2 The number of the optimized attributes 108

3.1 Identification result (%) 120
3.2 Identification result based on Euclidean distance (%) . . 121
3.3 Results of optimization 129
3.4 Identification result after the optimization 135
3.5 SN ratio before and after the optimization 135

4.1 Identification rate (%) 153
4.2 Identification rate (%) 155

5.1 Experimental results of face impression retrieval task . . . 174
5.2 Experimental results of similar face retrieval task 177

Preface

In this doctoral dissertation, the recognition methods for human face are focused on in order to implement flexible and adaptable human interface. It is clear that human face plays important roles in the face-to-face communication among humans. Human performs smooth communication by obtaining various kinds of information from face and utilizing them. The face can not only represent psychological state such as emotion, intention and interest but also represent social state such as age, gender and personality. The effective human interface can be implemented, if computer can recognize or understand the information which face represents. Therefore, it is required to develop effective recognition methods for the information which face represents, i.e., face recognition methods.

The information represented by face has the features of subjectivity, polysemy, ambiguity and condition dependence which are not treated in conventional computer science. Therefore, in the face recognition methods, the coding which can consider these features is needed. Moreover, in the face recognition, it is required to develop the recognition methods which have the consistency with the face cognitive structure of human. The recognition methods which can have the consistency with the face cognitive structure of human and can flexibly recognize face can be implemented by solving these problems.

This doctoral dissertation is related to a study on the face recognition methods to implement flexible and user-friendly human interface. In the face recognition, facial expression recognition, face identification, face cognitive model acquisition, face detection and so on are included as typical problems. In this dissertation, especially, the face recognition methods based on the Mahalanobis-Taguchi System (MTS) are applied to facial expression recognition and face identification. The facial expression recognition is practical problem to assume the human psychological state. The face identification is useful problem to identify the human social state. The experimental results show the effectiveness of the proposed methods for these problems. Moreover, in this dissertation, the face cognitive model acquisition method based on the Interactive Genet-

ic Algorithm (IGA) is proposed. The face cognitive model of human must be investigated in order to develop the face recognition methods which have the consistency with the face cognitive structure of human. The face cognitive acquisition method based on the IGA enables the computer to acquire the accurate face cognitive model of human. The experimental results reveal the effectiveness of the proposed face cognitive acquisition method and the accuracy of face cognitive model which is acquired by using the proposed method. This dissertation consists of 7 chapters. The brief content of each chapter is as follows.

Chapter 1 is related to the face recognition and the recognition methods. Some general problems in the face recognition are surveyed. Especially, the background, issues and recent works are surveyed in the facial expression recognition, the face identification and the face cognitive model acquisition which are taken up in this dissertation. Additionally, the MTS and IGA which are fundamental techniques for the recognition methods proposed in this dissertation are explained.

In chapter 2, I propose a facial expression recognition based on the MTS. The MTS is a pattern recognition method which can reflect human sensuous knowledge for object recognition. However, the MTS is a pattern recognition method for dividing data into two kinds. In the facial expression recognition, some facial expressions have to be treated. To apply the MTS to the facial expression recognition, I propose the extension method of the MTS. Moreover, in the MTS, the extraction of attributes which contribute to the recognition can be carried out. This operation is called the optimization of attribute. However, in the conventional optimization method, the relations among the attributes are not considered. To solve this problem, I also propose an optimization method which can consider the relations. Namely, the optimization method based on the Genetic Algorithm is proposed. The experimental results using some subjects show the effectiveness of this application and optimization method.

In chapter 3, I propose a face identification method based on the insertion of artificial fluctuations. As one of important problems which have to be solved in the face identification, there is the implementation of robustness against fluctuations in face photography such as lighting condition or face position. Many training facial images which are taken in various environments are required for the implementation of the robustness. Moreover, many training facial images have to be collected for implementing the recognition based on the MTS. It is burden for the user

to take all training facial images in real environment. To reduce user's burden, in this proposed method, the implementation of the robustness is performed by inserting artificial fluctuations to an original training facial image. The experimental results show the robustness and effectiveness of the proposed method.

In chapter 4, a face identification method based on the division of region on facial image is proposed. The photographic fluctuations which inevitably occur in taking the facial image are treated in chapter 3. In the face identification, the robustness against the appearance fluctuations such as background and hairstyle also has to be implemented. The implementation of the robustness against appearance fluctuations is harder problem than the photographic fluctuations from pattern recognition point of view. In the conventional face identification, the attributes used for the identification are extracted from a whole of facial image and then the identification is performed on the basis of an evaluation criterion which is calculated by using these attributes. However, it is difficult to implement the robustness against appearance fluctuations by using the conventional method because the extracted attributes include a large amount of noise. In the proposed method, the robustness against appearance fluctuations is implemented by dividing a facial image into some regions and then performing the identification based on the integration of some evaluation criteria which are obtained on each region. Some experimental results reveal that the proposed method is robust against appearance fluctuations in comparison with the conventional method.

In chapter 5, I propose a face cognitive model acquisition method using the IGA. The face cognitive model of human must be investigated in order to develop the face recognition methods which have the consistency with the face cognitive model of human. In this chapter, I attempt to acquire face cognitive model of human on the basis of the facial image database retrieval task method. Moreover, in the face cognitive model acquisition based on the facial image database retrieval task method, more accurate face cognitive model acquisition is performed by using the IGA which can consider the cognitive feature of human. Experimental results reveal that the proposed method can accurately acquire the face cognitive model of human.

Chapter 6 is concerned with the prospect for the development of comprehensive face recognition system. The recognition methods for the facial expression and the face identification are proposed in this dissertation. Moreover, the method to acquire the face cognitive model of

human is proposed. The comprehensive face recognition system which has the consistency with the face cognitive structure of human and can flexibly recognize face can be implemented by combining these methods. The prospect for the development of such face recognition system and application fields in which the face recognition system can be utilized are discussed.

Chapter 7 is general conclusion of this dissertation.

Chapter 1

Face Recognition and Recognition Methods

1.1 Introduction

In face-to-face communication, human obtains the information about other person's intention or state from face and decides own behavior on the basis of the information. Therefore, the information which face represents is important element in the communication among humans and can make the communication smooth. It is obvious that more flexible and user-friendly human interface can be implemented if computer can flexibly recognize or understand the information which face emits and then can adaptably decide own behavior. From this background, recently, the research on the recognition of the information which face represents has been carried out, and it has been required to develop the effective recognition methods for the information which face represents, i.e., face recognition methods.

However, the information represented by face differs from the information which is treated in the conventional computer science. For example, it is difficult to represent a face by using words or numerical values. The impression of face is different on each person and can not be uniquely determined. In addition, the interpretation of the information which face represents changes by the experience or situation of person. That is, this information has the features of subjectivity, polysemy, ambiguity and condition dependence. Therefore, in the face recognition, the coding which can consider these features has to be employed. Moreover, it is desirable that the recognition of face on computer is performed in the same way as the face recognition structure of human. Thus, in the face recognition, the recognition methods which have the consistency with the face cognitive structure of human have to be developed. The effective face

recognition methods can be implemented by solving these problems.

The purpose of this study is to develop effective face recognition methods for the implementation of flexible and adaptable human interface. In this dissertation, facial expression recognition, face identification and face cognitive model acquisition are treated as typical problems in the face recognition. In addition, the Mahalanobis-Taguchi System (MTS) and the Interactive Genetic Algorithm (IGA) are employed as fundamental techniques for the face recognition methods which are proposed for these problems.

In this chapter, at first, general problems in the face recognition are introduced. Especially, backgrounds, issues, recent works and related works are surveyed in the facial expression recognition, face identification and face cognitive model acquisition. Furthermore, the detail of the MTS and IGA used in the proposed face recognition methods are explained.

1.2 Face Recognition

This section is concerned with the introduction of face recognition. The face recognition and some problems included in face recognition are briefly introduced.

Recently, the research on the recognition of the information which face represents has been carried out [1]. This subject is called face recognition. The aim of face recognition is to implement effective human interfaces based on the interaction between human and computer. Therefore, the face recognition includes some problems for the implementation of such human interface; i.e., facial expression recognition, face identification, face cognitive model acquisition, face detection, facial image retrieval, caricature face. The facial expression recognition is the problem related to the presumption of emotion or psychological state of person. The face identification is useful to identify social state of person, and it can be utilized for security system or access control system. The researches on the face cognitive model acquisition and the face detection are carried out in order to support the development of effective methods for other problems. The caricature face is used to implement the entertainment system or to support the face cognitive model acquisition. These problems are related to each other. The development of effective methods in one problem leads to the improvement of the methods for other problems. For example, the performance of facial expression recognition can be improved by introducing the face identification function. That is, the personalization of

facial expression recognition can be performed. Therefore, the effective face recognition can be implemented by developing the effective method for each problem. The development of effective human interface based on the interaction between human and computer can be supported by the implementation of such face recognition. From this background, recently, the development of effective face recognition methods for these problems has been attempted. The brief explanation of each problem is indicated below.

1. **Facial expression recognition:** It is obvious that effective human interfaces can be implemented if computers can assume the user's intention or psychological state and then can adaptably decide own behavior. On the other hand, it has been reported that the facial expressions of human reflect about 55 % of human emotion. Thus, the user's intention or psychological state can be assumed if the facial expression which the user expresses can be correctly recognized. From such background, the research on the facial expression recognition has been carried out worldwide. In the facial expression recognition, the facial expressions of anger, disgust, surprise, sadness, fear and happiness are frequently used in order to evaluate the face recognition methods. These facial expressions are called six basic facial expressions. It is reported that in the six basic facial expressions, there is the facial expression pattern which does not depend on race, age, sex, area and so on. Therefore, in the facial expression recognition, it is important to find this facial expression pattern and then perform the facial expression recognition based on this facial expression pattern. The detail of the facial expression recognition is described in Section 3 of Chapter 1.

2. **Face Identification:** Recently, the identification utilizing the human face has been focused on in order to implement advanced and flexible security systems and access control systems. The identification using human face is called face identification. Comparing with the identification based on other information such as fingerprint, eyeprint and so on, the face identification has the advantages that it is natural for the user and it has efficiency of restraining crime. The major problem in the face identification is the implementation of robustness against various fluctuations. The various fluctuations such as the lighting, position and size fluctuations are included in the facial image when the facial image used for the identification is taken. The robustness against these fluctuations has to

be implemented in the face identification because these fluctuations inevitably occur when the facial image is taken. Moreover, in the face identification, the robustness against appearance fluctuations such as hairstyle, glasses and so on must be implemented in order to perform practical face identification. Therefore, a lot of studies on the face identification focus on the implementation of the robustness against such fluctuations. The detail of the face identification is described in Section 4 of Chapter 1.

3. **Face Detection:** The facial images which are used for the face recognition include various objects except the target face such as background. The face recognition can be performed by using such facial images. However, the region which includes only face has to be extracted from the facial image in order to improve the performance of the face recognition. Therefore, the research on the extraction of the region including target face, i.e., face detection, has been carried out. In the face detection, there are the face detections for two kinds of facial image. One is the face detection for stationary facial image, and another is the face detection for facial image sequence. In these types of facial image, the face detection is performed based on some clues; i.e., motion, color and shape. There are strong point and weak point in each clue. The clue which is used for the face detection is properly chosen in accordance with the purpose and the kind of facial image.

4. **Face cognitive model acquisition:** In the face recognition, the face recognition methods which have the consistency with the face cognitive structure of human have to be developed in order to utilize them for the implementation of the effective human interface. Therefore, the research on the investigation of face cognitive model which human has is busily carried out in engineering field. This subject is called face cognitive model acquisition. In the conventional face cognitive model acquisition, the face photograph is frequently utilized as the stimulus. However, the face cognitive model acquisition using the facial image as the stimulus came to be performed by the development of the computer graphics and animation technologies. Recently, various face cognitive models have been investigated by using the facial image, and the development of face recognition methods which have the consistency with the face cognitive structure of human has been attempted. The detail of the face cognitive model acquisition is explained in Section 5 of Chapter 1.

5. **Facial image retrieval:** The recent diversification of information has brought it with the problem that the image data must be treated in the information retrieval. However, the image data are subjective and ambiguous retrieval objects. That is, the impression received from image datum is different on each person and it is hard to express clearly the image data by verbal keywords. Therefore, it is difficult to implement the image retrieval by using the conventional method such as the retrieval method based on verbal keywords. Recently, the research on the image retrieval in which the facial image is employed as the retrieval object has been busily carried out because the facial image is considerably subjective and ambiguous retrieval object. In the facial image retrieval, the methods based on the contents of facial image and on adjective keywords have been proposed.

6. **Caricature face:** Recently, the creation of caricature face from facial image by computer has been attempted. The caricature face is drawn by emphasizing the characteristic feature on each person's face. Therefore, it is necessary to detect the characteristic feature on face. The face recognition methods such as the average face, the Karhunen-Loeve expansion (KL expansion) and so on which are utilized in other problems are employed for the caricature face. The development of face recognition techniques for other problems leads to the improvement of caricature face. The effective caricature face can allow the support of face cognitive model acquisition and entertainment system.

The research on the face recognition which is mentioned above has been carried out in U. S. since the 1990's middle. This is concerned with the development program called FERET (FacE REcognition Technology). The FERET is the program started in 1993 to develop the technology for the facial image retrieval which is used for criminal investigation and to turn face recognition technology to practical use. Strictly speaking, the purpose of FERET is the construction of facial image database which is used to develop and evaluate the algorithm for face recognition, and the subsidization of grant-in-aid and its evaluation through the contest of algorithm. The laboratories such as the MIT, Rockefeller University and so on which have excellent results for the face recognition participate in the contest of FERET. Hence, it is clear that the research on the face recognition will be carried out more busily and worldwide hereafter.

In this dissertation, especially, the facial expression recognition, the face identification and the face cognitive model acquisition are treated.

In the following section, the detail of each problem which are focused on in this dissertation is described.

1.3 Facial Expression Recognition

Recently, it has been required to implement effective and interactive communication between human and computer in order to develop natural and user-friendly computer environment [2]. The information in the communication among humans does not only include logical information but also include psychological information such as Kan-sei or emotion. Therefore, in the communication between human and computer, such information must be treated in the same way as the communication among humans. On the other hand, it is obvious that the emotion or psychological state of human is reflected to facial expression. Thus, the effective and interactive communication between human and computer can be implemented if computer can recognize or understand emotion or psychological state of human on the basis of facial expression. Recently, the research on facial expression recognition has been busily carried out in order to assume the emotion or psychological state of human from facial expression. In this section, the survey of facial expression recognition is described. First, the research on the relation between facial expression and emotion is introduced. Secondly, the Facial Action Cording System (FACS) which is frequently used in the study on facial expression is explained. Finally, the algorithms which are used for the facial expression recognition and the related works are surveyed.

1.3.1 Relation between Facial Expression and Emotion

The purpose of facial expression recognition is to assume the emotion or psychological state of human from facial expression. Therefore, the relation between facial expression and emotion has been investigated. The research on the investigation of the relation between facial expression and emotion is introduced below.

In the research related to the investigation of the relation between facial expression and emotion, two-dimensional space model proposed by Schlosberg is famous. The model proposed by Schlosberg is shown in Figure 1.1. In this model, the horizontal axis means "pleasant (P) - unpleasant (U)" and vertical axis means "attention (A) - rejection (R)", and six basic emotional categories are put in the periphery of circle. The facial expression is put in this circle in accordance with the criteria of

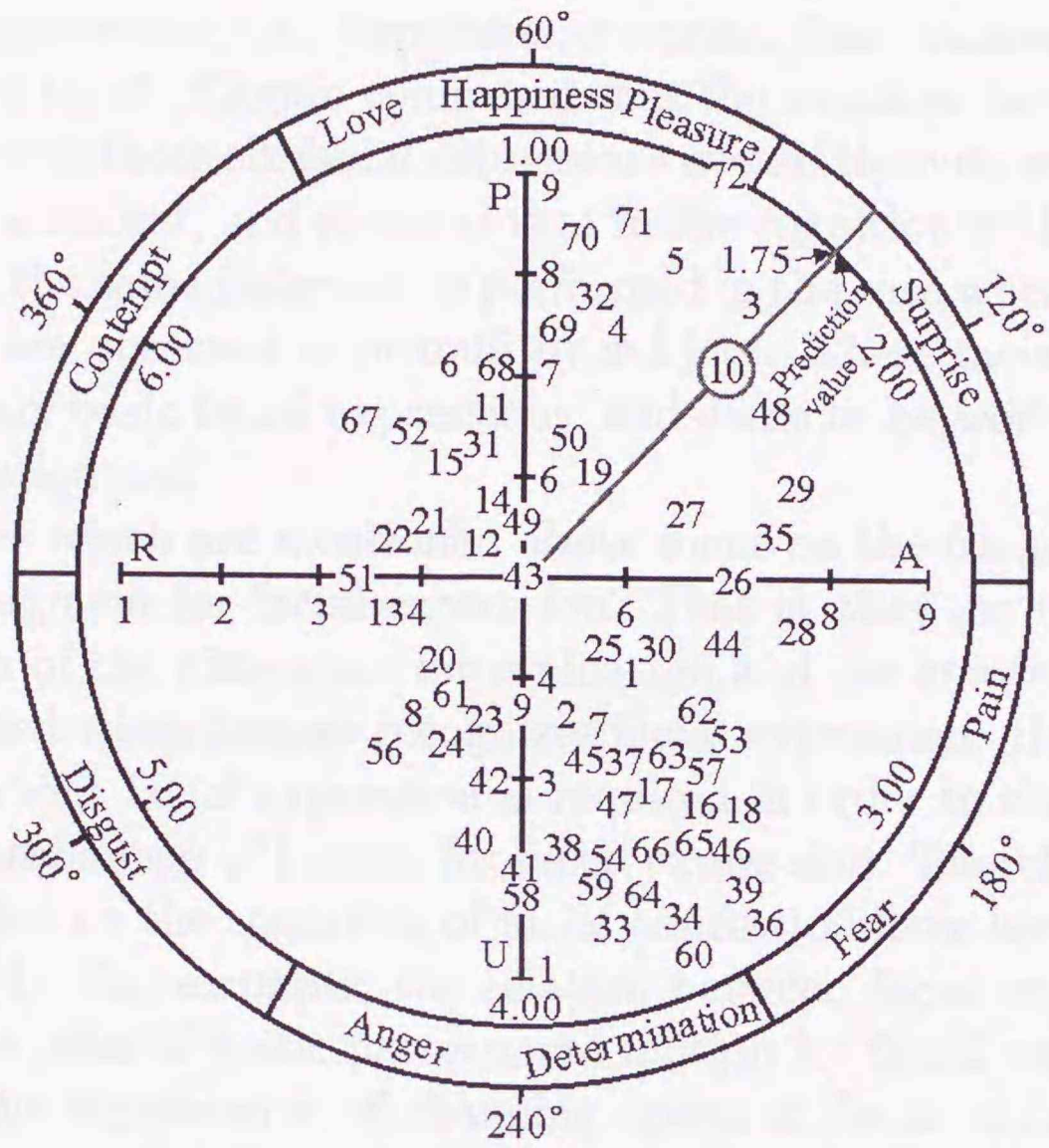


Figure 1.1: Relation model between facial expression and emotion

horizontal and vertical axes. The emotion corresponding to facial expression is assumed on the basis of the position where the facial expression is put. That is, the line between the center of circle and the position of facial expression is drawn, and then the emotion which is referred by extending this line is the emotion corresponding to facial expression. In Figure 1.1, the arrangement of the seventy-two facial expression images which were used to evaluate this model by Schlosberg is indicated as the numbers. Especially, the emotion for the facial expression image which is expressed as 10 is indicated. This model indicates that when human recognizes other person's facial expression, human evaluates it in these axes and the emotion corresponding to it is classified on the basis of this evaluation. A lot of the studies on the investigation of the relation between facial expression and emotion have been carried out after the proposal of this model. However, as the result, those studies brought the adequacy to this model.

On the other hand, the research on the basic emotion category which is recognized from facial expression has been energetically carried out

by Ekman and Izard. Especially, Ekman investigated the universality of six facial expressions, i.e., happiness, surprise, fear, sadness, anger and disgust. As a result, Ekman confirmed that the common facial expression pattern exists in these six facial expressions even if the culture background of subject is different, and revealed that in the cognition of these six facial expressions, the same judgment is performed in the case where these facial expressions are observed as stimuli by subject. These facial expressions were called six basic facial expressions, and came to be worldwide known by this investigation.

The studies which are mentioned above focus on the framework for the cognitive judgment for facial expression. That is, they are related to the investigation of the dimension for evaluation and the emotional category which are used when human recognizes facial expression. However, more knowledge about facial expression is required in order to clarify the psychological mechanism of human for facial expression. Therefore, recently, various studies on the cognition of facial expression have been busily carried out [3][4]. For example, the relation between facial expression and angle [5], the roles of static pattern and motion for facial expression cognition [6], the significance of changing speed of facial organs for facial expression cognition [7] and so on were investigated. It is likely that the investigation of facial expression cognition will be busily performed in the future because the investigation of facial expression cognition is needed to develop effective facial expression recognition methods on computer.

1.3.2 Facial Action Coding System

In the research on facial expression, the Facial Action Coding System (FACS) is frequently utilized. The FACS is a system to describe facial expression, and which was proposed by Ekman and Friesen [8]. Basically, in the FACS, facial expression is described by utilizing the movement of the muscle on face. The muscles on face are divided into 44 parts, and it is called Action Unit (AU). In each AU, the rule of change on face is described. The arbitrary facial expression can be described by using the combination of the AU and its strength which is the amount of movement. An example of AUs is shown in Table 1.1. The name, number and anatomical basis of each AU are described in Table 1.1. Most of AUs involves a single muscle. The number is arbitrary and does not have any significant except that 1 through 7 refer to brows, forehead or eyelids.

There are the facial expressions called six basic facial expressions. They are anger, disgust, fear, happiness, sadness and surprise. It has been re-

AU Number	FACS Name	Muscular Basis
1	Inner Brow Raiser	Frontalis, Pars Medialis
2	Outer Brow Raiser	Frontalis, Pars Lateralis
4	Brow Lowerer	Depressor Glabellae; Depressor Supercilli; Corrugator
5	Upper Lid Raiser	Levator Palpebrae Superioris
6	Cheek Raiser	Orbicularis Oculi, Pars Orbitalis
7	Lid Tightener	Orbicularis Oculi, Pars Palebralis
8	Lip Toward Each Other	Orbicularis Oris
9	Nose Wrinkler	Levator Labii Superioris, Alaeque Nasi
10	Upper Lip Raiser	Levator Labii Superioris, Caput Infraorbitalis
11	Nasolabial Furrow Deepener	Zygomatic Minor
12	Lip Corner Puller	Zygomatic Major
13	Cheek Puffer	Caninus
14	Dimpler	Buccinnator
15	Lip Corner Depressor	Triangularis
16	Lower Lip Depressor	Depressor Labii
17	Chin Raiser	Mentalis
18	Lip Puckerer	Incisivii Labii Superioris; Incisivii Labii Inferioris
20	Lip Stretcher	Risorius
22	Lip Funneler	Orbicularis Oris
23	Lip Tightner	Orbicularis Oris
24	Lip Pressor	Orbicularis Oris
25	Lips Part	Depressor Labii, or Relaxation of Mentalis or Orbicularis Oris
26	Jaw Drop	Masetter; Temporal and Internal Pterygoid Relaxed
27	Mouth Stretch	Pterygoids; Digastric
28	Lip Suck	Orbicularis Oris
38	Nostril Dilator	Nasalis, Pars Alaris
39	Nostril Compressor	Nasalis, Pars Transversa and Depressor Septi Nasi
41	Lid Droop	Relaxation of Levator Palpebrae Superioris
42	Slit	Orbicularis Oculi
43	Eyes Closed	Relaxation of Lavator Palpebrae Superioris
44	Squint	Orbicularis Oculi, Pars Palpebralis
45	Blink	Relaxation of Levator Palpebrae and Contraction of Orbicularis
46	Wink	Oculi, Pars Palpebralis Orbicularis Oculi

Table 1.1: An example of Action Units

Basic Emotion	Combination of AU parameters
Surprise	AU1-(40), 2-(30), 5-(60), 15-(20), 16-(25), 20-(10), 26-(60)
Fear	AU1-(50), 2-(10), 4-(80), 5-(60),15-(30), 20-(10), 26-(30)
Disgust	AU2-(60), 4-(40), 9-(20), 15-(60), 17-(30)
Anger	AU2-(30), 4-(60), 7-(50), 9-(20), 10-(10), 20-(15), 26-(30)
Happiness	AU1-(65), 6-(70), 12-(10), 14-(10)
Sadness	AU1-(40), 4-(50), 15-(40), 23-(20)

Table 1.2: Combination of AUs for six basic facial expressions

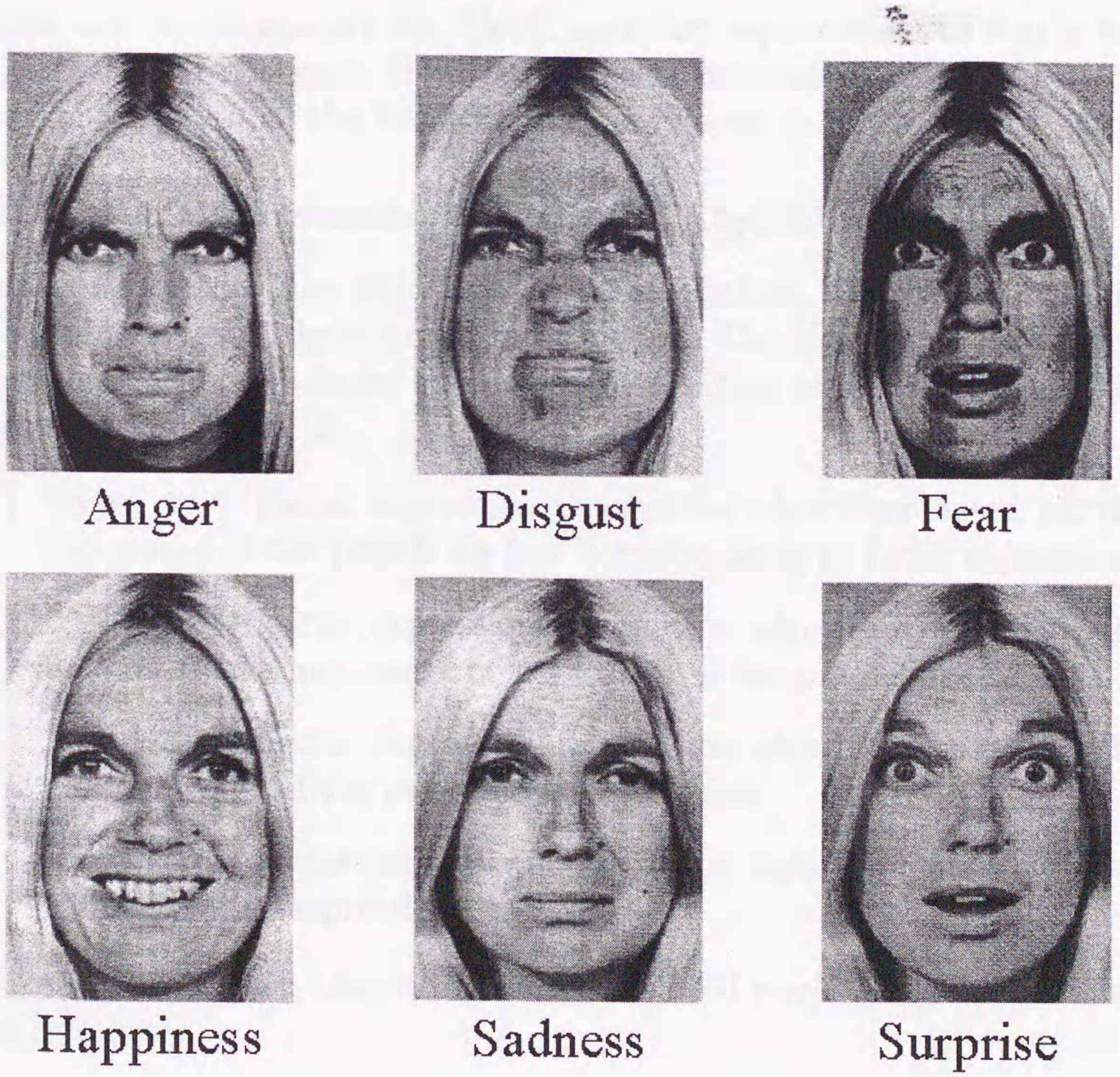


Figure 1.2: Six basic facial expressions based on FACS

ported that the six basic facial expressions have common facial expression pattern which does not depend on culture, sex, age and so on. These six basic facial expressions are often used for the studies on the facial expression recognition. The combination of AUs and the strength for each AU for these six basic facial expressions are shown in Table 1.2. Moreover, an example of six basic facial expressions based on FACS are illustrated in Figure 1.2.

The FACS is used in the various fields; e.g., facial expression recognition, facial expression composition, facial expression cognition and so on. However, human can recognize facial expression from the facial diagram which are drawn by combining simple lines and curves, or from cartoon. That is, human recognizes facial expression on the basis of the visual state of face and the change of face. Therefore, there are some objections which are raised against the FACS, and the improvement of FACS has been recently performed. However, in recent research on facial expression recognition method, the FACS is often employed.

1.3.3 Facial Expression Recognition Algorithm

In this subsection, the facial expression recognition algorithm and some studies using each algorithm are introduced. The algorithms for the implementation of the facial expression recognition are classified into the following four types [9].

1. **Type 1:** Facial expression recognition algorithm based on the movement of the muscle on face corresponding to facial expression
2. **Type 2:** Facial expression recognition algorithm based on the amount of the movement of the feature point pre-defined on face
3. **Type 3:** Facial expression recognition algorithm based on the global change of the internal feature on face
4. **Type 4:** Facial expression recognition algorithm utilizing time series of facial expression

The detail of each algorithm and the related works in each algorithm are as follows.

1.3.3.1 Type 1

As the facial expression recognition algorithm based on the movement of the muscle on face corresponding to facial expression, the facial expression

recognition method using the optical-flow has been proposed. Moreover, there is the method based on the FACS as the facial expression recognition using the algorithm of type 1. The facial expression recognition methods using these algorithms are introduced below.

Mase et al. proposed a facial expression recognition method using the optical-flow [10][11]. The optical-flow is one of analysis methods for moving object. The movement of object is analyzed by the optical-flow on the basis of the information of luminance in image data. The movement of object is represented as a vector. In the proposed method, the vector on the region which includes the muscle for facial expression is calculated by performing the optical-flow analysis using two frames in facial image sequence. The facial expression recognition is performed by the nearest neighbor method based on these vectors calculated on some regions. They designed the experiment to confirm the performance of the proposed method. It was found from the experimental results that the success rate for the facial expression recognition in the proposed method was over 80 %.

Moreover, Prochazka et al. also proposed a facial expression recognition method using the optical-flow [12]. However, in this method, the vectors used for the facial expression recognition are extracted by the optical-flow, and the Karhunen-Loeve expansion (KL expansion) is applied to these vectors in order to decrease the number of dimension of vector. The facial expression recognition is performed on the basis of the distance among the vectors in the KL space. In addition, they also proposed a facial expression recognition using the optical-flow and neural network.

In the FACS, the AUs which are used for each facial expression are defined. Therefore, the facial expression recognition can be implemented if the AUs which are used on facial expression image can be estimated. Choi et al. proposed a facial expression analysis method based on the FACS [13]. In the proposed method, two stationary facial images are used for the presumption of facial expression. One is the facial image of neutral face and another is the facial image in which an arbitrary facial expression is expressed. As first step of the facial expression analysis, the position of head in the facial image in which facial expression is expressed is adjusted to the position of head in the facial image of neutral face. The movement of the internal feature on face such as eyes can be correctly extracted by this operation. And then, the vector of the movement of the feature points for the internal feature which is predefined on face is calculated on the basis of the difference among these facial images.

The AUs on the facial expression image are estimated on the basis of the calculated vectors. This estimation is performed by utilizing tree structure. They designed the experiments with some subjects in order to confirm the effectiveness of the proposed method. The experimental results revealed that the AUs which were estimated by using the proposed method corresponded to the AUs judged by psychologist.

Moreover, several facial expression recognition methods based on the FACS have been proposed [14][15][16]. In the facial expression recognition method proposed by Otsuka et al., some feature points are put on the face [14]. The triangular regions on face are calculated by connecting these feature points. The node in each region is decided on the basis of feature points on each region. The correspondences between the movement of the nodes for these regions and the AU are calculated. The AUs used on the facial expression image are estimated on the basis of the correspondences. The facial expression recognition is performed by the AUs estimated.

In the facial expression recognition method proposed by Kawakami et al, the movement amount of the feature points which are put on facial organs is employed as the input data in neural network [15]. The AUs corresponding to the movement amount are obtained as the output of neural network. The facial expression on facial expression image is decided by the neural network which employs these AUs as the input data.

In the method proposed by Ando et al., seven parameters are used for facial expression recognition [16]. They are the tilt and position of eyebrow, the vertical length of eye, the position of upper lip and lower lip, the position of mouth ends, and the width of mouth. The AUs are decided by tree structure on the basis of the change of these parameters. The facial expression recognition is performed by comparing these AUs with the AUs for the six basic facial expressions.

1.3.3.2 Type 2

In the facial expression recognition method based on the amount of the movement of the feature point pre-defined on face, it is assumed that the feature points on face can be completely detected. In this method, the changing amounts of the geometric data for the internal features on face are employed as the parameters for the facial expression recognition. These data are detected on the basis of the feature points which are put on face in facial expression image. The facial expression recognition is carried out by comparing the data for unknown facial expression image with the data for each facial expression. An example of the facial expression

recognition methods based on the amount of the movement of the feature point predefined on face is introduced below.

Ishii et al. proposed a facial expression recognition method based on the change of parameters for facial organs [17]. In this method, four parameters for facial organs are employed; i.e., curvature of mouth, the ratio of length and breadth of mouth, the degree of eye open, and the tilt of eyebrow. The facial expression recognition is carried out on the basis of these parameters for facial expressions and the modified Euclidean distance. Through the experiment with some subjects, they confirmed that the facial expression recognition can be implemented by the proposed method.

Kobayashi et al. also proposed a facial expression recognition based on the movement of facial organs [18]. In the proposed method, at first, the contours of the facial organs in the facial image are automatically extracted. In the contour extraction of facial organs, elastic contour model and dynamic programming are employed. In the contour extraction based on the elastic contour model, some control points are put on the facial organs. These control points are connected to each other by spring model. The contour of facial organ is decided by minimizing the elasticity energy of this model. This minimization of elasticity energy is performed by using the dynamic programming. Moreover, the facial expression recognition is implemented by using the Categorizing and Learning Module (CALM) which is one of the self-organized neural networks. The amount of movement of control points which are decided by elastic contour model is used as the input data for the CALM. They designed the experiment with some subjects in order to confirm the effectiveness of the proposed method. The experimental results revealed that it is required to improve the proposed method because the success rate of facial expression recognition for six basic facial expressions was 75 % at most.

1.3.3.3 Type 3

As the facial expression recognition based on the global change of the internal feature of face, the facial expression recognition methods using the Potential Net, the spatial frequency, and the template matching have been proposed. In these methods, the facial expression recognition methods using the Potential Net are introduced below.

In the conventional facial expression recognition on computer, the shapes and positions of the internal features on face such as eye or nose are extracted from facial image and then the facial expression recognition is

performed on the basis of this information. However, it is difficult to extract stably the shapes and positions of the internal features from various facial images. Therefore, it is necessary to develop the facial expression recognition method which can treat a facial expression image as a pattern in order to implement robust facial expression recognition. Matsuno et al. proposed a facial expression recognition method using the Potential Net and the KL expansion [19]. The Potential Net is the physical model which can extract the characteristic of internal feature of face as a pattern from facial image. In the facial expression recognition using the Potential Net, some nodes which are connected to each other by spring model on two-dimensional grid are put on facial image. These nodes move to the positions of internal features in facial image on the basis of the external force and elastic force. The form of Potential Net considerably reflects the characteristic of face because the Potential Net is transformed by the movement of nodes. The vector which consists of the movement amount of nodes is calculated in the transformed Potential Net. Moreover, the KL expansion is applied to the vector in order to compose the low-dimensional space which is effective for the facial expression recognition. The facial expression recognition is performed on the basis of the distance in this space between the inputted vector and the vector for each facial expression. They made some experiments and revealed that in the proposed method, the success rate for the facial expression recognition of anger, happiness, surprise and sadness was over 90 %.

Bessho et al. also proposed facial expression recognition using the Potential Net [20]. However, in this method, the face detection based on the Potential Net in addition to the facial expression recognition based on the Potential Net is employed in order to implement more effective facial expression recognition. The face detection is performed by comparing the template Potential Nets prepared previously with the Potential Net which is calculated for the unknown facial image. They designed the experiment to confirm the effectiveness of proposed method. From the experimental results, they found that the proposed method could allow the effective implementation of facial expression recognition in the case where the facial images which involve background and are various sizes were employed as the recognition target.

1.3.3.4 Type 4

In the facial expression recognition utilizing time series of facial expression, the facial expression recognition method using the Hidden Markov

Model (HMM) has been proposed. Moreover, there is the facial expression recognition method based on the Recurrent Neural Network. The facial expression recognition method using the HMM is introduced below.

Facial expressions basically are related to the change of time. Therefore, it is difficult to maintain the stability of performance for the facial expression recognition in the case where the stationary facial images are employed. Sakaguchi et al. proposed a facial expression recognition method using the HMM in order to implement the facial expression recognition which can consider the sequential change of face for facial expression [21]. In the proposed method, the Wavelet filtering process is applied to each frame of the facial image sequence. The feature vectors are extracted on the basis of the average power in each frequency band in the Wavelet filtering. The vector quantization is performed in order to convert vector into the symbol which is inputted to HMM. The vector quantization is performed with the category-separated code book. The symbol sequence is inputted to the HMM, and the facial expression recognition is performed on the basis of the category which has highest probability. They designed the experiments with a specific person in order to confirm the effectiveness of the proposed method. The experimental results indicated that the average success rate for four facial expressions of surprise, sadness, smile and anger was 93.7 % in the facial expression recognition using the proposed method.

1.4 Face Identification

So far, password and ID card have been generally used to identify a person. However, there is the possibility that persons forget the password or lose the ID card in the case where the identification is performed by using them. To solve this problem, recently, the identification based on human biometrics have been focused on in order to implement more convenient and advanced person identification system. The biometrics indicates the biometrical characteristics of human. This biometrics has the following advantages for the identification.

- The biometrics is characteristic of each person.
- It does not change for a long time.
- It is difficult to fake or imitate it.

Some typical biometrics and its classification are shown in Table 1.3. As shown in Table 1.3, there are fingerprint, iris, face and so on as typical biometrics. The fingerprint is famous biometrics because it has been used for the criminal investigation from old times. Recently, in the identifications based on the biometrics, face identification is focused on because it has more effective advantages for the identification over the identification based on other biometrics [22][23]. The advantages in the face identification are as follows.

- When a person is identified by human, the identification is almost performed by using face. Therefore, the face identification is natural for human, and the identification which does not make the user uncomfortable can be implemented in comparison with other biometrics. Namely, it is easy to be accepted by people.

- The user does not need to do special actions for the identification. For example, in the identification using fingerprint, the user has to touch the apparatus to take the fingerprint used for the identification. In the face identification, the user has only to show own face to a camera.

- The face identification does not need special apparatus and environment. The apparatus needed for the face identification is a camera only. In addition, the face identification can be fundamentally performed on any environment.

- The face identification has the effect for restraining crimes because it is unnatural that the user shows own face to the identification system only and does not show it to people being around him/her.

- It is difficult to imitate or fake other person's face. Therefore, it is impossible to increase the success rate of funny business.

- The introduction cost is not so expensive because particular apparatuses are not required for the face identification.

- The registration of the user into the system is easy because the user's facial image is only taken in the registration for the face identification.

As described above, the development of effective face identification system is very useful for the progress of advanced security system or flexible user interface.

Classification	Biometrics
Morphological biometrics	Face, Fingerprint, Eyeprint, Lip, Palm, Iris
Functional biometrics	Voiceprint, Autography, Action
Biochemical biometrics	Blood, Saliva, DNA

Table 1.3: Typical biometrics and its classification

1.4.1 Face Identification Process

The process of face identification is illustrated in Figure 1.3. As shown in Figure 1.3, at first, the type of target face has to be decided. That is, the kind of the facial image employed for the face identification is decided. After this decision, the facial images based on the decided type are collected. The pre-processing is applied to the collected facial image in order to promote the performance of face identification. And then, the attribute used for the identification is extracted from the facial image. Finally, the identification is performed on the basis of the extracted attribute. The detail of each step is described below.

1.4.2 Target Facial Image

At first, the type of facial image employed in the face identification has to be decided. The following items must be judged for the decision of type of facial image.

- **Dimension:** In the dimension of facial image, there are two types; i.e., two-dimensional facial image and three-dimensional facial image. The face can be measured on three dimensions by using the three-dimensional measurement apparatus and technique. There are several studies on the implementation of face identification based on three-dimensional facial image. The accurate face identification may be able to be implemented by using three-dimensional facial image because the three-dimensional facial image can have more information about face than two-dimensional facial image. However, a three-dimensional scanner is necessary to measure face on three dimensions, and data size is considerably large. Therefore, the environment to install the face identification system is strongly restricted. In addition, there is the problem in terms of the computational cost. On the other hand, in the two-dimensional facial image, the loss of face information arises because face is originally three-dimensional objec-

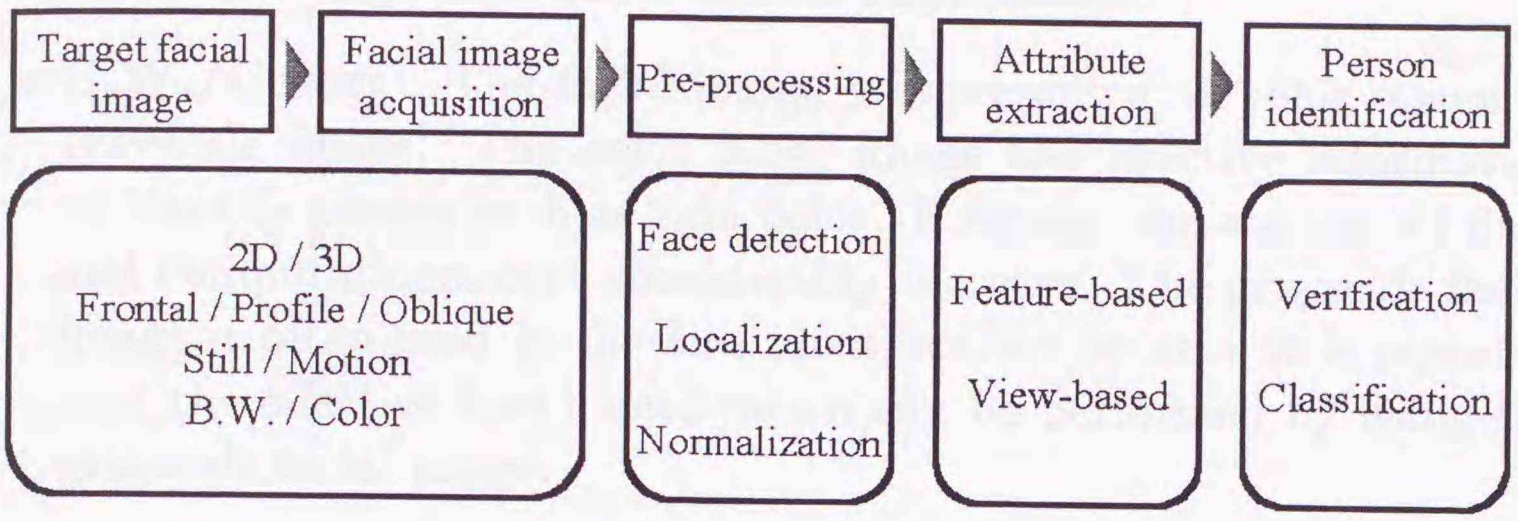


Figure 1.3: Process of face identification

t. However, two-dimensional facial image has enough information to perform the face identification, as human can identify a person in the case where the two-dimensional facial image is used. Moreover, there are some advantages in two-dimensional facial image in comparison with three-dimensional facial image; i.e., the computational cost is low, the environment to install is not restricted, and the facial image can be easily taken by a camera. For these advantages, two-dimensional facial image is frequently used in many studies on the face identification.

- **Angle:** In the face identification, it is necessary to decide the viewpoint for face; i.e., the angle of face. It makes hard the face identification to allow the user to input his/her face in various angles. Therefore, the research on the face identification is carried out on the basis of the assumption that the user is allowed to input the face in a particular angle. As the angle of face, there are frontal facial image, profile image and oblique facial image. Several studies on the face identification based on profile have been reported from latter half of 1970s. In addition, the studies on the face identification using the combination of frontal face and profile for the face identification has been reported. The facial image used in the recent face identification is almost frontal face although some studies on the face identification based on profile are reported. The identification using the oblique face has been not almost reported.

- **Still/Motion:** In the type of facial image, there are still and motion. The stationary facial image is usually employed in the face identification while the facial image sequence is frequently used in the facial expression recognition. The face identification based on

the facial image sequence is seldom implemented.

- **B.W./Color:** The facial image is represented as color image or grayscale image. The color facial image has effective information to identify person such as skin color. However, the amount of data and computational cost considerably increase. The grayscale facial image is often used in the face identification because it is reported that the effective face identification can be performed by using the grayscale facial image.

The type of facial image employed in the face identification is decided by combining these items; e.g., two-dimensional, frontal, stationary and grayscale facial image.

1.4.3 Pre-processing

The face identification can be accomplished by using the facial image involving the objects except face such as background. However, pre-processing is required to improve the performance of face identification. The face detection, localization and normalization are employed as the typical pre-processing for the face identification.

The face detection is the operation to find and extract the region of face in the image. As clues to implement the face detection, there are motion [24], color [25] and shape [26]. Table 1.4 indicates the strong and weak points of each clue.

The motion is effective clue for the face detection in the image sequence. In addition, the face detection based on motion is useful in the case where the real-time face detection is required. However, the accuracy for the face detection is not so high. Therefore, the face detection based on the motion is used to catch broadly the region in which the face exists.

It is known that the skin color is the effective clue for the detection of face from the image in which the complicated background is involved. Therefore, the face detection based on color is performed on the basis of the skin color in the image. However, the performance of this face detection depends on the state of background. Namely, the face detection based on the color is hard when the color of background is similar to the color of skin. Therefore, the face detection in which the relationship between feature of background and feature of face can be considered has to be employed.

As the approach which focuses on the shape of face, the methods which utilize the symmetry of face, the circle model for contour of face and so

Clue	Strong pint	Weak point
Motion	Motion is suitable for the face detection in the image sequence.	The face detection is hard when there are several moving objects in the image sequence.
Color	The face detection based on color is robust against the change of distance and angle. The processing speed is fast.	It is difficult to detect only a face from human body. The accurate decision of face position is difficult.
Shape	It is possible to decide correctly the face position on the basis of the contour of face or the internal features on face.	The face detection based on shape is often influenced by the change of angle and distance. The computational process is complicated.

Table 1.4: Strong and weak points of each clue for face detection

on have been proposed.

Moreover, several face detection methods based on other clues have been proposed; e.g., combination of color and shape [27], luminance [28], wavelet [29], pixel value [30][31] and so on.

The localization and normalization are also needed to improve the performance of face identification. The performance of face identification can be improved by adjusting the face size to same size of facial image registered in facial image database. This operation is called the normalization. The normalization is performed on the basis of the position of internal features on face such as eyes. Therefore, the operation to find the internal features on face, i.e., localization, is needed. The more accurate face identification can be implemented by adopting these pre-processing.

1.4.4 Attribute Extraction

The attributes used for the face identification are extracted from the facial image after the pre-processing. The attributes used for the face identification are classified into the following three types.

1. Geometric attribute

2. Attribute based on global pixel pattern

3. Attribute based on statistical technique

In the geometric attribute, the internal features on face are extracted by utilizing the knowledge about the structure of human face. The shape and size of the extracted internal features are converted to numerical data. The numerical data are employed as the attributes for the face identification. For example, the edge extraction is applied to the facial image and then the extracted shape and size of internal features on face are used as the attributes for the face identification. This attribute is often influenced by lighting condition because the accuracy of edge extraction depends on the lighting condition.

On the other hand, in the attribute based on the global pixel pattern, the gradation value on each pixel of the facial image is simply used as the attributes for the face identification. This attribute has the weakness that the size of attribute is very large. For example, in the case where the size of facial image is 128×128, the size of attributes is 16,384. In addition, this attribute is often influenced by the fluctuations on the environment where the facial image is taken. The mosaic attributes have been proposed in order to decrease the size of attribute and implement the robustness against the fluctuations on the environment. In the mosaic attribute, the facial image is divided into some blocks. The average value of pixels in each block is calculated. These average values are employed as the attributes for the face identification.

Recently, the attributes based on the statistical technique have been frequently utilized for the face identification. As the representative method of the attributes based on the statistical technique, there is the eigenface method. The eigenface method is introduced below.

In this method, the Principal Component Analysis is applied to a set of pixel vectors which are extracted from many facial images, and then the orthonormal basis called eigenface is calculated. The pattern in which dimension is compressed is obtained by applying Karhunen-Loeve expansion (KL expansion) based on this eigneface to the pixel vector of facial images. This pattern is employed as the attributes for the face identification. The above process is formalized as follows.

Suppose that a pixel vector of facial image is N dimensions and there are M pixel vectors; X_m; $m = 1, 2, \cdots, M$. These pixel vectors of facial image are used as a sample set. The average vector μ is calculated from this set. And then, the covariance matrix $\mathbf{R}$ for the sample set is calculated by the following formula.

$$\mathbf{R} = \frac{1}{M}\sum_{m=1}^{M}(X_m - \mu)(X_m - \mu)^t \tag{1.1}$$

Suppose that in K large eigenvalues for this matrix, the eigenvectors which satisfy the following formula (1.2) and (1.3) is ψ_k; $k = 1, 2, \cdots, K$.

$$R\psi_k = \lambda_k \psi_k \tag{1.2}$$

$$\psi_k^t \psi_k = 1 \tag{1.3}$$

where,

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K \geq \cdots \geq \lambda_N$$

Using this orthonormal basis $\psi_k; k = 1, 2, \cdots, K$, the result that an arbitrary vector X is developed by the formula (1.4) is calculated. The K dimensional vector Y which is obtained by using this result in formula (1.5) is the pattern in which the dimension is compressed.

$$y_k^{(m)} = (X_m - \mu)^t \psi_k \qquad k = 1, 2, \cdots, K \tag{1.4}$$

$$Y_m = (y_1^{(m)}, \cdots, y_K^{(m)}) \tag{1.5}$$

It is well known that in this method, the minimization of mean square error in the formula (1.7) is attempted by the approximation of the formula (1.6) for each vector on a sample set.

$$\hat{X}_m = \sum_{k=1}^{K} y_k^{(m)} \psi_k + \mu \tag{1.6}$$

$$J(\{\Psi_k\}) = \frac{1}{M}\sum_{m=1}^{M} \|X_m - \sum_{k=1}^{K} y_k^{(m)} \psi_k - \mu\|^2 \tag{1.7}$$

The eigenvector ψ_k obtained from a sample facial image is N dimensional vector, and it can be regarded as a facial image. Using the formula (1.6), an arbitrary facial image is approximated by the linear association of them. Therefore, the ψ_k is called eigneface because it is the basis of facial image.

The type of attribute used for the face identification are decided, and the attributes based on the decided type are extracted from the facial image. Following the attribute extraction, the face identification is performed on the basis of the matching process using these attributes. In next subsection, the detail of identification method is described.

1.4.5 Person Identification

In the face identification, there are two types of person identification; i.e., verification and classification. In the verification, the user simultaneously presents a certificate to identify him/her and the face. The matching between the face corresponding to the certificate and his/her face is performed. It is judged whether the user is owner of the certificate or not by using the matching result. In the classification, the user presents face only, and it is judged which person the user is in the facial image database. The classification needs the face information of all users in the facial image database when the identification is performed, while the face information of one user only is needed in the verification. Therefore, the verification can be regarded as a sub-problem of classification.

In the face identification, it is most desirable that the face on each person completely separates in the attribute space. However, there is the possibility that the erroneous identification occurs because the overlap among the users arises in the attribute space by various fluctuations caused when the facial image is taken. Therefore, the development of the face identification in which the rate of erroneous identification is low is required.

1.4.6 Issues on Face Identification

There are many issues which have to be solved in order to implement practical face identification systems. One of important issues is the implementation of the robustness against various fluctuations. The example of fluctuations is as follows.

1. **Photographic fluctuation**

   - Lighting condition (Color of lighting, influence of daylight, etc.)
   - Condition of background (Color of background, overlap with other persons, etc.)
   - Photography condition (Angle of camera-work, object distance, rotation of object, etc.)
   - Posture condition

2. **Appearance fluctuation**

   - Bangs and whiskers
   - Glasses, cap/hat and apparel
   - Change of face caused by suntan, injury, bruise and so on
   - Change of face shape caused by growing older, putting on weight, growing thin and so on
   - Adornments (make-up, painting, etc.)
   - Facial expressions

The fluctuations are classified into two types. One is the fluctuation caused by change of environment where the facial image is taken, called photographic fluctuation, and another is the fluctuation caused by change of appearance, called appearance fluctuation. The former is the fluctuations which inevitably occur when the facial image is taken. The latter can be avoided by restricting the condition of user when the face identification is performed. The latter is harder problem than former in the object recognition point of view. An example of the photographic fluctuations is shown in Figure 1.4, and an example of the appearance fluctuations is shown in Figure 1.5. Many studies on the implementation of the robustness against these fluctuations have been carried out for the face identification. In next subsection, several studies are introduced.

1.4.7 Related Works and Applications

There are numerous successful implementation of the face identification and practical applications on the face identification are developed. Several related works and applications on the face identification are described in Table 1.5. The major problem in the face identification is to implement the robustness against various fluctuations. In numerous studies related to the face identification, the implementation of the robustness

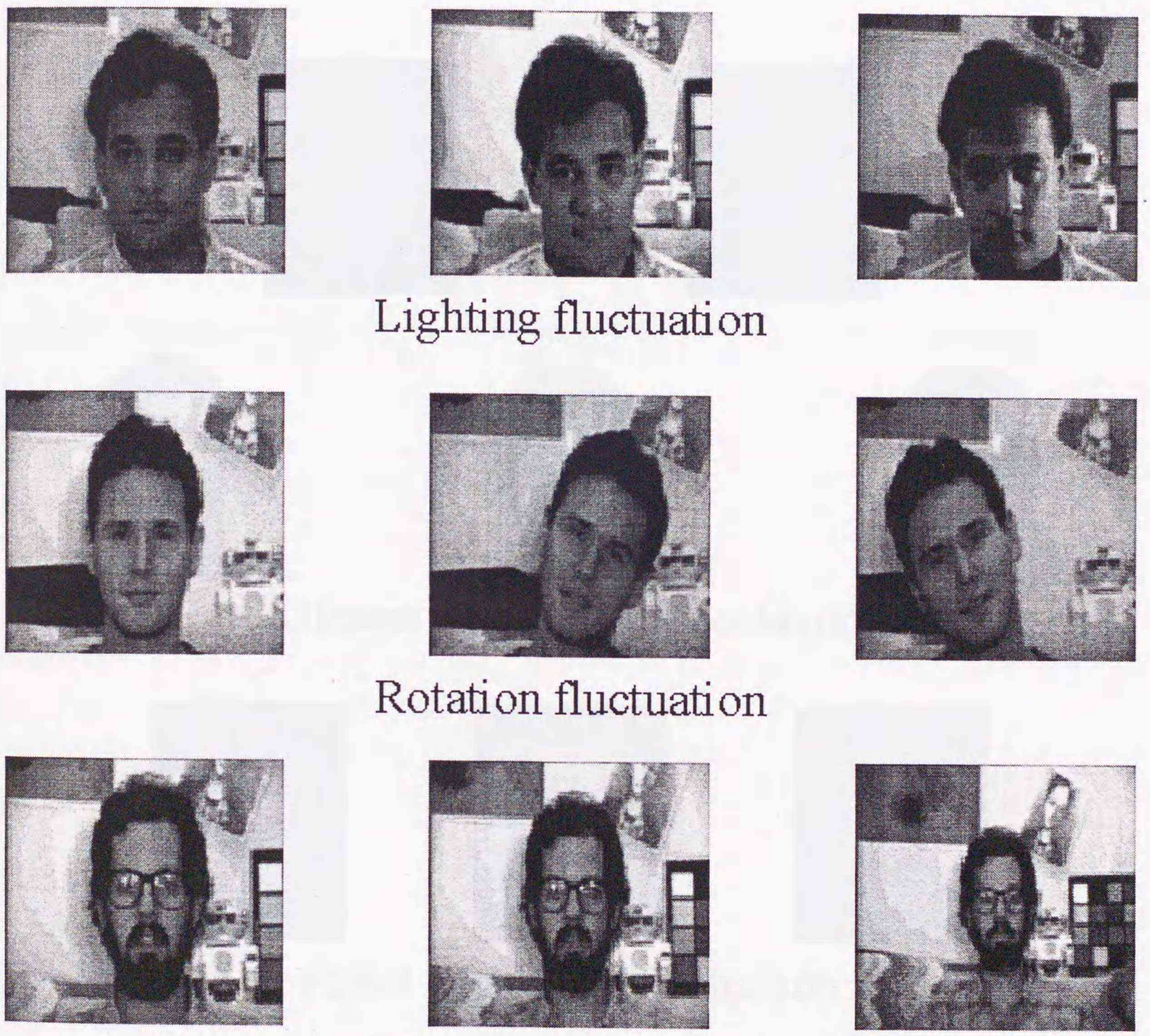
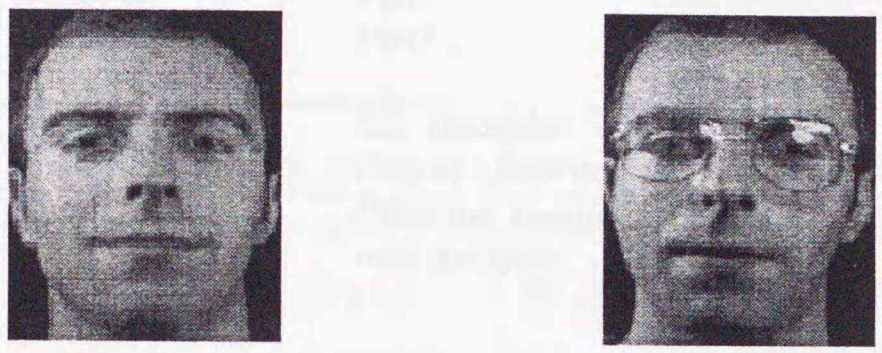
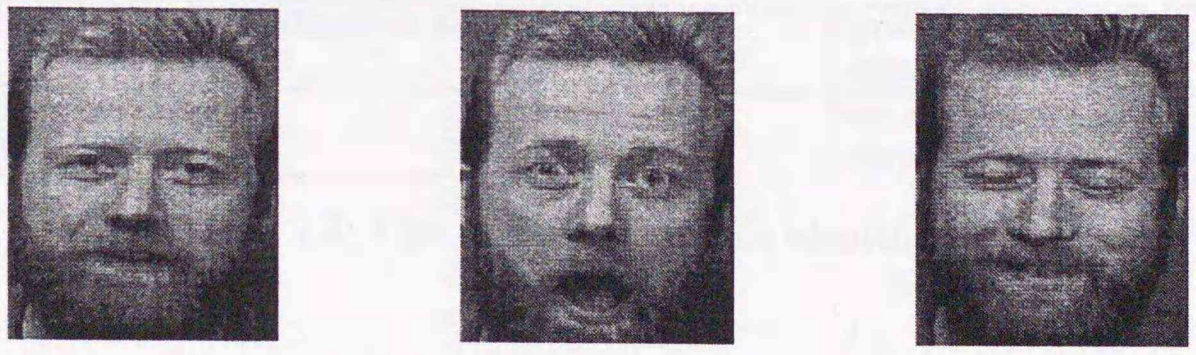


Figure 1.4: An example of photographic fluctuations (From MIT face database)





Glasses and apparel fluctuation



Facial expression fluctuation

Figure 1.5: An example of appearance fluctuations (From ORL face database and Purdue University face database)

Fluctuation	Author	Attribute extraction	Recognition method	Main reference	Year
Posture	Inakoshi et al.	Principal component analysis	Euclidean distance	[32]	1999
	Imaoka et al.	Luminance	Correlation	[33]	1999
	Kuriyama et al.	high order local auto-correlation of pixels	Discriminant analysis	[34]	1999
Lighting	Sohma et al.	Mosaic	Mahalanobis correlations	[35]	1997
	Doi et al.	Pixel	Correlation	[36][37]	1998
	Yamamoto et al.	Pixel	Virtual subspace method	[38]	2000
Posture & Lighting	Akamatsu et al.	KL expansion of Fourier spectrum	Euclidean distance	[39]	1993
Angle	Kurita et al.	Principal component analysis	Neural network based on gating network	[40]	1999
Facial expression	Kitada et al.	Morphological analysis	Matching	[41]	1998
	Umeyama et al.	Supervised independent component analysis	Correlation	[42]	1999
	Hirayama et al.	Isodensity maps	Template matching	[43]	1999
	Sakano et al.	Mosaic	Kernel mutual subspace method	[44]	1999
	Futamura et al.	Independent Component Analysis	Euclidean distance	[45]	1999
Posture & Facial expression	Pramadhanto et al.	Gabor wavelet	Dynamic link architecture	[46]	1997
Facial expression, Glasses & Age	Kosugi	Mosaic	Neural network	[47]	1993

Table 1.5: List of studies on face identification

against various fluctuations is attempted. In the Table 1.5, they are classified by the fluctuation which is focused on, and the attribute and the recognition method which are used for the identification are described. Moreover, some related works and applications in which the fluctuations are not treated in the experiment are described below.

It has been reported that the accuracy of identification is high in the identification based on fingerprint and eyeprint, so that practical systems have been developed. In the identification based on fingerprint and eyeprint, a particular action of the user is required in order to obtain fingerprint or eyeprint when the identification is performed. However, the user's burden is large in such identification. Although the identifications based on face and voice in which a particular action of the user

is not required have been proposed, the accuracy for the identification is not enough at present. Maeda et al. proposed an identification system based on the integration of facial image, walking image and voice in order to implement an identification in which a particular action is not required and which has high accuracy for the identification [48]. The face identification in this system is performed by using the Mosaic attributes. The experimental results revealed that the accuracy for the identification could be considerably improved by performing the identification based on the integration of each identification result.

Hisatomi et al. proposed a multi-agent-based face identification system [49]. This system consists of some agents in which different tasks are assigned; e.g., extraction of attributes, template matching and so on. The face identification is performed by the coordination among these agents. It was found from the experimental results that the proposed system could perform the accurate identification by cooperative behavior among these agents in the case where the mis-identification occurred on the one agent.

In conventional face identification, the identification has been performed by using the stationary facial image. However, the impersonated problem may arise because it is difficult to identify the similar facial images in the case where the attributes used for the identification are extracted from the stationary facial image. To solve this problem, Ito et al. proposed a face identification based on the change of facial expression [50]. The facial image sequence is employed in this face identification. The process in which each user expresses a facial expression is recorded, and this process is used as the password for the face identification. The possibility of the face identification using the facial expression as a password was confirmed through the identification experiment.

Tanaka et al. proposed a face identification using average cosine as the attributes [51]. Moreover, the effective regions on the facial image for the identification based on the average cosine attributes were searched. It was found from experimental results that the searched regions did not depend on the kind of facial image used for the identification.

Sukthankar et al. proposed a multi-agent-based face identification system for apartment complex [52]. This identification system consists of some agents which a particular task is assigned, e.g., visitor identification, interactive labeling of unknown visitor, evaluation of face recognition algorithm and so on. The face identification for apartment complex is implemented by the interaction among these agents. They confirmed the

possibility of the face identification for apartment complex through the experiment on the outdoor environment.

1.5 Face Cognitive Model Acquisition

As one of evaluation criteria for the face recognition methods, there is the consistency with the face cognitive structure of human [53]. For example, in the facial image retrieval, it is desirable that the facial images which can be judged to be similar by human can be retrieved by computer. Therefore, the research on how human processes face has been carried out. Namely, the face cognitive model of human has been investigated. This subject is called face cognitive model acquisition. In this section, the research on the face cognitive model acquisition and the face cognitive model acquisition based on the facial image database retrieval task method are introduced.

1.5.1 Face Cognitive Model

The study on how human processes face is the interesting subject, and it has been carried out of old. So far, the investigation of face cognitive model of human has been mainly performed in the psychological field. However, the research on the investigation of face cognitive model of human has come to be busily carried out in the engineering field because the knowledge about the face cognitive model of human is useful for the face recognition which is utilized to implement effective human interface. In the face recognition, the knowledge about the face cognitive model of human is required in order to implement the face recognition methods which have the consistency with the face cognitive structure of human. This subject is called face cognitive model acquisition, and the various face cognitive models have been recently investigated. An example of the studies on the face cognitive model acquisition is shown in Table 1.6. As shown in Table 1.6, the face cognitive models for similarity, age, sex and so on have been investigated.

In the conventional face cognitive model acquisition, the face photographs are used as the stimuli. That is, the investigator presents the face photographs to the subject, and then the subject evaluates the face photographs. The face cognitive model of human is investigated on the basis of the relation between the face photographs and the evaluations. So far, the face cognitive model acquisition using the face photograph stimuli has been often carried out, and significant knowledge has been

Face Cognitive Model	Authors	Main Reference	Year
Similarity judgment among faces	A. Morisaki et al.	[55]	1995
Age and sex	T. Takasaki et al.	[56]	1996
False smile	M. Nomura et al.	[57]	1999
Average face	M. Nishitani et al.	[58]	1999
Attractive face	T. Kato et al.	[59]	1998
Childish, adult, manly and womanish face	M. Yamaguchi et al.	[60]	1996
Relation between facial expressions and viewpoint	T. Ogawa et al.	[5]	1997
Relation between configural processing and surface-based processing of face	T. Tujii	[61]	1998
Relation between facial expressions and face identification	R. Nagayama et al.	[62]	1998
Relation between facial expressions and music & talk	T. Kato et al.	[63]	1998
Relation between facial expressions and face Identification	M. Miyoshi	[64]	1998
Motion in facial expressions	N. Matsuzaki et al.	[6]	1999
Facial expressions	M. Kamachi et al.	[7]	1998
	T. Kato et al.	[65]	1998
	S. Shibui et al.	[66]	1998
	M. Oda et al.	[67]	1999
	N. Wada et al.	[68]	1999
	S. Yoshikawa et al.	[69]	1999
	T. Kato et al.	[70]	1999
	N. Wada et al.	[71]	1999
	A. Miura et al.	[72]	1999
	M. Yamaguchi et al.	[73]	1999
	H. Umemura et al.	[74]	1999

Table 1.6: List of face cognitive model acquisitions

obtained. However, in the face cognitive model acquisition using the face photograph stimuli, the face photographs which are taken in the real environment are used. Namely, the investigator can not directly adjust the face photograph stimuli. Moreover, in the face cognitive model for face expressions, the accurate face cognitive model can be acquired by using the image sequence because the facial expressions are mainly recognized by human on the basis of the change of face. Therefore, it is difficult to acquire more accurate face cognitive model of human in the case where the face photograph stimuli are utilized.

To solve this problem, the face cognitive model acquisition has come to be performed by using the image processing technology [54]. Namely, the facial images are used as the stimuli for the face cognitive model acqui-

sition. In the case where the facial images are utilized as the stimuli for the face cognitive model acquisition, the investigator can directly adjust or change the facial images. Moreover, the face cognitive model can be acquired as numerical data without performing the direct measurement of the stimuli because the facial images are created from the numerical pixel values. The various and accurate face cognitive models can be acquired by using the image processing technology in comparison with the face cognitive model acquisition using the face photograph stimuli.

The image processing technology which is frequently used in the face cognitive model acquisition is Morphing technology. The Morphing technology is often used for the face cognitive model acquisition for the average face and facial expressions. The different facial images can be combined in arbitrary composition rate by using the Morphing technology. Therefore, the facial images which do not actually exist can be created by using the Morphing technology. The Morphing technology can increase the kind of the face cognitive model which can be investigated in the face cognitive model acquisition.

Moreover, the face composition technology is also useful for the face cognitive model acquisition in addition to the image processing technology. It is possible to compose the true-to-life face by the development of various methods in the computer graphics and animation technologies. The face cognitive model acquisition can be also performed by using the face which is composed by using the computer graphics and animation technologies. In the face cognitive model acquisition using the face composition technology, it is easier to adjust or change the composed face than the facial image based on the image processing because the composed face is created on the basis of the numerical data on the computer. Furthermore, the face image sequence can be created by using this face composition technology. Thus, the face composition technology is effective for the face cognitive model acquisition for the face which does not actually exist and the facial expressions.

As mentioned above, the development of the image processing and face composition technologies is significant for the research on the face cognitive model acquisition because it is possible to operate quantitatively the face stimuli. That is, the relation between the features in the facial image and the psychological behavior can be quantitatively investigated by using the facial image or the composed face stimuli and operating directly them. Therefore, the development of image processing and face composition technologies yield the paradigm shift in the face cognitive

model acquisition. In next subsection, the computer-based face cognitive model acquisition method using facial image stimuli which can adapt to various face cognitive models is introduced.

1.5.2 Face Cognitive Model Acquisition based on Facial Image Database Retrieval Task Method

In the conventional information retrieval, sentence or word data have been frequently used as the retrieval target. However, the diversification of information has brought with it the problem that it is necessary to treat image data as the retrieval target. So far, in the image retrieval, the image retrieval method based on the verbal keywords has been often employed. In this method, the database manager attaches the verbal labels to the image data which are stored in the image database. The user inputs verbal keywords when the image retrieval is performed. The matching processing between the keywords inputted by the user and the labels which are attached to the image data is performed. The image data having the labels which are similar to the inputted keywords are retrieved. In this method, it is difficult to retrieve the image data which are suitable for each user's subjectivity because the database manager attaches the labels on the basis of his/her subjectivity. That is, it is difficult to adapt to each user's subjectivity because the labels which are attached to the image data reflect the database manager's subjectivity.

To solve this problem, recently, the image retrieval method based on the contents of image data has been proposed [75][76]. In this method, the image data are retrieved based on the raw information such as pixel values. Therefore, the user can perform the image retrieval which reflects his/her subjectivity because the other person's subjectivity is not considered in the image data. The retrieval method based on the contents of image data is often employed in the recent image retrieval because this method can adapt to each person's subjectivity and implement flexible image retrieval. In this image retrieval method, the target image datum is needed when the image retrieval is started because the image data are retrieved on the basis of the raw information which image data have. Therefore, it is difficult to treat the ambiguous target in the user's mind as the retrieval target. However, it is obvious that there is the case where the ambiguous target in the user's mind such as the impression has to be treated as the retrieval target. Thus, the image retrieval also must be able to adapt to such retrieval target.

Oda proposed the image retrieval method which can perform the re-

trieval in the case where the retrieval target is subjective and ambiguous [77][78][79]. In the proposed method, the facial images are employed as the retrieval object. The facial image is subjective and ambiguous retrieval target. The impression received from face is different on each person. Moreover, it is difficult to represent face by verbal keywords and perform the drawing. However, the user can judge whether the facial images which are presented to him/her are suitable for own retrieval target or not. Therefore, in this method, the image retrieval is performed on the basis of the user's judgment for the facial images and the interaction between the user and the computer. The retrieval process in this method is shown in Figure 1.6. As shown in Figure 1.6, at first, some facial images which are randomly chosen from facial image database are presented to the user. The user selects the facial images which are suitable for the retrieval target in the presented facial images. The contents-based facial image retrieval is performed on the basis of the selected facial images by the user. That is, new facial images are retrieved from facial image database by using the raw information such as pixel values in the facial images selected by the user. The user concretes the retrieval target which was ambiguous in the initial retrieval step through the interaction, and retrieves the satisfying facial images for the retrieval target. Additionally, in this method, the image retrieval based on the contents of facial images is performed by using average face. The average values are calculated from raw information which the user-selected facial images have, and then the average face is created by combining these average values. The facial images are retrieved from the facial image database on the basis of this average face. Recently, this interactive contents-based image retrieval method is often used in the image retrieval because the user can retrieve subjective and ambiguous images. Furthermore, the retrieval method in which other algorithm instead of the method based on the average face is employed has been proposed [80].

This retrieval method has the advantage which can retrieve subjective and ambiguous retrieval target in mind. In this type of image retrieval, the user selects the facial image in accordance with the evaluation criterion for the retrieval target. Therefore, it is clear that the user selects the facial images on the basis of the face cognitive structure for the retrieval target. Namely, the face cognitive model of the user can be acquired by investigating the information which the selected facial images have. Oda et al. proposed the facial image database retrieval task method for the face cognitive model acquisition by utilizing this feature. In the face

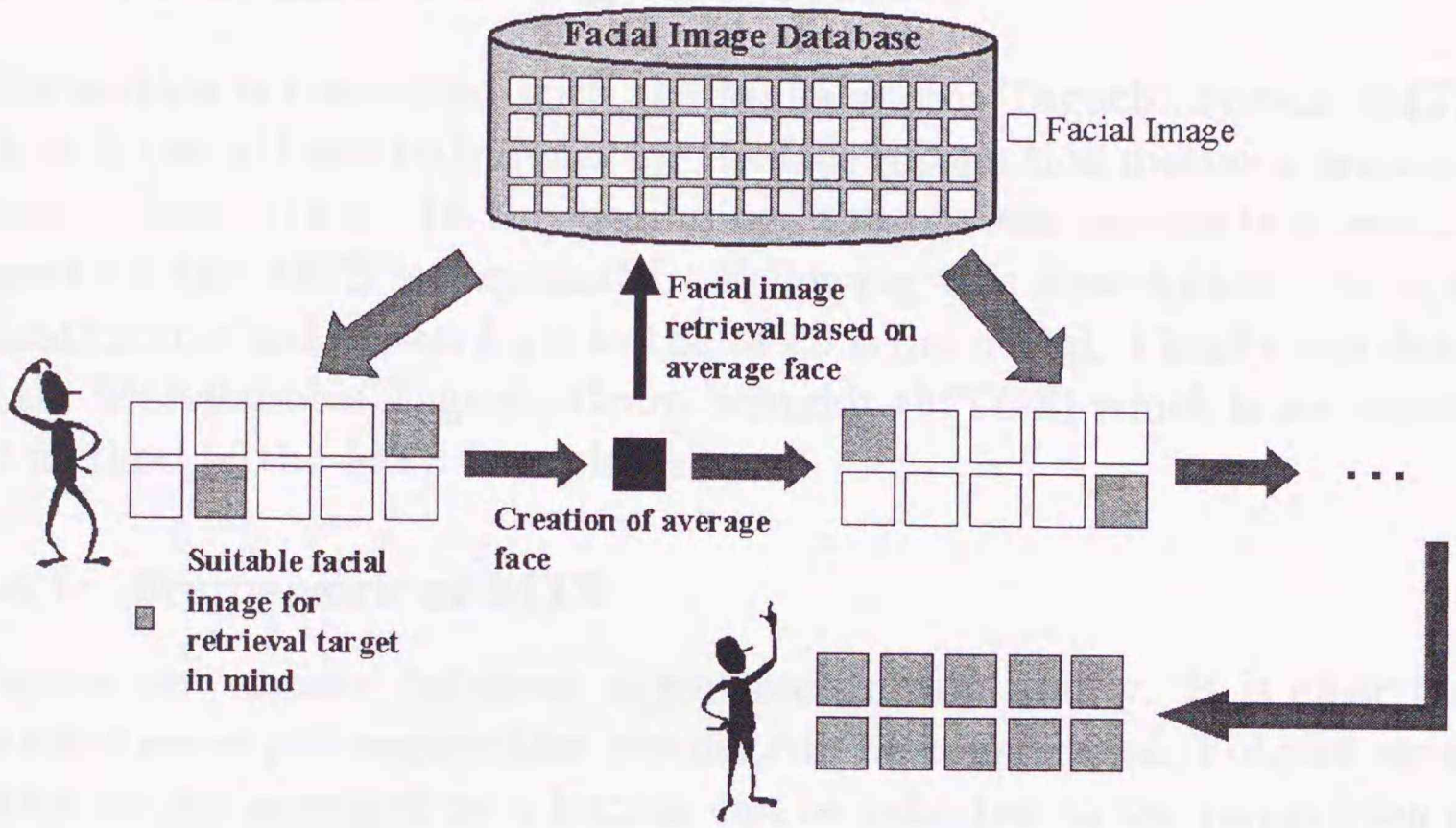


Figure 1.6: Retrieval process for the ambiguous retrieval target

cognitive model acquisition based on this method, various kinds of face cognitive model can be easily acquired because the user only selects the facial images which are suitable for the retrieval target. Oda et al. investigated some face cognitive models of human by utilizing the facial image database retrieval task method. As first step, Oda et al. performed the face cognitive model acquisition using line-draw facial image in order to simplify the evaluation of performance of the proposed method and the construction of facial image database. In the literatures of [81], [82] and [83], the face cognitive models for the favorite face and similar face were investigated. The face cognitive model acquisitions of human for favorite、distasteful, gentle and fierce face were attempted in the literature [84]. The face cognitive model for the forbidding face was investigated in the literature [85]. Moreover, Oda et al. performed the face cognitive model acquisition using the photographic facial images [86]. On the other hand, the cognitive model acquisition for the diagram was also performed on the basis of this method [87].

As mentioned above, the facial image database retrieval task method is effective method for the face cognitive model acquisition. Moreover, various face cognitive models can be easily acquired by using this method. Recently, the face cognitive acquisition using the image processing technology and the facial image database task method has often been carried out.

1.6 Mahalanobis-Taguchi System

This section is concerned with the Mahalanobis-Taguchi System (MTS) which is one of basic techniques for the face recognition methods proposed in this dissertation. In the beginning, the pattern recognition method based on the MTS is explicated. Following this description, the optimization method of attribute in the MTS is described. Finally, the detail of the Mahalanobis-Taguchi-Gram-Schmidt (MTGS) which is an extended method of the MTS is explained.

1.6.1 Framework of MTS

Human can acquire excellent object recognition ability. It is clear that an effective object recognition system can be constructed if object recognition ability acquired by a human can be reflected to the recognition on computer.

How does human acquire an excellent object recognition ability? As shown in Figure 1.7, when a new object is given, a human recognizes an object on the basis of all kinds of information which can be obtained from it. This information includes characters such as shape, color, and so on. Human accumulates the recognition experience for the object, and gradually acquires sensuous knowledge of the effective recognition for the object. This sensuous knowledge indicates the kinds of character which are effective for the recognition, border value, relationship among the characters and so on. Therefore, an effective recognition system can be implemented if this sensuous knowledge which is acquired from the experience can be reflected to the recognition on computer. However, it is considerably difficult to represent the sensuous knowledge which human has as rules or numerical values.

The MTS has been proposed by Taguchi in order to implement a pattern recognition which can reflect human sensuous knowledge for object recognition [88][89][90][91]. In the MTS, a human's sensuous knowledge for the object recognition is acquired by modeling the sensuous knowledge acquisition process in object recognition. Moreover, in this modeling, the human object recognition and the pattern recognition by computer are combined. Therefore, the MTS has the following three features. First, a set of training data for the recognition is collected by a human. A set of training data consists of the objects which are considered to be correct as the recognition target. In addition, this collection is performed by evaluating the object itself. That is, a human does not evaluate the attributes

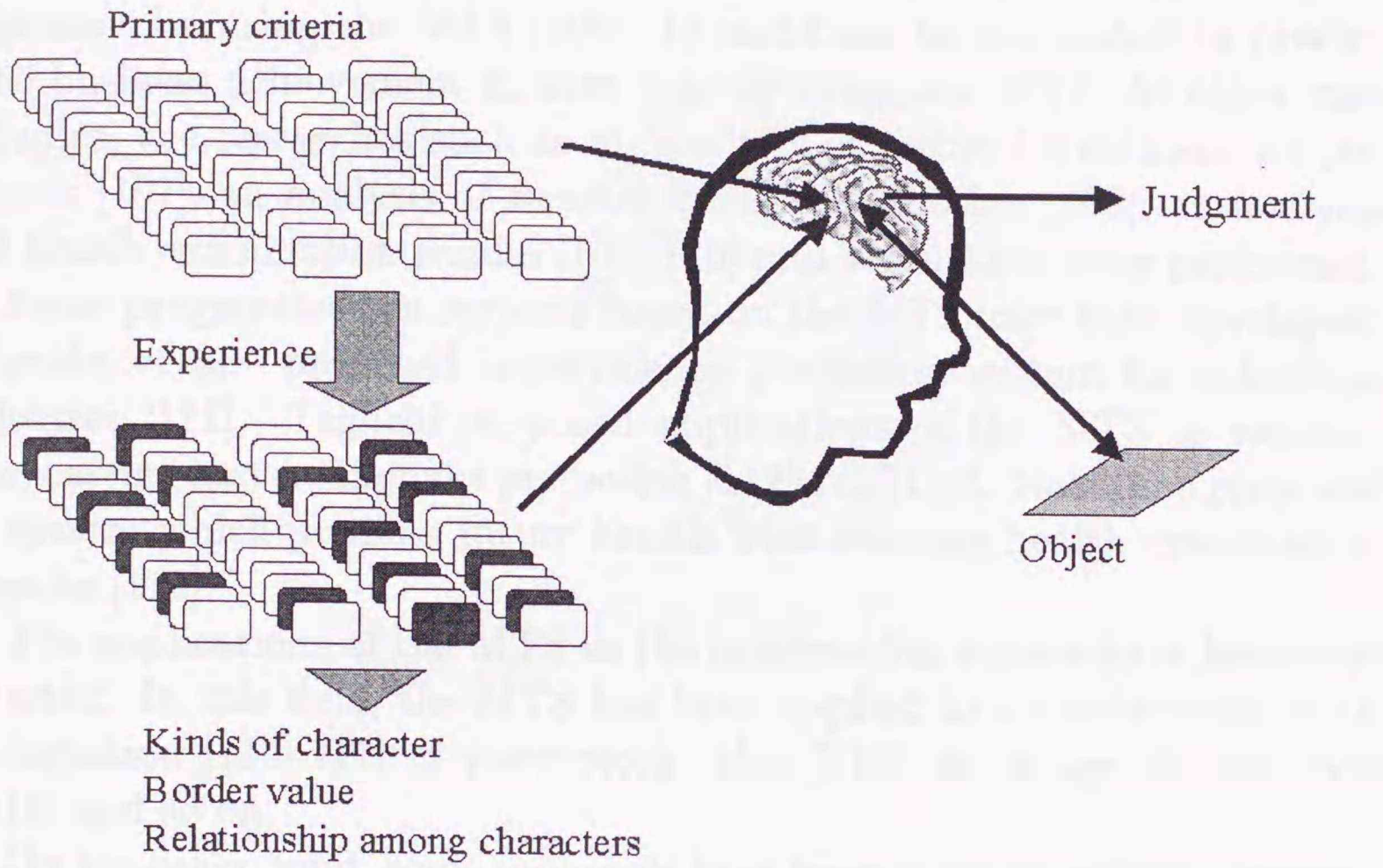


Figure 1.7: Image of sensuous recognition ability acquisition process

which are extracted from the object. Secondly, the correlations among the attributes are automatically extracted from the collected training data, and then the recognition based on these correlations is performed. Finally, the effective attributes for the recognition can be re-extracted by using recognition result. Namely, the optimization of attribute can be carried out. From these features, recently, the MTS has been applied to various problems. Especially, the applications for sensuous defects inspection in the industrial products are developed. Some applications and related works using the MTS are described below.

As applications for industrial products, a defects detection system in circuit pattern [92][93], an infrared absorption spectrum analysis system for quality control of materia medica [94], a decision system of positions and kinds of sensors in fire alarm system [95][96], a sensing system for prevention of driving accidents [97], a recognition system for false banknote [98], a fault diagnosis system for sensors in future spacecraft [99][100][101], a printed letter inspection system [102][103], a defects detection system of clutch disk [104] and so on have been proposed.

Moreover, some applications have been proposed to analyze multidimensional information. Suzuki et al. utilized the MTS to analyze the relationship between the programmer's programming ability and the

questionnaire result [105]. Shibano analyzed corporate condition in management by using the MTS [106]. In addition, he attempted to predict the business achievement in next year by using the MTS. As other case studies, the researches such as an analysis of medical treatment for patients [107], an analysis of wearing comfort of clothes [108], an analysis of health examination results [109][110] and so on have been performed.

Some prognostication systems based on the MTS have been developed. Tanaka et al. proposed a prevalence prediction system for infectious diseases [111]. Taguchi proposed applications of the MTS to weather forecasting and earthquake prediction [112][113][114]. Nakajima proposed a system which predicts future health from existing health examination results [115].

The applications of the MTS to the information science have been also started. In this field, the MTS has been applied to a handwritten word recognition [116][117], a voice recognition [118], an image classification [119] and so on.

On the other hand, some proposals have been made in order to improve the performance of the MTS [120][121][122].

I explain the detail of pattern recognition based on the MTS in the following subsection.

1.6.2 Pattern Recognition based on MTS

In the MTS, the pattern recognition based on the Mahalanobis space and the Mahalanobis distance is performed. The pattern recognition based on the MTS is illustrated in Figure 1.8. The Mahalanobis space is a space which can consider the correlations among the attributes in multi-dimensional data. This Mahalanobis space is constructed from a set of multi-dimensional training data. The Mahalanobis distance is one-dimensional evaluation criterion which is calculated in the Mahalanobis space. Strictly speaking, it indicates the distance between the origin of the Mahalanobis space and an unknown datum. The MTS judges whether the unknown datum belongs to the Mahalanobis space or not, in accordance with the Mahalanobis distance. The threshold for this judgment is decided by a human as the high recognition performance can be obtained. The recognition based on a human sensuous knowledge is implemented by combining this pattern recognition and some operations. The recognition process which is from the collection of training data to the decision of threshold is as follows.

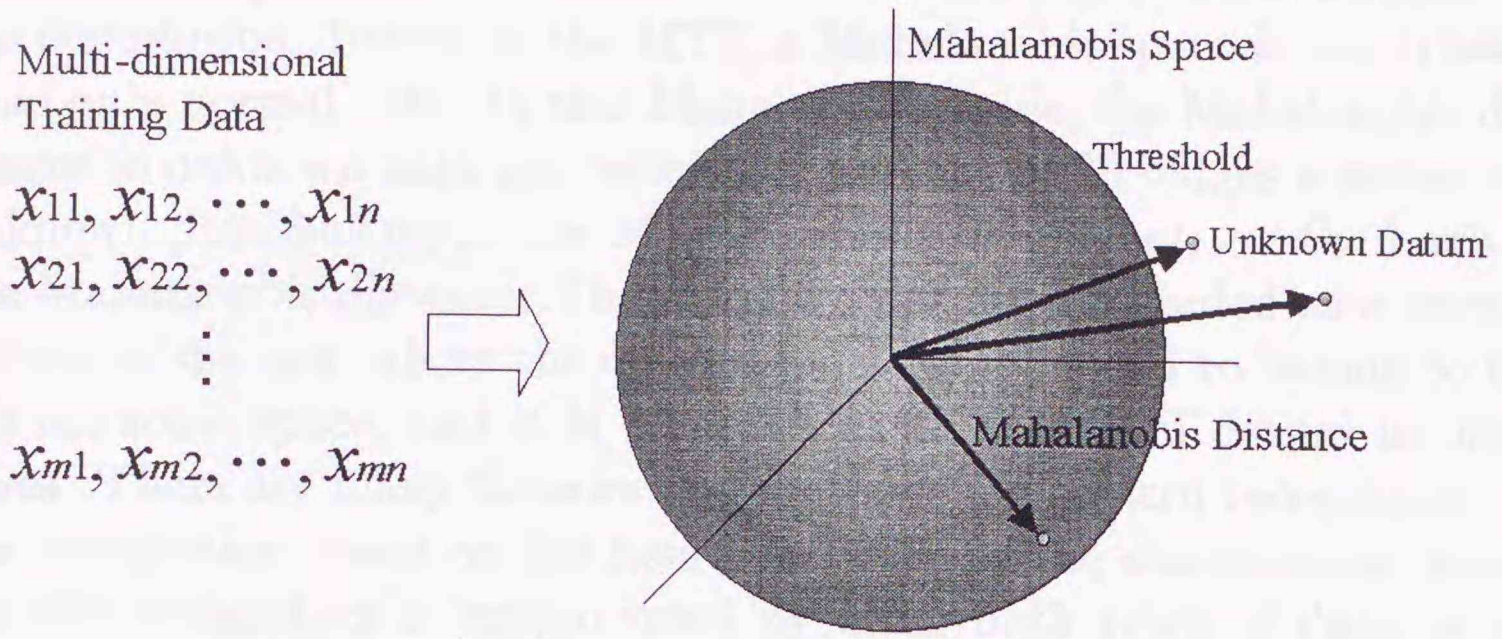


Figure 1.8: Pattern recognition based on the MTS

1. Collection of training data
2. Extraction of attribute
3. Construction of Mahalanobis space
4. Calculation of Mahalanobis distance to unknown datum
5. Decision of threshold

Basically, the MTS is a pattern recognition method to divide data into two kinds, namely normal data and abnormal data. The normal data indicate the correct data as the recognition target, and the abnormal data mean other data except normal data. For instance, in the case where a particular person's face is identified, the faces of a particular person are normal data and all faces of other persons are abnormal data. As known from this example, it is likely that the normal data have homogeneous distribution on the attribute space. On the other hand, it is likely that the abnormal data have unknown distribution. In this case, we can say with fair certainty that the recognition of the normal data is easier than the recognition of the abnormal data because there are too many kinds of abnormal data. Thus, the MTS implements dichotomy of data by recognizing the normal data correctly and regarding the data except the normal data as the abnormal data rather than recognizing both of normal data and abnormal data. Furthermore, the concept of this recognition is similar to the human recognition. That is, humans do the object recognition by judging whether the object is correct as the recognition

target or not, in the case where there are too many other objects for the comparison. Hence, in the MTS, a Mahalanobis space is constructed from only normal data. In this Mahalanobis space, the Mahalanobis distances to unknown data are calculated, and the MTS judges whether the unknown data belong to the Mahalanobis space or not, on the basis of the Mahalanobis distance. The unknown datum is regarded as a normal datum in the case where the unknown datum is judged to belong to the Mahalanobis space, and it is regarded as an abnormal datum in other cases. There are many theories and methods for pattern recognition. In the recognition based on the neural network or the discriminant analysis, the recognition is implemented by using both kinds of data as the training data. The MTS has prominent feature that a recognition space is constructed only by normal data.

Consequently, as a first step, a set of normal data is collected as training data. These normal data are decided by a human sensuous judgment in order to implement a pattern recognition which can reflect human sensuous knowledge.

The attributes are extracted from these collected training data. It is suggested that all kinds of attribute which can be extracted from the object are used for the recognition in the MTS. For example, as the attributes for recognition of image datum, the value on each pixel can be taken. However, in the case where the size of image is large, too many attributes are extracted.

In the case of image and wave form, Taguchi has proposed the attributes based on differential characteristic and integral characteristic. The differential characteristic indicates the changes of element in image or wave form, similar to the concept of frequency. The integral characteristic indicates the distance of element in image or wave form. The extraction method of the attributes based on these characteristics is illustrated in Figure 1.9. As shown in Figure 1.9, in the differential characteristic, some lines are drawn on image or wave form. The attribute based on the differential characteristic indicates the number of the points that each line intersects the pixel or wave form. In the integral characteristic, some lines and center line are drawn. This center line means the line which halves the image or wave form. The Euclidean distances between the center line and the intersecting points on each line are calculated. On each line, the sum of Euclidean distances is taken as the attribute based on the integral characteristic. It is reported that the effective pattern recognition can be performed by using the attributes based on these differential and integral

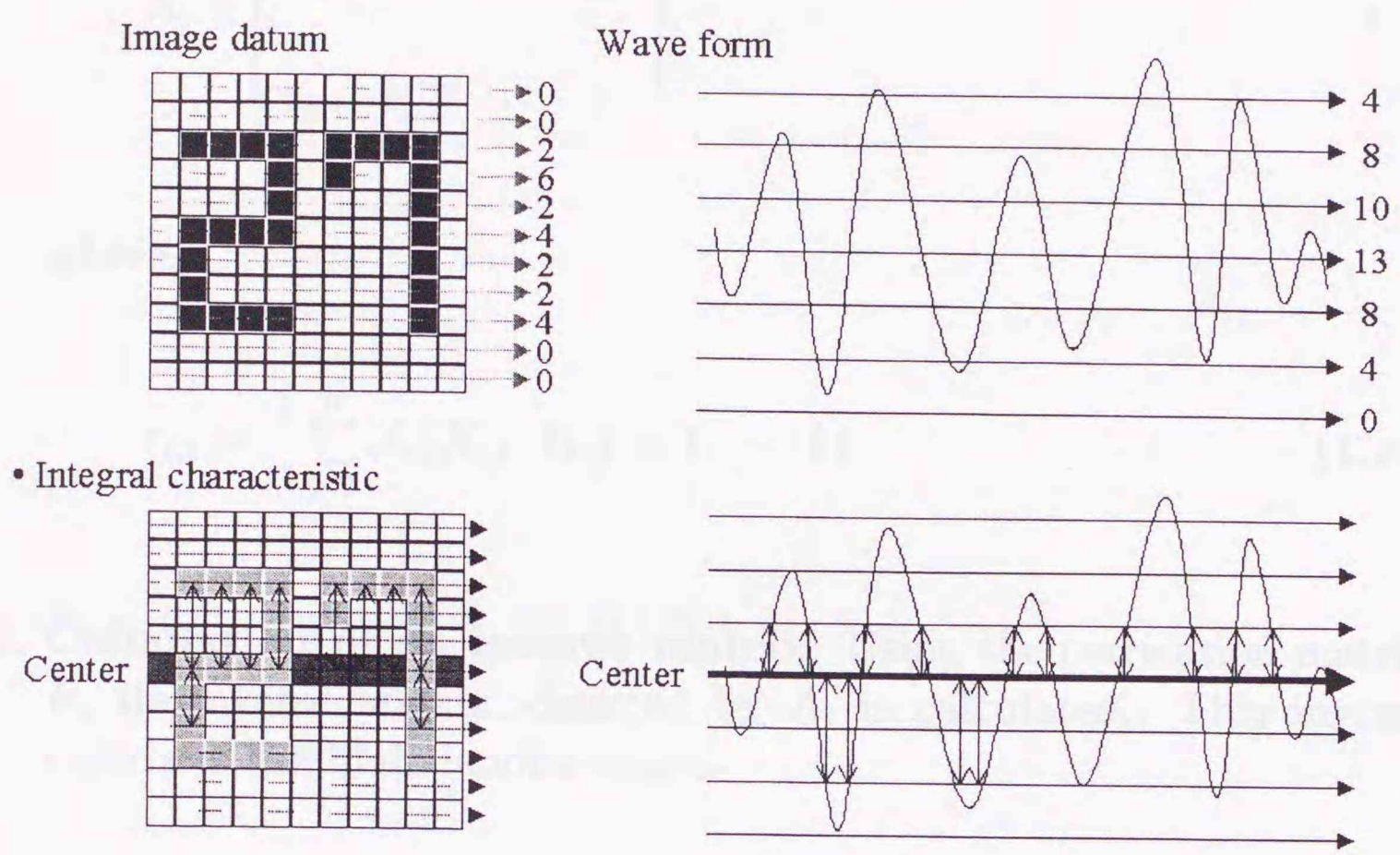


Figure 1.9: Extraction method of the attribute based on differential and integral characteristics

characteristics [102][103][104]. Moreover, the number of attributes can be changed to an arbitrary number by using the attributes based on these characteristics.

After the extraction of attributes, the Mahalanobis space is constructed from the extracted attributes. The Mahalanobis space is constructed by the following process.

1. **Normalization of a set of training data**: k denotes the "dimension" of the datum (i.e., the number of attributes), and n denotes the number of data in the set. The set of training data is transformed by the average m_j and the standard deviation δ_j, and then $X_{i,j}$ is calculated. This operation is called normalization, and the resultant $X_{i,j}$ is called the normalized data.

$$X_{i,j} = \frac{x_{i,j} - m_j}{\delta_j} \quad (i = 1, \cdots, n; j = 1, \cdots, k) \tag{1.8}$$

2. **Calculation of a correlation matrix**: The following correlation matrix R is calculated from a set of the normalized training data.

$$R = \begin{bmatrix} 1 & r_{1,2} & \cdots & r_{1,k} \\ r_{2,1} & 1 & \cdots & r_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k,1} & r_{k,2} & \cdots & 1 \end{bmatrix} \tag{1.9}$$

where,

$$r_{i,j} = \frac{1}{n} \sum_{l=1}^{n} X_{l,i} X_{l,j} \quad (i, j = 1, \cdots, k) \tag{1.10}$$

3. **Calculation of an inverse matrix**: Using the correlation matrix R, its inverse matrix, denoted by A, is calculated. This inverse matrix is the Mahalanobis space.

$$A = R^{-1} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & \cdots & a_{k,k} \end{bmatrix} \tag{1.11}$$

The Mahalanobis distance is calculated by using this inverse matrix. The datum z to be recognized is normalized by using the average m_j and the standard deviation δ_j of Equation (1.8) to obtain Z. The Mahalanobis distance, denoted by D^2, is calculated as

$$Y = \{Z_1, Z_2, \cdots, Z_k\} \tag{1.12}$$

$$D^2 = \frac{Z^T A Z}{k} \tag{1.13}$$

The Mahalanobis distance theoretically takes from zero to infinity. When an unknown datum is very similar to the normal data pattern, the Mahalanobis distance takes a value close to 1. When the unknown datum differs from the normal data pattern, its value rapidly increases. Final recognition result is decided by using this Mahalanobis distance.

Generally, the probability of a datum being similar to the normal data pattern is very small if this value exceeds 4. However, it is recommended that the threshold is properly decided by a human in accordance with the purpose. For instance, in the case of defects inspection, the economic loss is large when the products having defects are recognized as the normal products which do not have defects. Therefore, it is desirable that the threshold to recognize unknown data is changed on each application.

The pattern recognition can be implemented by above operations. Furthermore, in the MTS, the optimization of attribute is provided in order to model the re-extraction of effective characters in human recognition process. The detail of this optimization method is described below.

1.6.3 Optimization of Attribute

In the MTS, the attributes that do not contribute to the recognition can be screened out by performing the optimization of attribute. Fundamentally, this optimization is performed by using abnormal data, Signal/Noise ratio (SN ratio), orthogonal array and factorial effect. However, there is a case where the normal data is used instead of abnormal data in accordance with the SN ratio which is employed in the optimization. This optimization process is as follows.

1. **Collection of abnormal data:** In the optimization process of the MTS, abnormal data that do not belong to the Mahalanobis space of the optimization target are needed because the aim of this optimization is to search the combination of attributes which can increase the Mahalanobis distances of abnormal data to decrease erroneous recognition of abnormal data. Therefore, the Mahalanobis distances between the Mahalanobis space which is the optimization target and the abnormal data are utilized in this optimization as the evaluation criterion for the Mahalanobis space.

2. **Construction of Mahalanobis space using attributes based on orthogonal array:** The Mahalanobis space is constructed from the attributes that correspond to the combinations in the orthogonal array. The orthogonal array is a matrix consisting of the combinations of attributes which is produced to reduce the experiment times effectively. There are L_8, L_{16}, L_{32}, L_{64}, L_{128} and so on. The adequate orthogonal array is employed in accordance with the number of attributes which are used before the optimization. That is, the orthogonal array L_{128} is used for 120 attributes. Figure 1.10 shows the

Attribute number

	1	2	3	4	5	6	...	121	122	123	124	125	126	127		S/N ratio
1	1	1	1	1	1	1		1	1	1	1	1	1	1		39.42
2	1	1	1	1	1	1		2	2	2	2	2	2	2		23.06
3	1	1	1	1	1	1		2	2	2	2	2	2	2		34.88
4	1	1	1	1	1	1	...	1	1	1	1	1	1	1		29.70
5	1	1	1	1	1	1		2	2	2	2	2	2	2		32.85
6	1	1	1	1	1	1		1	1	1	1	1	1	1		32.15
⋮	⋮							⋮								⋮
123	2	2	1	2	1	1		2	2	1	2	1	1	2		31.99
124	2	2	1	2	1	1		1	1	2	1	2	2	1		35.97
125	2	2	1	2	1	1		2	2	1	2	1	1	2		33.41
126	2	2	1	2	1	1	...	1	1	2	1	2	2	1		31.87
127	2	2	1	2	2	1		1	1	2	1	2	2	1		33.38
128	2	2	1	2	1	1		2	2	1	2	1	1	2		33.14

Figure 1.10: Orthogonal array L_{128} and SN ratio

L_{128} orthogonal array. In the orthogonal array, Level 1 (1 in Figure 1.10) indicates that its attribute is used for constructing the Mahalanobis space. On the contrary, Level 2 (2 in Figure 1.10) shows that its attribute is not used for constructing the Mahalanobis space. Thus, in the L_{128}, one hundred twenty-eight Mahalanobis spaces are constructed for the optimization.

3. **Calculation of SN ratio:** First, the Mahalanobis distances between the abnormal data and the Mahalanobis space which is constructed from the attributes based on the orthogonal array are calculated. The SN ratio is computed by using these Mahalanobis distances. The SN ratio is one-dimensional evaluation criterion for a set of the Mahalanobis distances. There are three types in the SN ratio. The type of SN ratio which is employed for the optimization is changed in accordance with the purpose. The detail and calculation method of each type of SN ratio are as follows.

   (a) **Larger-the-better type:** The larger-the-better type is employed to increase the Mahalanobis distances between the Ma-

halanobis space which is the optimization target and abnormal data. Since the SN ratio is an evaluation criterion which can consider the distribution of the Mahalanobis distances, the increase of the SN ratio in this type means that the Mahalanobis distances between the abnormal data and the Mahalanobis space become large on the average. It is possible to increase the degree of the separation between the target Mahalanobis space and abnormal data by using this SN ratio, so that it is possible to decrease the possibility that the erroneous recognition of abnormal data occurs. The SN ratio in the larger-the-better type is calculated by the following formula.

$$\eta = -10\log\frac{1}{d}\left(\frac{1}{D_1^2} + \frac{1}{D_2^2} + \cdots + \frac{1}{D_d^2}\right) \tag{1.14}$$

where,

η : SN ratio (db)
d : The number of abnormal data

(b) **Smaller-the-better type:** The smaller-the-better type is used for the opposite purpose with the larger-the-better type. This type is used in the purpose to decrease the separation between the Mahalanobis space and the data. Therefore, the normal data have to be employed for the optimization. Namely, this type is utilized when the optimization is performed to construct the Mahalanobis space which can ensure that the normal data can be recognized correctly. The SN ratio in the smaller-the-better type is calculated by the following formula.

$$\eta = -10\log\frac{1}{d}\left(D_1^2 + D_2^2 + \cdots + D_d^2\right) \tag{1.15}$$

(c) **Approximation type:** The approximation type is used in the case where there is a particular target value of the Mahalanobis distance in abnormal datum. The calculation of this SN ratio is carried out by the following formulas.

$$S_T = D_1^2 + D_2^2 + \cdots + D_d^2 \tag{1.16}$$

$$S_m = \frac{\left(\sum_{i=1}^{d} D_i\right)^2}{d} \tag{1.17}$$

$$S_e = S_T - S_m \tag{1.18}$$

$$V_e = \frac{1}{d-1} S_e \tag{1.19}$$

$$\eta = \frac{\frac{1}{d}(S_m - V_e)}{V_e} \tag{1.20}$$

In this type, the SN ratio will be large in the case where the distribution of the Mahalanobis distances of abnormal data is small. Thus, the Mahalanobis distances become close to a particular target value by changing the combination of the attributes which are used for the recognition.

4. **Calculation of factorial effect:** The factorial effect of each attribute is calculated on the basis of SN ratio. The total of the SN ratios of Level 1 and Level 2 is computed on each attribute. The factorial effect of each attribute is calculated by subtracting the average SN ratio of Level 2 from the average SN ratio of Level 1.

5. **Decision of attributes used for the recognition:** The attributes which have a large positive factorial effect are the attributes that significantly contribute to the recognition. It is generally established that attributes having the factorial effect of over 3 db are effective for the recognition. On the contrary, the attributes which have a negative factorial effect are the attributes that include a large amount of noise and reduce the performance of the recognition. Human selects the attributes used for the recognition on the basis of these factorial effects. After the selection of attributes, the Mahalanobis space is constructed from the selected attributes. The attributes used for the recognition are finally decided by confirming the performance of the constructed Mahalanobis space.

As described above, in the MTS, the sensuous recognition which human can perform is implemented by combining three factors; i.e., the

pattern recognition based on the Mahalanobis space and distance, the optimization of attributes using SN ratio and the operation by a human. It is likely that although the introduction of human operation into the system is effective, the system performance depends on the human recognition ability and the human load is large. To solve these problems, the research has been carried out on the improvement of the utility in the MTS. Nakatsugawa et al. proposed a threshold decision algorithm which is based on the distribution of normal data in order to implement a threshold decision which does not depend on the application [123]. It is obvious that the performance and utility in the MTS will be improved by such study in the future.

1.6.4 Mahalanobis-Taguchi-Gram-Schmidt

In previous subsection, the recognition method and the optimization method in the MTS were described. As described above, the calculation of an inverse matrix is required for the construction of a Mahalanobis space. Moreover, the construction of many Mahalanobis spaces is performed in the optimization of attribute. The calculation of inverse matrix is time consuming, and its computational cost is high. In addition, there is the case where the inverse matrix can not be calculated. That is, the Mahalanobis space can not be constructed. Recently, the Mahalanobis-Taguchi-Gram-Schmidt (MTGS) has been proposed to reduce the computational cost and to avoid the inability to construct an Mahalanobis space [124][125]. The MTGS is implemented by introducing the Gram-Schmidt orthogonalization process into the construction of the Mahalanobis space and the calculation of Mahalanobis distance. Moreover, in the MTGS, the optimization of attribute and the recognition can be simultaneously carried out. The recognition method of the MTGS is as follows.

k denotes the number of attributes, and n denotes the number of training data in the set. The set of training data is normalized by the average m_j and the standard deviation δ_j, and then $X_{i,j}$ is calculated.

$$X_{i,j} = \frac{x_{i,j} - m_j}{\delta_j} \quad (i = 1, \cdots, n; j = 1, \cdots, k) \tag{1.21}$$

Here, the variables $c_{i,j}$ is prepared instead of $X_{i,j}$. Moreover, the attributes are arranged in the order which is considered to be effective for the recognition by a human. The attribute which is considered to be most

effective for the recognition is treated as $X_{i,1}$. After this arrangement, $c_{i,j}$ is calculated as follows.

$$\begin{cases} c_{i,1} &= X_{i,1} \\ c_{i,2} &= X_{i,2} - r_{2,1}c_{i,1} \\ c_{i,3} &= X_{i,3} - r_{3,1}c_{i,1} - r_{3,2}c_{i,2} \\ &\vdots \\ c_{i,k} &= X_{i,k} - r_{k,1}c_{i,1} - r_{k,2}c_{i,2} - \cdots - r_{k,k-1}c_{i,k-1} \end{cases} \tag{1.22}$$

where,

$$r_{i,j} = \frac{1}{nV_j}\sum_{l=1}^{n} X_{l,i}c_{l,j} \tag{1.23}$$

$$\begin{cases} V_1 = &= \frac{1}{n}(c_{1,1}^2 + c_{2,1}^2 + \cdots + c_{n,1}^2) \\ V_2 = &= \frac{1}{n-1}(c_{1,2}^2 + c_{2,2}^2 + \cdots + c_{n,2}^2) \\ &\vdots \\ V_k = &= \frac{1}{n-k+1}(c_{1,k}^2 + c_{2,k}^2 + \cdots + c_{n,k}^2) \end{cases} \tag{1.24}$$

The Mahalanobis distance is calculated by using V_i and $r_{i,j}$. First, the datum z to be recognized is normalized by using the average m_j and the standard deviation δ_j of Equation (1.21) to obtain Z.

$$Z = \{Z_1, Z_2, Z_3, \cdots, Z_k\} \tag{1.25}$$

The variable e_i is prepared, and the e_i is calculated on the basis of Z.

$$\begin{cases} e_1 &= Z_1 \\ e_2 &= Z_2 - r_{2,1}e_1 \\ e_3 &= Z_3 - r_{3,1}e_1 - r_{3,2}e_2 \\ &\vdots \\ e_k &= Z_k - r_{k,1}e_1 - r_{k,2}e_2 - \cdots - r_{k,k-1}e_{k-1} \end{cases} \tag{1.26}$$

y_i is computed by using e_i and V_i of the formula (1.24).

$$\begin{cases} y_1 &= e_1 \\ y_2 &= \frac{e_2}{\sqrt{V_2}} \\ y_3 &= \frac{e_3}{\sqrt{V_3}} \\ &\vdots \\ y_k &= \frac{e_k}{\sqrt{V_k}} \end{cases} \tag{1.27}$$

The Mahalanobis distance is calculated as

$$D^2 = \frac{1}{k}\sum_{i=1}^{k} y_i^2 \tag{1.28}$$

The calculation of the Mahalanobis distance and the optimization can be simultaneously performed because the Mahalanobis distance of the datum can be calculated by using the arbitrary number of attributes without the reconstruction of the Mahalanobis space. In the MTGS, the optimization of attribute is also performed by using the SN ratio. In the optimization of the MTGS, the orthogonal array and factorial effect are not used because the attributes are sorted in effective order for the recognition which is decided by a human.

Furthermore, in the MTGS, the direction of abnormal data can be decided. If the Mahalanobis space is constructed from the training data which consists of standard normal data, the excellent normal datum may take a large Mahalanobis distance. Namely, the normal datum is regarded as an abnormal datum. To avoid this problem, Rajesh et al. have proposed a method to distinguish excellent normal data from abnormal datum by using the MTGS [126]. However, it is hard to perform this method in the case where the datum has many attributes. The further improvement is expected.

1.7 Interactive Evolutionary Computing

What kind of impression do you receive when you look at a picture? Probably, the impression which you receive may differ from the impressions which other people receive, because each person has subjectivity or Kan-sei. However, in conventional system, it is assumed that the evaluation of person is unique. Namely, the subjectivity or Kan-sei is not considered in the conventional system. Therefore, in conventional system, it is difficult to treat the problem such as creation or design based on the user's subjective preference.

Recently, the paradigm called Interactive Evolutionary Computing (IEC) has been proposed in order to implement the system which can consider the subjectivity or Kan-sei on each person. This paradigm is employed on various fields such as engineering, art, computer graphics, education and so on, and many applications are developed. Moreover, the improvement of the performance of the IEC is carried out. The IEC becomes important technique in artificial intelligence.

In this section, the framework and feature of the IEC which is fundamental technique for the face recognition method in this dissertation are explained and several applications using the IEC on some fields are surveyed. In addition, some attempts to solve the problems in the IEC are introduced.

1.7.1 Framework and Feature of IEC

Since the word of "Kan-sei" came to be used in the engineering field, the research which focuses on human is carried out. As an example of such research focusing on human, there is the problem such as creation, adjustment or design which is based on subjective preference on each person. To deal with the problem, the approach which constructs human mental model instead of human is employed in conventional artificial intelligence. However, it is considerably difficult to construct the human mental model completely.

Recently, a new paradigm has been proposed to deal with this problem. This paradigm is called the Interactive Evolutionary Computing (IEC) since this paradigm is implemented by combining the evolutionary computing such as the Genetic Algorithm or the Genetic Programming and the interaction between human and computer. Especially, it is called the Interactive Genetic Algorithm (IGA) in the case where the Genetic Algorithm (GA) is employed as the evolutionary computing, and it is called the Interactive Genetic Programming (IGP) in the case where the Genetic Programming (GP) is used, respectively. The IEC is utilized on various fields because the IEC has the advantage that the user can obtain the solution which reflects the subjective preference on each person [127]. The framework and feature of the IEC are described below.

The user's operation and the internal behavior in the system based on the IEC is shown in Figure 1.11. As shown in Figure 1.11, at first, the evolutionary computing randomly generates a population which consists of some individuals, namely genotype, because there is no information about the user's subjectivity. The individual consists of the parameters

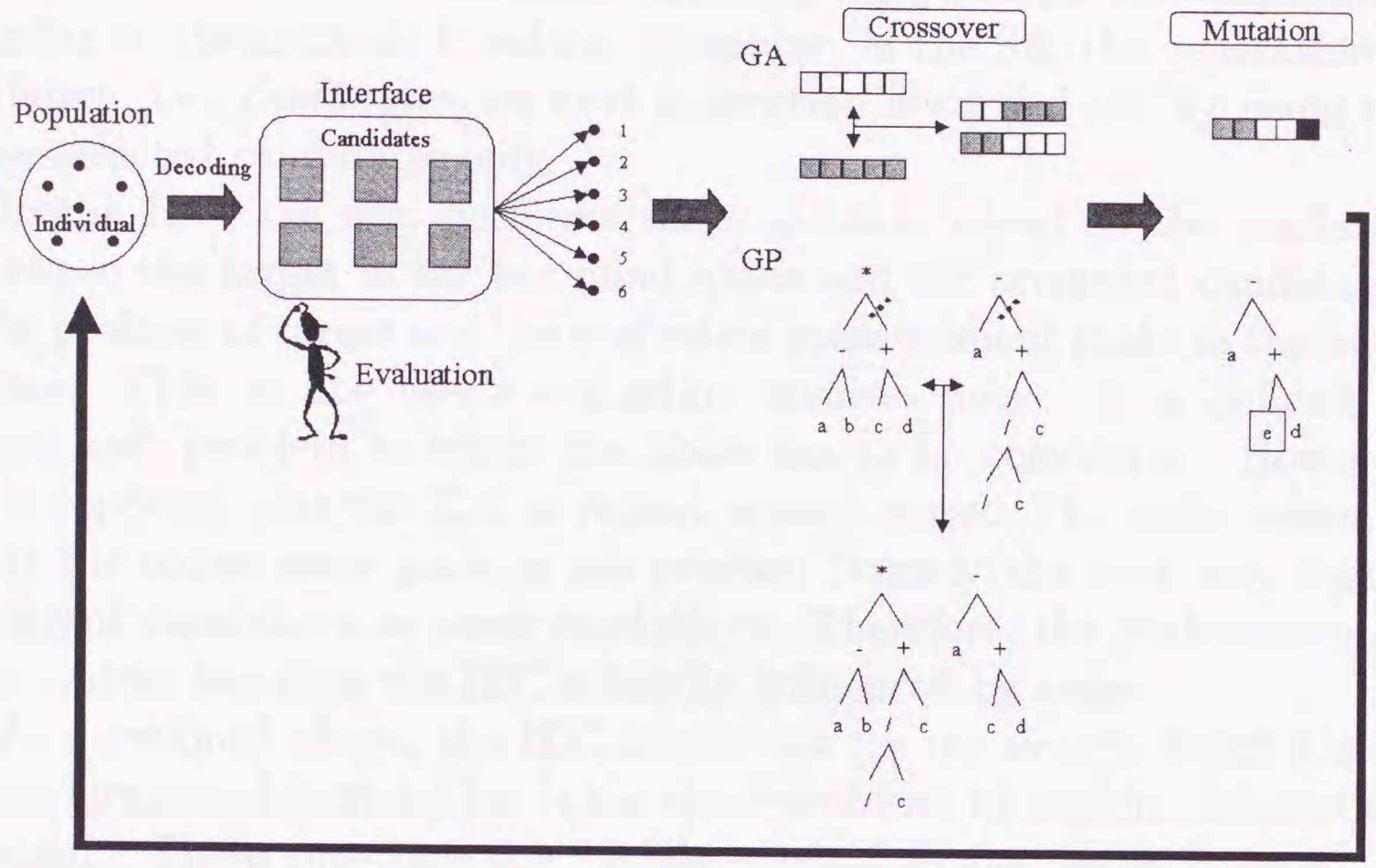


Figure 1.11: User's operation and internal behavior in the system based on the IEC

which the image, sound and so on have. Next, the individuals in this population are decoded. That is, images, sounds and so on are created. These created objects are called candidates. These candidates are displayed on the interface. The user performs the evaluation of candidates on the interface in accordance with the distance between the target in his/her mind space and candidate. The evaluation is carried out by scalar value. That is, the user gives high evaluation value to the candidate which is similar to the target on his/her mind space. The evaluation value which is given by the user is treated as the fitness value on the evolutionary computing. The crossover operation and the mutation operation are applied to a set of individuals on the basis of the fitness value. The candidates which are more suitable for the user's subjective preference can be generated since the genetic operators are applied to based on the user's subjective evaluation. After the individuals are returned to the population, the individuals are decoded and then they are presented to the user. The user evaluates them again. The user can obtain the satisfiable candidates by repeating this operation.

Although the evaluation is performed by giving evaluation value to the candidates in the IEC, there is a case where the evaluation is performed by binary value. Namely, it is the evaluation method that the user selects the candidates only which are suitable for own preference. In this

case, the IEC is called Simulated Breeding (SB) because the operation is similar to the artificial breeding. Moreover, in the SB, the generation of children, i.e., candidates, on next generation is carried out by using the user-selected candidates only.

In the IEC, the user performs the evaluation based on the similarity between the target in his/her mind space and the presented candidates. The position of target and the evaluation measurement shake in the mind space. That is, the user's evaluation involves noise. It is difficult to treat such problem in which the noise has to be considered. However, it is reported that the IEC is robust against noise. The main reason is that the convergence point is not precise. Namely, the user may regard different candidates as same candidates. Therefore, the performance of the system based on the IEC is hardly influenced by noise.

As mentioned above, the IEC is effective for the system which has to treat human subjectivity but it has some problems to implement practical system. These problems are mainly related to the user's fatigue. In the system based on the IEC, The user has to evaluate the presented candidates repeatedly. This is the backbreaking task for the user and fatigues the user. Therefore, in the IEC, the following items have to be considered. One is that the population size and the generation size have to be small in comparison with the standard GA. The small population size causes the problem of initial convergence. The high convergence speed is required in the case where the generation size is small. Another is that the simple evaluation method is needed in order to reduce the user's fatigue. The evaluation method in the SB is easier for the user than standard IEC. However, the evaluation method in the SB also causes the initial convergence because new children are generated by the selected candidates only. The detail of these problems and the improvement of the IEC to solve these problems are described in next subsection.

1.7.2 Applications and Related Works

The studies on the IEC are classified into three types. One is applications of the IEC to art such as image or music creation. Another type is an application in engineering. In engineering, sound processing, image retrieval, image processing, knowledge acquisition and so on are involved as the fields of application. Final one is the study to improve the performance of the IEC. The brief introduction of the studies on each type is described below.

1.7.2.1 Applications in Art

In initial research on the IEC, image creation was main application. As first application of the IEC, R. Dwakins proposed the Biomorph [129]. Thereafter, K. Sims proposed a system which interactively generates CG images. This CG image is drawn by mathematical formula, and this formula is evolved by the IGP. Unemi proposed SBART for CG creation by same way [130]. The Biomorph by R. Dawkins and the Genetic Images by K. Sims are shown in Figure 1.12 and Figure 1.13, respectively. The concern with the IEC has grown in the field of art after these proposals. Several applications based on the IEC in art field are briefly introduced below.

Aoki et al. proposed a design support system for lighting design in computer graphic [131]. In the system, the decisions of the type, the position, the ON/OFF and the strength for three lights are supported by the IGA. They designed an experiment in order to compare the IEC with trial and error. In this experiment, a design concept was given to the user, and the user created CG image by each method. The user compared these images which are created by both methods and then selected either CG image which was more suitable for the design concept. The experimental result statistically revealed that the proposed system was not effective for expert CG designers but was effective for amateur CG designers. The most important point to note is that the subjective evaluation test and the statistical test were introduced into the experiment. The conventional studies in the art have focused on the implementation of the system. They explicated the effect of the system statistically. Their experiment was meaningful in this respect.

Nakanishi proposed a fashion design aid system using the IEC [132]. In this system, a ladies' one-piece design is drawn in accordance with eight parameters. The individuals are generated by taking one of eleven values on each parameter. The decision of this value is performed on the basis of probability which each value has. The user chooses a most favorite design from the candidates which are created by decoding the individuals. The probability for each value is changed on the basis of the chosen design. The individual consisting of the values which have highest probability on each parameter means most suitable design for the user. Moreover, this system can obtain the Kan-sei information of a user by using the GP and operation history. He made an experiment based on a simulation using the mathematical function obtained by the GP and indicated the effectiveness of the proposed system.

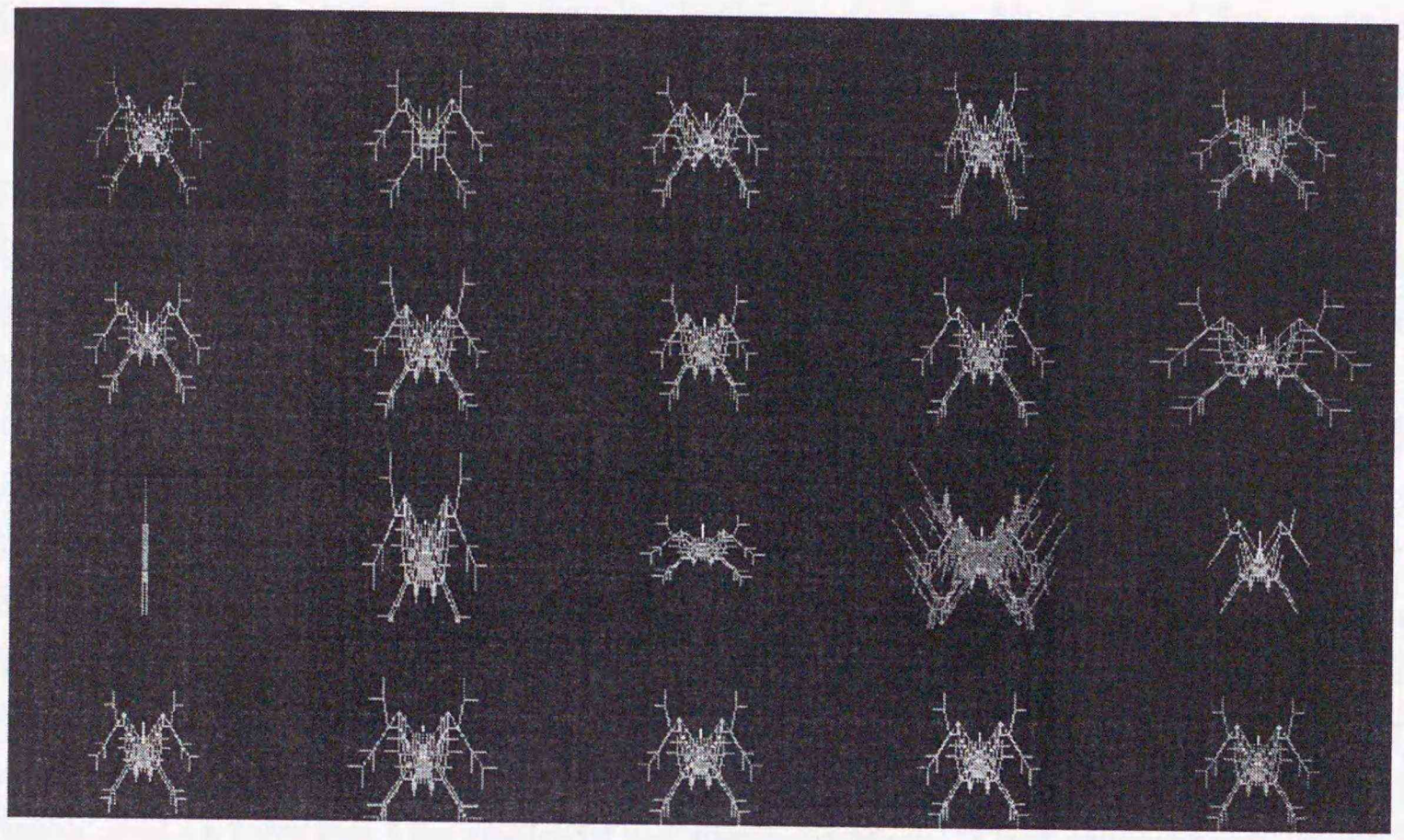


Figure 1.12: Biomorph by R. Dawkins

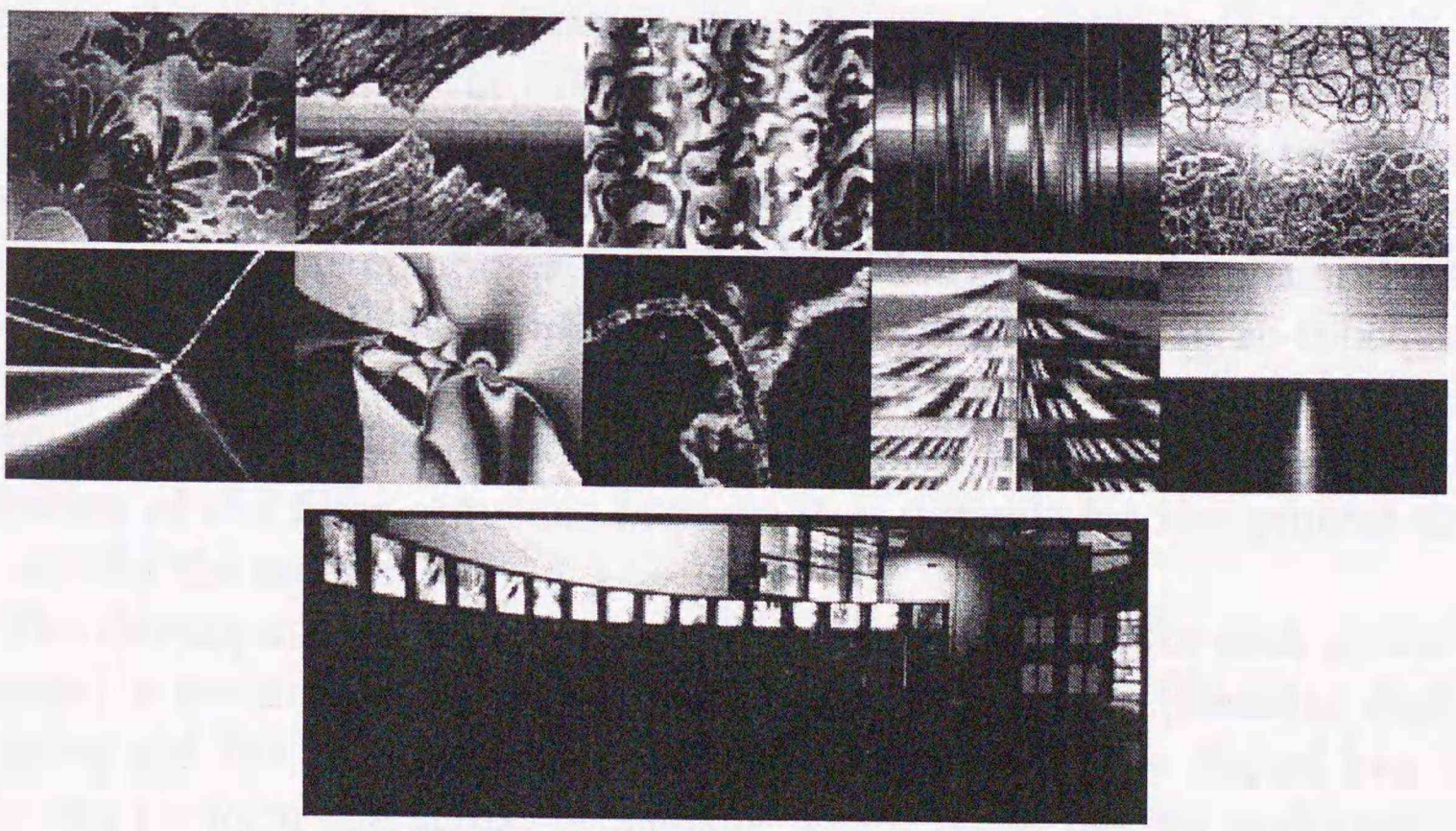


Figure 1.13: Genetic Images by K. Sims

Hee et al. proposed a 3D fashion design aid system using the IGA [133]. In this system, they improved the efficiency of the IGA by introducing domain specific knowledge. In the fashion design, there exist impractical designs. This system avoids generating such designs by using domain specific knowledge. That is, some clothing parts which are treated in the system are combined in advance so that impractical designs are not created. They made the experiment using some subjects. The experimental results revealed that the developed system was effective for the fashion design aid.

The IEC was applied to the melody creation for jazz session by Biles [134]. This system is called GenJam. In this system, the user listens to the generated melody, and evaluates it by pushing a good button or a bad button at any time. Namely, the user can evaluate the melody at any time while it is played because it is time consuming and increases the user's fatigue to evaluate the melody after the play is finished. Therefore, the GenJam does not evolve the whole of music but it treats a phrase as a unit and it evolves the melody on each phrase. This evaluation method is effective for the IEC in the case where sequential signal is treated in the system.

1.7.2.2 Applications in Engineering

Recently, the IEC has also come to be utilized in engineering. In engineering, the following applications have been proposed.

Mutoh et al. applied the SB to the decision of the image filter processing sequence [135]. The image data came to be frequently used by general computer user in accordance with the development of digital camera, scanner, personal computer and World Wide Web. The general user uses a software which equips with several image filters to re-touch the image data. Generally, the different processing effect can be obtained by changing the sequence of these filters. The proposed system supports the decision of the filter sequence because it is difficult for the general user to decide the sequence.

The development of the hearing aid which is suitable for each senior is needed in accordance with the increase of senior citizen. Recently, digital hearing aid has been developed for such a demand. The digital hearing aid can perform the signal processing which could not be performed in conventional analogue hearing aid. In the digital hearing aid, the parameter tuning is required in order to make appropriate hearing aid for each person. In conventional tuning, the doctor measures the hearing condi-

tion of the user and then adjusts the parameter. However, it is difficult to adjust to most appropriate parameter by the doctor because the person which can evaluate goodness/badness of the hearing aid is the user only. Therefore, the user has to tune the parameter in order to make most appropriate digital hearing aid. This is hard task for the user who is not expert. Takagi et al. proposed a system based on the IEC for the parameter tuning of digital hearing aid [136]. In this system, the user can obtain most appropriate parameter for the signal processing of digital hearing aid by evaluating the easiness to hear the sounds which are presented by the system. Their experimental results statistically revealed that the proposed system could effectively support the parameter tuning for the digital hearing aid.

It is difficult to retrieve the image data which are suitable for each user's subjectivity from large-scale database or the internet. In the case where image data are treated, it is difficult to implement the retrieval using the verbal keywords because the user does not know the kinds of the images which are stored in the database or the internet. Moreover, there is a case where the subjectivity of the database manager who attaches the keywords to the image data differs from the subjectivity of each user. The IEC can be utilized to retrieve the image data which are suitable for each user's subjectivity. In the retrieval system based on the IEC, the raw information such as pixel value which image datum has is used as the parameter. The image data are retrieved by using this parameter. The user evaluates the presented image data on the basis of own subjectivity. This parameter is evolved by this evaluation and the parameter to retrieve more appropriate image is generated by the EC. There are some proposals for the image retrieval [137][138][139]. Cho implemented an image retrieval system using the wavelet and the IGA [137]. Kitamoto et al. proposed an image retrieval system using the queue-based genetic algorithm [138][139]. In the IEC, the user spends long time for the evaluation of candidates. Moreover, the GA operation is time consuming in treating photographic image data because the amount of information in photographic image data is considerably large. Kitamoto et al. improved the speed of the system by overlapping the human evaluation time and the GA operation time.

Caldwell proposed an application for tracking a criminal suspect, i.e., a montage system [140]. At first, some photofit facial images are displayed to the user in this system. The user evaluates all facial images which are displayed on the interface and then gives the rating to all of them.

This rating is regarded as the fitness value of individual. The genetic operations are applied to the facial images based on this value. The face of criminal suspect emerges by repeating this process.

Isono et al. proposed a 3D facial image generating system using the IGA [141]. In this system, the coding of individual is performed by using the PCA. It has been reported that the PCA is effective method for face recognition. The experimental results revealed that an effective facial image generating can be implemented by combining the PCA and the IGA.

Monmarche et al. proposed a HTML design support system by using the IGA [142]. In this system, background style, link color, font style and so on which are elements for designing HTML are coded as the individual. Moreover, the vector of probability is used instead of the population in this system. The individuals are generated in accordance with the probability which each value on each locus has. The probabilities are updated according to the values which the user-selected individuals have.

Kuriyama et al. proposed a composition support system for kids [143]. The decision of the contexture in composition is important and difficult task for kids. In their system, twenty-four pictures are presented on a display. The kid selects four pictures from the displayed pictures and then aligns the pictures. The kid makes two sequences by this operation. This sequence means the story of the composition. The selected sequences are employed as the parents, and the GA operation is applied to the parents. Some children are generated on the basis of the GA operation. The kid newly selects two sequences from the presented sequences. The decision of story in the composition is supported by repeating this operation.

To select preferable candidates is easier than to write program. The selection in the IEC can be regarded as the reward in the reinforcement learning. Therefore, the IEC can learn the rule or the mechanism for control as well as the reinforcement learning. As an application in this field, Pagliarini et al. proposed a moving robot based on Neural Network and the IEC [144]. The user selects robots which have preferable movement. The evolutionary computing evolves the weights in the Neural Network. The robot having the Neural Network which can perform the preferable movement is created by repeating the selection.

Fujii et al. introduced the concept of evolution and interaction in the IEC into a fuzzy classifier system [145]. They employed a robot control problem as an example and attempted to generate interactively rules which reflect the user's subjectivity.

Tabuchi et al. applied the IEC to knowledge acquisition for classifying eight kinds of defect data in stainless steel [146]. In this system, basically, the search is automatically performed. The user performs the evaluation at the interval of one hundred steps. It was found from the experimental result that the rule was considerably changed when the user evaluated it, and the robust rule could be obtained.

1.7.2.3 Improvement of the IEC

In the IEC, some technical problems have to be solved in order to implement practical system. These problems and the attempts for solving them are described below.

In the system based on the IEC, some candidates are presented on the interface. The user evaluates candidates by comparing them. Therefore, it is likely that the user's fatigue for the evaluation can be reduced by facilitating the comparison of candidates which are presented on the interface. Ohsaki et al. proposed an interface which changes the order of the presented candidates in order to facilitate the comparison [147][148]. This interface predicts the evaluation values of the candidates which will be given by the user. The candidates are displayed on the interface in accordance with the order of predicted evaluation value. As the prediction method, Ohsaki et al. employed a neural network and Euclidean distance. They designed some practical experiments in order to confirm the effectiveness of the proposed interface. From the experimental result, they concluded that the proposed interface was not effective statistically but the direction of future research could be obtained from the opinion of the subjects.

In the IEC, there is a problem of initial convergence. In the system based on the IEC, the number of candidates displayed on the interface is small because it makes the user's evaluation difficult to present a large number of candidates at the same time. This number of candidates means the population size in GA. In the case where the population size is small, the problem of initial convergence occurs. Namely, the individuals in the population are similar to each other in the initial generation, and the different kinds of individual are not generated. Therefore, the user can not obtain optimal candidates. Moreover, this problem frequently occurs in the system using the SB because the children of the next generation are generated on the basis of the user-selected candidates only. To avoid this problem, Unemi proposed a multi-field interface for the SB [149]. This interface mimics the island-model GA. The interface has two fields for

breeding the candidates. The user can breed different kinds of candidates on each field. Moreover, the candidates which are bred on one field can be transplanted into another field. The diversity can be increased by hybridizing the different kinds of candidates which are bred on each field.

In the IEC, there are two kinds of user's evaluation method for the candidates. Caldwell proposed the method that the user gives the evaluation values to all presented candidates in nine grades [140]. This method is called the rating-all method. On the other hand, the selective method such as SB has been proposed. In the selective method, the user selects one or some which is suitable for the purpose from the presented candidates. Although the rating-all method is effective in terms of the search performance because all individuals can get the fitness value, this task is backbreaking for the user and fatigues the user. On the other hand, the user's load is very light in the case of the selective method. However, there is the problem of the initial convergence in the selective method. The selective method is very useful evaluation method for the user if the initial convergence problem can be solved. To implement such an evaluation method, Sugimoto et al. proposed fitness assignment methods; i.e., bias method and fuzzy method [150]. In both methods, the user selects one suitable candidate for the purpose. The fitness is given to all candidates which are presented on the interface on the basis of the user-selected candidate. In the bias method, large fitness value is given to user-selected candidate and small fitness value is given to all candidates which are not selected by the user. In the fuzzy method, the fitness value is decided by the fuzzy reasoning which is based on the generation, the distance and the angle between the user-selected candidate and others. They designed the experiment using some subjects in order to confirm the effectiveness of their proposed methods. The experimental results revealed that the proposed methods were effective evaluation method for the reduction of user's fatigue in comparison with the rating-all method.

In the IEC, the candidates which are randomly generated are presented to the user in the first generation because the system does not have the information of the user's subjectivity or Kan-sei. The effectiveness of the system will increase if the system can store the information of user's subjectivity or Kan-sei in previous task and can reflect it to the candidates which are presented in the first generation. Kato et al. propose an image retrieval system utilizing personal history reflecting user's subjectivity [151]. The system utilizes the personal history in previous retrieval to generate the initial population. The experimental results revealed that

high evaluation values were given to the individuals in the initial step of retrieval in comparison with the system without the personal history.

The purpose of the conventional IEC is to find the suitable candidates for the user's subjective preference. Namely, only a user evaluates the candidates. This can be regarded as the optimization problem of an evaluation function which the user has. However, there are many problems which have multiple evaluation function, i.e., multi-objective optimization problem. Shibuya et al. proposed the extension of the IGA in order to solve multi-objective optimization problem [152]. They took up the motion animation design problem as the multi-objective optimization problem. In this motion animation problem, the purpose is to make the motion animation which is suitable for the user's subjectivity. However, there is the pre-defined evaluation for natural motion in the motion animation. That is, in this problem, this pre-defined evaluation function and the evaluation function which the user has must be optimized at the same time. Additionally, this problem has some constraints which must be satisfied. They introduced the concept of multi-objective GA into the IGA so as to treat this multi-objective optimization problem. The process of multi-objective interactive genetic algorithm is illustrated in Figure 1.14. At first, the individuals in the population are randomly generated. These individuals are divided into some groups on the basis of the similarity among the individuals. In each group, Paleto optimal sets are automatically found by the multi-objective GA. The groups are merged into the population, and some individuals are presented to the user for the evaluation. These individuals are sorted in accordance with the user's subjective preference. The crossover operator is applied to the candidates which high evaluation is given. The problem is optimized by repeating this process. They designed the experiment with some subjects. In the experimental results, it was shown that the proposed method could generate high quality motion animations with less stress.

1.7.3 Future of IEC

In this section, the framework and feature of the IEC were explained, and the applications on various fields were surveyed. In addition, some technical problems in the IEC and the methods to solve them were introduced. As mentioned above, the task domain which is suitable for applying the IEC is expanding under the influence of the development of Kan-sei engineering or human interface. It is likely that its domain is extending hereafter.

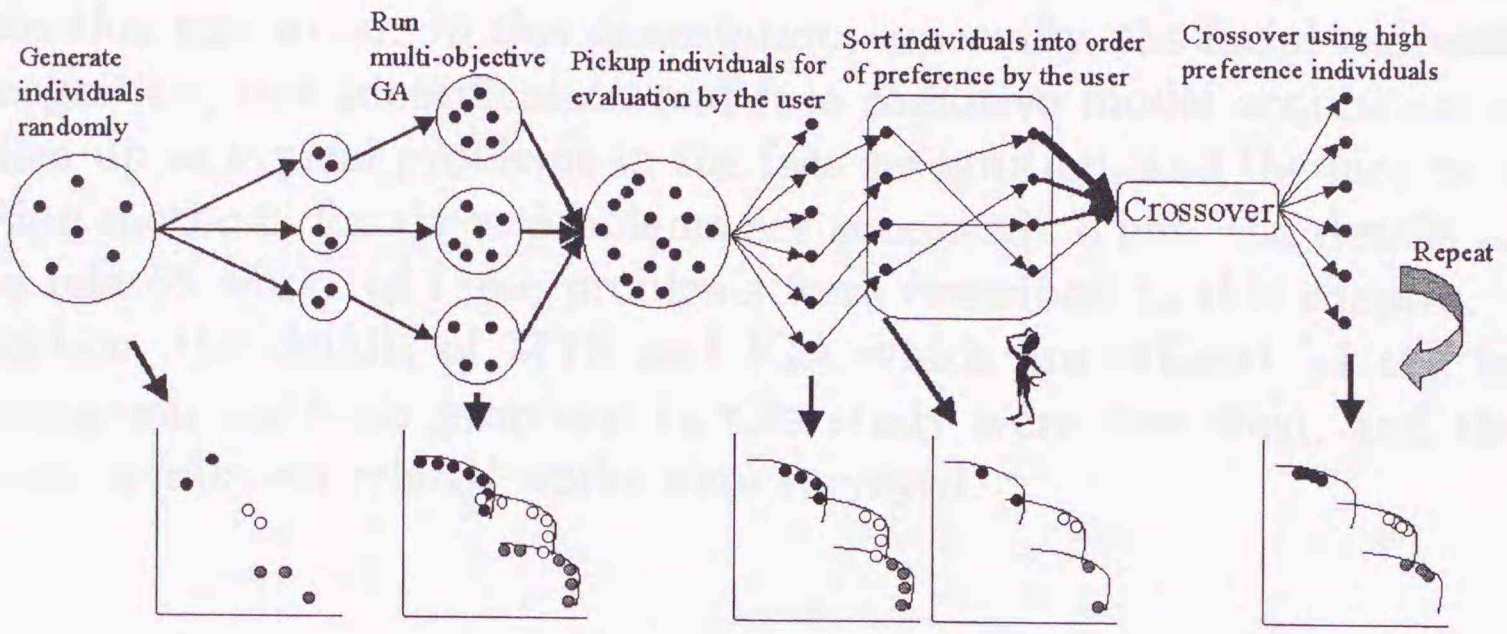


Figure 1.14: Multi-objective Interactive Genetic Algorithm

The paradigm of IEC is different from the approach of analysis type in the conventional artificial intelligence that the human excellent evaluation ability is replaced as a computing model. The IEC is an approach of composition type which is implemented by introducing the human excellent ability itself into the system. It is obvious that this approach which esteems the human excellent ability is effective for the problems such as the creation, adjustment or design based on human subjectivity or Kansei. If the user's fatigue problem which is one of important problems in the IEC for the development practical system can be solved, the IEC will become representative method in the artificial intelligence.

1.8 Conclusion

In this chapter, the face recognition and its recognition methods to implement flexible and adaptable human interface were focused on. Human face plays important roles in the face-to-face communication among humans. Humans can obtain from face various kinds of information such as intention, emotion, state and so on. The effective and user-friendly human interface can be implemented if computer can flexibly recognize or understand the information which human face represents. Therefore, recently, the research on the recognition of the information which face represents has been carried out, and this subject is called the face recognition. In this chapter, the survey on the face recognition was described. The face recognition consists of several problems such as the face expres-

sion recognition, face identification, face cognitive model acquisition, face detection and so on. In this dissertation, especially, the facial expression recognition, face identification and face cognitive model acquisition are taken up as typical problems in the face recognition, and the face recognition methods for these problems are proposed. Thus, the details and the related works of these problems were described in this chapter. In addition, the details of MTS and IGA which are utilized for the face recognition methods proposed in this study were described, and their recent works and related works were surveyed.

Bibliography

[1] S. Akamatsu, "Computer Recognition of Human Face - A Survey -", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J80-A, No.8, pp.1215-1230, Aug. 1997. (in Japanese)

[2] S. Morishima, "Technical Problems in Facial Expression Recognition and Synthesis", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU97-141, pp.83-92, Nov. 1997. (in Japanese)

[3] H. Yamada, "Psychological Model for Facial Expressions of Emotion Recognition", Journal of The Society of Instrument and Control Engineers, Vol.33, No.12, pp.1063-1069, Dec. 1994. (in Japanese)

[4] S. Yoshikawa, "Perspectives of research on facial expression", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU97-11, pp.69-74, Nov. 1997. (in Japanese)

[5] T. Ogawa, M. Oda, S. Yoshikawa and S. Akamatsu, "Evaluation of Facial Expression Differing in Face Angles -Constructing a Database of Facial Expressions-", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU97-136, pp.47-52, Nov. 1997. (in Japanese)

[6] N. Matsuzaki and T. Sato, "The role of the static pattern and the motion in the perception of facial expressions", Technical report of Institute of Electronics, Information and Communication Engineers, HCS99-12, pp.1-5, June 1999. (in Japanese)

[7] M. Kamachi, S. Yoshikawa and S. Akamatsu, "Does changing speed affect the recognition of facial expression?", Technical report of Institute of Electronics, Information and Communication Engineers, HCS98-34, pp.17-24, Oct. 1998. (in Japanese)

[8] P. Ekman and W. V. Friesen, "Facial Action Coding System", Consulting Psychologists Press, Palo Alto, US., 1978.

[9] O. Hasegawa, S. Morishima and M. Kaneko, "Processing of Facial Information by Computer", Journal of Institute of Electronics, Information and Communication Engineers, Vol.80-A, No.8, pp.1231-1249, Aug. 1997. (in Japanese)

[10] K. Mase, "Detection of Facial Muscle Motion by Optical-flow", Technical report of Institute of Electronics, Information and Communication Engineers, PRU89-128, pp.17-24, March 1989. (in Japanese)

[11] K. Mase, E. Maeda and Y. Suenaga, "Recognition of Emotion from Image Sequence of Facial Expression", Technical report of Institute of Electronics, Information and Communication Engineers, PRU91-24, pp.95-101, May 1991. (in Japanese)

[12] Z. Prochazka, T. Ito and T. Okamoto, "Facial Expression Recognition System Using Facial Motion", Technical report of Institute of Electronics, Information and Communication Engineers, HIP96-42, pp.19-24, March 1997. (in Japanese)

[13] C. S. Choi, H. Harashima and T. Takebe, "Analysis of Facial Expression Using Three-Dimensional Facial Model", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J74-D-II, No.6, pp.766-777, June 1991. (in Japanese)

[14] T. Otsuka and J. Ohya, "A Study of a Non-Verbal Information Transfer Method for a Communication Environment -Head Pose Estimation, Facial Expression Recognition and Conversion", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU98-224, pp.33-39, Feb. 1999. (in Japanese)

[15] F. Kawakami, H. Yamada, H. Harashima and S. Morishima, "Construction of Expression Analysis and Synthesis System based on 3-D Emotion Model", Technical report of Institute of Electronics, Information and Communication Engineers, HCS95-27, pp.7-14, March 1996. (in Japanese)

[16] Y. Ando and T. Yonekura, "A Study of Emotional Interaction System Using Facial Expression", Technical report of Institute of Electronics, Information and Communication Engineers, MVE96-72, pp.29-36, March 1997. (in Japanese)

[17] T. Ishii and H. Iwata, "Automatic Recognition of Human Facial Expression by Computer Image Processing", Transactions of the

Japan Society of Mechanical Engineers (Series C), Vol.52, No.483, pp.2989-2992, Sept. 1986. (in Japanese)

[18] H. Kobayashi, H. Takahashi and K. Kikuchi, "Automatic Extraction of Facial Organs and Recognition of Facial Expressions", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-138, pp.37-42, Nov. 1999. (in Japanese)

[19] K. Matsuno, C. Lee and S. Tsuji, "Recognition of Facial Expression Using Potential Net and KL Expansion", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J77-D-II, No.8, pp.1591-1600, Aug. 1994. (in Japanese)

[20] H. Bessyo, Y. Iwai and M. Yachida, "Detecting Human Face and Recognizing Facial Expressions Using Potential Net", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-139, pp.43-48, Nov. 1999. (in Japanese)

[21] T. Sakaguchi, J. Ohya and F. Kishino, "Facial Expression Recognition from Image Sequence Using Hidden Markov Model", Journal of The Institute of Image Information and Television Engineers, Vol.8, No.49, pp.1060-1067, Aug. 1995. (in Japanese)

[22] H. Sakano, "The State of art of Biometric Person Authentication", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-29, pp.75-82, June 1999. (in Japanese)

[23] M. Yoshino, "Facial Image Identification in Criminal Investigation", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-28, pp.67-74, June 1999. (in Japanese)

[24] H. Sato, N. Terashima and H. Tominaga, "Human Face Pose Estimation for Creating Avatars", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-153, pp.23-30, Nov. 1999. (in Japanese)

[25] K. Sugiyama, K. Yamada, Y. Yonamine and H. Nakagawa, ""Topical Face" Detection Using Decision Tree Photograph Newspaper", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-154, pp.31-38, Nov. 1999. (in Japanese)

[26] T. Yokoyama, Y. Yagi and M. Yachida, "Active Contour Models for Extracting Facial Contour", Journal of Information Processing Society of Japan, Vol.40, No.3, pp.1127-1137, March 1999. (in Japanese)

[27] R. Nagata and T. Kawaguchi, "Detection of Human Faces in a Color Image with Complex Background", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-111, pp.77-82, Nov. 1999. (in Japanese)

[28] Y. Li and H. Kobatake, "Face Detection Based On Dynamic Symmetry Operator", Technical report of Institute of Electronics, Information and Communication Engineers, HIP98-22, pp.3-8, Sept. 1998. (in Japanese)

[29] S. Clippingdale and T. Ito, "Feature selection experiments on the FAVRET face tracking and recognition system", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-109, pp.61-68, Sept. 1999. (in Japanese)

[30] M. Oshida, T.Yamaguchi, M. Tominaga, H. Koshimizu, T. Endou, T. Toriu and K. Murakami, "On the method for face and facial parts recognition in which Hough transform and the related procedures were utilized -Generating facial data for facial caricaturing system PICASSO-", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU98-225, pp.41-48, Feb. 1999. (in Japanese)

[31] O. Hori, "Face Detection in an Image Based on a Statistics Approach", Journal of Information Processing Society of Japan, Vol.40, No.8, pp.3281-3288, Aug. 1999. (in Japanese)

[32] N. Inakoshi, M. Maekawa and K. Nakamura, "Large-Scale Individual Human verification using face image", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU98-23, pp.77-84, May 1999. (in Japanese)

[33] H. Imaoka and S. Sakamoto, "Pose-Independent Face Recognition Method", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-26, pp.51-58, June 1999. (in Japanese)

[34] S. Kuriyama, K. Yamamoto, H. Hongo and K. Kato, "Face Recognition by using hyper feature fields", Technical report of Institute of

Electronics, Information and Communication Engineers, PRMU99-115, pp.105-110, Nov. 1998. (in Japanese)

[35] M. Sohma and K. Nagao, "Robust Face Recognition under Drastic Changes of Conditions of Image Acquisition", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J80-D-II, No.8, pp.2225-2231, Aug. 1997. (in Japanese)

[36] M. Doi, Q. Chen, A. Matani, O. Oshiro, K. Sato and K. Chihara, "Lock Control System Based on Face Identification", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J80-D-II, No.8, pp.2203-2208, Aug. 1997. (in Japanese)

[37] M. Doi, K. Sato and K. Chihara, "A Robust Face Identification against Lighting Fluctuation", Journal of the Institute of Systems, Control and Information Engineers, Vol.11, No.10, pp.546-553, Oct. 1998. (in Japanese)

[38] N. Yamamoto, K. Sigeneri and T. Shakunage, "Virtual Subspace Method for Robust Face Recognition Free with Lighting Conditions", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-208, pp.31-38, Jan. 2000. (in Japanese)

[39] S. Akamatsu, T. Sakaki, H. Fukamachi and Y. Suenaga, "Robust Face Identification by Pattern Matching Based on KL Expansion of the Fourier Spectrum", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J76-D-II, No.7, pp.1363-1373, July 1993. (in Japanese)

[40] T. Kurita, H. Shimai, T. Takahashi and T. Shigehara, "Integration of Viewpoint Dependent Classifiers for Viewpoint Invariant Face Recognition", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-170, pp.35-42, Dec. 1999. (in Japanese)

[41] M. Kitada and K. Kotani, "Face Identification Based on Morphology Analysis of Face Image", The Special Interest Group Notes of Information Processing Society of Japan, 98-CVIM-109, No.5, pp.65-72, 1998. (in Japanese)

[42] S. Umeyama, T. Kurita and S. Akaho, "A New Face Identification Method Using Supervised Independent Component Analysis", Technical report of Institute of Electronics, Information and

Communication Engineers, PRMU99-27, pp.59-66, June 1999. (in Japanese)

[43] Y. Hirayama and O.Nakamura, "A Robust Person Identification System for Appearance Change with Expression", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU97-13, pp.17-24, Nov. 1999. (in Japanese)

[44] H. Sakano and N. Mukawa, "Kernel Mutual Subspace Method and It's Application to Face Recognition", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-116, pp.111-116, Nov. 1999. (in Japanese)

[45] N. Futamura, K. Okada, S. Akamatsu, K. Mori and Y. Suenaga, "An ICA based on representation of facial images and its application for face recognition systems", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-180, pp.21-28, Dec. 1999. (in Japanese)

[46] D. Pramadhanto, H. Wu and M. Yachida, "Face Identification under Varying Pose Using a Single Example View", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J80-D-II, No.8, pp.2232-2238, Aug. 1997. (in Japanese)

[47] M. Kosugi, "Human-Face Recognition Using Mosaic Pattern and Neural Networks", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J76-D-II, No.6, pp.1132-1139, June 1993. (in Japanese)

[48] S. Maeda, M. Okamoto, T. Kawahara, M. Minoh, K. Ikeda and S. Doshita, "Individual Identification by Integrating Facial Image, Walking Image and Voice Feature", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J79-D-II, No.4, pp.600-607, April 1996. (in Japanese)

[49] K. Hisatomi and S. Hashimoto, "Face Recognition Using Distributed and Cooperative Agent System", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU97-133, pp.25-30, Nov. 1997. (in Japanese)

[50] N. Ito and H. Tominaga, "A Method of Identify Verification by Temporal Facial Expression", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J78-D-II, No.9, pp.1315-1324, Sept. 1995. (in Japanese)

[51] Y. Tanaka, Y. Mekada, M. Kasuga and H. Yoshikawa, "A Feature based on effective divisions for Personal Identification", Technical report of Institute of Electronics, Information and Communication Engineers, HCS99-50, pp.69-74, Oct. 1999. (in Japanese)

[52] R. Sukthankar and R. Stockton, "ARGUS: An Automated Multi-Agent Visitor Identification System", Proceedings of The 1999 National Conference on Artificial Intelligence (AAAI'99), pp.208-213, July 1999.

[53] S. Akamatsu, "Computational Studies on KANSEI Information Conveyed by Human Face", Journal of Japan Society for Fuzzy Theory and Systems, Vol.10, No.2, pp.225-235, April 1998. (in Japanese)

[54] H. Yamada, "Research on the Recognition of Facial Expressions of Emotion and the Image Processing Technology", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-140, pp.49-56, Nov. 1999. (in Japanese)

[55] A. Morisak and T. Inui, "Representation of Faces in Similarity Judgment", Technical report of Institute of Electronics, Information and Communication Engineers, PRU95-87, pp.37-42, July. 1995. (in Japanese)

[56] T. Takasaki, M. Yamaguchi, S. Yoshikawa and S. Akamatsu, "Validity of gender classification and gender judgment of faces - Comparison between faces of different age groups-", Technical report of Institute of Electronics, Information and Communication Engineers, HIP95-34, pp.49-54, Jan. 1996. (in Japanese)

[57] M. Nomura and K. Kakehi, "The effects of partial and configural information on recognition of a false smile", Technical report of Institute of Electronics, Information and Communication Engineers, HCS98-45, pp.43-50, Jan. 1999. (in Japanese)

[58] M. Nishitani, S. Yoshikawa and S. Akamatsu, "Characteristics of average faces", Technical report of Institute of Electronics, Information and Communication Engineers, HCS98-42, pp.23-30, Jan. 1999. (in Japanese)

[59] T. Kato, D. Ama, K. Morioka and S. Akamatsu, "Is facial attractiveness determined by any particular facial part?", Technical re-

port of Institute of Electronics, Information and Communication Engineers, HIP97-53, pp.17-22, March 1998. (in Japanese)

[60] M. Yamaguchi, T. Kato and S. Akamatsu, "Relationship Between Physical Traits and Subjective Impression of the Face -Age and Sex Information", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J79-A, No.2, pp.279-287, Feb. 1996. (in Japanese)

[61] T. Tsujii, "Configural and surface-based processing in face recognition", Technical report of Institute of Electronics, Information and Communication Engineers, HCS98-32, pp.1-8, Oct. 1998. (in Japanese)

[62] R. Nagayama, H. Yoshida, M. Miyatani and T. Toshima, "The study of relationship between facial expression analysis and person identity processing in face recognition: Analysis using spatial filtering and priming paradigm", Technical report of Institute of Electronics, Information and Communication Engineers, HCS97-23, pp.9-16, Jan. 1998. (in Japanese)

[63] T. Kato and S. Akamatsu, "Multimodal Characteristics of Facial-Expression Judgment", Technical report of Institute of Electronics, Information and Communication Engineers, HIP98-24, pp.17-22, Sept. 1998. (in Japanese)

[64] M. Miyoshi, "The event-related brain potential study of face recognition processes", Technical report of Institute of Electronics, Information and Communication Engineers, HCS98-35, pp.33-39, Oct. 1998. (in Japanese)

[65] T. Kato, M. Nishimura, Y. Matsubara, M. Samukawa, K. Hattori and S. Akamatsu, "Feature Saliency in Facial-Expression Judgment", Technical report of Institute of Electronics, Information and Communication Engineers, HCS98-3, pp.13-18, June 1998. (in Japanese)

[66] S. Shibui, H. Yamada, Y. Itsukushima and T. Sato, "Categorization Methods of Facial Expressions", Technical report of Institute of Electronics, Information and Communication Engineers, HCS97-22, pp.1-8, Jan. 1998. (in Japanese)

[67] M. Oda, M. Yamaguchi, T. Ogawa and S. Akamatsu, "Facial expression recognition and synthesis based on photographs", Technical

report of Institute of Electronics, Information and Communication Engineers, HIP99-35, pp.13-18, Nov. 1999. (in Japanese)

[68] N. Wada and K. Maiya, "Cognitive structure of facial expressions of emotions in Japanese: Re-experiment with Japanese facial expressions", Technical report of Institute of Electronics, Information and Communication Engineers, HCS98-44, pp.39-42, Jan. 1999. (in Japanese)

[69] S. Yoshikawa and A. Morisaki, "Sending emotional message by facial expressions", Technical report of Institute of Electronics, Information and Communication Engineers, HCS98-43, pp.31-38, Jan. 1999. (in Japanese)

[70] T. Kato, M. Ishikawa and S. Akamatsu, "The Mere Exposure Effect for Facial Expression", Technical report of Institute of Electronics, Information and Communication Engineers, HCS99-13, pp.7-12, June 1999. (in Japanese)

[71] N. Wada and K. Maiya, "A Study on Japanese Facial Expressions: On the Influence of Observers and Monitor TV upon Expressions of Emotions", Technical report of Institute of Electronics, Information and Communication Engineers, HCS99-15, pp.17-20, June 1999. (in Japanese)

[72] A. Miura and K. Maiya, "A Study on Japanese Facial Expressions: An Experiment on the Cognition of Facial Expressions in Western Movies", Technical report of Institute of Electronics, Information and Communication Engineers, HCS99-14, pp.13-16, June 1999. (in Japanese)

[73] M. Yamaguchi, M. Oda, N. Ogawa and S. Akamatsu, "Psychological judgment on composed facial expressions by PCA", Technical report of Institute of Electronics, Information and Communication Engineers, HIP98-53, pp.9-15, March 1999. (in Japanese)

[74] H. Umemura, H. Watanabe and K. Matsuoka, "An investigation on recognition of facial expressions using 3-D animation stimuli", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-179, pp.15-20, Dec. 1999. (in Japanese)

[75] M. Kosugi, K. Ohnuki and H. Ashiya, "Retrieval of Similar Faces by Using Coarse and Fine Gray-Level Mosaics", Journal of Institute of

Electronics, Information and Communication Engineers, Vol.J80-D-II, No.8, pp.2218-2224, Aug. 1997. (in Japanese)

[76] M. Sakauchi, "Image Retrieval Technology", Journal of Institute of Electronics, Information and Communication Engineers, Vol. 71, No.9, pp.911-914, Sept. 1988. (in Japanese)

[77] M. Oda, "An Image Retrieval System that Uses Human Cognitive Properties on Image Concept Formation", Journal of Information Processing Society of Japan, Vol.35, No.7, pp.1449-1456, July 1994. (in Japanese)

[78] M. Oda and T. Kato, "An Efficient Mechanism for the Face Image Retrieval System", Proceedings of 9th Symposium of Human Interface, pp.83-90, Oct. 1993. (in Japanese)

[79] M. Oda, S. Akamatsu and H. Fukamachi, "Adaptability of K-L Expansion Technique for Ambiguous Face Image Retrieval", Journal of Institute of Electronics, Information and Communication Engineers, Vol. J79-A, No.2, pp.288-297, Feb. 1996. (in Japanese)

[80] M. Yamashiro, S. Endo, K. Yamada, A. Ohuchi and M. Oda, "Face image retrieval method based on extended C.4.5", Journal of Institute of Electronics, Information and Communication Engineers, Vol. J79-D-II, No.12, pp.1-3, Dec. 1996. (in Japanese)

[81] M. Oda and T. Kato, "The Role of Facial Features in the Retrieval of Facial Image", Technical report of The Institute of Image Information and Television Engineers, Vol.17, No.38, pp.37-42, June 1993. (in Japanese)

[82] T. Kato and M. Oda, "Indirect Measurement of Feature Saliency in Face Processing", Journal of The Institute of Image Information and Television Engineers, Vol. 49, No.8, pp.1068-1077, Aug. 1995.

[83] M. Oda and T. Kato, "What kinds of Facial Features Are Used in Face Retrieval?", Proceedings of IEEE International Workshop on Robot and Human Communication, pp.265-270, July 1993.

[84] M. Oda and T. Kato, "AN EVALUATION OF FEATURE SALIENCY OF FACES FOR IMPROVING THE AMBIGUOUS-FACE RETRIEVAL SYSTEM", Proceedings of 15th International Symposium on Human Factors in Telecommunications, pp.415-417, March 1995.

[85] M. Oda, "Are ambiguous images uncertain?", Technical report of Institute of Electronics, Information and Communication Engineers, HIP96-10, pp.53-58, June 1996. (in Japanese)

[86] M. Oda, S. Akamatsu and H. Fukamachi, "Similarity judgment of a face by K-L expansion technique based on a comparison to human judgment", Proceedings of JAPAN-ISRAEL international Workshop on Computer Vision and Visual Communication, pp.25-30, Oct. 1994.

[87] M. Oda, "Context Dependency Effect in the Formation Image Concepts and its Application", Proceedings of 1991 IEEE International Conference on Systems, Man and Cybernetics, pp.1673-1678, 1991.

[88] G. Taguchi, "Diagnosis of Signal-to-Noise Ratio", Journal of Quality Engineering Society, Vol.2, No.4, pp.2-7, Aug. 1994. (in Japanese)

[89] G. Taguchi, "Diagnosis of Signal-to-Noise Ratio (2) Selection of Measurements", Journal of Quality Engineering Society, Vol.2, No.5, pp.2-5, Oct. 1994. (in Japanese)

[90] G. Taguchi, "Multi-Variate Sensing System Design", Journal of Quality Engineering Society, Vol.3, No.1, pp.2-7, Feb. 1995. (in Japanese)

[91] G. Taguchi, "Mathematics for Quality Engineering -Quality Engineering as a general measure of functionality", Journal of Quality Engineering Society, Vol.6, No.6, pp.5-10, Dec. 1998. (in Japanese)

[92] M. Nakatsugawa, D. Jin, S. Teshima and A. Ohuchi, "An Application of MTS for Detect Inspection of Circuit Pattern Model", Proceedings of International Workshop on Soft Computing in Industry, pp.192-197, June 1999.

[93] M. Nakatsugawa, M. Yamamoto, K. Suzuki, A. Ohuchi, D. Jin and S. Teshima, "Prediction of Kan-sei Model using MTS", Proceedings of The Australia Society for Operations Research Inc. Queensland Branch and ORSJ Hokkaido Chapter Joint Workshop, pp.914-926, July 1999.

[94] H. Asabara, Y. Ota, S. Takeda, M. Marumoto and T. Sato, "Pattern Recognition of Infrared Absorption Spectrum Analysis Using Ma-

halanobis Distance", Journal of Quality Engineering Society, Vol.7, No.2, pp.53-59, April 1999. (in Japanese)

[95] T. Kamoshita, K. Tabata, H. Okano, K. Takahashi and H. Yano, "Optimization of a Multi-dimensional Information System Using Mahalanobis Distance - The Case of Fire Alarm System-", Journal of Quality Engineering Society, Vol.4, No.3, pp.54-68, June 1996. (in Japanese)

[96] T. Kamoshita, "Functionary Evaluation of Fire Alarm System", Journal of Quality Engineering Society, Vol.7, No.3, pp.87-92, June 1999. (in Japanese)

[97] K. Mizoguchi, A. Kawai, K. Hamada and T. Kamoshita, "A Research on A Sensing System Using Mahalanobis Distance for Prevention of Driving Accidents", Journal of Quality Engineering Society, Vol.7, No.3, pp.79-86, June 1999. (in Japanese)

[98] S. Sakano, "A Study of Bills Recognition using MTS", Journal of Quality Engineering Society, Vol.8, No.3, pp.58-64, June 2000. (in Japanese)

[99] R. Matsuda, Y. Ikeda, T. Kamoshita and K. Tobara, "Application of Mahalanobis Taguchi System to the Fault Diagnosis Program - A Software Study for the Future Spacecraft (Part I)-", Journal of Quality Engineering Society, Vol.7, No.5, pp.55-61, Oct. 1999. (in Japanese)

[100] R. Matsuda, Y. Ikeda, T. Kamoshita and K. Tobara, "Application of Mahalanobis Taguchi System to Sensor Fault Detection -A Software Study for the Future Spacecraft (Part II)-", Journal of Quality Engineering Society, Vol.8, No.1, pp.65-69, Feb. 2000. (in Japanese)

[101] R. Matsuda, Y. Ikeda, T. Kamoshita and K. Tobara, "Application of Orthogonalization Technique to the Fault Diagnosis Program -A Software Study for the Future Spacecraft (Part III)-", Journal of Quality Engineering Society, Vol.8, No.2, pp.61-67, April 2000. (in Japanese)

[102] S. Teshima and H. Kubo, "A Study on Pattern Recognition using Mahalanobis Space", Proceedings of The Australia Society for Operations Research Inc. Queensland Branch and ORSJ Hokkaido Chapter Joint Workshop, pp.1269-1279, July 1999.

[103] S. Teshima, D. Jin, T. Bando and H. Toyoshima and H. Kubo, "Development of Printed Letter Inspection Technique Using Mahalanobis-Taguchi System", Journal of Quality Engineering Society, Vol.8, No.2, pp.68-74, April 2000. (in Japanese)

[104] S. Teshima, T. Bando and D. Jin, "A Research of Defects Detection Using the Mahalanobis-Taguchi System Method", Journal of Quality Engineering Society, Vol.6, No.6, pp.74-81, Oct. 1997.

[105] T. Suzuki, K. Takada, M. Takabashi and H. Yano, "A Technique for the evaluation of programming ability based on MTS", Journal of Quality Engineering Society, Vol.8, No.4, pp.57-64, Aug. 2000. (in Japanese)

[106] H. Shibano and H. Yasunaga, "Application of Mahalanobis Taguchi System for Corporate Performance Diagnosis", Journal of Quality Engineering Society, Vol.7, No.4, pp.41-48, Aug. 1999. (in Japanese)

[107] G. Taguchi, "Application of Mahalanobis Distance for Medical Treatment", Journal of Quality Engineering Society, Vol.2, No.6, pp.2-6, Dec. 1994. (in Japanese)

[108] K. Ikeda, "Functionality Evaluation of Wearing Comfort", Journal of Quality Engineering Society, Vol.7, No.5, pp.73-76, Oct. 1999. (in Japanese)

[109] R. Hasegawa, I. Kabe and K. Noji, "Application of MTS on the Reliability Improvement of Final Diagnosis in Health Examination", Journal of Quality Engineering Society, Vol.7, No.2, pp.67-74, April 1999. (in Japanese)

[110] R. Hasegawa, I. Kabe, K. Noji and T. Kanetaka, "Relativity Improvement and Economic Evaluation of a Health Examination Using Schmidt Orthogonalization Method", Journal of Quality Engineering Society, Vol.8, No.1, pp.45-50, Feb. 2000. (in Japanese)

[111] S. Tanaka, M. Yamamoto, K. Suzuki and A. Ohuchi, "Application of Mahalanobis-Taguchi System to the Prediction of the Prevalence of Infectious Diseases", Proceedings of 17th Annual ASI's Taguchi Methods Symposium, pp.143-157, Oct. 1999.

[112] G. Taguchi, "Weather Forecasting System with Mahalanobis Distance (1) Short Term Forecasting", Journal of Quality Engineering Society, Vol.3, No.6, pp.3-7, Dec. 1995. (in Japanese)

[113] G. Taguchi, "Weather Forecasting and Quality Engineering The case of a long-term prediction", Journal of Quality Engineering Society, Vol.4, No.1, pp.2-6, Feb. 1996. (in Japanese)

[114] G. Taguchi, "Earthquake Prediction and Quality Engineering", Journal of Quality Engineering Society, Vol.3, No.3, pp.2-6, June 1995. (in Japanese)

[115] H. Nakajima, K. Takada, H. Yano, Y. Shibamoto, I. Takagi, M. Yamauchi and G. Toda, "Forecasting of the Future Health From Existing Medical Examination Results Using Mahalanobis Taguchi System", Journal of Quality Engineering Society, Vol.7, No.4, pp.49-57, Oct. 1999. (in Japanese)

[116] G. Taguchi, "Quality Engineering for Pattern Recognition", Journal of Quality Engineering Society, Vol.3, No.2, pp.2-5, April 1995. (in Japanese)

[117] G. Taguchi, "Pattern Recognition and Quality Engineering", Journal of Quality Engineering Society, Vol.3, No.4, pp.2-5, Aug. 1995. (in Japanese)

[118] G. Taguchi, "Discussions on Voice Pattern Recognition Systems", Journal of Quality Engineering Society, Vol.3, No.5, pp.3-7, Oct. 1995. (in Japanese)

[119] D. D. Frey, "Application of Wavelets and Mahalanobis Distances to Robust Design of An Image Classification System", Proceedings of 17th Annual ASI's Taguchi Methods Symposium, pp.198-208, Oct. 1999.

[120] G. Taguchi and R. Jugulum, "Role of S/N ratios in Multivariate Diagnosis", Journal of Quality Engineering Society, Vol.7, No.6, pp.63-69, Dec. 1999.

[121] G. Taguchi and R. Jugulum, "Quality Loss Function in Multivariate Diagnosis", Journal of Quality Engineering Society, Vol.8, No.4, pp.47-52, Aug. 2000.

[122] M. Nakatsugawa and A. Ohuchi, "An Evaluation of the Simple Effect's Dispersion for the Term-optimization in MTS", Proceedings of The 2nd Joint International Workshop ORSJ-The Operations

Research Society of Japan, Hokkaido Chapter and ASOR- The Australian Society for Operations Research Inc., Queensland Branch, pp.293-300, June 2000.

[123] M. Nakatsugawa, M. Yamamoto, K. Suzuki and A. Ohuchi, "A Study on The Construction of Effective Mahalanobis Space", Proceedings of 17th Annual ASI's Taguchi Methods Symposium, pp.110-123, Oct. 1999.

[124] G. Taguchi, "Application of Schmit's Orthogonalization in Mahalanobis Space", Journal of Quality Engineering Society, Vol.5, No.6, pp.11-15, Dec. 1997. (in Japanese)

[125] R. Jugulum, S. Taguchi and K. Yang, "New Developments in Multivariate Diagnosis: Comparison between Two Methods", Journal of Quality Engineering Society, Vol.7, No.5, pp.62-68, Oct. 1999.

[126] R. Jugulum, "Some Interesting Thoughts in Mautivariate Diagnosis", Proceedings of Quality Engineering Symposium, pp.65-68, June 1999.

[127] H. Takagi, T. Unemi and T. Terano, "Perspective on Interactive Evolutionary Computing", Journal of Japanese Society for Artificial Intelligence, Vol.13, No.5, pp.693-703, Sept. 1998. (in Japanese)

[128] H. Takagi and K. Aoki, "Interactive Evolutionary Computing from creativity support to engineering applications", Proceedings of Interactive Evolutionary Computing Workshop, pp.1-6, March. 1998. (in Japanese)

[129] R. Dawkins, Y. Nakajima, A. Endo, T. Endo and T. Hikida (trans.), "The Blind Watchmaker", HAYAKAWA-SHOBOU, Tokyo, 1993. (in Japanese)

[130] T. Unemi, "Genetic Algorithms and Computer Graphics Arts", Journal of Japanese Society for Artificial Intelligence, Vol.9, No.4, pp.518-523, July 1994. (in Japanese)

[131] K. Aoki and H. Takagi, "Interactive GA-Based Design Support System for Lighting Design in 3-D Computer Graphics", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J81-D-II, No.7, pp.1601-1608, July 1998. (in Japanese)

[132] Y. Nakanishi, "Control and Evaluation of Interactive Evolutionary System with a Preference Function: Applying Genetic Programming to Design Aid System", Journal of Japanese Society for Artificial Intelligence, Vol.13, No.5, pp.704-711, Sept. 1998. (in Japanese)

[133] H. Kim and S. Cho, "Development of an IGA-based Fashion Design Aid System with Domain Specific Knowledge", Proceedings of 1999 IEEE International Conference on Systems, Man, and Cybernetics, Vol.III, pp.663-668, Oct. 1999.

[134] J. A. Biles, "Life with GenJam", Proceedings of 1999 IEEE International Conference on Systems, Man, and Cybernetics, Vol.III, pp.652-656, Oct. 1999.

[135] T. Mutoh, N. Komagata and K. Ueda, "An Experimental Study for Automatically Generating Image Filter Sequence by Using Simulated Breeding", Proceedings of Interactive Evolutionary Computing Workshop, pp.7-12, March. 1998. (in Japanese)

[136] H. Takagi and M. Ohsaki, "IEC-based Hearing Aid Fitting", Proceedings of 1999 IEEE International Conference on Systems, Man, and Cybernetics, Vol.III, pp.657-662, Oct. 1999.

[137] S. Cho, "Applying Interactive Genetic Algorithm to Content-based Image Retrieval: A Preliminary Result", Proceedings of Interactive Evolutionary Computing Workshop, pp.19-24, March. 1998. (in Japanese)

[138] A. Kitamoto and M. Takagi, "Interactive Image Browsing Using Pipeline-type Genetic Algorithm", Proceedings of Interactive Evolutionary Computing Workshop, pp.31-36, March. 1998. (in Japanese)

[139] A. Kitamoto and M. Takagi, "Interactive Image Browsing Using Queue-Based Genetic Algorithm", Journal of Japanese Society for Artificial Intelligence, Vol.13, No.5, pp.728-738, Sept. 1998. (in Japanese)

[140] C. Caldwell and V. S. Johnston, "Tracking a Criminal Suspect through Face-Space with a Genetic Algorithm", Proceedings of International Conference on Genetic Algorithms, pp.416-421, 1991.

[141] K. Isono, M. Oda and S. Akamatsu, "A Facial Image Generating System Based on a Genetic Algorithm", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J82-D-II, No.3, pp.483-493, March 1999. (in Japanese)

[142] N. Monmarche, G. Nocent, M. Slimane, G. Venturini and P. Santini, "Imagine: a tool for generating HTML style sheets with an interactive genetic algorithm based on genes frequencies", Proceedings of 1999 IEEE International Conference on Systems, Man, and Cybernetics, Vol.III, pp.640-645, Oct. 1999.

[143] K. Kuriyama, T. Terano and M. Numao, "Authoring Support by Interactive Genetic Algorithm and Case Based Reasoning", Proceedings of Interactive Evolutionary Computing Workshop, pp.13-18, March. 1998. (in Japanese)

[144] L. Pagliarini, H. H. Lund, O. Miglino and D. Parisi, "Artificial Life: A New Way to Build Educational and Therapeutic Games.", ALIFE-V, pp.152-156, May 1996.

[145] M. Fujii and T. Furuhashi, "A Study of Interaction between Humans and Systems Using Fuzzy Classifier System", Proceedings of Interactive Evolutionary Computing Workshop, pp.37-41, March. 1998. (in Japanese)

[146] M. Tabuchi and T. Taura, "Methodology for Interactive Knowledge Acquisition between Genetic Learning Engine and Human", Journal of Japanese Society for Artificial Intelligence, Vol.11, No.4, pp.600-607, July 1996. (in Japanese)

[147] M. Ohsaki and H. Takagi, "Reduction of the Fatigue of Human Interactive EC Operators -Improvement of Present Interface by Prediction of Evaluation Order-", Journal of Japanese Society for Artificial Intelligence, Vol.13, No.5, pp.712-719, Sept. 1998. (in Japanese)

[148] H. Takagi, M. Ohsaki and T. Ingu, "The Methods to Reduce of Human Interactive EC Operators", Proceedings of Interactive Evolutionary Computing Workshop, pp.47-52, March. 1998. (in Japanese)

[149] T. Unemi, "A User Interface for Simulated Breeding using Multifield", Proceedings of Interactive Evolutionary Computing Workshop, pp.42-45, March. 1998. (in Japanese)

[150] F. Sugimoto, K. Nishio and N. Honda, "A Method of Fitness Assignment in Interactive Genetic Algorithm Using Fuzzy Reasoning Based on Psychological Measure", Journal of Japanese Society for Artificial Intelligence, Vol.13, No.5, pp.739-745, Sept. 1998. (in Japanese)

[151] S. Kato and S. Iisaku, "A Proposal of How to Make Good Use of the Personal History Information Reflecting User Subjectivity in An Image Retrieval Technique using GAs", Technical report of The Institute of Electronics, Information and Communication Engineers, AI97-10, pp.71-78, May 1997. (in Japanese)

[152] M. Shibuya, H. Kita and S. Kobayashi, "Integration of Multiobjective Genetic Algorithm and its Application to Animation to Animation Design", Proceedings of 1999 IEEE International Conference on Systems, Man, and Cybernetics, Vol.III, pp.646-651, Oct. 1999.

Chapter 2

Facial Expression Recognition based on Mahalanobis-Taguchi System

2.1 Introduction

Facial expression of human is a medium to communicate with other person smoothly and an indicator to represent psychological state [1][2]. Human adaptably changes own behavior by inferring other person's psychological state from facial expression. In addition, human sends own information to other person by using facial expression. It is obvious that flexible and adaptable computer environment can be implemented if facial expressions can be used in the communication between computer and human.

From such a background, the research to implement effective communication between computer and human by providing computer with the function which can recognize facial expression of human has been worldwide carried out [3][4][5]. However, the high recognition performance in facial expression recognition is not achieved because ambiguity and personal equation have to be considered in facial expression recognition. Therefore, it is required to establish effective method for facial expression recognition.

On the other hand, human has the ability which can flexibly recognize facial expression. For this fact, the research that clarifies human's cognitive structure for facial expression is carried out in order to develop effective facial expression recognition method [6][7]. As indicated by this research, it is quite likely that an effective method for facial expression recognition can be implemented if the excellent ability of human for facial expression recognition can be reflected to the recognition on computer.

Recently, the Mahalanobis-Taguchi System (MTS) has been proposed by Taguchi [8]. The MTS is one of statistical pattern recognition methods

developed to implement the pattern recognition which can reflect human's sensuous knowledge for object recognition. The MTS has the following features to implement the pattern recognition which can reflect human's sensuous knowledge. The multi-dimensional training data for the pattern recognition are collected by a human. The correlations among the attributes which are common in a set of the collected multi-dimensional training data are automatically extracted. The Mahalanobis space which can consider the extracted correlations is constructed. The pattern recognition of unknown datum is performed by using the Mahalanobis distance which is one-dimensional evaluation criterion calculated on the basis of the Mahalanobis space. That is, the pattern recognition which can reflect human's sensuous knowledge is implemented by combining the pattern recognition on computer with the sensuous object recognition which is performed by a human. Moreover, in the MTS, the effective attributes for the recognition can be re-extracted by using recognition result. Namely, the optimization of attributes can be performed. For this effectiveness, recently, the MTS has been applied to handwritten word recognition and voice recognition on unspecified user [9][10][11].

In this chapter, I propose a facial expression recognition based on the MTS. Especially, I treat the facial expression recognition in the stationary facial image. The facial expression is not expressed by one internal feature on face such as eyes, nose and mouth but is expressed by the relation of some internal features. It is likely that the MTS is effective for the facial expression recognition because the MTS can consider the correlation among the attributes of the multi-dimensional data. In a facial expression, there exists difference on each person. However, it is reported that there exists a common facial expression pattern on a facial expression. The MTS can perform the recognition considering this common facial expression pattern because the MTS can automatically find the correlations among the attributes which are common in a set of multi-dimensional data. Moreover, this facial expression pattern can be confirmed because the MTS can re-extract effective attributes for the recognition. The optimization of attributes can also reduce the computing cost.

In the conventional MTS, the optimization of the attributes using the orthogonal array is performed. The orthogonal array is the combinations of attributes to decrease the number of experiments. The optimization based on main effects, i.e., the average of simple effects, is performed when the optimization is performed by using orthogonal array. That

is, the optimization which does not consider the relation among the attributes in the multi-dimensional data is performed. To solve this problem, in this chapter, I also propose an optimization method based on the Genetic Algorithm. Finally, I confirm the effectiveness of the proposed facial expression recognition and optimization method through practical experiment using some subjects.

This chapter is organized as follows: In Section 2, the facial image and the image processing used in my study will be explained. In Section 3, I will propose a facial expression recognition based on the MTS. I will describe the experimental method using some subjects in Section 4. The experimental result will be mentioned in Section 5. In Section 6, I will propose an optimization method based on the Genetic Algorithm. Finally, I will finish this chapter with the conclusions.

2.2 Facial Image Processing

2.2.1 Facial Image

In this study, I employ the stationary facial image as the recognition target. The studies on the facial expression recognition using the facial image sequence are carried out [12][13]. However, I use the stationary facial image because a large amount of image processing time and the special device are needed to use the facial image sequence. The facial image used in this study is the frontal facial image taken by a monochrome camera. The detail of facial image is as follows.

- The size of facial image is 300 $\times$ 300 pixels.
- The gradation of pixel is 256.
- The background color is white.
- It does not include hair and glasses.

Figure 2.1 shows an example of facial image used in my study. The facial images are taken in my laboratory with special lighting device. Moreover, the facial images are taken by using the baselines for the position of eyes, nose and mouth (white lines in Figure 2.1). These facial images include a little fluctuation of face size and position. The MTS can implement robust pattern recognition by giving noise to the training data because the pattern recognition based on the MTS is performed on the basis of the correlations which are common in a set of multi-dimensional training

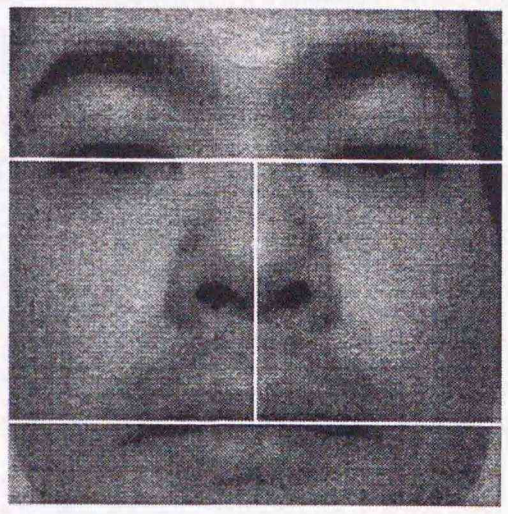


Figure 2.1: An example of facial image

data. Thus, I do not carry out the perfect normalization of the size and position of facial image in order to implement the robustness against the fluctuations on the size and position.

The facial expressions treated in this study are 7 facial expressions; anger, disgust, fear, happiness, neutral, sadness and surprise. It has been reported that each of six facial expressions (anger, disgust, fear, happiness, sadness and surprise) has a common facial expression pattern. In addition, these are used frequently in the facial expression recognition.

I employ the Facial Action Cording System (FACS), which is used in the psychological field, when I take the facial image on each facial expression [14]. The FACS has been proposed by Ekman. The FACS is a system to describe facial expression by combining the movement of muscles on face. In the FACS, the muscles on the surface of the face are divided into 44 parts called Action Unit (AU). An arbitrary facial expression is described by combining some AUs. I use it to express the facial expressions but do not treat it as the attribute for the recognition, since facial expression in the FACS is expressed based on the amount of movement of muscle, i.e., AU. The FACS can make it easy to express the facial expressions, and each person can express the similar facial expression on each facial expression. That is, the difference of facial expression on each person can be reduced.

2.2.2 Edge Extraction

It seems reasonable to suppose that the size and position of the internal features on the face are effective for the facial expression recognition. I extract the edge in the facial image to obtain the information of these sizes and positions.

I adopt the Global Range Edge Detection Method in the edge extraction [15]. This method decides the edge in the image data on the basis of the deviation of gradation value in pixel. The edge extraction using the Global Range Edge Detection Method is performed by the following formulas.

$$\partial_{ij}^{1} = \frac{1}{|I_{ij}|}\frac{1}{3}\{(p_{i-1,j-1} + p_{i,j-1} + p_{i+1,j-1}) - (p_{i-1,j+1} + p_{i,j+1} + p_{i+1,j+1})\} \tag{2.1}$$

$$\partial_{ij}^{2} = \frac{1}{|I_{ij}|}\frac{1}{3}\{(p_{i-1,j-1} + p_{i-1,j} + p_{i-1,j+1}) - (p_{i+1,j-1} + p_{i+1,j} + p_{i+1,j+1})\} \tag{2.2}$$

$$\partial_{ij}^{3} = \frac{1}{|I_{ij}|}\frac{1}{3}\{(p_{i-1,j-1} + p_{i,j-1} + p_{i-1,j}) - (p_{i+1,j} + p_{i,j+1} + p_{i+1,j+1})\} \tag{2.3}$$

$$\partial_{ij}^{4} = \frac{1}{|I_{ij}|}\frac{1}{3}\{(p_{i,j-1} + p_{i+1,j-1} + p_{i+1,j}) - (p_{i-1,j} + p_{i-1,j+1} + p_{i,j+1})\} \tag{2.4}$$

$$|I_{ij}| = \sqrt{\frac{1}{9}\sum_{r=i-1}^{i+1}\sum_{s=j-1}^{j+1} p_{r,s}^{2}} \tag{2.5}$$

$$|\partial_{ij}| = max(|\partial_{ij}^{1}|, \cdots, |\partial_{ij}^{4}|) \tag{2.6}$$

$$|\partial_{ij}| \geq \mu + \sigma \tag{2.7}$$

$$\mu = \frac{1}{MN}\frac{1}{4}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\sum_{4}^{k=1}|\partial_{ij}^{k}| \tag{2.8}$$

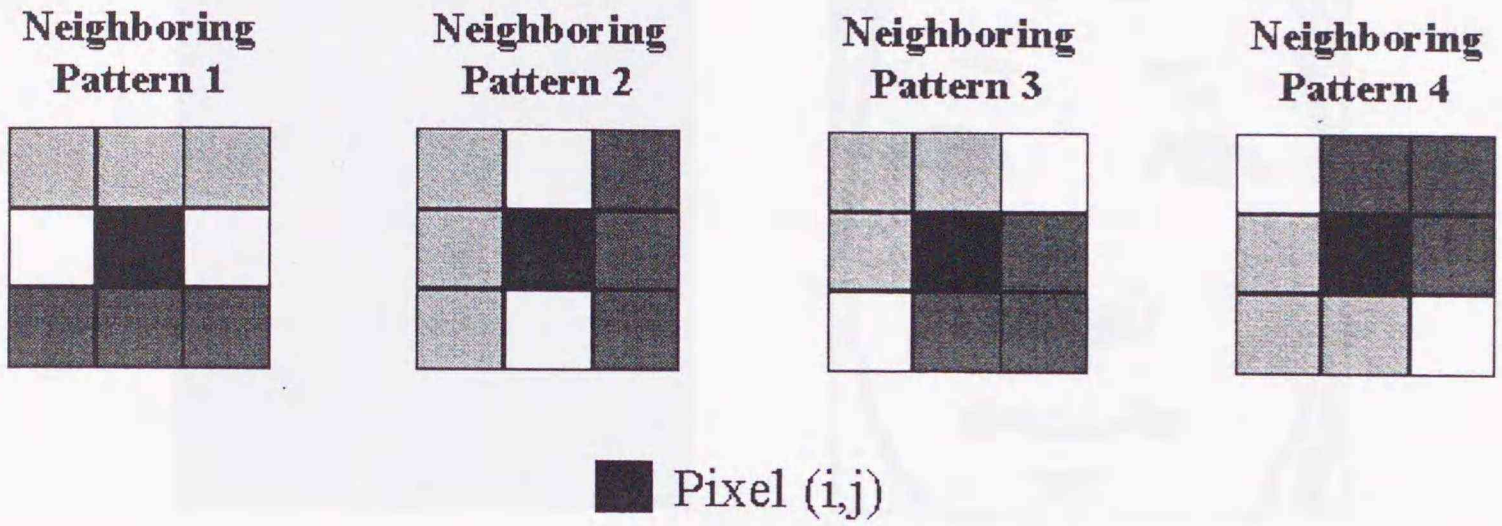


Figure 2.2: Neighboring patterns

$$\sigma = \sqrt{\frac{1}{MN}\frac{1}{4}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\sum_{k=1}^{4}(\partial_{ij}^{k} - \mu)^2} \tag{2.9}$$

$p_{i,j}$: The gradation value at pixel (i, j)

I_{ij} : The average of gradation value in the neighborhood at pixel (i, j)

∂_{ij} : The deviation of the gradation value in the neighboring pattern of pixel (i,j)

μ : The average of the deviation in the neighboring pattern at all pixels of the image

σ : The standard deviation of the deviation in the neighboring pattern at all pixels of the image

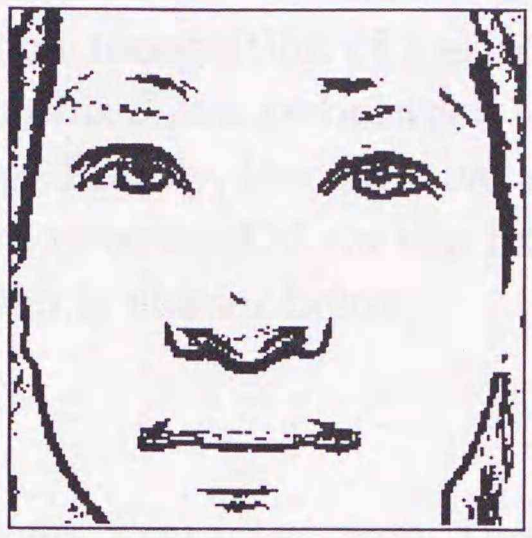
M : The width of the image

N : The height of the image

The Global Range Edge Detection Method investigates the deviations at four neighboring patterns of a certain pixel, which are described in formula (2.1) - (2.4) and are shown in Figure 2.2. If the maximum value of the deviation at a certain pixel exceeds the sum of the average and the standard deviation of the deviation in the neighboring pattern at all pixels of the image, this method judges the pixel as the edge. Figure 2.3 shows an example of the edge extracted facial image.



Original Facial Image



Edge Extracted Facial Image

Figure 2.3: An example of the edge extracted facial image

2.3 Facial Expression Recognition based on the MT-S

The detail of facial expression recognition based on the MTS is described below.

2.3.1 Recognition Process

The process of facial expression recognition based on the MTS is as follows:

1. Collection of the facial images on each facial expression
2. Edge extraction of the collected facial images
3. Extraction of attributes from the edge extracted facial images
4. Construction of the Mahalanobis space on each facial expression
5. Facial expression recognition of unknown facial images based on the Mahalanobis distance
6. Optimization of attributes based on the recognition result

As mentioned in Section 2, the facial images are collected on each facial expression. The edge extraction is performed on the collected facial images by using the Global Range Edge Detection Method. Then, the attributes are extracted from the edge extracted facial images. The

Mahalanobis space is constructed on each facial expression from these extracted attributes. The facial expression recognition of unknown facial image is performed by using the Mahalanobis distance between the Mahalanobis spaces and unknown facial image. Finally, the effective attributes for recognizing each facial expression are re-extracted on the basis of the recognition result. The detail of each step is shown below.

2.3.2 Extraction of Attribute

The attributes used for the recognition are extracted from the edge extracted facial images. In the case where binary image datum or wave form is employed as the recognition target, the attributes based on differential characteristic and integral characteristic are frequently utilized in the recognition based on the MTS. The differential characteristic indicates the changes of element in the data, similar to the concept of frequency. The integral characteristic indicates the distance of element in the data. It is reported that the effective pattern recognition can be implemented by using the attributes based on these characteristics [16]. The extraction methods of the attributes based on these characteristics for the facial image are shown below.

1. **Differential characteristic:** In the extraction of the attributes based on the differential characteristic, h equally spaced parallel lines are drawn on the facial image, which is shown in Figure 2.4. The number of changes in pixel value along each line is taken as the attribute based on the differential characteristic. If the facial image is a complete line-draw image as shown in Figure 2.4, the attribute equals the number of the intersecting points on each line. Figure 2.4 shows the extraction method of this attribute along the row of facial image. The attributes based on this characteristic are also calculated down the columns on the facial image. It is ideal that the lines are drawn on each row and column of pixels in the facial image. However, the number of lines can be modified in order to reduce the computational cost. The attributes based on the differential characteristic along the rows are calculated by the following formula.

$$DF_h = \frac{\sum_{j=1}^{M-1} c_{i,j}}{2} \tag{2.10}$$

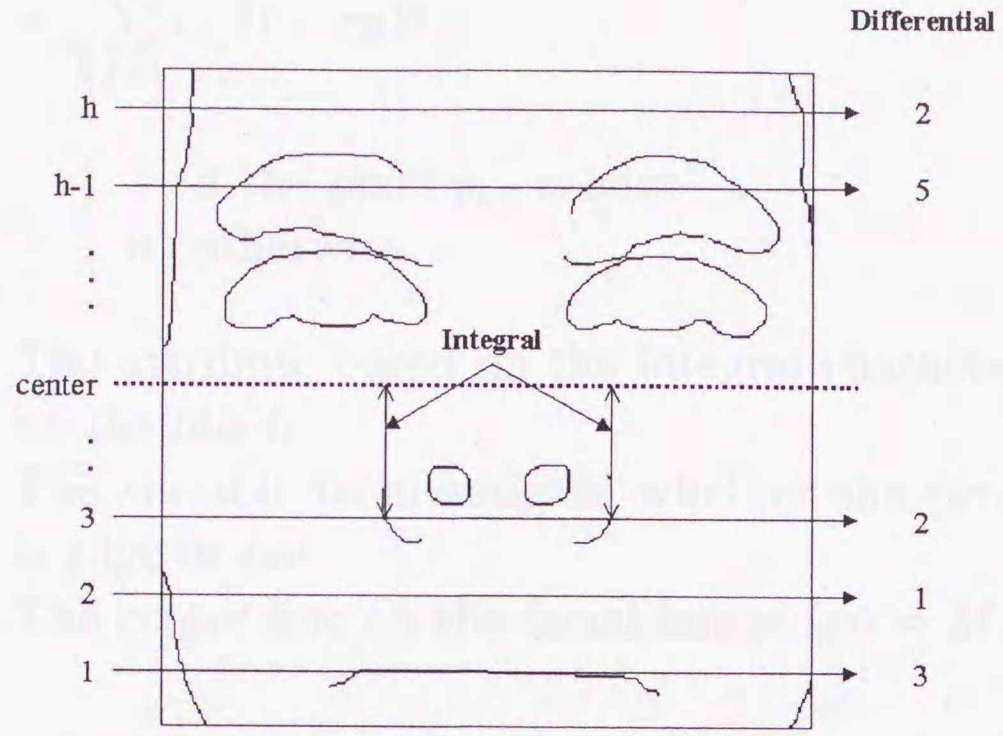


Figure 2.4: An example of the attributes based on the differential characteristic and the integral characteristic

$$c_{i,j} = \begin{cases} 1 & \text{if } p_{i,j} \neq p_{i,j+1} \\ 0 & \text{otherwise} \end{cases}$$

DF_h : The attribute based on the differential characteristic on the line h

$c_{i,j}$: The variable to investigate the change of the pixel value on the facial image.

$p_{i,j}$: The value of pixel (i, j) on the line h in the edge extracted facial image

M : The width of the facial image

2. **Integral characteristic:**

   As shown in Figure 2.4, h parallel lines and a center line are drawn for the extraction of the attributes based on the integral characteristic. The center line is a line to halve the facial image along the rows or down the columns. The Euclidean distances between the edges on line h and the center line are computed. The sum of these Euclidean distances is taken as the attribute based on the integral characteristic. The attributes based on the integral characteristic along the rows are calculated by using the following formula.

$$IN_h = \sqrt{\sum_{j=1}^{M} e_{i,j}(i - cn)^2} \tag{2.11}$$

$$e_{i,j} = \begin{cases} 1 & \text{if the pixel } p_{i,j} \text{ is edge} \\ 0 & \text{otherwise} \end{cases}$$

IN_h : The attribute based on the integral characteristic on the line h

$e_{i,j}$: The variable to investigate whether the pixel $p_{i,j}$ is edge or not

cn : The center line on the facial image ($cn = M/2$)

The attribute based on these characteristics are much easier to collect and deal with than KL expansion which is frequently used in face recognition [17]. Moreover, the amount of calculation is much smaller than KL expansion.

2.3.3 Construction of Mahalanobis Space

The Mahalanobis space is constructed from the attributes extracted with the differential and the integral characteristics. The Mahalanobis space is a space which can consider the correlations among the attributes in multi-dimensional data. The recognition based on the MTS is performed by judging whether unknown datum belongs to the Mahalanobis space or not. The Mahalanobis distance between the unknown datum and the Mahalanobis space, and a threshold are employed in order to make this judgment. The construction method of the Mahalanobis space from training data and the calculation method of the Mahalanobis distance between unknown datum and the Mahalanobis space are omitted here because they are described in Section 6 of Chapter 1.

The Mahalanobis distance takes a value close to 1 or smaller than 1, when the datum to recognize is very similar to the average of the training data which are used to construct the Mahalanobis space. On the other hand, its value rapidly increases when the datum differs from the average of the training data which are used to construct the Mahalanobis space. Generally, the probability of belonging to the Mahalanobis space is very small if this value exceeds 4.

As mentioned above, the MTS is a pattern recognition method to divide the data into two kinds which are the data which belong to the Mahalanobis space and the data which does not belong to the Mahalanobis

space. However, seven Mahalanobis spaces have to be treated in this study because there are 7 facial expressions. There is the possibility that the Mahalanobis space which the unknown datum belongs to can not be decided, if the threshold is employed in order to recognize the unknown datum. Therefore, the method based on minimum Mahalanobis distance is employed in this study instead of the threshold. Namely, The Mahalanobis distance between each Mahalanobis space and unknown facial image is calculated and then the unknown facial image is judged to belong to the Mahalanobis space which has shortest Mahalanobis distance.

2.4 Experimental Method

I designed the experiment to confirm the effectiveness of the proposed facial expression recognition. In this experiment, ten students were recruited from Hokkaido University. The subjects were divided into two groups. The facial images for the training data were taken from one group. The facial images taken from another group were used as the unknown facial images.

First of all, I asked the subjects to express a facial expression. Next, I showed the photograph on the facial expression which is inserted in the Ekman's literature to the subject. Following this, I explained the AUs and the change of face on the facial expression. After this explanation, the subject practiced expressing the facial expression with a mirror. The subject continued the practice until the subject could express the facial expression based on the FACS. I started to take the facial images on the facial expression after this practice. The subjects did not use a mirror in taking the facial images. However, I gave the explanation to the subjects again if the facial expression which the subject expressed considerably differed from the facial expression based on the FACS. I restarted to take the facial image after the subject practiced expressing the facial expression again.

I took 20 facial images from a subject on each facial expression. The reason why I took 20 facial images from a subject is that many training data are required for the construction of a Mahalanobis space. Moreover, the robustness against difference on facial expression can be implemented by taking various facial images on a facial expression. Therefore, a Mahalanobis space on each facial expression is constructed from 100 facial images. In addition, I used 700 facial images as the unknown facial images, which are taken from other 5 subjects. Figure 2.5 shows an ex-

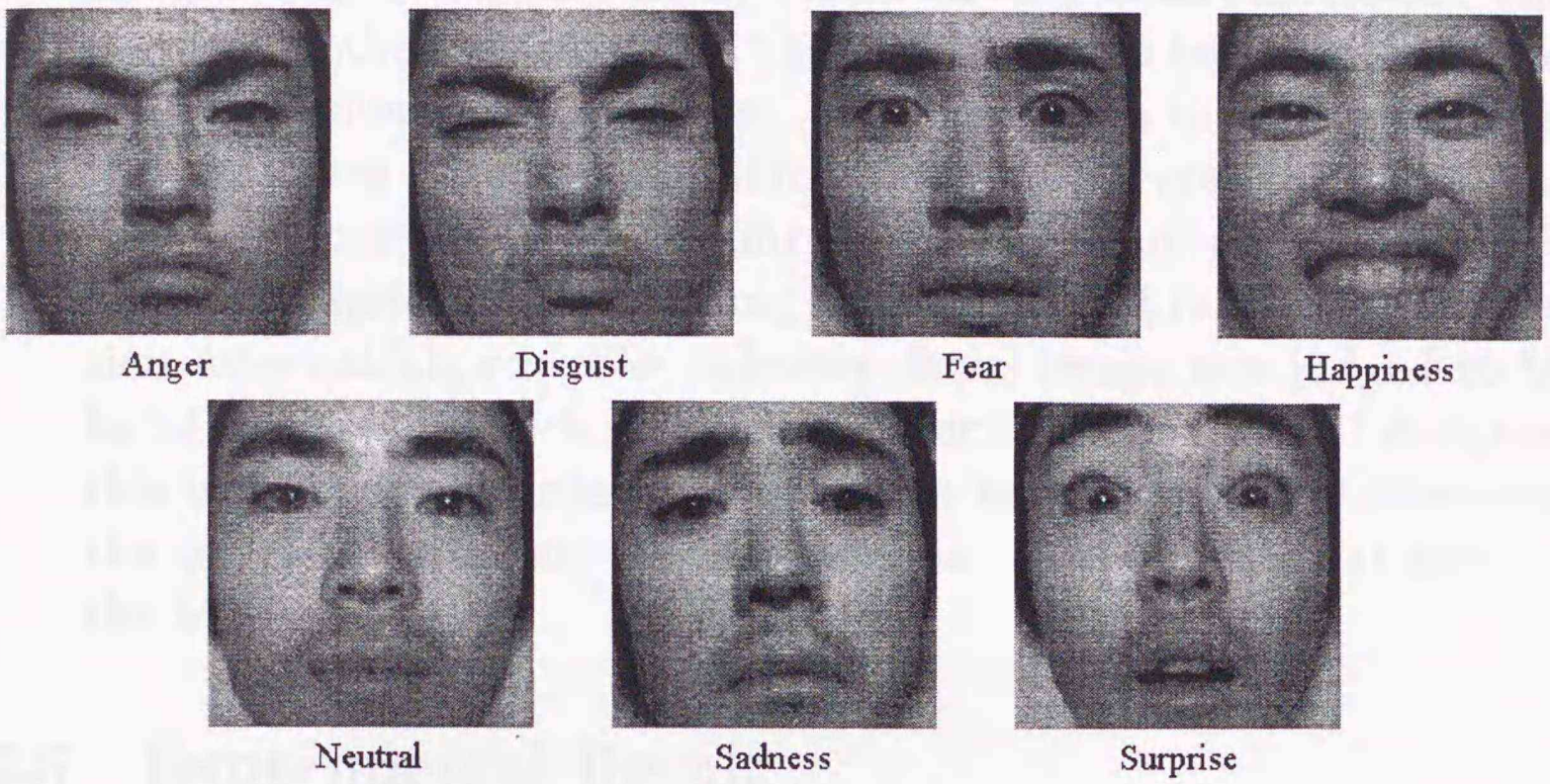


Figure 2.5: An example of the facial image on each facial expression

ample of the facial image used as the training facial image on each facial expression.

I set the lines at intervals of 12 pixels for the extraction of the attributes based on the differential and integral characteristics. The number of attributes along the rows was 48 and that down the columns was also 48. Therefore, I constructed the Mahalanobis space from 96 attributes.

Moreover, I designed two kinds of experiment in order to compare with the proposed facial expression recognition. These experiments are as follows.

1. **Facial expression recognition using Euclidean distance and Mosaic attribute (MOSAIC):** The facial images on each facial expression were converted into 100 x 100 mosaic facial images. The average of gradation value was calculated in each region of training facial images on each facial expression. The average of this average on each facial expression is calculated on each region. The Euclidean distance between the unknown datum and the average of each facial expression was calculated. The unknown datum was judged to belong to the facial expression which had the smallest Euclidean distance.

2. **Facial expression recognition with Euclidean distance and the attributes based on the differential and integral characteristics (D&I):** The attributes based on the differential and

integral characteristics were extracted on each facial expression. The averages of these attributes in the training facial images on each facial expression were calculated. The Euclidean distances between the attributes based on the differential and integral characteristics which are extracted from the unknown facial image and the averages of the attributes on the training facial images of each facial expression were calculated. The unknown facial image was judged to be facial expression which had shortest Euclidean distance. I designed this experiment in order to confirm the effectiveness of considering the correlations among the attributes in multi-dimensional data on the MTS.

2.5 Experimental Result

The experimental results are shown in Table 2.1 and Figure 2.6.

Table 2.1 indicates the rate which unknown facial images on each facial expression were correctly recognized. In MOSAIC, the recognition rate was low excluding Happiness and Surprise. The position and shape of the internal features on face considerably change in Happiness and Surprise. That is, it is easy to distinguish the facial expression of Happiness and Surprise from other facial expressions. Therefore, it is likely that the high recognition performance could be implemented in Happiness and Surprise of MOSAIC. However, the rate was low in Anger, Disgust, Fear, Neutral and Sadness. The shape and position of one or some internal features on face are similar to them on other facial expressions. Therefore, it might be difficult to recognize these facial expressions by using the Euclidean distance which could not consider the relations among the internal features on face. On the other hand, the results on MTS revealed that the proposed facial expression recognition could correctly recognize the facial images on 7 facial expressions. From the results of I&D, it seemed that the feature of considering the correlations among the attributes in multi-dimensional data was main cause of high recognition performance.

The rate of correct recognition in MOSAIC is higher than I&D. Therefore, it may be that the facial expression recognition based on the combination of the MTS and the Mosaic attribute is effective in comparison with the recognition based on the combination of the MTS and I&D. However, ten thousand attributes are required for the construction of a Mahalanobis space in the case of 100 x 100 Mosaic image. This number of attributes is too many to construct the Mahalanobis space. Thus, it is

	Anger	Disgust	Fear	Happiness	Neutral	Sadness	Surprise
MOSAIC	52	38	55	73	43	65	81
D&I	70	53	10	82	42	48	59
MTS	100	100	100	100	100	100	100

Table 2.1: Recognition result (%)

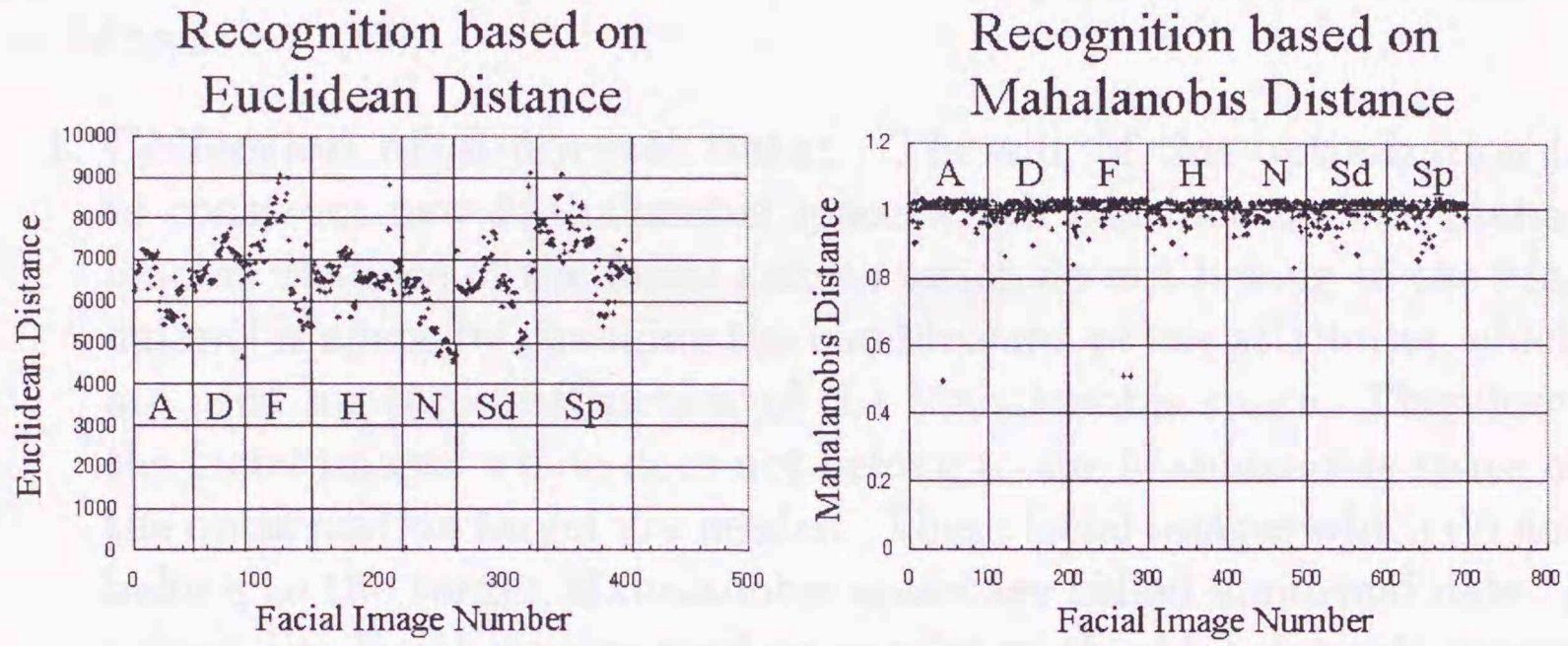


Figure 2.6: Mahalanobis and Euclidean distances of unknown facial images

likely that more effective facial expression recognition can be implemented by combining some Mosaic attributes with some attributes based on the differential and integral characteristics.

Moreover, Figure 2.6 shows the Mahalanobis distances and Euclidean distances of the unknown facial image which were correctly recognized in MTS and MOSAIC, respectively. In MOSAIC, the wide distribution of Euclidean distances on each facial expression was indicated. On the other hand, I found that the Mahalanobis distances of all unknown facial images were smaller than 1.1 in MTS. Judging from this result, it is likely that facial expression using threshold might be able to be implemented in the facial expression recognition based on the MTS.

2.6 Optimization of Attributes

I used 96 attributes for the recognition in the above experiment based on the MTS. In this Section, I attempt to re-extract the attributes which significantly contribute to recognition on each facial expression. This op-

timization can reduce the computational cost. In the conventional MTS, the optimization of attribute is performed by using orthogonal array, Signal/Noise ratio (SN ratio) and factorial effect. In the optimization using them, the effectiveness of attribute for the recognition is estimated based on the main effect; i.e., the average of simple effect. Thus, the relation among attributes is not considered in the estimation of the effectiveness. To solve this problem, I propose an optimization method using the Genetic Algorithm [18]. The optimization process of this method is as follows.

1. **Collection of abnormal data:** The aim of this optimization is to construct new Mahalanobis space which can increase the Mahalanobis distance of the facial images which do not belong to the Mahalanobis space by changing the combination of the attributes which are used for the construction of the Mahalanobis space. Therefore, the facial images which does not belong to the Mahalanobis space of the optimization target are needed. These facial images which do not belong to the target Mahalanobis space are called abnormal data. I collect the facial images used to construct the Mahalanobis spaces except the optimization target Mahalanobis space and the facial images judged to belong to the Mahalanobis space except optimization target as the abnormal data. For example, the facial images which are used to construct the Mahalanobis spaces of Anger, Disgust, Fear, Neutral, Sadness and Surprise and the facial images judged to belong to these Mahalanobis spaces are treated as the abnormal data when the optimization target is the Mahalanobis space of Happiness. Thus, the total number of abnormal data is 1200.

2. **Creation of the population:** In this optimization, the combination of the attributes is searched in order to construct the effective Mahalanobis space. Therefore, the combination of attributes to construct the Mahalanobis space is coded to an individual. In the individual, 0 represents that its attribute is not used to construct the Mahalanobis space and 1 represents that its attribute is used, respectively. Therefore, the length of the individual is 96. The population which consists of 50 individuals is created in this optimization.

3. **Reconstruction of the Mahalanobis space based on each individual:** New Mahalanobis spaces are constructed by using the attributes corresponding to the individuals in the population. Therefore, fifty Mahalanobis spaces are reconstructed for a population.

4. **Calculation of the fitness on the individual:** The SN ratio is employed to evaluate the effectiveness of the Mahalanobis spaces which are constructed in accordance with the individuals. Namely, I use the SN ratio as the fitness value on the individual. There are three kinds of SN ratio. The different kind of SN ratio is properly used in accordance with the purpose. The larger-the-better type of SN ratio is used in the case where the abnormal data are used in the optimization (cf. Section 6 of Chapter 1). The larger-the-better type of the SN ratio is calculated on the basis of the Mahalanobis distances of the abnormal data. The larger-the-better type of SN ratio is the criterion which represents the disjunction between the abnormal data and the Mahalanobis space. Therefore, the increase of this SN ratio indicates that the abnormal data does not belong the Mahalanobis space. That is, the correct recognition of abnormal data can be performed. The SN ratio is calculated by the following formula.

$$\eta = -10\log\frac{1}{n}\left(\frac{1}{D_1^2} + \frac{1}{D_2^2} + \cdots + \frac{1}{D_n^2}\right) \tag{2.12}$$

η : SN ratio
n : The number of the abnormal data collected in the process 1

5. **Selection:** The individuals which become the parents in next generation are decided on the basis of its fitness. I employ the roulette strategy and allow the duplication. In the roulette strategy, the individuals are stochastically selected as the parents in accordance with its fitness value.

6. **Crossover:** In the crossover operation, new two individuals are generated from the parents. I use one point crossover as the crossover operator in order to promote the convergence speed. The crossover rate is 0.1.

7. **Mutation:** The mutation operation inverts stochastically the value on the locus of the individual. I employ 0.01 as the mutation rate.

This process is repeated until 50 generations. The individual which has largest fitness, namely the SN ratio, is treated as the optimized attributes.

Table 2.2 shows the number of the optimized attributes on each facial expression. This result indicated that the number of attributes can be

Anger	Disgust	Fear	Happiness	Neutral	Sadness	Surprise
59	54	49	53	51	52	51

Table 2.2: The number of the optimized attributes

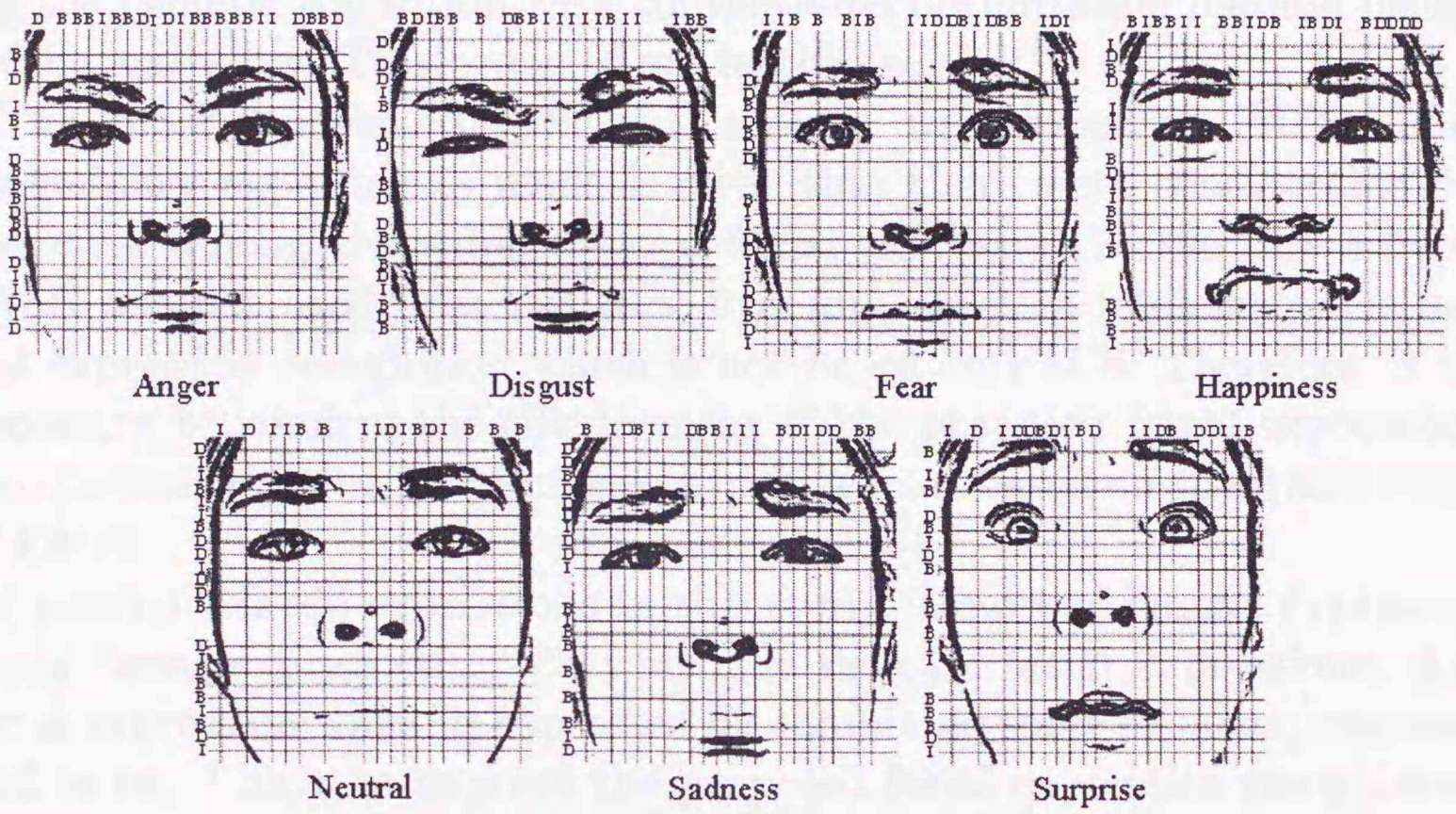


Figure 2.7: Optimization result

reduced to about 50 from 96. Moreover, I found from Figure 2.7 that the effective attributes for each facial expression were different from other facial expressions.

2.7 Conclusion

In this chapter, I proposed a facial expression recognition based on the MTS and confirmed the effectiveness of the proposed facial expression recognition through the practical experiment with some subjects. From the experimental results, I found that the MTS was effective for facial expression recognition in comparison with facial expression recognition based on the Euclidean distance. It is likely that the consideration of

the correlations among the attributes in multi-dimensional data was effective for the facial expression recognition. Moreover, I proposed an optimization method using the Genetic Algorithm. The optimization result revealed that the proposed optimization method could reduce the number of attributes to about half. However, the computational cost is high in the proposed optimization method because many Mahalanobis spaces have to be constructed. This problem may be improved by combining the Genetic Algorithm with conventional optimization method using orthogonal array. This is one of my feature works.

The facial expressions used in this study were based on FACS. The preliminary experiments made it clear that there were considerable differences among the subjects on a facial expression in the case where FACS was not used. For this fact, it is important to implement the facial expression recognition which is not based on FACS. Therefore, it is necessary to confirm the effectiveness of the proposed facial expression recognition with the facial images which are not based on the framework of FACS.

I treated 7 facial expressions in this study. However, human expresses many facial expressions. They include delicate facial expressions, the facial expression which is expressed by combining some facial expressions and so on. I have to improve the proposed facial expression recognition to adapt to such facial expressions.

Finally, I treated the facial expression recognition with the frontal facial images. If the profile can be treated in the facial expression recognition, more effective facial expression recognition can be implemented. The development of such a facial expression recognition is the aim of my future work.

Bibliography

[1] S. Yoshikawa, M. Masuya and M. Nakamura (eds.), "Face and Mind - An Introduction to Psychology of Face-", SAIENSU-SHA Publishers, Tokyo, Japan, 1993. (in Japanese)

[2] L. A. Zebrowitz, S. Hata and Y. Nakao (trans.), "Reading Face", Taishukan-Shoten Publishers, Tokyo, Japan, 1999. (in Japanese)

[3] S. Morishima, "Technical Problems in Facial Expression Recognition and Synthesis", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU97-141, pp.83-92, Nov. 1997. (in Japanese)

[4] O. Hasegawa, S. Morishima and M. Kaneko, "Processing of Facial Information by Computer", Journal of Institute of Electronics, Information and Communication Engineers, Vol.80-A, No.8, pp.1231-1249, Aug. 1997. (in Japanese)

[5] H. Harashima, "Invitation of the face learning", Iwanami Shoten Publishers, Tokyo, Japan, 1998. (in Japanese)

[6] T. Ogawa, M. Oda, S. Yoshikawa and S. Akamatsu, "Evaluation of Facial Expression Differing in Face Angles -Constructing a Database of Facial Expressions-", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU97-136, pp.47-52, Nov. 1997. (in Japanese)

[7] T. Kato, M. Nishimura, Y. Matsubara, M. Samukawa, K. Hattori and S. Akamatsu, "Feature Saliency in Facial-Expression Judgment", Technical report of Institute of Electronics, Information and Communication Engineers, HCS98-3, pp.13-18, June 1998. (in Japanese)

[8] G. Taguchi, "Mathematics for Quality Engineering -Quality Engineering as a general measure of functionality", Journal of Quality Engineering Society, Vol.6, No.6, pp.5-10, Dec. 1998. (in Japanese)

[9] G. Taguchi, "Quality Engineering for Pattern Recognition", Journal of Quality Engineering Society, Vol.3, No.2, pp.2-5, April 1995. (in Japanese)

[10] G. Taguchi, "Pattern Recognition and Quality Engineering", Journal of Quality Engineering Society, Vol.3, No.4, pp.2-5, Aug. 1995. (in Japanese)

[11] G. Taguchi, "Discussions on Voice Pattern Recognition Systems", Journal of Quality Engineering Society, Vol.3, No.5, pp.3-7, Oct. 1995. (in Japanese)

[12] Z. Prochazka, T. Ito and T. Okamoto, "Facial Expression Recognition System Using Facial Motion", Technical report of Institute of Electronics, Information and Communication Engineers, HIP96-42, pp.19-24, March 1997. (in Japanese)

[13] Y. Ando and T. Yonekura, "A Study of Emotional Interaction System Using Facial Expression", Technical report of Institute of Electronics, Information and Communication Engineers, NVE96-72, pp.29-36, March 1997. (in Japanese)

[14] P. Ekman and W. V. Friesen, "Facial Action Coding System", Consulting Psychologists Press, Palo Alto, US., 1978.

[15] Y. Gong, "INTELLIGENT IMAGE DATABASES Towards Advanced Image Retrieval", KLUWER ACADEMIC PUBLISHERS, Massachusetts, US, 1998.

[16] S. Teshima, T. Bando and D. Jin, "A Research of Defects Detection Using the Mahalanobis-Taguchi System Method", Journal of Quality Engineering Society, Vol.6, No.6, pp.74-81, Oct. 1997.

[17] M. Oda, S. Akamatsu and H. Fukamachi, "Adaptability of K-L Expansion Technique for Ambiguous Face Image Retrieval", Journal of Institute of Electronics, Information and Communication Engineers, Vol. J79-A, No.2, pp.288-297, Feb. 1996. (in Japanese)

[18] D. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison-Wesley Publishing Co., 1989.

Chapter 3

A Robust Face Identification against Lighting and Face Position Fluctuations

3.1 Introduction

Recently, the research on identification using human biometrics has been carried out in order to develop convenient and advanced security systems or access control systems [1]. The biometrics is human features such as fingerprint, iris and so on which do not change during lifetime and are different among each person. Comparing with the conventional identification using password or card, effective identification systems can be implemented by using the biometrics because there is not a possibility that it is lost and forgotten. In the identification using human biometrics, the identification based on human face has been focused on because this type of identification has the advantage that it can inhibit crime and does not require special device for the identification in comparison with the identification using other biometrics.

In the face identification, the robustness against various fluctuations must be implemented in order to develop practical identification system. These fluctuations include photographic fluctuations such as lighting fluctuation, pose fluctuation and so on which necessarily arise when facial image is taken and appearance fluctuations such as glasses, hairstyle, facial expressions and so on. Many studies to implement robustness against the photographic fluctuations are carried out in recent research on the face identification because the implementation of robustness against appearance fluctuations is more difficult problem than photographic fluctuations in object recognition point of view [2][3][4][5].

One of methods to implement the robustness against photographic fluctuations is the identification based in many facial images involving various

photographic fluctuations which are taken on real environment. However, it is burden for the user to take many facial images in real environment. Moreover, the convenience which is one of advantages in the identification using human biometrics is impaired.

In this chapter, I propose a robust face identification against lighting and face position fluctuations. In the proposed face identification, a few facial images of the user are taken. The robustness against lighting and face position fluctuations in real environment can be implemented by inserting artificial fluctuations into the facial images taken. Moreover, the edge extraction which is not influenced much by lighting fluctuation is employed in the pre-processing of facial image. On the other hand, I use the Mahalanobis-Taguchi System (MTS) as the pattern recognition method for the identification [6]. The MTS is one of statistical pattern recognition methods which was proposed by Taguchi. The MTS can perform robust pattern recognition by giving noise to the training data because the recognition in the MTS is carried out based on the correlations among attributes which is common in a set of training data [7]. In addition, the MTS can reduce attributes to only those that significantly contribute to recognition by using orthogonal array, Signal/Noise ratio (SN ratio) and factorial effect. Therefore, the time required for identification can be decreased, and the amount of data in the facial image database can be reduced. I confirmed that an effective facial expression recognition could be implemented by using the MTS [8][9]. Therefore, it is quite likely that the MTS is an effective method for the proposed face identification. I confirm the effectiveness of the proposed face identification method through practical experiment with some subjects.

This chapter is organized as follows: In Section 2, I will propose a robust face identification against lighting and face position fluctuations. The experimental method using some subjects will be described in Section 3. In Section 4, the experimental result will be mentioned. Finally, I will finish this chapter with the conclusions.

3.2 Robust Face Identification against Lighting and Face Position Fluctuations

The detail of the proposed face identification is described below.

3.2.1 Face Identification Process

The identification process of the proposed face identification is as follows.

1. Collection of facial images on each person
2. Insertion of artificial fluctuations into the collected facial images
3. Extraction of attributes from the facial images involving the artificial fluctuations
4. Construction of Mahalanobis space on each person from the extracted attributes
5. Face identification of unknown facial image based on the Mahalanobis distance
6. Optimization of attributes using orthogonal array, SN ratio and factorial effect

First, the facial images are collected on each person. The artificial fluctuations of lighting and face position are inserted into the facial images collected. Then, the attributes based on differential and integral characteristics are extracted from the facial images involving artificial fluctuations. The Mahalanobis space is constructed on each person from these attributes. The face identification of unknown facial image is performed by using the Mahalanobis distances between the Mahalanobis spaces and unknown facial image. Finally, the effective attributes to recognize each person are re-extracted based on orthogonal array, SN ratio and factorial effect. The detail of each step is shown below.

3.2.2 Facial Image

In this study, I take facial images by using a monochrome camera. The facial image is 300×300 pixels. The gradation value of pixel is 256. The background color is white. It does not include hair, glasses and facial expression. The facial images are taken in my laboratory with special lighting device. Moreover, the facial images are taken by using baselines for the position of eyes, nose and mouth. Therefore, these facial images include a little photographic fluctuation.

3.2.3 Artificial Fluctuations

The lighting and face position fluctuations which are artificially created are inserted into the facial image on each person in order to implement the robustness against lighting and face position fluctuations in real environment. The details of these artificial fluctuations are as follows.

1. **Lighting fluctuation:** In this study, the edge extraction is employed as the pre-processing of facial image. The extraction state of edge changes in accordance with lighting fluctuation. That is, the edge can be clearly extracted when the lighting condition is good. On the contrary, it is difficult to extract clearly the edge when the lighting condition is bad. Therefore, I create the artificial lighting fluctuations by changing the strength of edge extraction. In the edge extraction, I use the Global Range Edge Detection Method [10]. The detail of the edge extraction based on this method is described in Section 2 of Chapter 2. I change a formula (2.7) in this method to the following formula.

$$|\partial_{ij}| \geq \alpha(\mu + \sigma) \tag{3.1}$$

In the Global Range Edge Detection Method, the edge is judged on the basis of the deviation of gradation value in pixel. Basically, the threshold to judge the edge in the image data is decided by the deviation of gradation value in the whole of image. I add a constant α to formula (2.7) in order to change the extraction state of edge purposely. The extraction state of edge can be changed by varying this α. Figure 3.1 shows an example of the edge extracted facial image for each value of α. As shown in Figure 3.1, the edge can be clearly extracted when the value of α is small, and the accuracy of edge extraction is lost when the value of α is large.

2. **Face position fluctuation:** The artificial face position fluctuation is produced by moving the edge extracted facial image in four directions. The directions are up, down, left and right. An example of face position fluctuation is illustrated in Figure 3.2.

The face identification is performed on the basis of the facial images involving these artificial fluctuations.

3.2.4 Face Identification based on MTS

After the insertion of artificial fluctuations, the attributes based on differential and integral characteristics are extracted from the facial images involving the artificial lighting and face position fluctuations [11][12]. The Mahalanobis space on each person is constructed from the extracted attributes. In the face identification, there is two types of identification.

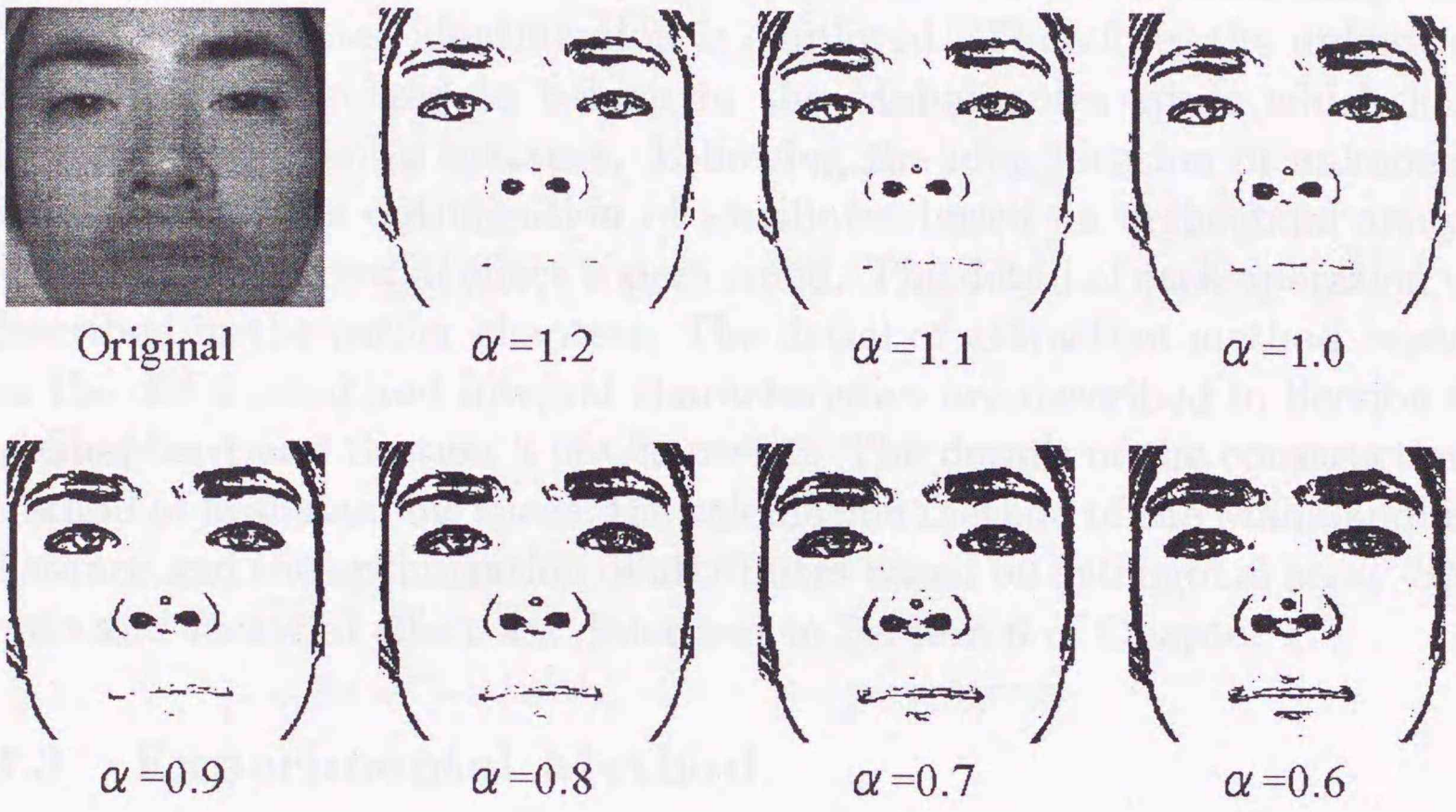


Figure 3.1: An example of the edge extracted facial image for each value of α

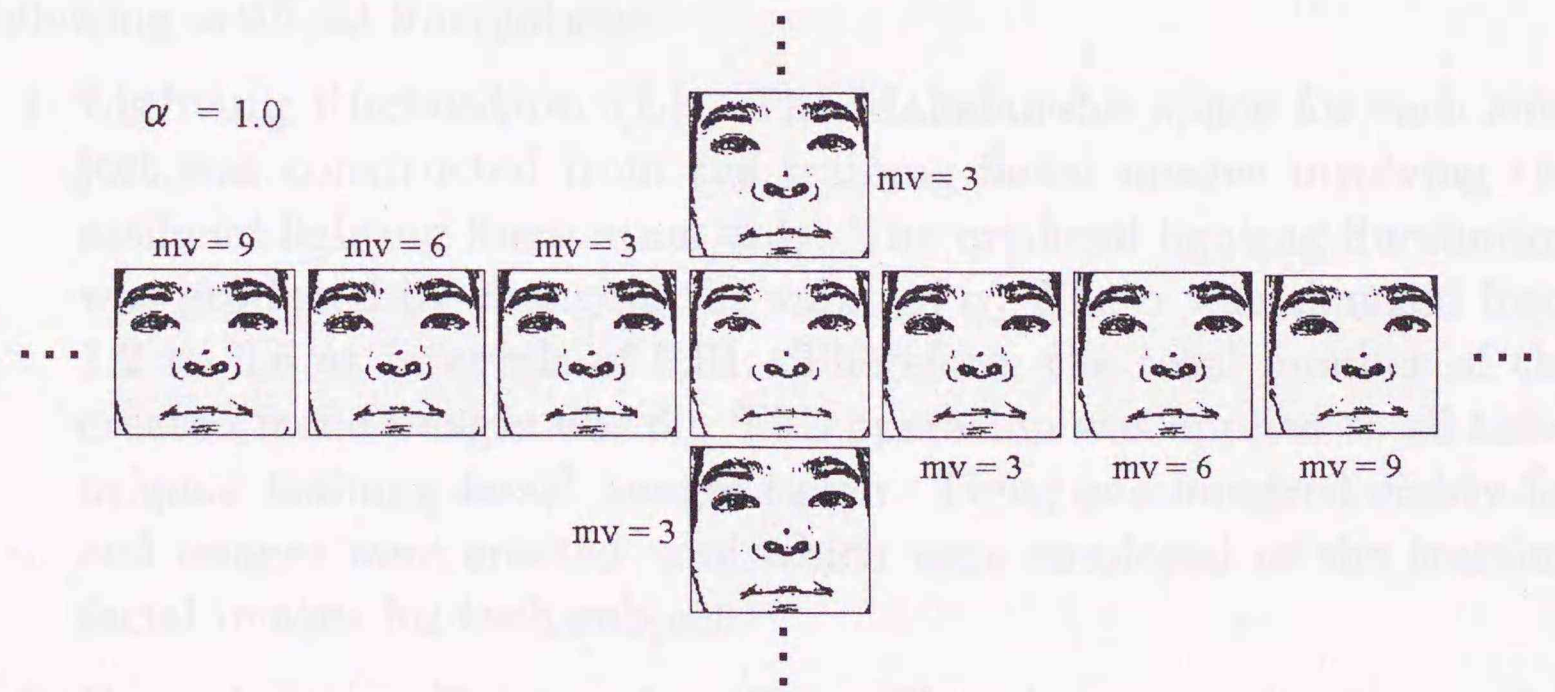


Figure 3.2: An example of face position fluctuations

One is the face identification which judges whether unknown facial image is a specified person or not. Another is the face identification that determines which person in a face database the unknown facial image is. In this study, latter identification is employed. Therefore, the unknown facial image is judged to belong to the Mahalanobis space which has shortest Mahalanobis distance. Following the identification of unknown facial images, the optimization of attributes based on orthogonal array, SN ration and factorial effect is performed. The detail of each operation is described in the earlier chapters. The detail of extraction method based on the differential and integral characteristics are described in Section 6 of Chapter 1 and Section 3 of Chapter 2. The details of the construction method of Mahalanobis space, the calculation method of the Mahalanobis distance and the optimization of attributes based on orthogonal array, SN ratio and factorial effect are described in Section 6 of Chapter 1.

3.3 Experimental Method

I designed the following experiments in order to confirm the effectiveness of the proposed face identification. In this experiment, ten students were recruited as subjects from Hokkaido University. I took four facial images of each subject. I employed three facial images of each subject as the original training facial images, and used one facial image as the original unknown facial image. In this experiment, I constructed three kinds of Mahalanobis spaces from different training facial images involving the following artificial fluctuations.

1. **Lighting fluctuation (L):** The Mahalanobis space for each subject was constructed from the training facial images involving the artificial lighting fluctuation only. The artificial lighting fluctuation was produced by changing the value of α. The α was changed from 1.2 to 0.6 at intervals of 0.01. Therefore, the total number of the created facial images was 60. This operation was applied to all three original training facial images taken. Thus, one hundred eighty facial images were created, and which were employed as the training facial images for each subject.

2. **Face position fluctuation (P):** The edge extraction in $\alpha = 1.0$ was applied to the facial image. The edge extracted facial image was moved 15 pixels at intervals of 1 pixel in four directions (up, down, left and right). This operation was performed on three original training facial images for each subject. Therefore, one hundred

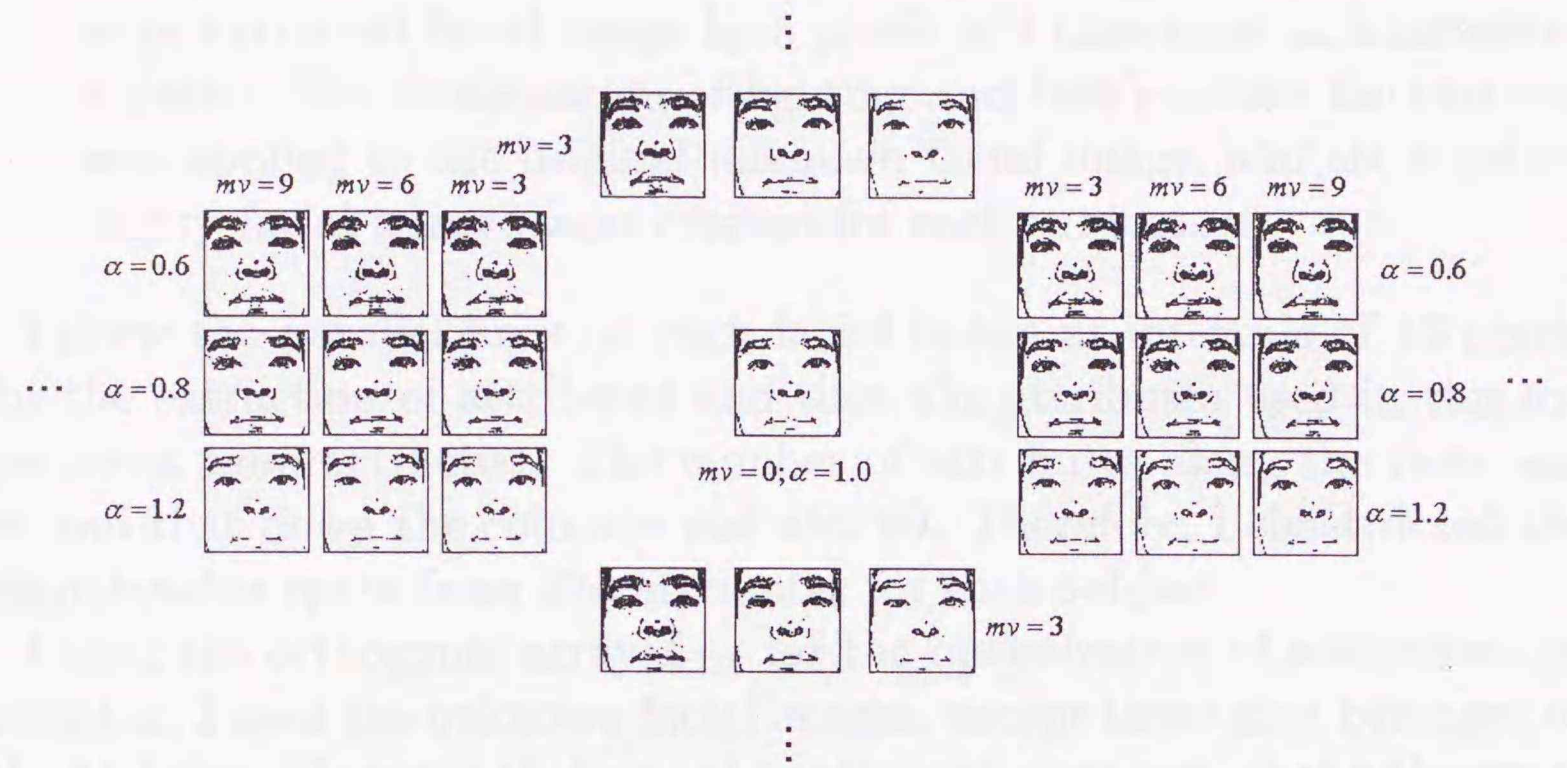


Figure 3.3: An example of training facial images in LP

seventy-one training facial images involving face position fluctuation were created.

3. **Lighting and face position fluctuations (LP):** The combination of the lighting fluctuation and the face position fluctuation described above was applied to the original training facial image. This operation was performed on only 1 original training facial image for each subject. Therefore, three thousand four hundred twenty training facial images involving the lighting and face position fluctuations were created for each subject. Figure 3.3 shows an example of training facial images in LP.

Moreover, I applied two kinds of fluctuation to the original unknown facial image taken.

1. **Lighting fluctuation (S):** These unknown facial images included lighting fluctuation only, which were created by varying the value of α between 1.1 and 0.7 at intervals of 0.01. Therefore, forty unknown facial images involving the artificial lighting fluctuation were created for each subject.

2. **Lighting and face position fluctuation (B):** The lighting and face position fluctuations were applied to the original unknown facial image taken. The lighting fluctuation was created by shifting the value of α between 1.1 and 0.7 at intervals of 0.01. Moreover, the

face position fluctuation was created by shifting the position of the edge extracted facial image by 5 pixels in 4 directions at intervals of 1 pixel. This combination of lighting and face position fluctuations was applied to the original unknown facial image, and six hundred eighty facial images were created for each subject.

I drew the parallel lines on each facial image at intervals of 10 pixels for the extraction of attributes and then the attributes used in this experiment were extracted. The number of attributes along the rows was 60 and that down the columns was also 60. Therefore, I constructed the Mahalanobis space from 120 attributes for each subject.

I used the orthogonal array L_{128} for the optimization of attributes. In addition, I used the unknown facial images, except those that belonged to the Mahalanobis space that was the optimization target, as the abnormal data.

3.4 Experimental Result

3.4.1 Identification Result

The identification results are shown in Table 3.1, Table 3.2, Figure 3.4, Figure 3.5, Figure 3.6, Figure 3.7, Figure 3.8, Figure 3.9 and Figure 3.10. Table 3.1 shows the rates which the correct identification was made. In Table 3.1, s1 - s10 represent the subject number, and L_S, L_B, P_S, P_B, LP_S and LP_B represent each experiment. The left side of each combination of letters represents the kind of the training facial images used and the right side represents the kind of the unknown facial images used. In all experiments, effective face identification was achieved in spite of the use of the unknown facial images that included lighting and face position fluctuations. The results of L and P revealed that the effective Mahalanobis space differed on each subject. I think that this occurred by the photographic fluctuations which might have occurred during the initial photography. However, the rate of correct identification was highest in LP. Therefore, I found that the Mahalanobis space constructed from the training facial images which included both of artificial fluctuations was more effective in the implementation of robust face identification than the Mahalanobis space constructed using one artificial fluctuation.

Three different original training facial images were used for constructing the Mahalanobis space in L and P, and only one facial image was used as the training facial image in LP. In the face identification system, it

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	Total
L_S	100	97.5	100	30	100	100	62.5	100	100	100	89.0
L_B	96.0	70.7	100	37.6	100	100	37.8	100	100	93.0	83.7
P_S	75	100	100	77.5	100	95	20	100	100	100	86.8
P_B	73.7	100	100	76.3	100	92.5	27.5	96.5	96.6	100	86.3
LP_S	100	100	100	90	100	100	100	100	100	100	99.0
LP_B	100	100	100	88.5	100	100	100	100	100	100	98.9

Table 3.1: Identification result (%)

is unusual to repeatedly photograph a user in order to implement the robustness against various fluctuations, because such photography is a burden to the user. With this in mind, it is likely that the inclusion of the artificial fluctuations in the training facial images can be employed effectively in the face identification.

The total rate of correct identification was 89.0% in L_S, and it was 99.0 % in LP_S, respectively. The rate of identification using LP was higher than that using L facial images. In these experiments, the unknown facial images did not include artificial face position fluctuation. Therefore, these unknown facial images included only the natural fluctuation of the face position that occurred during the initial photographing of the subjects. This result indicated that the Mahalanobis space created from the training facial images which included the artificial fluctuation of the face position was also effective in the identification of facial images which included natural face position fluctuation which might have occurred during the initial photography.

In the identification using L of s4 and the identification using P of s7, the rate of correct identification using B unknown facial images was higher than S. Normally, the rate of correct identification will decrease in B unknown facial images in comparison with S unknown facial images because L and P include only either artificial fluctuation. This may be due to the initial face position. That is, the unknown facial image was moved to the appropriate position for the identification by artificial face position fluctuation. Therefore, accurate face identification might be achieved by identifying the person from various unknown facial images in which the artificial fluctuations of lighting and face position are included.

Moreover, in order to confirm the effectiveness of identification based on the correlations among attributes in the multi-dimensional training

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	Total
LP_B	100	60.7	46.4	0.4	80.6	100	2.5	23.4	100	100	61.4

Table 3.2: Identification result based on Euclidean distance (%)

data in the MTS, I performed face identification based on the Euclidean distance and the average of the attributes extracted on the basis of differential and integral characteristics. The results of this experiment are shown in Table 3.2. Compared with the result obtained using the MTS, the rate of identification in the LP_B experiment was poor (98.9% vs. 61.4%). From this result, I can be fairly certain that the MTS, which is based on consideration of the correlations among the attributes in the multi-dimensional data, is effective for the face identification.

Figure 3.4 and Figure 3.5 illustrate the correlation matrix on each subject. In these figures, the correlation pattern among the attributes on each subject was similar to each other. Basically, face resembles other's face in terms of the shape and position of internal facial features. However, there is a little difference among each person, namely individuality. Therefore, it is likely that the basic state of the correlations was similar to each other but had a little difference among each subject.

Figure 3.6, Figure 3.7, Figure 3.8 and Figure 3.9 represent the Mahalanobis distances between the unknown facial images which were created from original unknown facial image on each subject and the Mahalanobis spaces on all subjects in LP_B experiment. These Mahalanobis distances to each Mahalanobis space are sorted in the order of small Mahalanobis distances between the unknown facial images and the Mahalanobis space which these unknown facial images should belong to. In these figures, the Mahalanobis distances between the Mahalanobis space which these unknown facial images should belong to and the unknown facial images were very short in comparison with the Mahalanobis distance between the other's Mahalanobis space and these unknown facial images. These results suggest that the face identification based on the MTS has high classification ability.

However, the rate of the correct identification of s4 was lower than other subjects in spite of the use of training facial images that included the lighting and face position fluctuations. Figure 3.10 shows the result that the correct identification was not achieved. In Figure 3.10, the facial images on the top side of this figure indicate the original edge extracted facial images. The facial images on the down side show the facial images

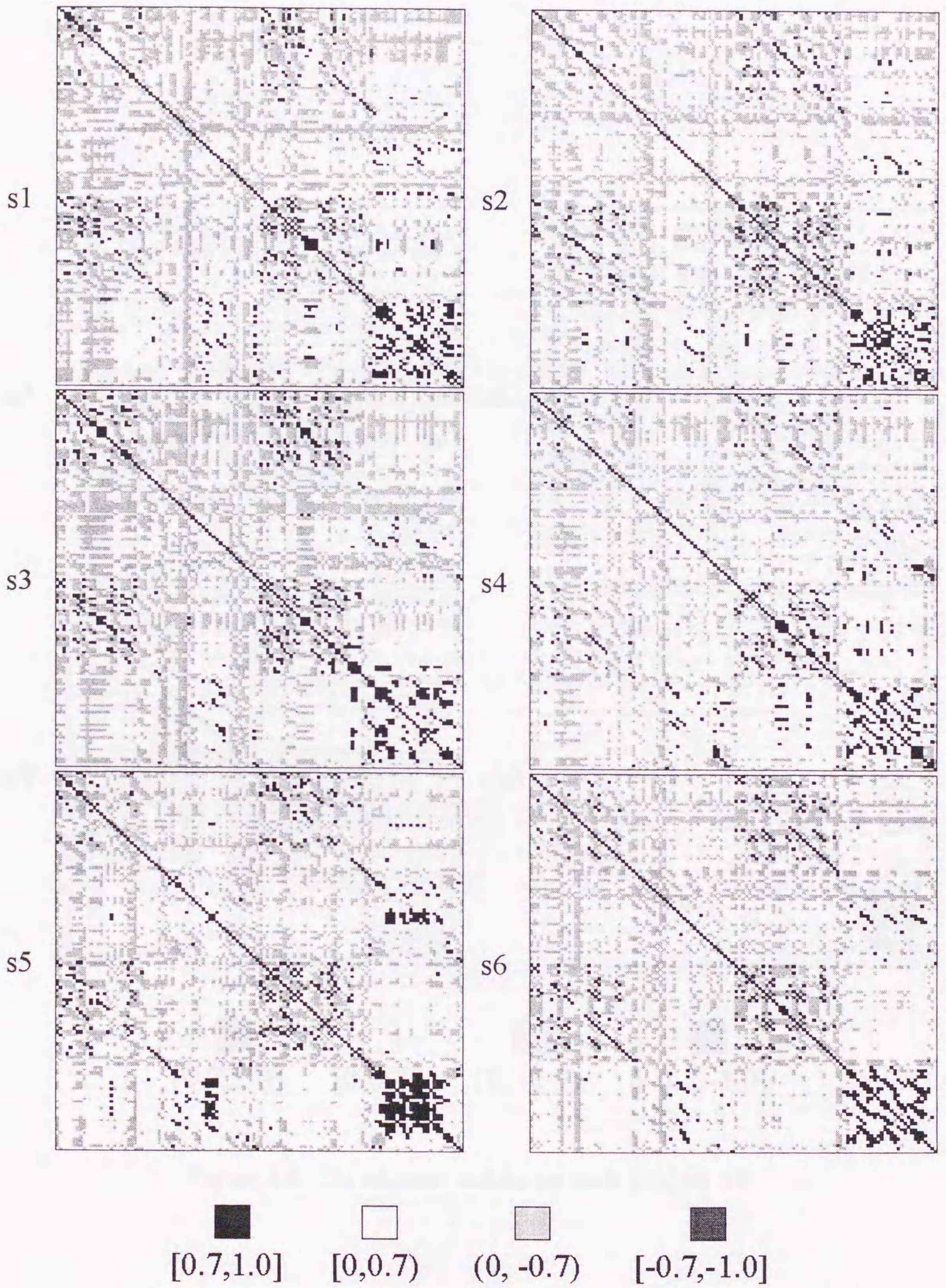


Figure 3.4: Correlation matrix on each subject (A)

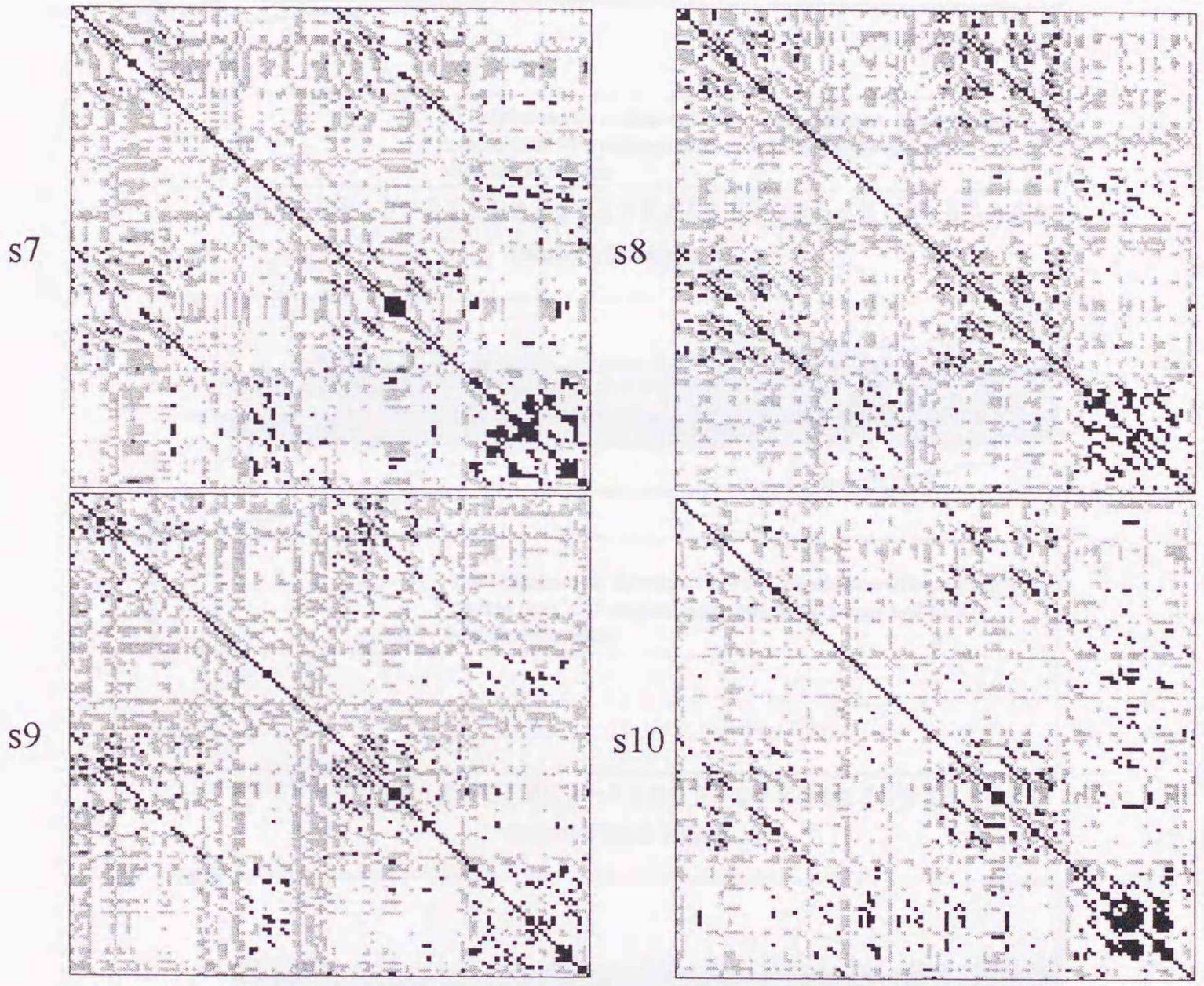


Figure 3.5: Correlation matrix on each subject (B)

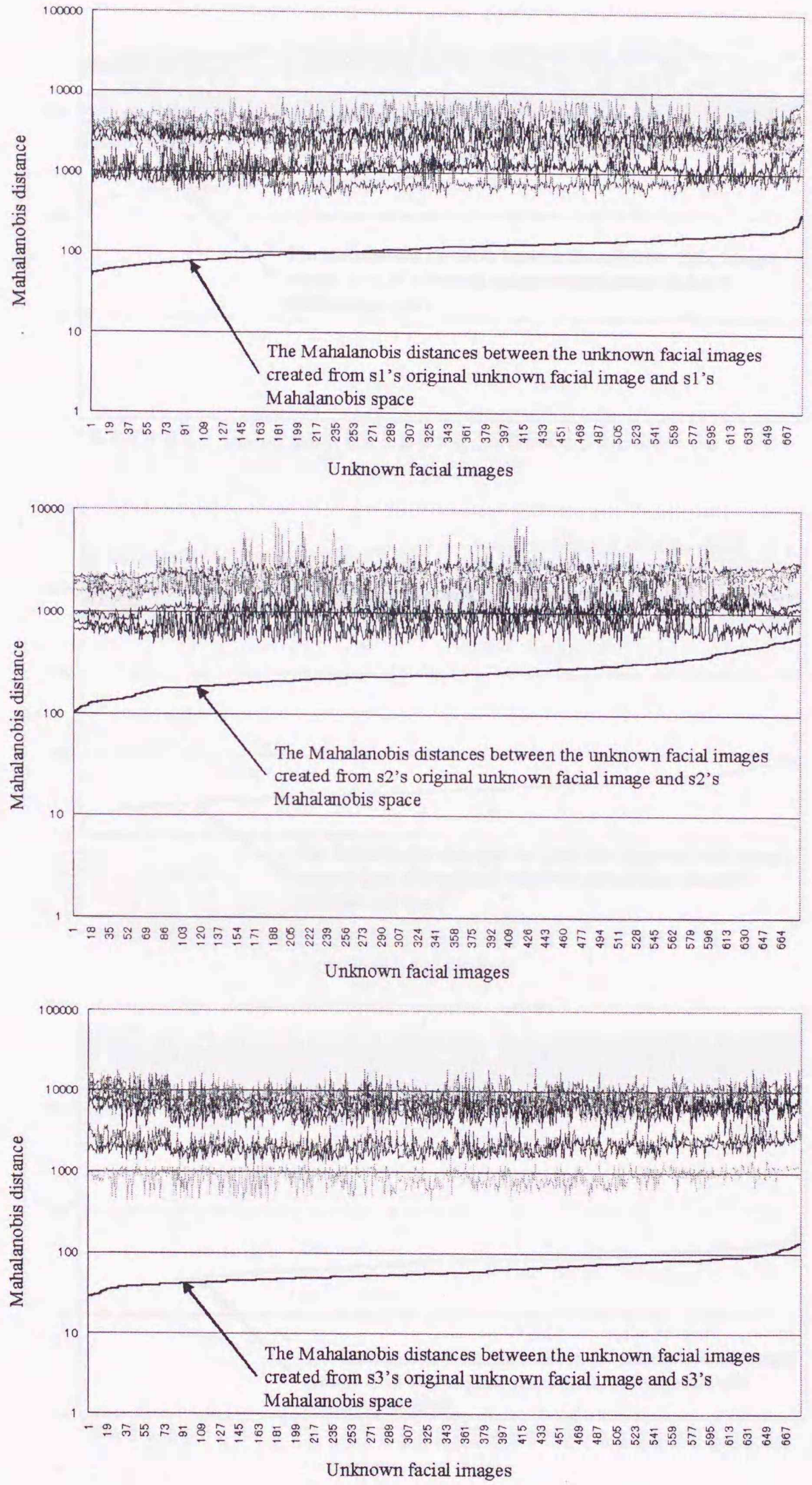


Figure 3.6: Mahalanobis distances between each Mahalanobis space and the unknown images (A)

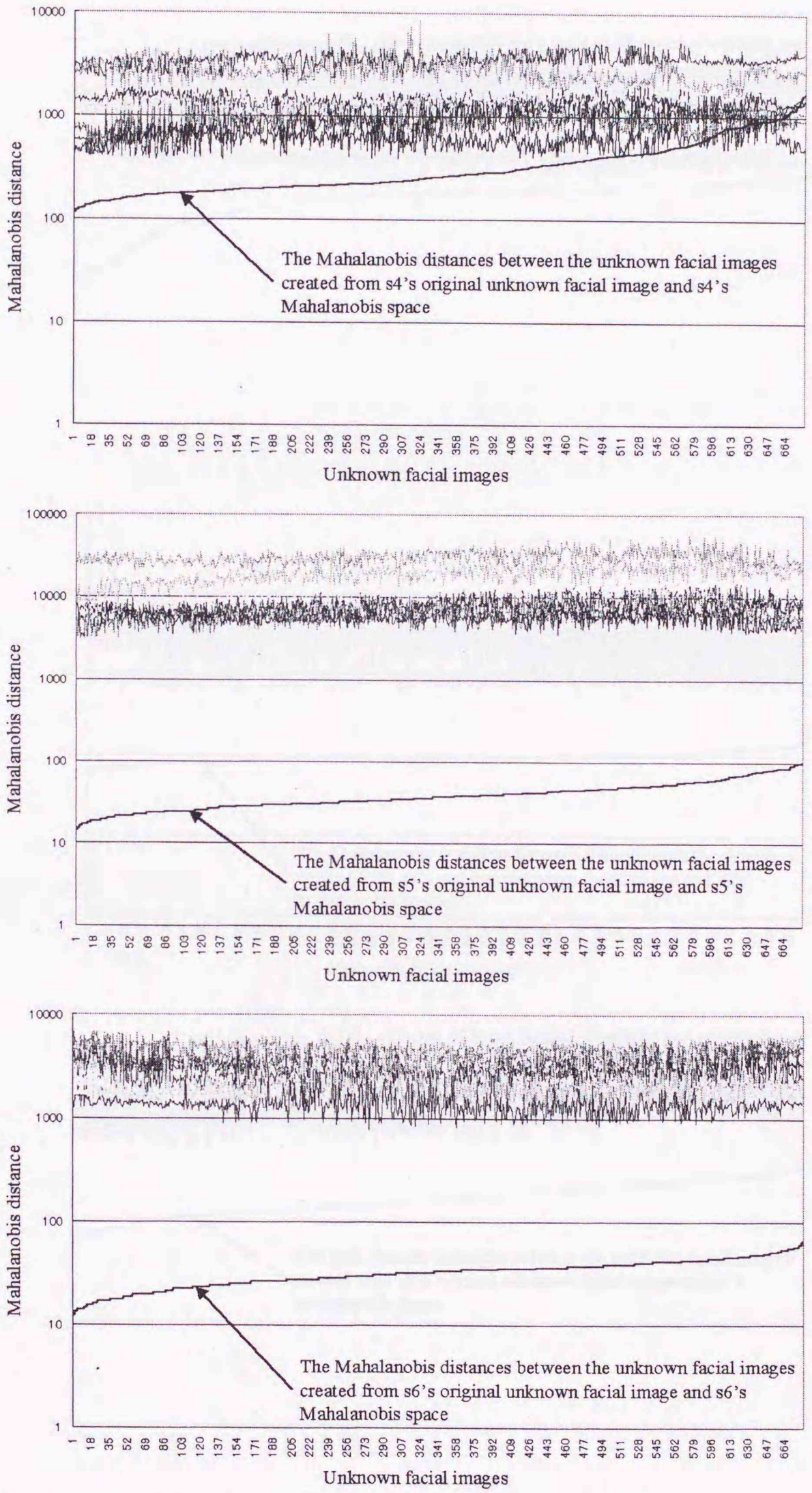


Figure 3.7: Mahalanobis distances between each Mahalanobis space and the unknown images (B)

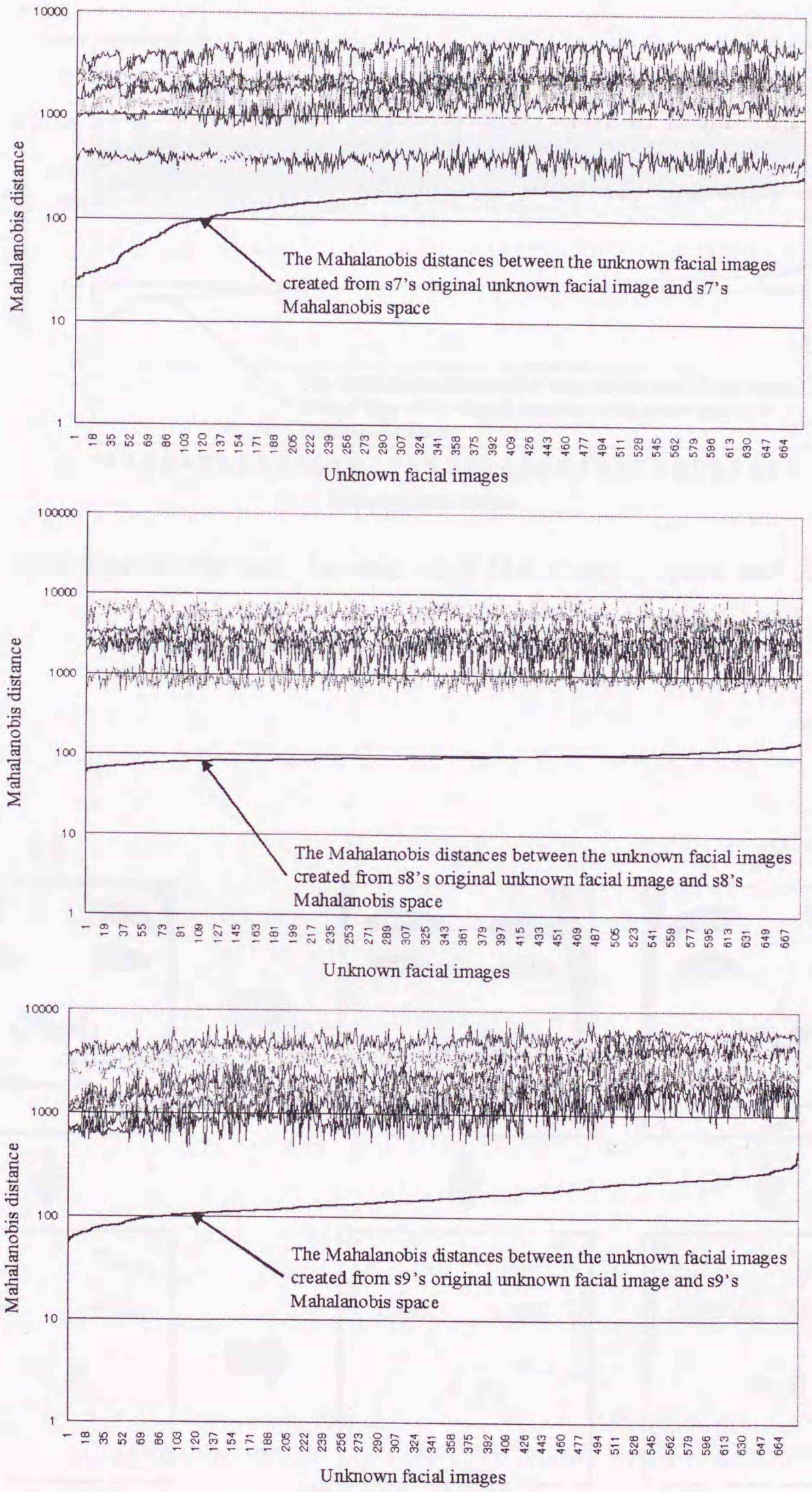


Figure 3.8: Mahalanobis distances between each Mahalanobis space and the unknown images (C)

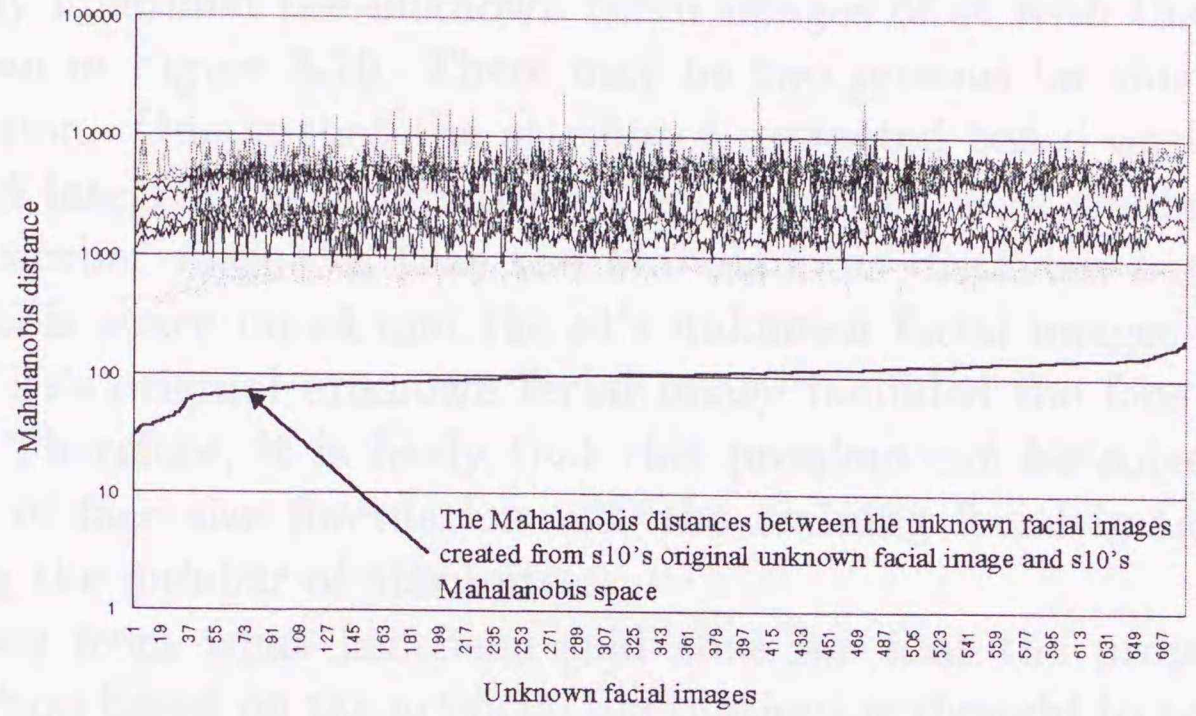


Figure 3.9: Mahalanobis distances between each Mahalanobis space and the unknown images (D)

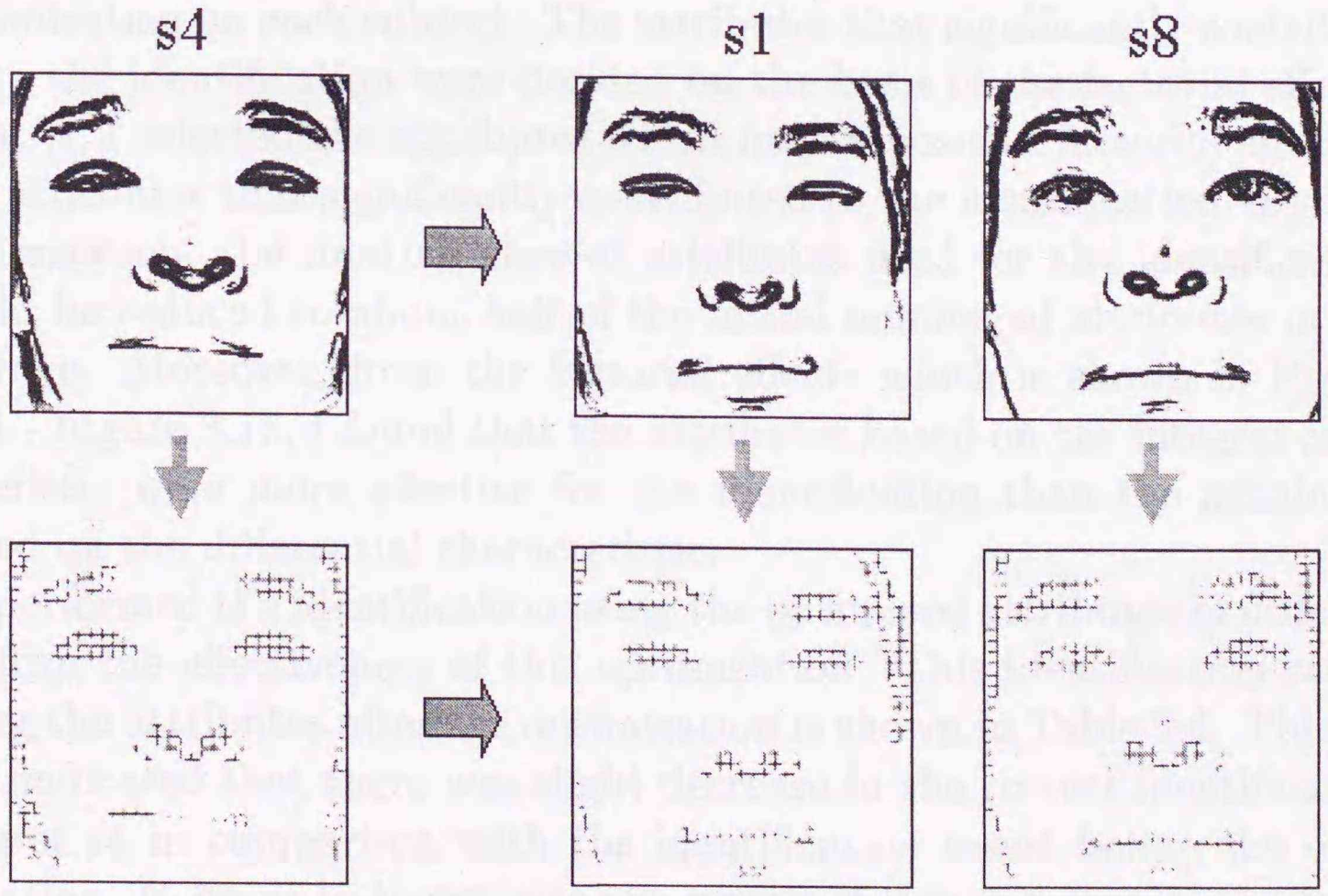


Figure 3.10: Identification result of s4

based on the lines which were used to extract the attributes based on the differential and integral characteristics. The proposed face identification incorrectly identified the unknown facial images of s4 with the two persons shown in Figure 3.10. There may be two reasons for this incorrect identification. One is that the attributes extracted based on the differential and integral characteristics on each subject were similar to each other. Another reason is that the Mahalanobis distances between the Mahalanobis space on s4 and the s4's unknown facial images increased since the s4's original unknown facial image included the face size fluctuation. Therefore, it is likely that this problem can be solved by the inclusion of face size fluctuation into the training facial image and by increasing the number of attributes.

It follows from what has been said thus far that the proposed face identification based on the artificial fluctuations is thought to be effective for implementing a robust face identification.

3.4.2 Optimization Result

Table 3.3, Table 3.4, Table 3.5, Figure 3.11, Figure 3.12, Figure 3.13, Figure 3.14, Figure 3.15, Figure 3.16, Figure 3.17 and Figure 3.18 show the optimization results. Table 3.3 shows the number of attributes after optimization on each subject. The attributes that significantly contributed to the identification were decided on the basis of the factorial effects. That is, I selected the attributes which have a positive factorial effect as the attributes that significantly contributed to the identification. By this optimization, the final number of attributes used for the identification could be reduced to about half of the initial number of attributes in the average. Moreover, from the factorial effects which is shown in Figure 3.11 - Figure 3.15, I found that the attributes based on the integral characteristic were more effective for the identification than the attributes based on the differential characteristic.

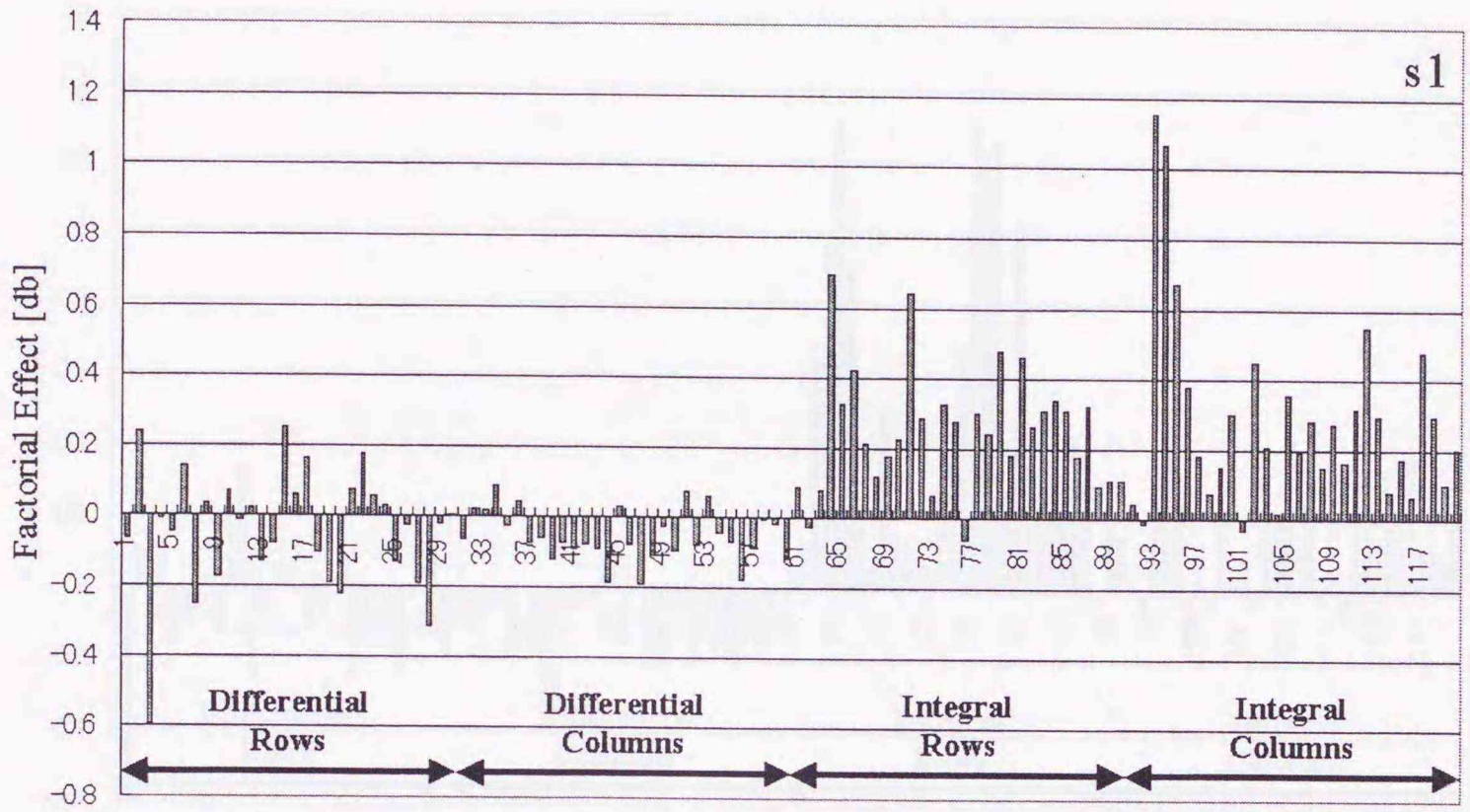
I performed the identification using the optimized attributes in order to confirm the effectiveness of this optimization. This identification result using the attributes after the optimization is shown in Table 3.4. This result indicated that there was slight decrease in the correct identification rate of s4 in comparison with the identification result before the optimization. In order to investigate the reason of this decrease, I calculated the SN ratio before and after the optimization between the Mahalanobis space which was constructed based on the optimized attributes and the abnormal facial images which were used for the optimization on each sub-

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10
The number of attributes	76	68	65	71	75	69	69	76	68	70

Table 3.3: Results of optimization

ject. This result is shown in Table 3.5. The SN ratio of the Mahalanobis space which is constructed using the optimized attributes increased on the subjects except s1, s5 and s8. This result indicates that the possibility of incorrect identification for abnormal facial images decreases. The increase of the Mahalanobis distance after the optimization on the abnormal facial images is also shown in Figure 3.16. Therefore, the decrease of correct identification rate in s4 is not due to the Mahalanobis distance of abnormal facial images. On the contrary, I calculated the Mahalanobis distances of the unknown facial images which should belong to the optimized Mahalanobis space. Figure 3.17 shows the Mahalanobis distances before and after the optimization between the Mahalanobis space which is optimization target and the unknown facial images which should belong to this Mahalanobis space on some subjects. This result indicated that there is the case where the Mahalanobis distances of unknown facial images which should belong to the Mahalanobis space which is the optimization target increase although the Mahalanobis distances of abnormal facial images decrease. Figure 3.18 indicates the Mahalanobis distances before and after the optimization between the unknown facial images which are created from the original s4's unknown facial image and the Mahalanobis space on each subject. From this result, I found that the decrease of the correct identification rate in s4 was caused by the increase of the Mahalanobis distance of the unknown facial images that should belong to the Mahalanobis space which is the optimization target.

The above optimization results revealed that there were the following two problems in the optimization based on the orthogonal array, SN ratio and factorial effect. There is the case where the SN ratio for the Mahalanobis space after optimization decreases. In the optimization based on the orthogonal array and factorial effect, the effectiveness of the attributes is calculated based on main effect of SN ratio. That is, the optimization which does not consider the correlations among the attributes is performed. Moreover, there is a case where the correct identification rate decrease by the increase of the Mahalanobis distance of the unknown facial image which should belong to the optimization target Mahalanobis space even if the possibility that abnormal facial images are incorrect-



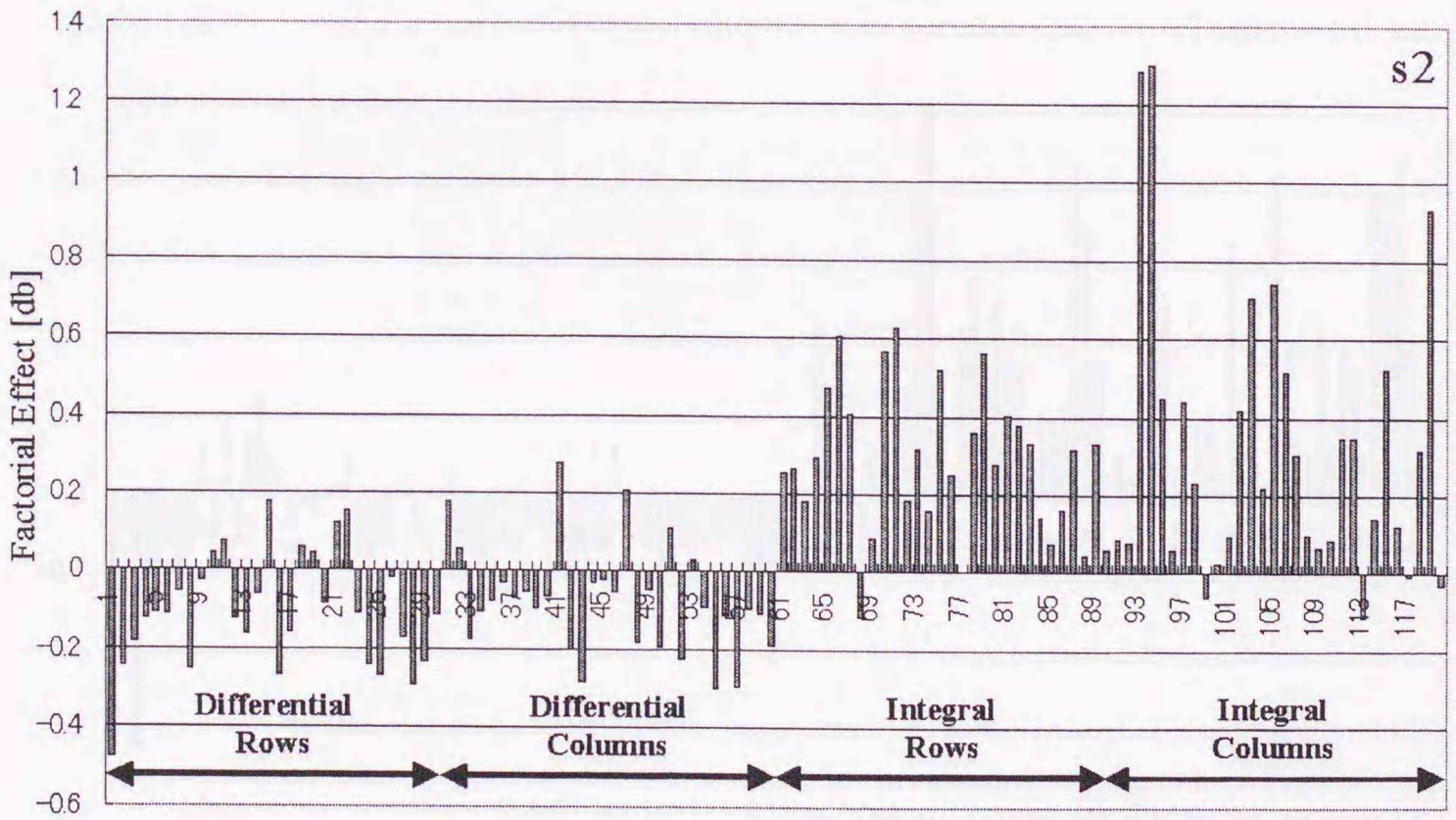


Figure 3.11: Factorial effects on each subject (A)

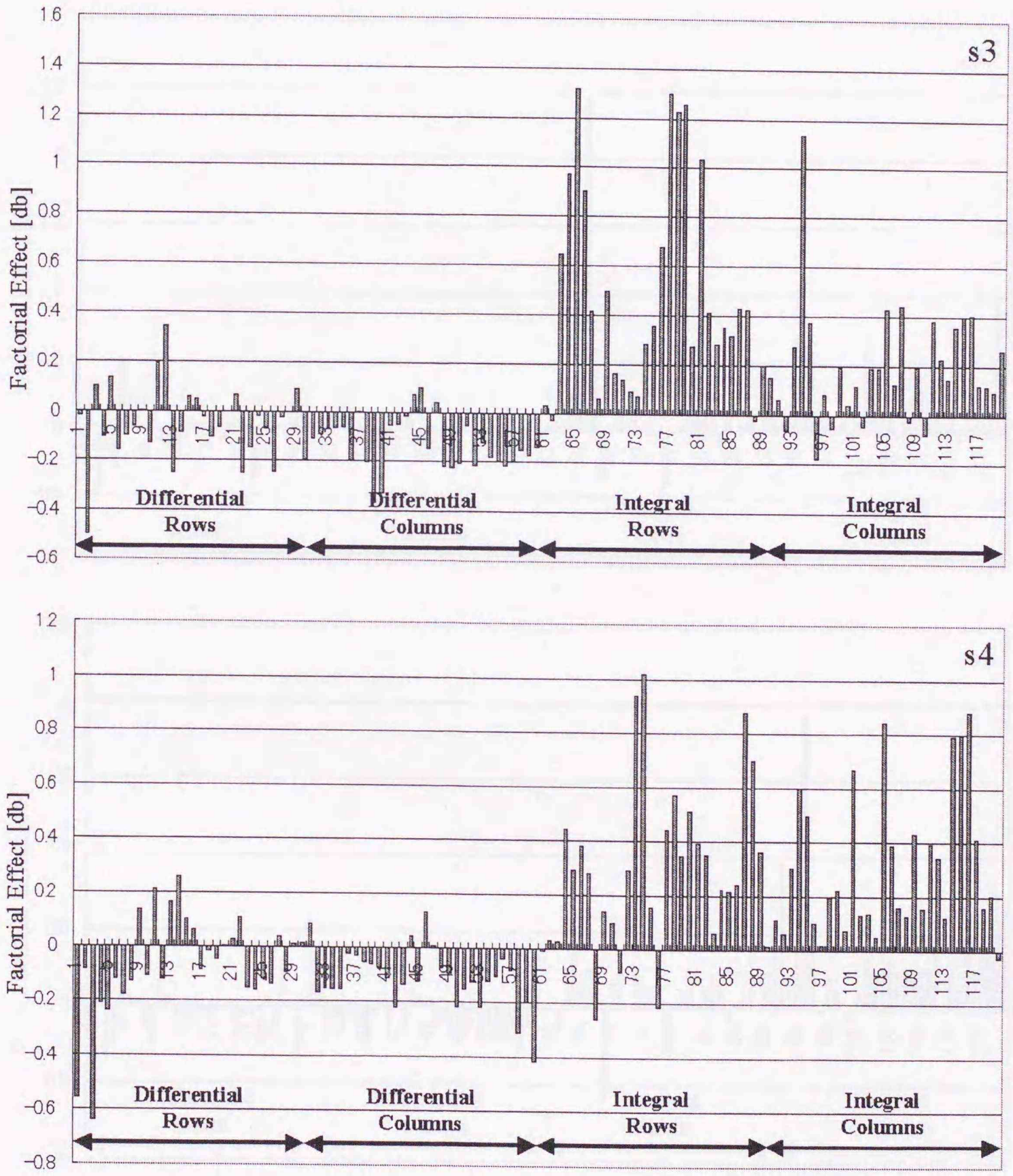
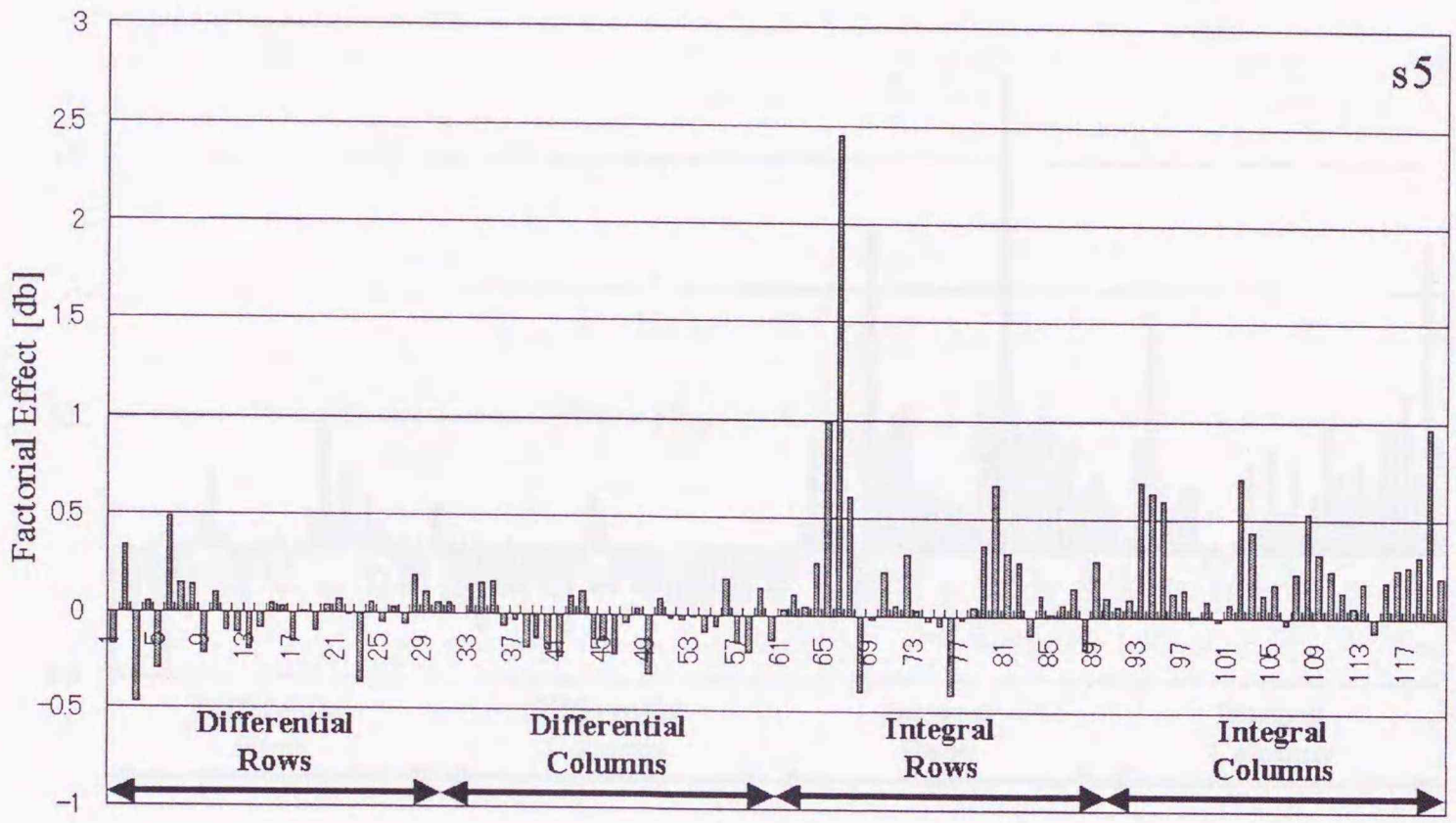


Figure 3.12: Factorial effects on each subject (B)



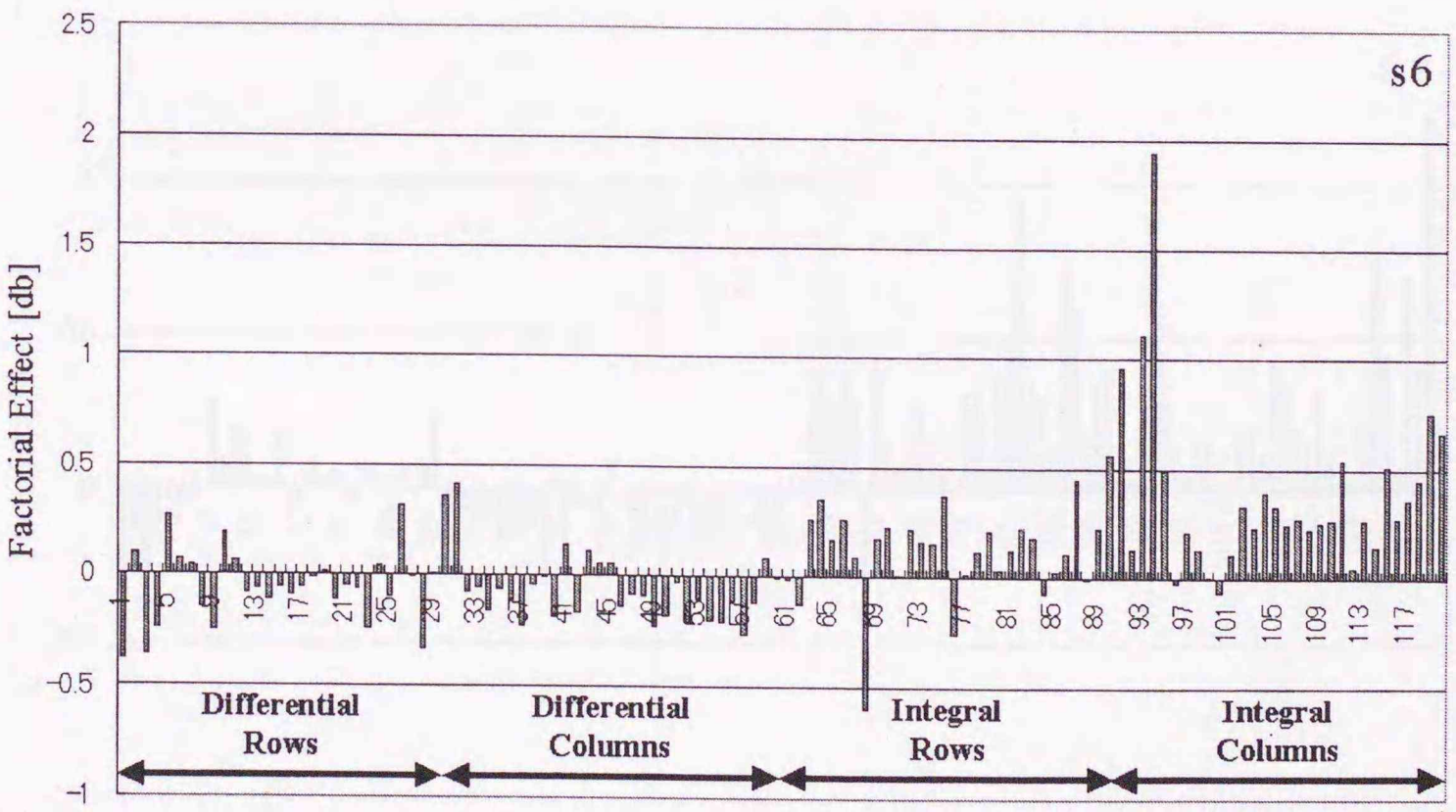


Figure 3.13: Factorial effects on each subject (C)

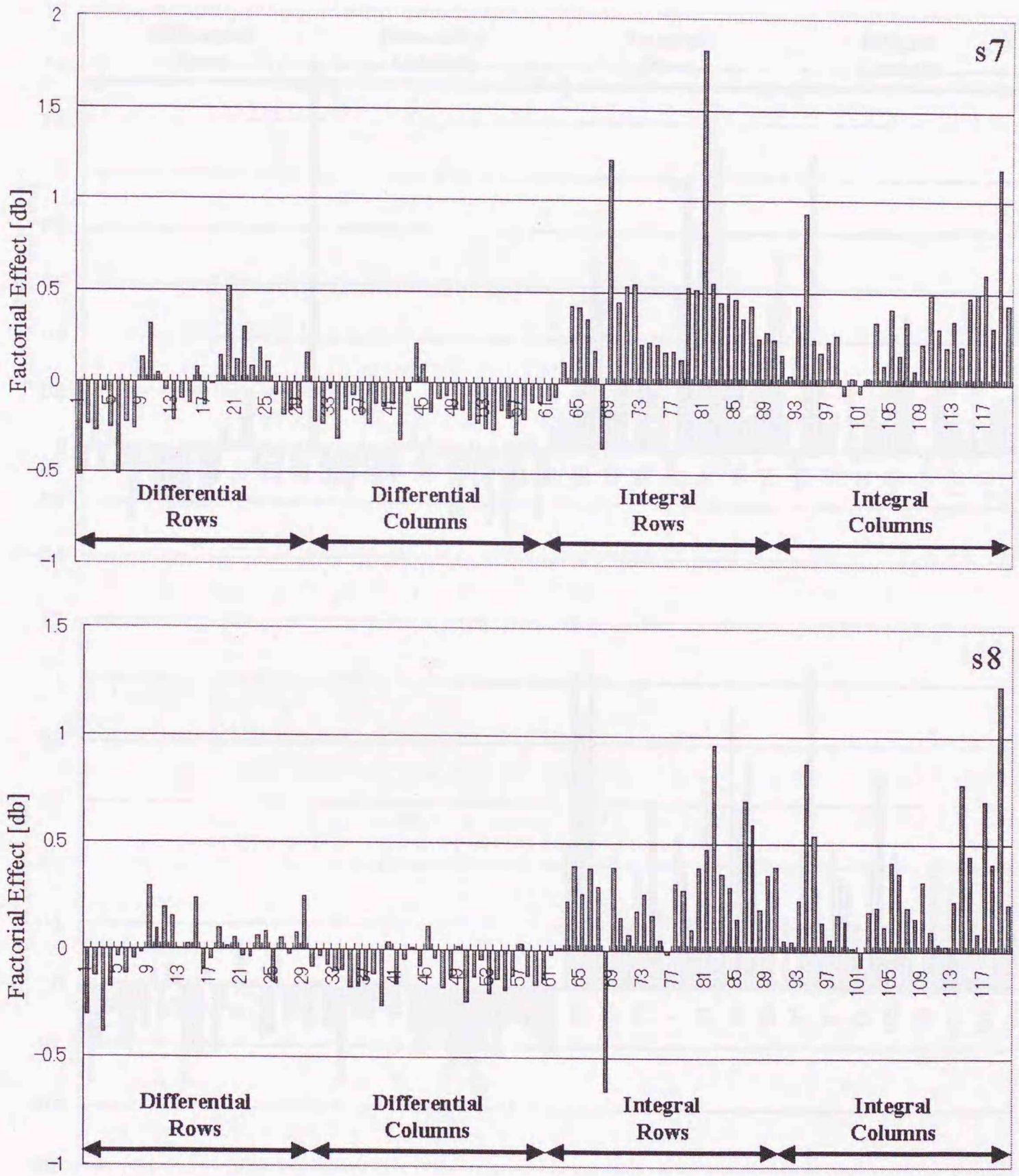


Figure 3.14: Factorial effects on each subject (D)

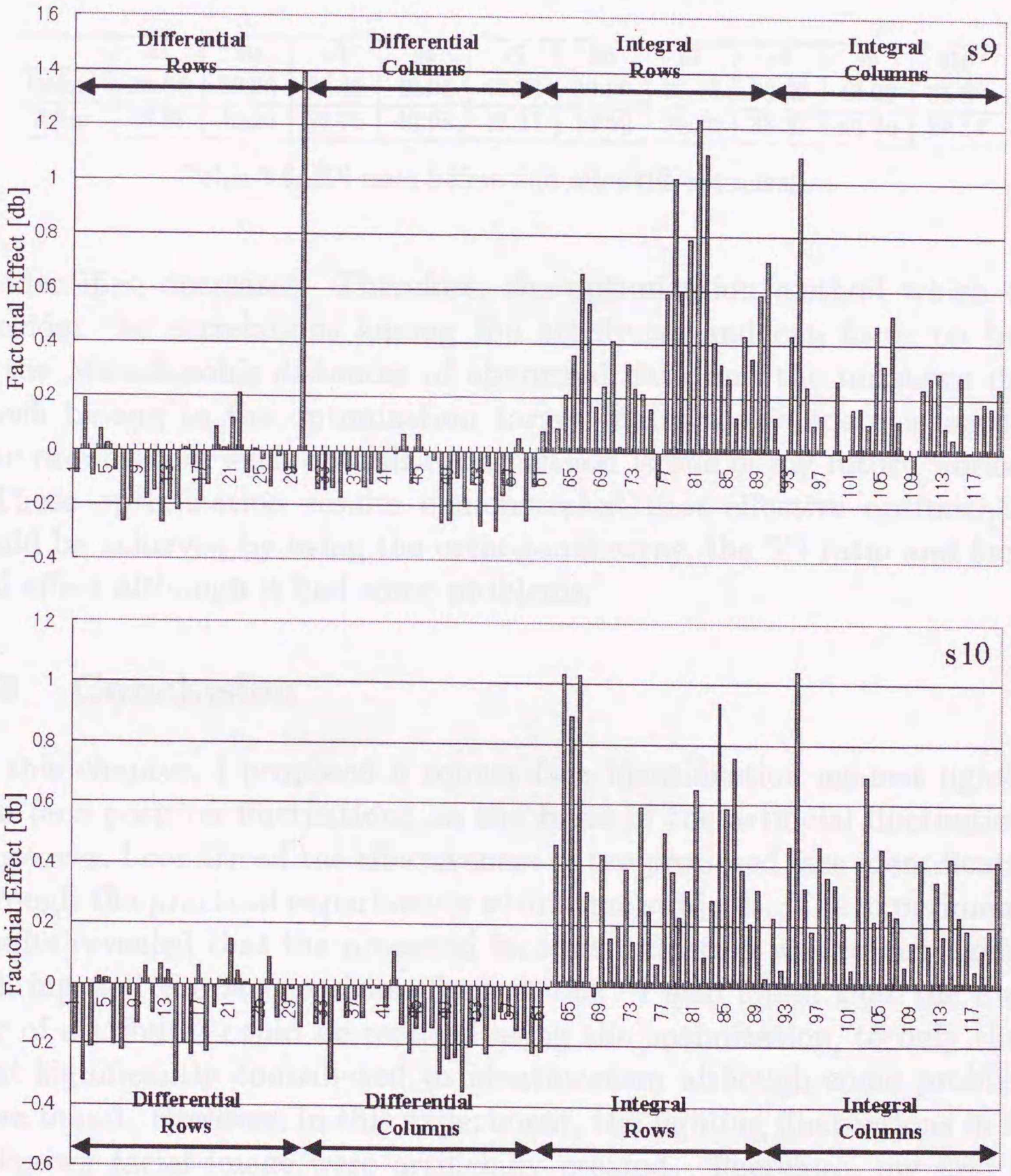


Figure 3.15: Factorial effects on each subject (E)

s1	s2	s3	s4	s5	S6	s7	s8	s9	s10
100	100	100	75.7	100	100	100	100	100	100

Table 3.4: Identification result after the optimization

	s1	s2	s3	s4	s5	S6	s7	s8	s9	s10
Before	38.86	39.96	38.89	40.91	42.26	39.09	39.13	38.49	43.23	38.88
After	36.61	40.20	39.75	40.98	42.21	39.50	39.80	38.37	44.20	39.34

Table 3.5: SN ratio before and after the optimization

ly identified decreases. Therefore, the optimization method which can consider the correlations among the attributes and can focus on both of the Mahalanobis distances of abnormal data and the unknown data which belong to the optimization target Mahalanobis space is needed. The proposal for such optimization method is one of my future works.

These optimization results demonstrated that effective optimization could be achieved by using the orthogonal array, the SN ratio and factorial effect although it had some problems.

3.5 Conclusion

In this chapter, I proposed a robust face identification against lighting and face position fluctuations on the basis of the artificial fluctuations. Moreover, I confirmed the effectiveness of the proposed face identification through the practical experiments with some subjects. The experimental results revealed that the proposed face identification was robust against lighting and the face position fluctuations. I also found that the number of attributes could be reduced using the optimization, to only those that significantly contributed to identification although some problems were found. However, in this experiment, the lighting fluctuations in the unknown facial image were artificially created. Therefore, my first future work is to confirm the robustness of the proposed face identification against natural lighting fluctuations.

In this chapter, I employed one hundred twenty attributes on a facial image. It is likely that the correct identification rate can be improved by concentrating the lines for the extraction of attributes to the effective region for the identification which was obtained by the optimization. In addition, the robustness against face size fluctuation has to be im-

plemented in the face identification. A more effective face identification system could be designed by taking this fluctuation into account. The development of such a face identification is the aim of my future work.

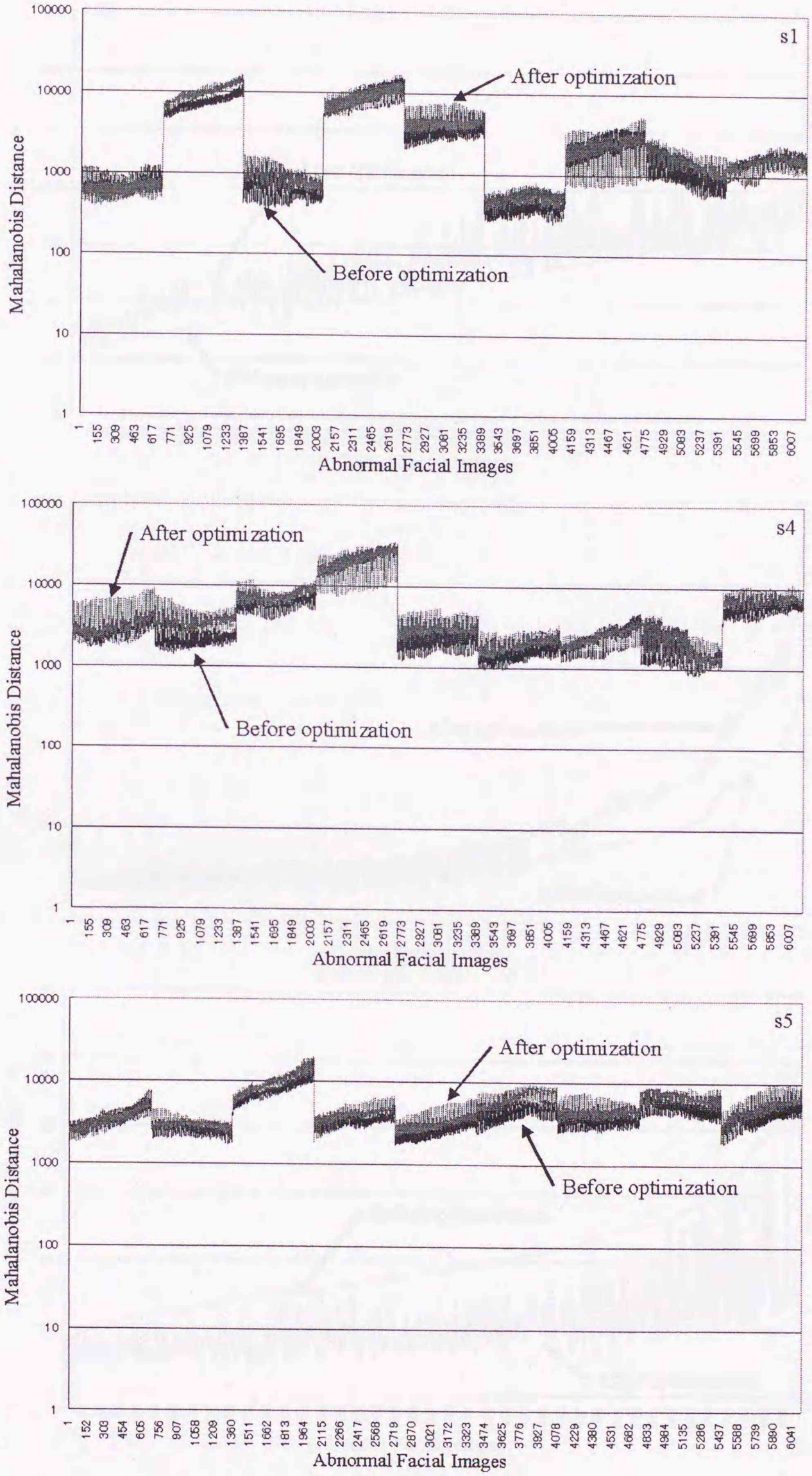


Figure 3.16: Mahalanobis distances before and after the optimization between abnormal facial images and the optimization target Mahalanobis space

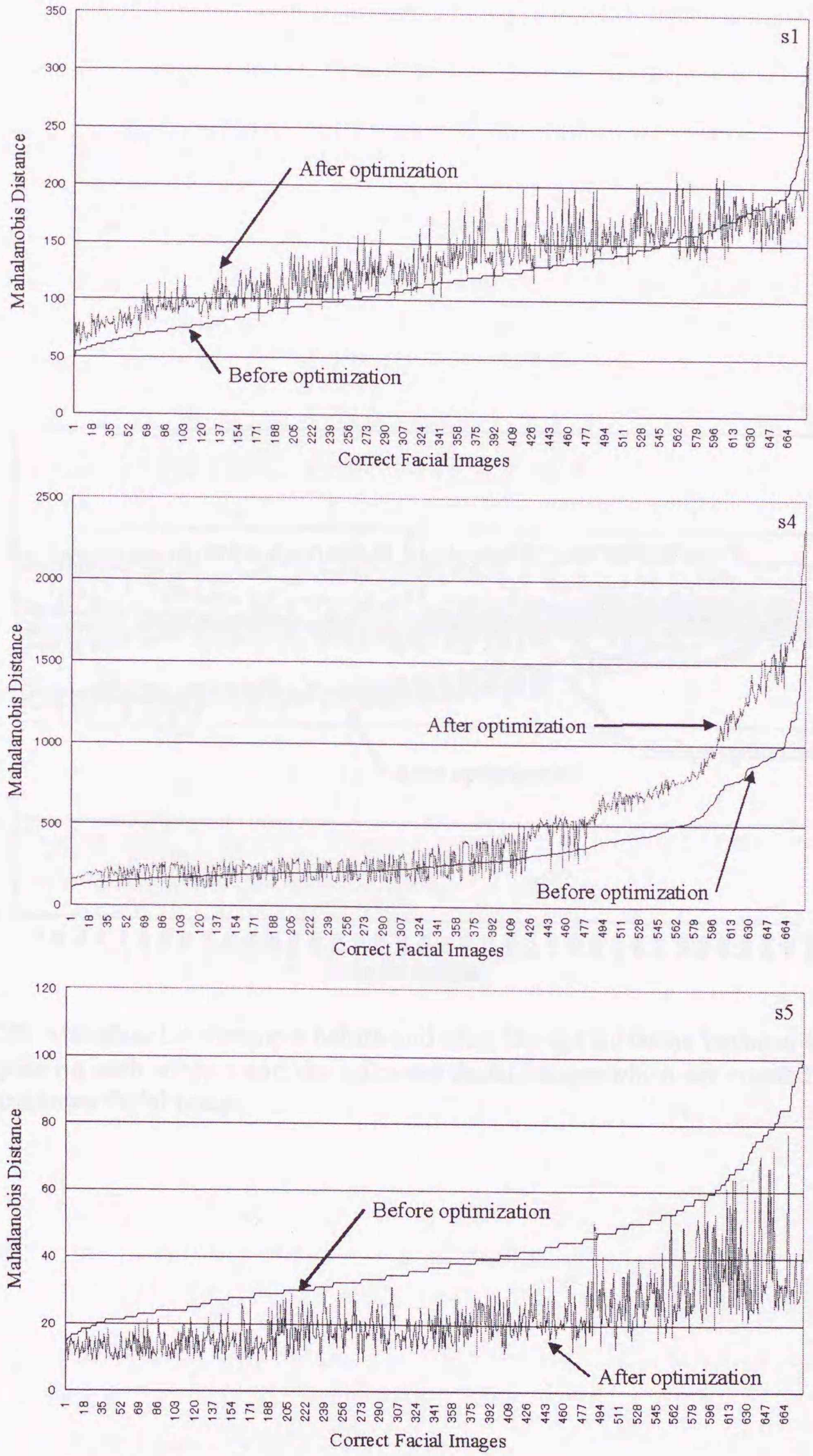


Figure 3.17: Mahalanobis distances before and after the optimization of unknown facial images which should belong to the optimization target Mahalanobis space

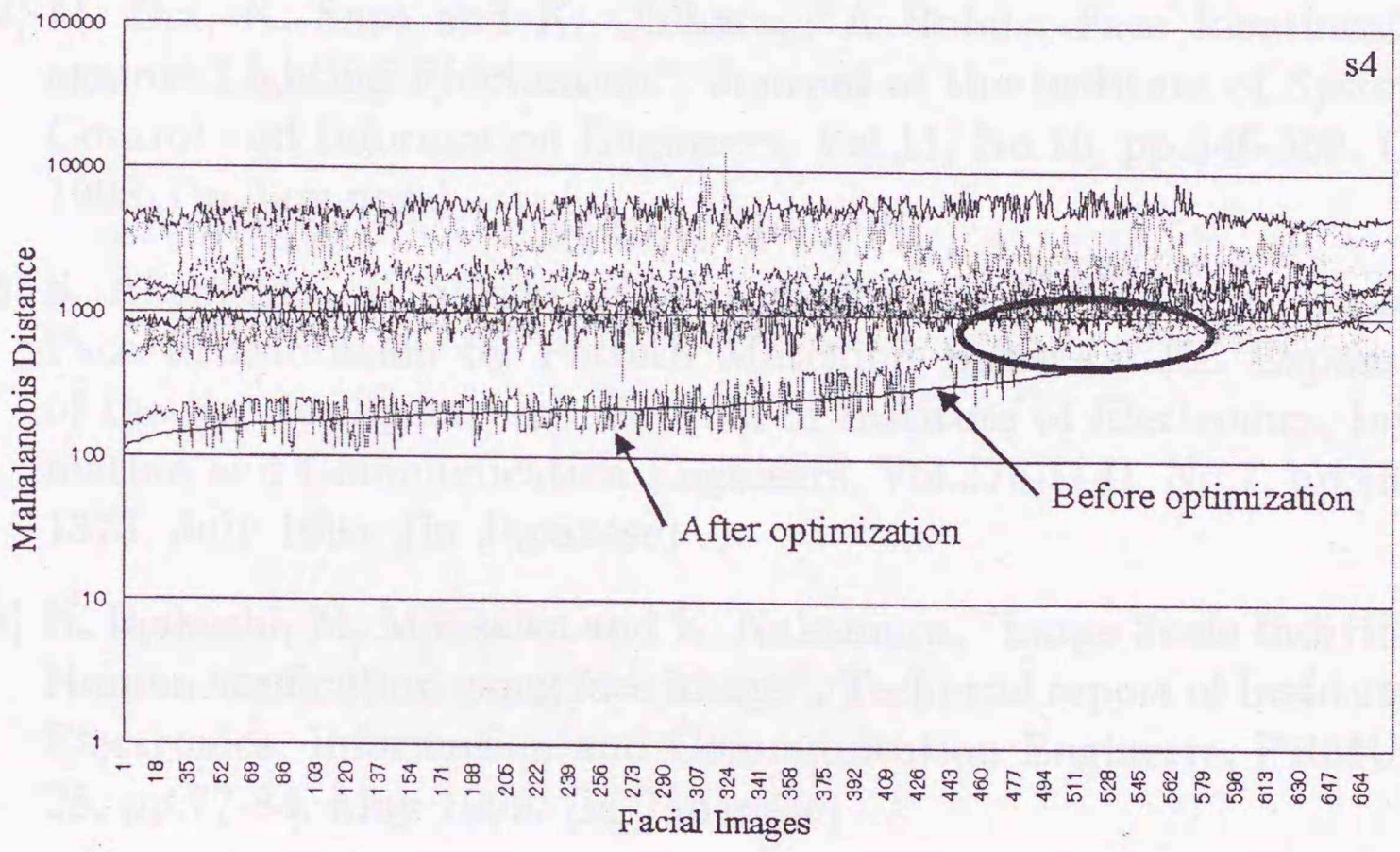


Figure 3.18: Mahalanobis distances before and after the optimization between the Mahalanobis space on each subject and the unknown facial images which are created from s4's original unknown facial image

Bibliography

[1] H. Sakano, "The State of art of Biometric Person Authentication", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-29, pp.75-82, June 1999. (in Japanese)

[2] M. Doi, K. Sato and K. Chihara, "A Robust Face Identification against Lighting Fluctuation", Journal of the Institute of Systems, Control and Information Engineers, Vol.11, No.10, pp.546-553, Oct. 1998. (in Japanese)

[3] S. Akamatsu, T. Sakaki, H. Fukamachi and Y. Suenaga, "Robust Face Identification by Pattern Matching Based on KL Expansion of the Fourier Spectrum", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J76-D-II, No.7, pp.1363-1373, July 1993. (in Japanese)

[4] N. Inakoshi, M. Maekawa and K. Nakamura, "Large-Scale Individual Human verification using face image", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU98-23, pp.77-84, May 1999. (in Japanese)

[5] H. Imaoka and S. Sakamoto, "Pose-Independent Face Recognition Method", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-26, pp.51-58, June 1999. (in Japanese)

[6] G. Taguchi, "Mathematics for Quality Engineering -Quality Engineering as a general measure of functionality", Journal of Quality Engineering Society, Vol.6, No.6, pp.5-10, Dec. 1998. (in Japanese)

[7] M. Nakatsugawa, M. Yamamoto, K. Suzuki, A. Ohuchi, D. Jin and S. Teshima, "Prediction of Kan-sei Model using MTS", Proceedings of The Australia Society for Operations Research Inc. Queensland Branch and ORSJ Hokkaido Chapter Joint Workshop, pp.914-926, July 1999.

[8] M. Nagao, M. Yamamoto, K. Suzuki and A. Ohuchi, "Acceleration of Evolutionary Image Retrieval System -Introduction of Facial Expression Recognition", Proceedings of The National Conference incorporating ASOR Queensland Branch and ORSJ Hokkaido Chapter Joint Workshop (ASOR'99), pp.898-913, July 1999.

[9] M. Nagao, M. Yamamoto, K. Suzuki and A. Ohuchi, "Application of the MTS Method for Facial Expression Recognition", Journal of The Institute of Electrical Engineers of Japan, Vol.120-C, No.8/9, pp.1157-1164, Aug. 2000. (in Japanese)

[10] Y. Gong, "INTELLIGENT IMAGE DATABASES Towards Advanced Image Retrieval", KLUWER ACADEMIC PUBLISHERS, Massachusetts, US, 1998.

[11] G. Taguchi, "Quality Engineering for Pattern Recognition", Journal of Quality Engineering Society, Vol.3, No.2, pp.2-5, April 1995. (in Japanese)

[12] G. Taguchi, "Pattern Recognition and Quality Engineering", Journal of Quality Engineering Society, Vol.3, No.4, pp.2-5, Aug. 1995. (in Japanese)

Chapter 4

Face Identification based on Division of Facial Image

4.1 Introduction

Recently, the research on face identification has been carried out in order to develop convenient security systems or access control systems [1][2][3]. In the face identification system, one of important problems is the implementation of robustness against various fluctuations. These fluctuations are classified into two kinds, and which are photographic fluctuations such as lighting, pose and so on, and appearance fluctuations such as hairstyle, glasses and so on. The implementation of robustness against photographic fluctuations is frequently attempted in the recent studies on the face identification because the implementation of robustness against appearance fluctuations is more difficult problem than the photographic fluctuations in terms of object recognition. However, the problem of the robustness against appearance fluctuations has to be solved in order to develop practical face identification systems. The purpose of my study is to implement a robust face identification against appearance fluctuations.

In conventional face identification, the attributes used for the identification are extracted from a whole of a facial image. The pattern recognition to measure the similarity between these extracted attributes and the attributes which are extracted from unknown facial image is performed. The identification result is decided on the basis of the pattern recognition result. However, it is difficult to implement robust face identification against appearance fluctuations because the attributes extracted from a whole of a facial image include a large quantity of noise which is caused by glasses, hairstyle, background and so on.

In order to solve this problem, a face identification based on the internal features on face such as mouth, nose and eye has been proposed [4][5].

In this method, the internal features on face are extracted from a facial image. The pattern recognition is performed for each internal feature. The identification is carried out on the basis of the pattern recognition result of each internal feature. However, it is difficult to extract stably the internal features from face, and the identification result may be influenced by the accuracy of extraction of internal features on face. Moreover, it is ineffective for the face identification systems which require real-time processing in terms of computational cost.

In this chapter, I propose a face identification method based on the division of facial image in order to implement the robustness against appearance fluctuations. In the proposed method, the internal features are not extracted but facial image is divided into some simple regions. The pattern recognition is performed on each region. The identification is performed by integrating the recognition result obtained on each region. It is likely that the proposed face identification method can implement the robustness against appearance fluctuations because facial image can be divided into the regions which include a large quantity of noise and the regions which do not include so much noise. That is, the face identification based on the accurate information which does not include so much noise can be carried out.

Moreover, in the proposed method, I employ the Mahalanobis-Taguchi System (MTS) as the pattern recognition method [6]. The MTS is one of statistical pattern recognition methods which can perform the recognition based on the correlations among the attributes in the multi-dimensional data. In addition, the robustness against noises in unknown datum can be implemented by giving artificial noises to the training data. It is likely that these advantages are effective for the face identification which requires the consideration of relations among the internal features on face and the robustness against the fluctuations in real environment. Furthermore, I confirmed that the MTS could allow the effective implementation of the robust face identification against lighting fluctuation and face position fluctuation [7][8]. Therefore, it is likely that the MTS will be effective for this proposed method. Finally, in this chapter, I confirm the effectiveness of the proposed face identification method through practical experiment with some subjects.

This chapter is organized as follows: In Section 2, I will propose a face identification method based on the division of facial image. In this section, facial image used in this study, artificial photographic fluctuations for the implementation of robustness, identification process based on the

division of facial image, pattern recognition on each region and identification method based on the integration of some pattern recognition results will be explained. The experimental method will be described in Section 3. In Section 4, the experimental result will be discussed. In Section 5, I will mention the features and future works of the proposed method. Finally, I will finish this chapter with the conclusions.

4.2 Face Identification Method based on Division of Facial Image

4.2.1 Facial Image

The facial images used in this study are the frontal and stationary facial images taken by a monochrome camera. The details of the facial images are as follows.

- The size of facial images is 300 $\times$ 300 pixels.
- The gradation of images is 256 gray scale.
- The background color is white.
- The images include hair and glasses.
- The images do not express any facial expressions.

Figure 4.1 shows an example of facial images used in this study. These facial images are taken in my laboratory without special lighting devices. Moreover, these facial images are taken by using baselines for the position of eyes, nose and mouth (white lines in Figure 4.1). However, these facial images include the fluctuations of face size and position as shown in Figure 4.1 because the accuracy of photography is different on each person.

4.2.2 Identification Process based on Division of Facial Image

The identification process based on the division of facial image is illustrated in Figure 4.2. First of all, the facial images are collected on each person. Secondly, the artificial photographic fluctuations are inserted into the facial image collected so as to implement the robustness against photographic fluctuations in real environment. The artificial photographic fluctuations are lighting, face position and face size fluctuations. The

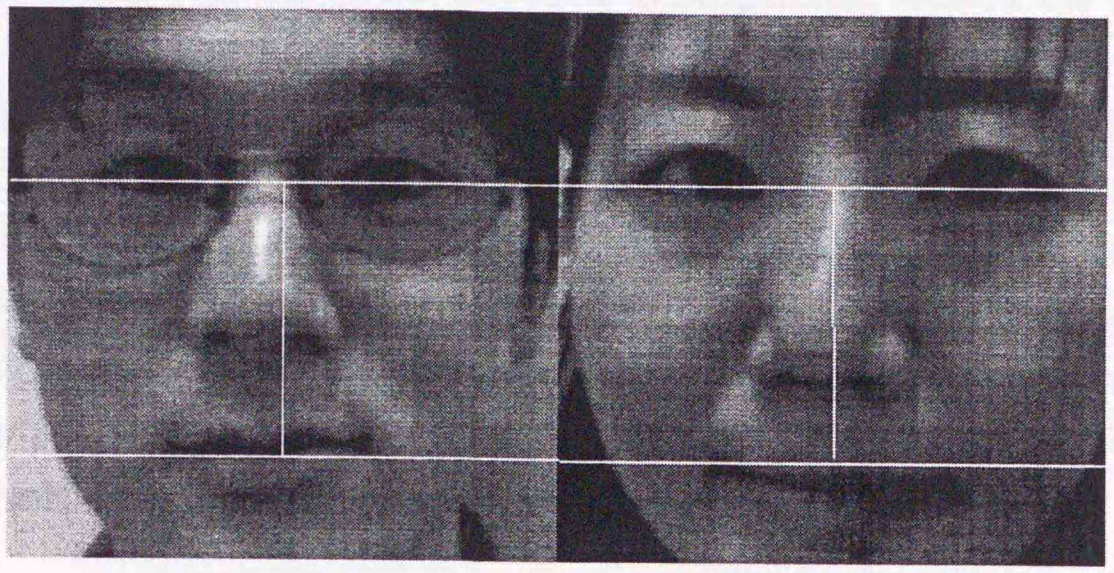


Figure 4.1: An example of facial images

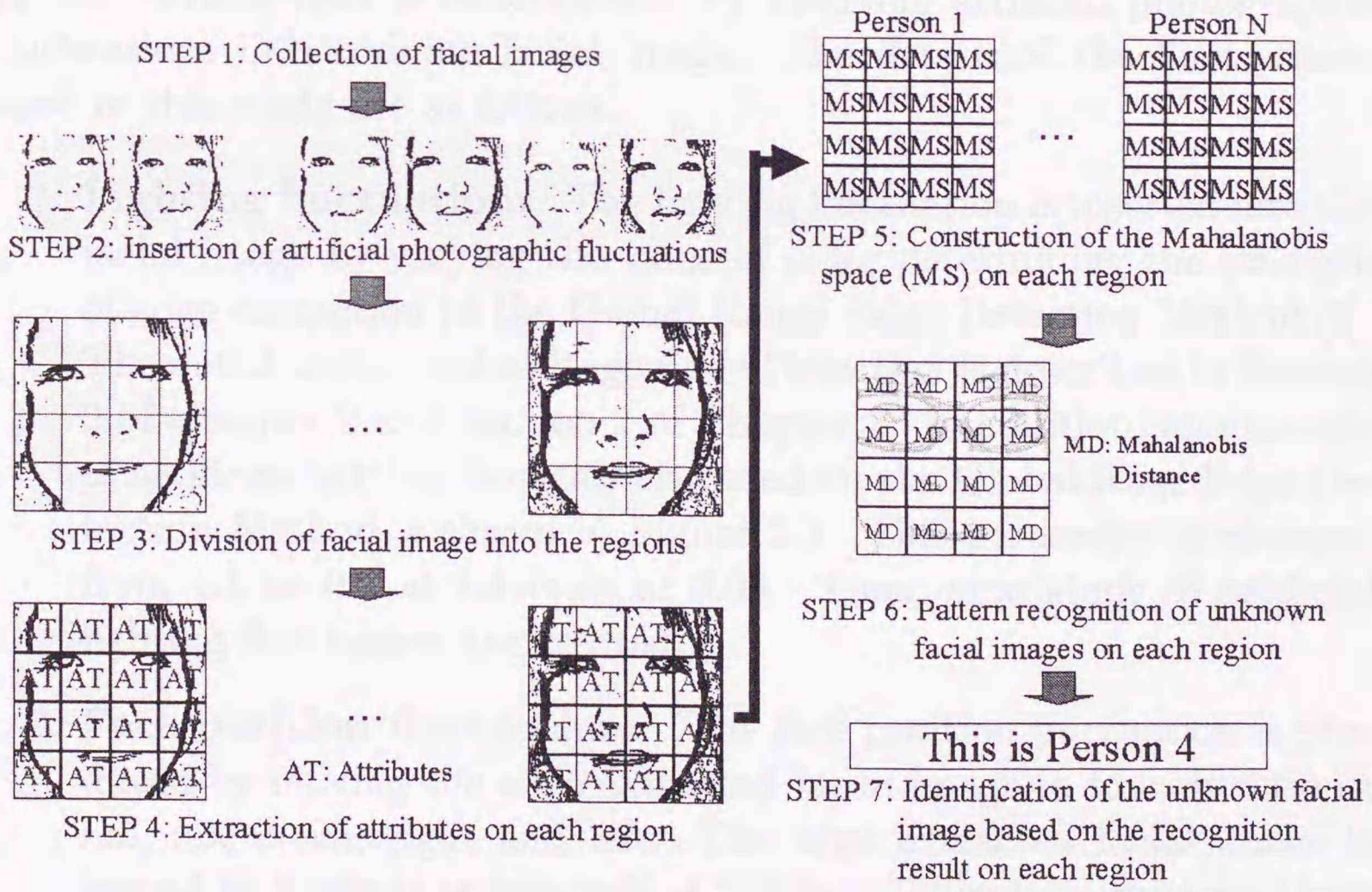


Figure 4.2: Identification process based on the division of facial image

facial image which includes the artificial photographic fluctuations is divided into some regions, and then the attributes for the identification are extracted on each region. The pattern recognition based on the MTS is carried out by using the Mahalanobis space which is constructed from the multi-dimensional training data and the Mahalanobis distance which is the evaluation criterion of similarity. Therefore, the Mahalanobis space is constructed from the attributes on each region of each person. The pattern recognition of unknown facial image is executed on each region of each person. The method judges which person the unknown facial image belongs to on the basis of the Mahalanobis distance on each region of each person. The detail of each step is shown below.

4.2.3 Artificial Fluctuation

In the pattern recognition based on the MTS, the robustness against noises can be implemented by giving artificial noises to the training data because the MTS performs the pattern recognition based on the correlations among attributes which is common in the multi-dimensional training data. In this study, the robustness against photographic fluctuations in real environment is implemented by inserting artificial photographic fluctuations into training facial image. The details of the fluctuations used in this study are as follows.

1. **Lighting fluctuation:** The lighting fluctuation is inserted into the facial image by varying the value of α for determining the strength of edge extraction in the Global Range Edge Detection Method [9]. The detail of the Global Rage Edge Detection is described in Section 2 of Chapter 2 and Section 2 of Chapter 3. In addition, an example of artificial lighting fluctuations based on the Global Rage Edge Detection Method is shown in Figure 3.1. This parameter is changed from 1.1 to 0.7 at intervals of 0.05. Thus, nine kinds of artificial lighting fluctuation are created.

2. **Face position fluctuation:** The face position fluctuation is produced by moving the edge extracted facial image in four directions; i.e., up, down, right and left. The edge extracted facial image is moved to 6 pixels at intervals of 2 pixels. Therefore, thirteen kinds of face position fluctuation are produced.

3. **Face size fluctuation:** The edge extracted facial image is expanded and shrunk [10]. The parameter for expansion is varied between

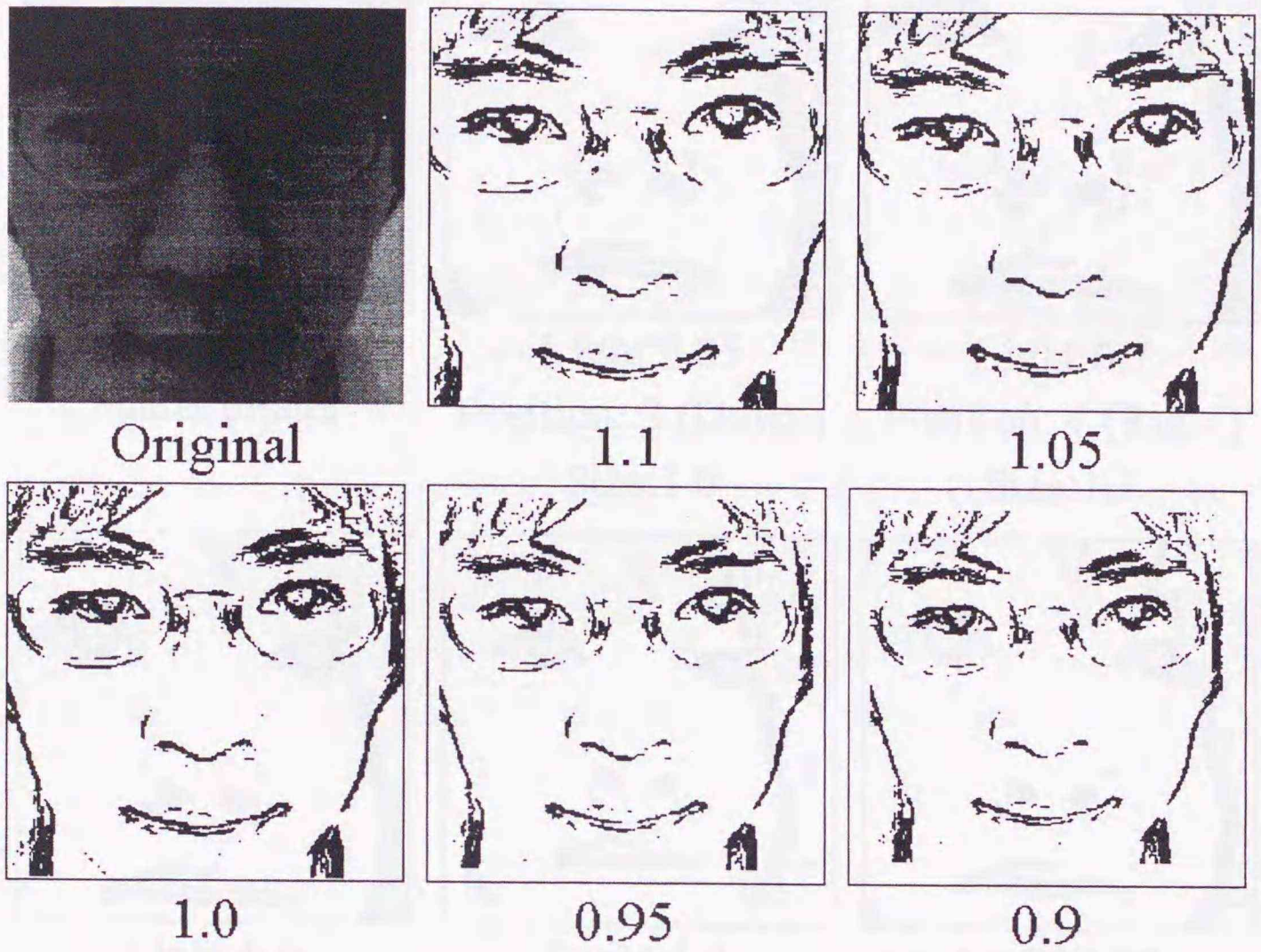


Figure 4.3: An example of artificial face size fluctuations

1.1 and 0.9 at intervals of 0.05. Thus, five kinds of face size fluctuation are created. An example of face size fluctuations is shown in Figure 4.3.

The combination of these fluctuations is applied to one original training facial image. Therefore, five hundred eighty-five training facial images involving the artificial photographic fluctuations are created from an original training facial image. An example of training facial images which include the combination of these artificial photographic fluctuations is shown in Figure 4.4.

4.2.4 Pattern Recognition based on MTS

The training facial images which include the artificial photographic fluctuations are divided into some regions. After the division, the attributes used for the pattern recognition are extracted on each region of the training facial images. In this study, the attributes based on differential characteristic and integral characteristic are employed. The Mahalanobis space is constructed on each region from these attributes extracted. And



Original Training
Facial Image



Light:0.85
Position: 2 (Down)
Size:1.0



Light:0.7
Position: 4 (Right)
Size:1.1



Light:0.9
Position: 6 (Down)
Size:1.1



Light:1.0
Position: 2 (Right)
Size:0.9



Light:1.05
Position: 2 (Top)
Size:1.05

Figure 4.4: An example of training facial images involving the combination of three artificial photographic fluctuations

then, unknown facial image is divided into same region as the training facial image. The pattern recognition based on the MTS is performed on each region of each person. Therefore, Mahalanobis distance is obtained on each region of each person as the result of pattern recognition for an unknown facial image. It is judged which person in facial image database the unknown facial image is by integrating some Mahalanobis distances for each person. The detail of each operation is described in the earlier chapters. The detail of extraction method based on the differential and integral characteristics are described in Section 6 of Chapter 1 and Section 3 of Chapter 2. The details of the construction method of Mahalanobis space and the calculation method of the Mahalanobis distance are described in Section 6 of Chapter 1.

4.2.5 Identification based on Integration of Mahalanobis distances

In the proposed face identification method, the Mahalanobis distance is calculated on each region of each person. The identification of unknown facial image is performed by integrating some Mahalanobis distances for each person. In this study, two methods are employed in order to integrate some Mahalanobis distances. The details of these methods are as follows.

1. **PROP(SUM):** The Mahalanobis distance between the unknown facial image and the Mahalanobis space is calculated on each region of each person. The sum of Mahalanobis distance on each region is calculated for each person. The unknown facial image is judged to belong to the person which has the minimum total of these Mahalanobis distances. This process represented by formula is as follows.

$$D^{(p)} = \sum_{i=1}^{M} \sum_{j=1}^{N} d_{ij}^{(p)} \tag{4.1}$$

$$\min(D^{(1)}, \cdots, D^{(p)}) \tag{4.2}$$

$D^{(p)}$	:	Sum of Mahalanobis distance on each region for person p
$d_{ij}^{(p)}$	:	Mahalanobis distance on region (i, j) of person p
p	:	The number of persons in facial image database
M,N	:	The number of division ($M \times N$)

2. **PROP(MIN):** The person which has the minimum Mahalanobis distance between the unknown facial image and the Mahalanobis space is determined on each region. The number of region which has the minimum Mahalanobis distance is counted on each person. The unknown facial image is judged to belong to the person which has the maximum number of the regions that have the minimum Mahalanobis distance. This process is formulated as follows.

$$R^{(p)} = \sum_{i=1}^{M} \sum_{j=1}^{N} r_{ij}^{(p)} \tag{4.3}$$

$$\max(R^{(1)}, \cdots, R^{(p)}) \tag{4.4}$$

$$r_{ij}^{l} = \begin{cases} 1 & : \quad d_{ij}^{(l)} = \min(d_{ij}^{(1)}, \cdots, d_{ij}^{(p)}) \\ 0 & : \quad \text{otherwise} \end{cases}$$

$R^{(p)}$: The number of region which has the minimum Mahalanobis distance for person p

$r_{ij}^{(l)}$: The variable to find the region which has minimum Mahalanobis distance

4.3 Experimental Method

I designed the following experiments with some subjects in order to confirm the effectiveness of the proposed method. The detail of experimental method is described below.

In this experiment, twenty-six students were recruited from Hokkaido University. I took two facial images on each subject. One facial image was used as the original training facial image, and another was used as the original unknown facial image. I took the original unknown facial image at intervals of several days after having taken the original training facial image in order to take the original unknown facial image involving natural appearance fluctuations. Figure 4.5 shows an example of original training facial images and original unknown facial images on some subjects. As shown in Figure 4.5, the original unknown facial images have a little difference from the original training facial images in terms of photography and appearance.

Five hundred eighty-five training facial images were created on each subject by inserting the artificial photographic fluctuations, described in Section 2, into one original training facial image. The Mahalanobis space on each subject was constructed on the basis of these training facial images involving the artificial photographic fluctuations.

In this experiment, I divided the facial image into 4×4 regions (16 regions). It is likely that the noises can be removed from the effective region for the identification in the case where facial image is divided into many regions. However, the division of facial image into too many regions produces the region which does not include any edges, which has bad effect upon the identification. Therefore, I employed 4×4 as the number of division so that such region is not created.

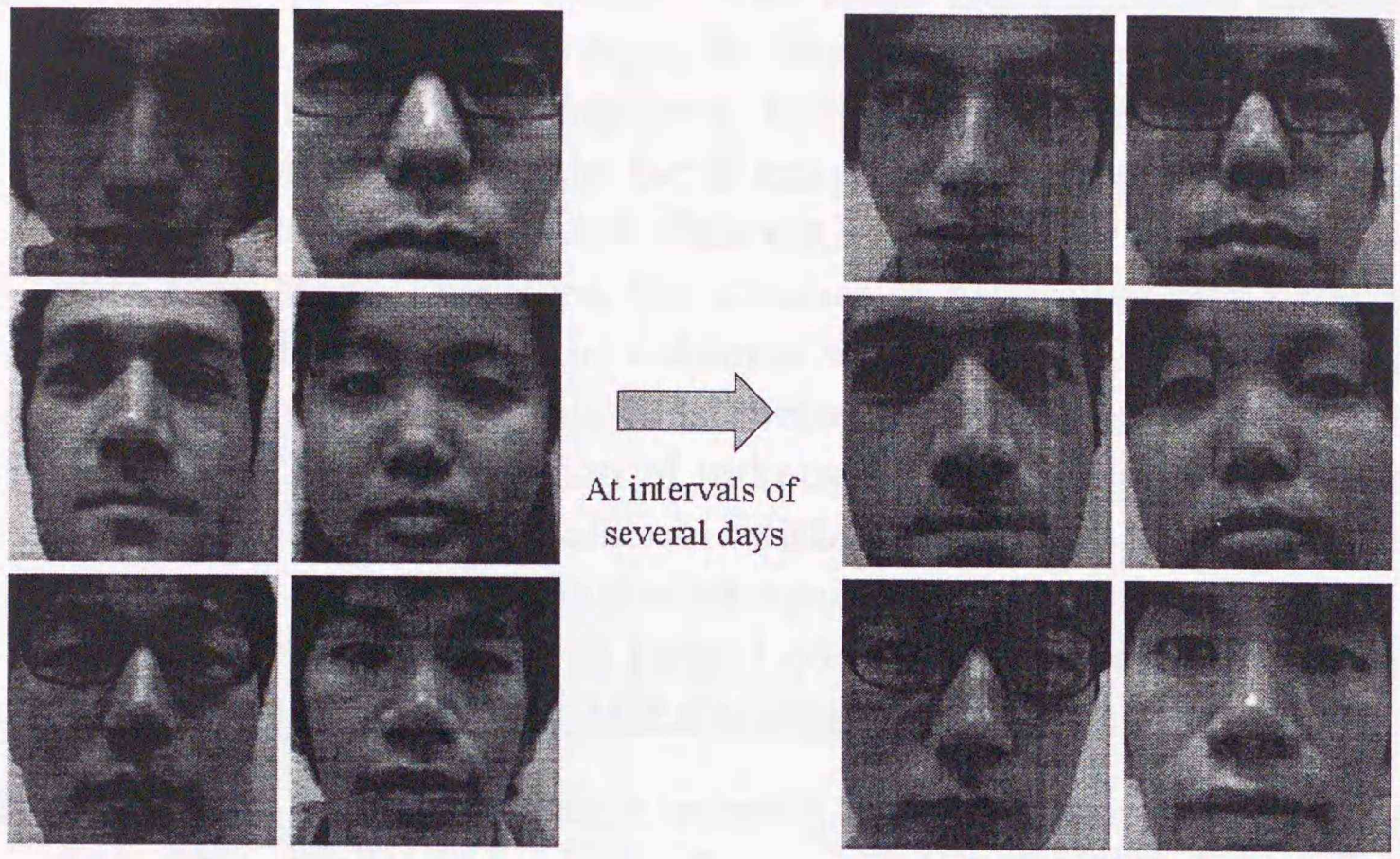


Figure 4.5: An example of original training facial images and original unknown facial images

I drew the parallel lines on each region of the facial image at intervals of 5 pixels for the extraction of attributes based on the differential and integral characteristics. Therefore, the number of attributes along the rows was 30 and that down the columns was also 30 on each region. I constructed the Mahalanobis space from 60 attributes for each region.

Moreover, in this experiment, I used the facial images which included the artificial photographic fluctuations as the unknown facial images, which are created by inserting the artificial photographic fluctuations into the original unknown facial image. These artificial fluctuations are described in Section 2. However, the amount of face position fluctuation was changed to 4 pixels from 6 pixels. Therefore, the number of unknown facial images for the identification was 10530 (26 persons$\times$405 unknown facial images).

In the proposed method, the Mahalanobis distance which was calculated on each region was treated as 1000 if the Mahalanobis distance exceeds 1000 because too large Mahalanobis distance is ineffective for PROP(SUM).

Moreover, I designed two experiments using the conventional method in order to compare with the proposed face identification method.

1. **CNV(120):** In this method, the same training facial images involving the artificial photographic fluctuations that were used in the proposed method were employed. In this experiment, I drew the parallel lines on a whole of the facial image at intervals of 10 pixels, and the attributes based on the differential and integral characteristics were extracted. Therefore, the number of attributes along the rows was 60 and that down the columns was also 60. The Mahalanobis space on each subject was constructed on the basis of these 120 attributes. The identification of unknown facial image was performed on the basis of the Mahalanobis distance between the unknown facial image and the Mahalanobis space on each person. Thus, the unknown facial image was judged to belong to the Mahalanobis space having the shortest Mahalanobis distance.

2. **CNV(240):** In the same training facial images as CNV(120), I drew the parallel lines on a whole of facial image at intervals of 5 pixels. Thus, in this experiment, the number of attributes was 240. The identification method was equal with CNV(120).

4.4 Experimental Result

The experimental results are shown in Table 4.1 and Figure 4.6. Table 4.1 shows the rate of correct identification on each method. In table 4.1, the results of conventional system were 51.0 % in CNV(120) and 59.0 % in CNV(240), respectively. On the other hand, the rate was 61.1 % in PROP(SUM) and 65.1 % in PROP(MIN). Especially, the identification rate of CNV(240) was lower than PROP(MIN) by 6.1 % though the lines for the extraction of attributes were drawn on same positions of facial image as PROP(MIN). This result revealed that the division of facial image could allow effective implementation of robust face identification.

Moreover, PROP(MIN) was superior to PROP(SUM) for the proposed method. In the pattern recognition based on the MTS, the Mahalanobis distance rapidly increases when the unknown datum differs from the training data which are used to construct the Mahalanobis space. Namely, the sum of Mahalanobis distance will become considerably large when the region in unknown datum differs from the region in the training data. Therefore, it is likely that PROP(MIN) could get high rate of correct identification in comparison with PROP(SUM).

Furthermore, I investigated the position of the region which had shortest Mahalanobis distance in unknown facial images which were correctly

CNV(120)	CNV(240)	PROP(SUM)	PROP(MIN)
51.0	59.0	61.1	65.1

Table 4.1: Identification rate (%)

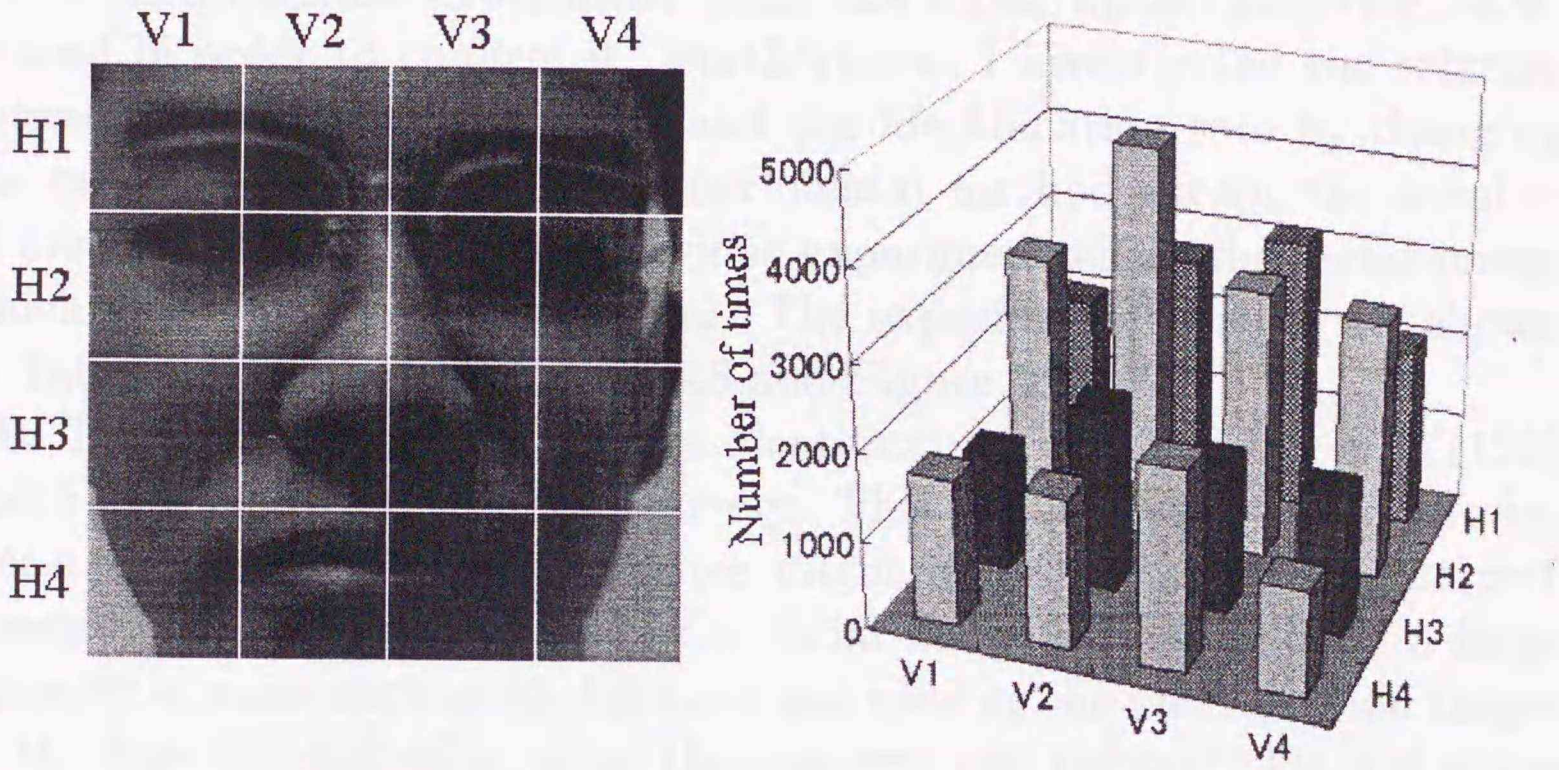


Figure 4.6: Frequency of shortest Mahalanobis distance on each region of unknown facial images which were correctly identified

identified. Figure 4.6 indicates the number of times that shortest Mahalanobis distance was indicated on each region of unknown facial images which were correctly identified. In Figure 4.6, the frequency on the regions of H2 and H1 was larger than other regions. It is likely that the frequency on these regions became large because these regions included the eyes and eyebrows which were effective information for the face identification. However, the frequency on the regions of H1 was lower than the regions of H2. It is reasonable to suppose that the frequency on the regions of H1 was small in comparison with the regions of H2 because there were many cases where the hair, i.e., noise, were included in the regions of H1. In addition, the average number of the regions which had the shortest Mahalanobis distance in an unknown facial image which was identified correctly was 4.9.

These results revealed that the proposed method could implement effective face identification by performing the correct recognition on the regions which did not include so much noise.

4.5 Discussion

In the previous experiment, I designed the experiment with the facial images which did not include background. However, it is likely that the proposed method is also effective in the case where the facial image involving background is used for the identification. Therefore, I designed a face identification experiment using the facial image involving background in order to confirm it. Furthermore, I investigated the relation between the number of division and the identification rate by changing the number of division. The experimental method except the number of division was same as the previous experiment using the facial image which did not include background. The experimental results are shown in Table 4.2, Figure 4.7, Figure 4.8 and Figure 4.9.

In Table 4.2, the rate of correct identification was 7.0 % in CNV(120) and 5.5 % in CNV(240), respectively. This result revealed that the increase of the number of attributes might decrease the rate of correct identification in the case where the facial image which included a large quantity of noise such as background was used as the identification target in the face identification using the conventional method. On the other hand, the rates of correct identification in the proposed method were higher than the conventional method. Therefore, it is obvious that the proposed method is more effective than the conventional method in the case where the facial images which include background are used for the identification. Moreover, the rate of correct identification increased in accordance with the increase of the number of division, and the highest rate of correct identification was indicated in the case of the division of 6×6. What this result makes clear is that the noises were removed from the region which included the internal features by increasing the number of division. Namely, the internal features on face could be effectively used for the identification.

Figure 4.7, Figure 4.8 and Figure 4.9 illustrate the number of times that shortest Mahalanobis distance was indicated on each region of unknown facial images which were correctly identified on each number of division. From these figures, I found that the frequency on the regions which included eyes and eyebrows was large in comparison with other regions. Therefore, it is likely that eyes and eyebrows were important internal features to identify person. Moreover, in these figures, the frequency on the regions which included face was large in comparison with the regions which included background. This result revealed that the proposed method avoided using the regions including background and

Conventional Method		Division 2×2	
CNV(120)	CNV(240)	PROP(SUM)	PROP(MIN)
7.0	5.0	7.9	8.0
Division 4×4		Division 6×6	
PROP(SUM)	PROP(MIN)	PROP(SUM)	PROP(MIN)
12.8	18.1	11.0	26.5

Table 4.2: Identification rate (%)

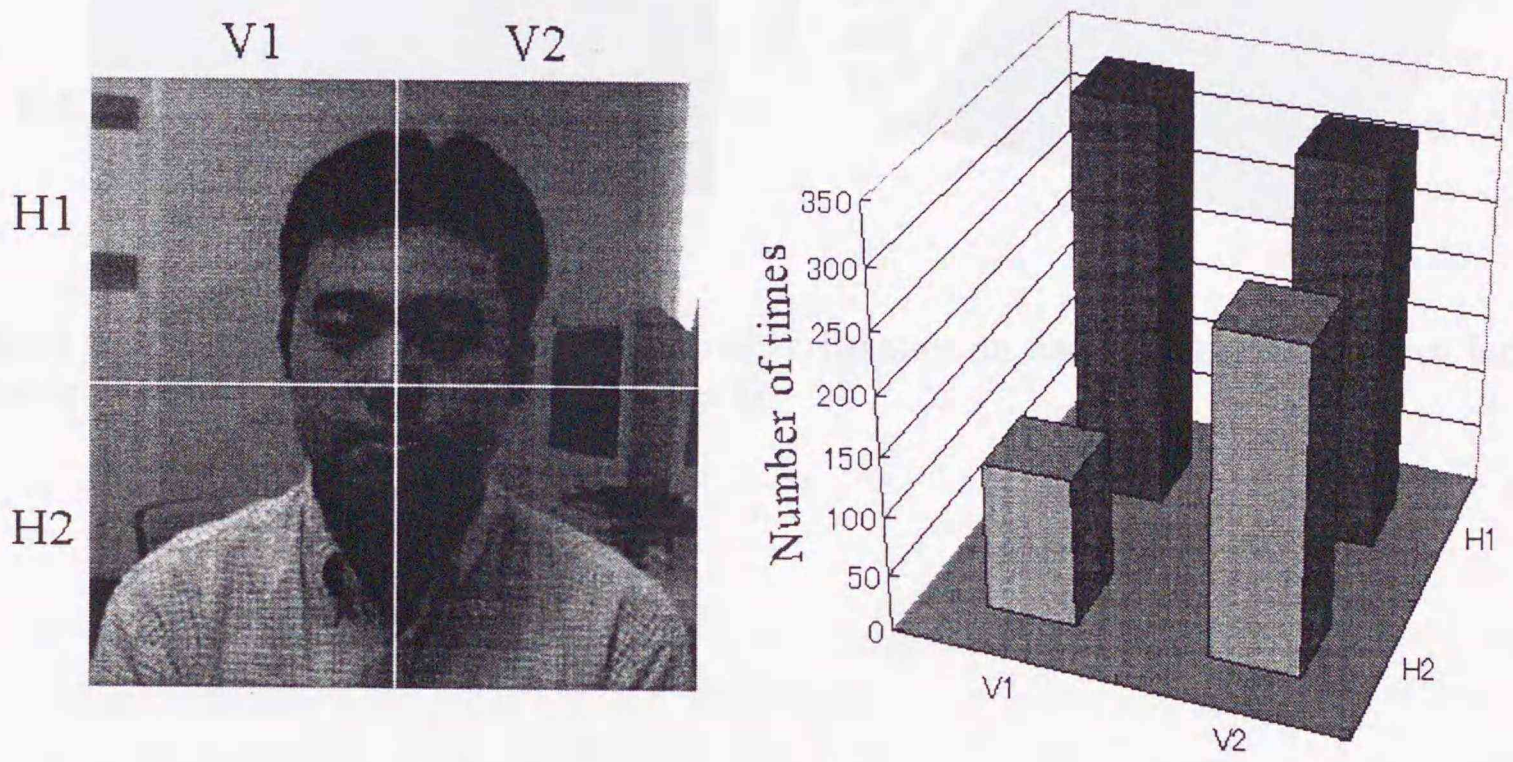


Figure 4.7: Frequency of shortest Mahalanobis distance on each region of unknown facial images which were correctly identified (2×2)

used effectively the region including face for the identification when the correct identification was performed.

In the case where the facial image including background was used, the highest rate was indicated in the division of 6×6. However, the regions which do not include any edges will be produced when the facial image will be divided into more regions. The region which does not include any edges is ineffective for the identification, so that may decrease the rate of correct identification. Furthermore, the region which includes only one internal feature will be also produced. The MTS is the pattern recognition method which can consider the correlations among the attributes. Therefore, it is likely that the identification which considers the relations among the internal features on face can not be performed when such regions are produced. In addition, the computational cost is high when

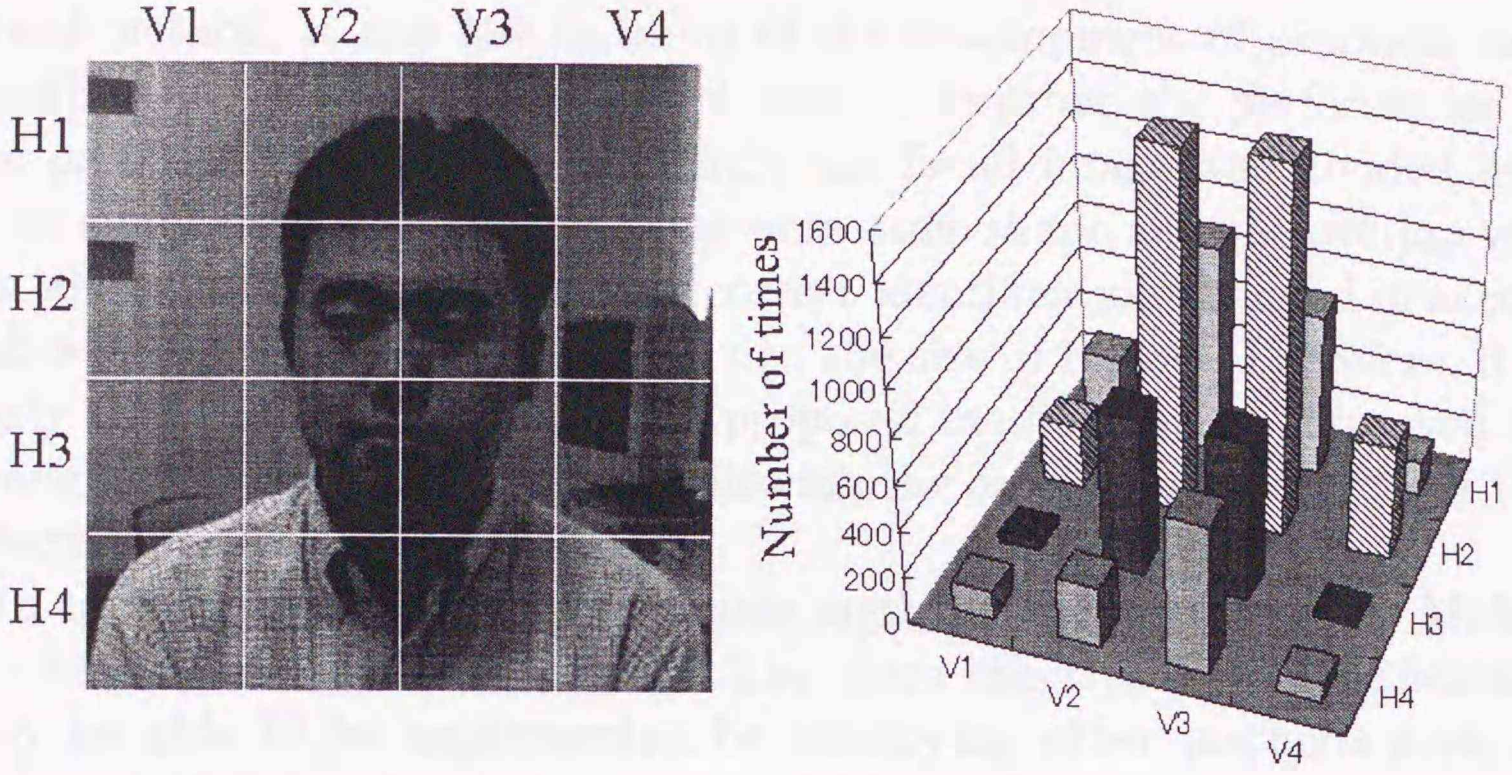


Figure 4.8: Frequency of shortest Mahalanobis distance on each region of unknown facial images which were correctly identified (4×4)

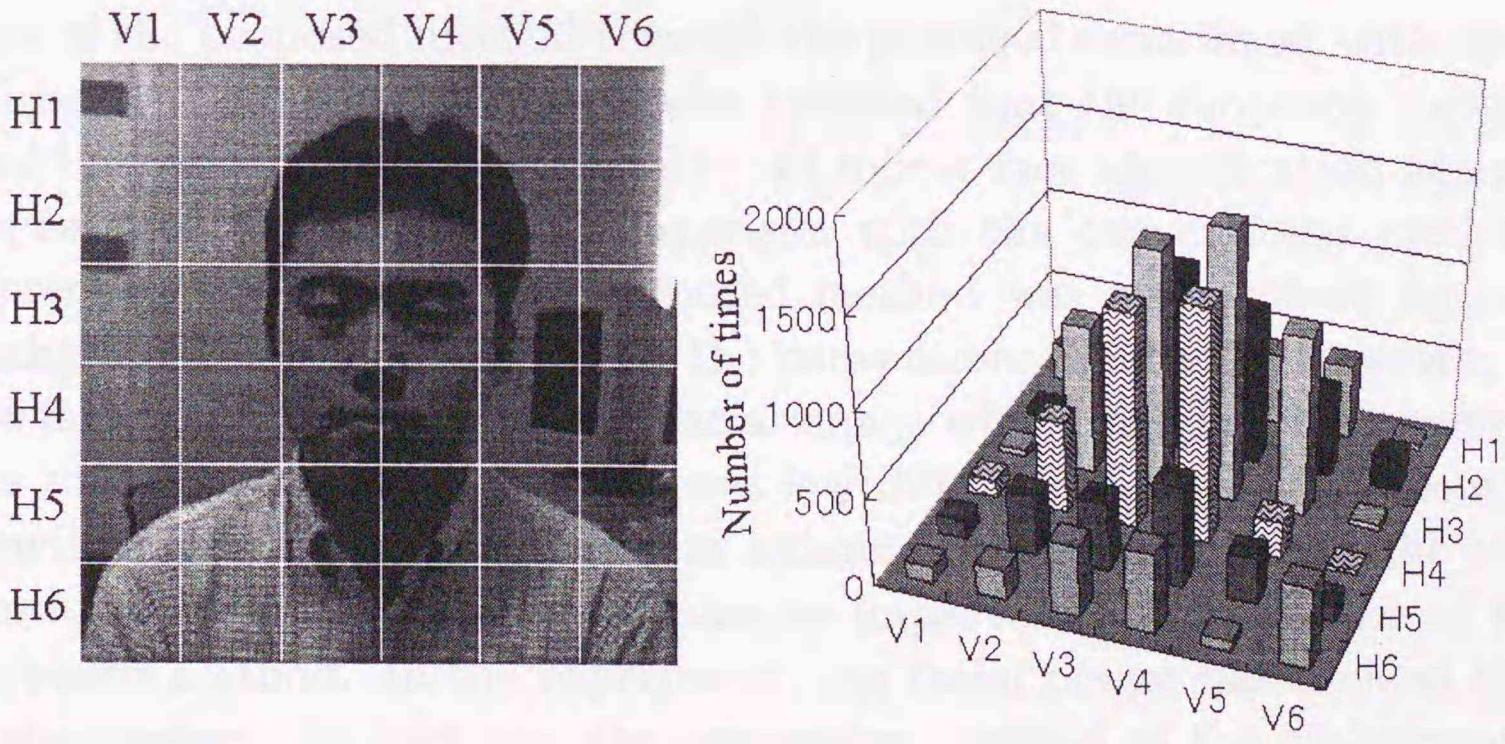


Figure 4.9: Frequency of shortest Mahalanobis distance on each region of unknown facial images which were correctly identified (6×6)

the number of regions is large. Therefore, appropriate number of division has to be decided in accordance with the purpose.

Although the rate of correct identification was higher than the conventional method, it was low in terms of the development of practical face identification system. Therefore, I have to improve the performance of the proposed method. In this study, the facial image was divided into some simple regions. These regions were same shape and the overlap was not allowed. However, the rate of correct identification changed in accordance with the number of regions, i.e., the size of region. Therefore, it is likely that the performance of the proposed method can be improved by changing the shape of region, by allowing the overlap or by searching the effective position of region.

In the proposed method, two simple methods to integrate some Mahalanobis distances were employed. The more effective face identification may be able to be implemented by employing other methods such as weighted Mahalanobis distances or considering the relation among these Mahalanobis distances. These are my future works.

4.6 Conclusion

In this chapter, I proposed a face identification method based on the division of facial image in order to implement robust face identification against appearance fluctuations. Moreover, I confirmed the effectiveness of the proposed method through the practical experiment with some subjects. The experimental results revealed that the proposed method could allow effective implementation of robust face identification against appearance fluctuations in comparison with the conventional method. Moreover, I found that the proposed method was more robust against background in comparison with the conventional method. However, in the face identification using the facial image which included background, the rate of correct identification was low in terms of the development of practical face identification system although it was higher than the conventional method. Therefore, I have to improve the performance of the proposed method. In the experiment, the facial image was divided into simple region. In addition, the integration method of the Mahalanobis distances which were calculated on these regions was simple. Therefore, it is likely that the performance of the proposed method can be improved by employing effective method for the division and the integration. These are my future works.

Bibliography

[1] H. Sakano, "The State of art of Biometric Person Authentication", Technical report of Institute of Electronics, Information and Communication Engineers, PRMU99-29, pp.75-82, June 1999. (in Japanese)

[2] H. Sakano and I. Nakamura, "A Current State of Art of Authentication Technology and a View of Future", Journal of Information Processing Society of Japan, Vol.41, No.7, pp.816-822, July 2000. (in Japanese)

[3] R. Sukthankar and R. Stockton, "ARGUS: An Automated Multi-Agent Visitor Identification System", Proceedings of The 1999 National Conference on Artificial Intelligence (AAAI'99), pp.208-213, July 1999.

[4] M. Doi, Q. Chen, A. Matani, O. Oshiro, K. Sato and K. Chihara, "Lock Control System Based on Face Identification", Journal of Institute of Electronics, Information and Communication Engineers, Vol.J80-D-II, No.8, pp.2203-2208, Aug. 1997. (in Japanese)

[5] M. Doi, K. Sato and K. Chihara, "A Robust Face Identification against Lighting Fluctuation", Journal of the Institute of Systems, Control and Information Engineers, Vol.11, No.10, pp.546-553, Oct. 1998. (in Japanese)

[6] G. Taguchi, "Mathematics for Quality Engineering -Quality Engineering as a general measure of functionality", Journal of Quality Engineering Society, Vol.6, No.6, pp.5-10, Dec. 1998. (in Japanese)

[7] M. Nagao, M. Yamamoto, K. Suzuki and A. Ohuchi, "MTS Approach to Facial Image Recognition", Proceedings of The 1999 IEEE Systems, Man, and Cybernetics Conference (SMC'99), Vol.IV, pp.937-942, Oct. 1999.

[8] M. Nagao, M. Yamamoto, K. Suzuki and Azuma Ohuchi, "A Robust Face Identification System using MTS", Proceedings of 17th Annual ASI's Taguchi Methods Symposium, pp.124-142, Oct. 1999.

[9] Y. Gong, "INTELLIGENT IMAGE DATABASES Towards Advanced Image Retrieval", KLUWER ACADEMIC PUBLISHERS, Massachusetts, US, 1998.

[10] N. Yagi, S. Inoue, M. Hayashi, K. Nakasu, K. Mitani, M. Okui, S. Suzuki and Y. Kanatsugu, "Practical Image Processing using C Programming Language", Ohmsha Publishers, Tokyo, Japan, 1993. (in Japanese)

Chapter 5

Face Cognitive Model Acquisition using Interactive Genetic Algorithm

5.1 Introduction

As a research theme of the engineering, human face has been focused on in order to develop effective and user-friendly human interfaces [1][2]. Especially, the development of the face recognition methods to recognize or understand the information which face represents in the face-to-face communication is required. As one of evaluation of face recognition methods, the development of face recognition methods which have the consistency with human cognitive structure of face is taken up. Therefore, it is necessary to acquire the scientific knowledge about face cognitive model of human and reflect the acquired knowledge to the face recognition methods. Namely, the structure to recognize face which human has must be acquired. From this background, the research on the face cognitive model acquisition of human has been carried out.

Oda et al. proposed a facial image database retrieval task method as a computer-based method to acquire the face cognitive model of human [3][4]. In this method, some facial images stored in a facial image database are simultaneously presented to a subject. The subject is asked to select many facial images which are suitable for a particular retrieval goal from the presented facial images. The subject evaluates the presented facial images and then selects the facial images which satisfy the retrieval criterion. The set of facial images judged to satisfy the retrieval criterion by the subject is retrieved by repeating this operation. The retrieved facial images reflect the face cognitive model of the subject because the facial images are retrieved on the basis of the face cognitive model for a particular retrieval goal. Therefore, the face cognitive model can be acquired by investigating the factors which are common in these facial

images retrieved. In this method, the subject only selects the suitable face images for a particular retrieval goal. Thus, this method has the advantage that it is easy to acquire the accurate face cognitive model for various face impressions or targets by its indirect or implicit approach to the acquisition of face cognitive model.

In this method, it is ideal that the subject decides the suitable facial images for the retrieval goal after evaluating all facial images which are stored in the facial image database in order to implement the acquisition of accurate face cognitive model because the subject does not know the kinds of facial images stored in the facial image database. However, it is considerably difficult to evaluate all facial images stored in the facial image database. Therefore, if the subject evaluates the facial images stored in a part of the facial image database with compromise, it is difficult to acquire the accurate face cognitive model. Moreover, in the case where the subject does not know the facial images stored in the facial image database, it is likely that the subject gradually forms a new face cognitive model which is suitable for the facial images stored in facial image database during the retrieval task. Therefore, the function to assume the face cognitive model of the subject on each step and to present appropriate facial images for the recent face cognitive model of the subject is required in order to acquire more accurate face cognitive model of the subject in the face cognitive model acquisition based on the facial image database retrieval task method.

In this chapter, I propose a face cognitive model acquisition method using the Interactive Genetic Algorithm (IGA) in order to acquire accurate face cognitive model in the facial image database retrieval task method. The IGA is a method to generate the individuals which reflect human subjectivity or Kan-sei by introducing the artificial selection by a human instead of the natural selection in the conventional Genetic Algorithm (GA). The IGA is focused on the recent development of Kan-sei engineering or human interface, and it is applied to various fields [5][6]. In addition, the research on the facial image retrieval based on the IGA is also carried out and it is reported that the facial images reflecting human subjectivity or Kan-sei can be retrieved by using the IGA [7]. However, in the facial image retrieval using the conventional IGA, only facial images selected by the user in each evaluation step are employed as the parents in next generation. The children are generated by applying the genetic operators to these parents and then the facial images are retrieved on the basis of these children generated. In this case, it is likely that the

facial images presented to the user may be limited in the facial image database. Namely, the initial convergence problem in the search problem based on the GA may arise. Therefore, it is difficult to acquire accurate face cognitive model in the case where the conventional IGA is merely applied to the problem of the face cognitive model acquisition. In the proposed method, I improve the conventional IGA so that the facial images which reflect the face cognitive model of the subject and have the diversity can be presented to the subject. In the face cognitive model acquisition based on facial image database retrieval task method, the subject gradually forms an appropriate face cognitive model for the facial images stored in the facial image database through the evaluation of the facial images. It has been reported that the context of stimuli in successive presentation influences the cognitive model formation process in the human brain. Therefore, in the proposed method, the fitness function in which the retrieval history can be considered is used to assume appropriate face cognitive model of the subject on each step. Moreover, in the proposed method, the natural selection based on this fitness function is employed in order to maintain the diversity of facial images. It is likely that the accurate face cognitive model of the subject can be acquired by using the combination of the retrieval history, natural selection and artificial selection. Finally, I confirm the effectiveness of the proposed method through the retrieval experiment using some subjects.

This chapter is organized as follows: In Section 2, the facial image database retrieval in this study will be explained. In Section 3, I will propose a face cognitive model acquisition method using the IGA. The experimental method using some subjects will be described in Section 4. In Section 5, the experimental result will be mentioned. In Section 6, the features of the proposed method and future works are discussed. Finally, I will finish this chapter with the conclusions.

5.2 Facial Image Database Retrieval

The facial image database retrieval used in this study is described below.

5.2.1 Facial Image

In this study, I employ a line-draw facial image as the retrieval object. The purpose of this study is to acquire the face cognitive model of human. It is clear that human receives various impressions from a line-draw facial image and has the face cognitive model for a line-draw facial image.

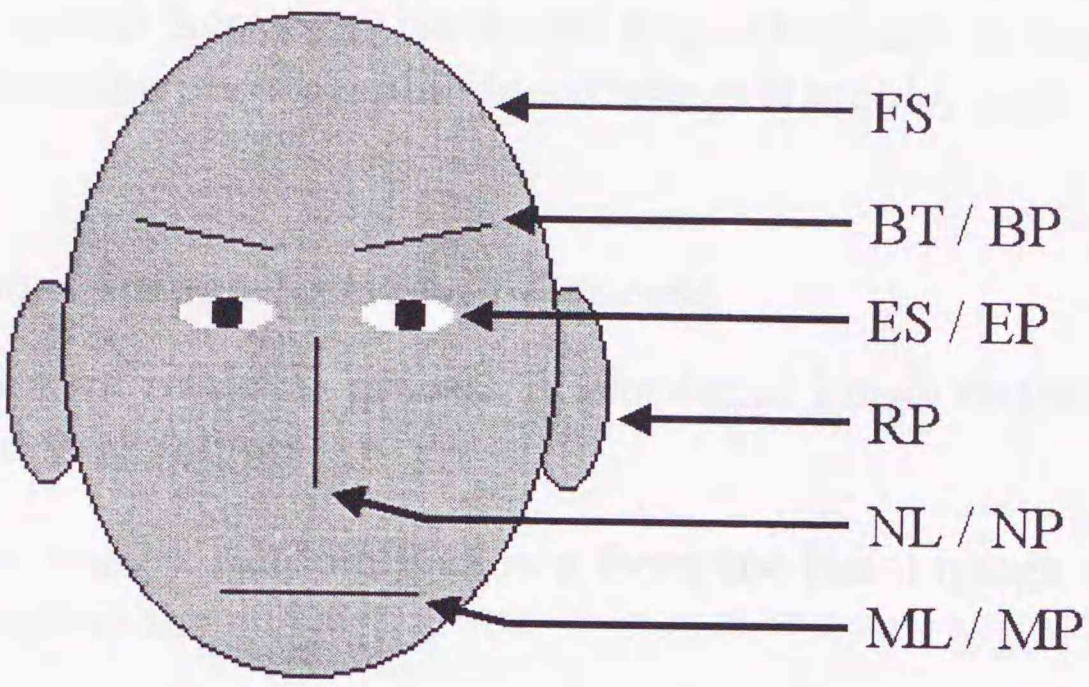


Figure 5.1: An example of facial image

Therefore, I use the line-draw facial image as a toy problem in the face cognitive model acquisition in order to simplify the construction of facial image database and the experiment to confirm the effectiveness of the proposed method.

The facial image database in this study contains nearly 60,000 line-draw facial images, each of choosing particular values for 10 facial features. The facial features used are face shape (FS), eyebrow tilt (BT), eyebrow position (BP), eye shape (ES), eye position (EP), nose length (NL), nose position (NP), mouth length (ML), mouth position (MP) and ear position (RP). There are three possible values {-1,0,1} for each facial feature, thus producing 59,049 (i.e., 3^{10}) possible faces. However, it is supposed that the subject does not know that the facial image is created by combining the value on each facial feature when the subject performs the facial image database retrieval task. An example of the facial image used in this study is shown in Figure 5.1.

It is desirable that the photographic facial image is used as the retrieval object to acquire the face cognitive model which is utilized for the practical face recognition. However, it has been reported that the face cognitive model acquisition using the photographic facial image can be implemented by using the features based on the KL expansion instead of the shape features on face which are difficult to extract [8]. That is, the photographic facial image can be treated as the combination of the

numerical datum in the same way as the line-draw facial image. Therefore, the proposed face cognitive model acquisition method can be used in the case where the photographic facial image is employed as the retrieval object.

5.2.2 Facial Image Retrieval Process

The facial image retrieval process in the facial image database retrieval task method is as follows.

1. N facial images randomly chosen from the facial image database are presented to the subject.
2. The subject selects one or some facial images which are suitable for the retrieval goal.
3. New facial images are retrieved from the facial image database on the basis of the presumption of the subject's face cognitive model and then they are presented to the subject.
4. The subject repeats this process when the requirement to finish the retrieval task is not satisfied.

When the subject starts a retrieval task, ten line-drawn facial images randomly chosen from the facial image database are simultaneously presented on a computer display because there is no information to assume the face cognitive model of the subject. The subject evaluates them and selects face or faces which meet his or her selection criterion of the retrieval goal. The subject is allowed to select as many facial images as he or she wishes, so long as they meet his or her selection criteria. The facial images selected by the subject are stored in the background buffer. The facial images presented to the subject in next step are decided on the basis of the presumption of the subject's face cognitive model. The suitable facial images for the retrieval goal are retrieved by repeating this operation.

The background buffer can hold up to 10 of the most recently selected facial images. The subject can at any time call up and re-evaluate the currently-held facial images and can decide whether to retain or discard any of them. Moreover, when the number of the currently-held facial images exceeds ten, the facial image is automatically discarded in the old order of subject's selection. Figure 5.2 shows the interface for the facial image database retrieval task method.

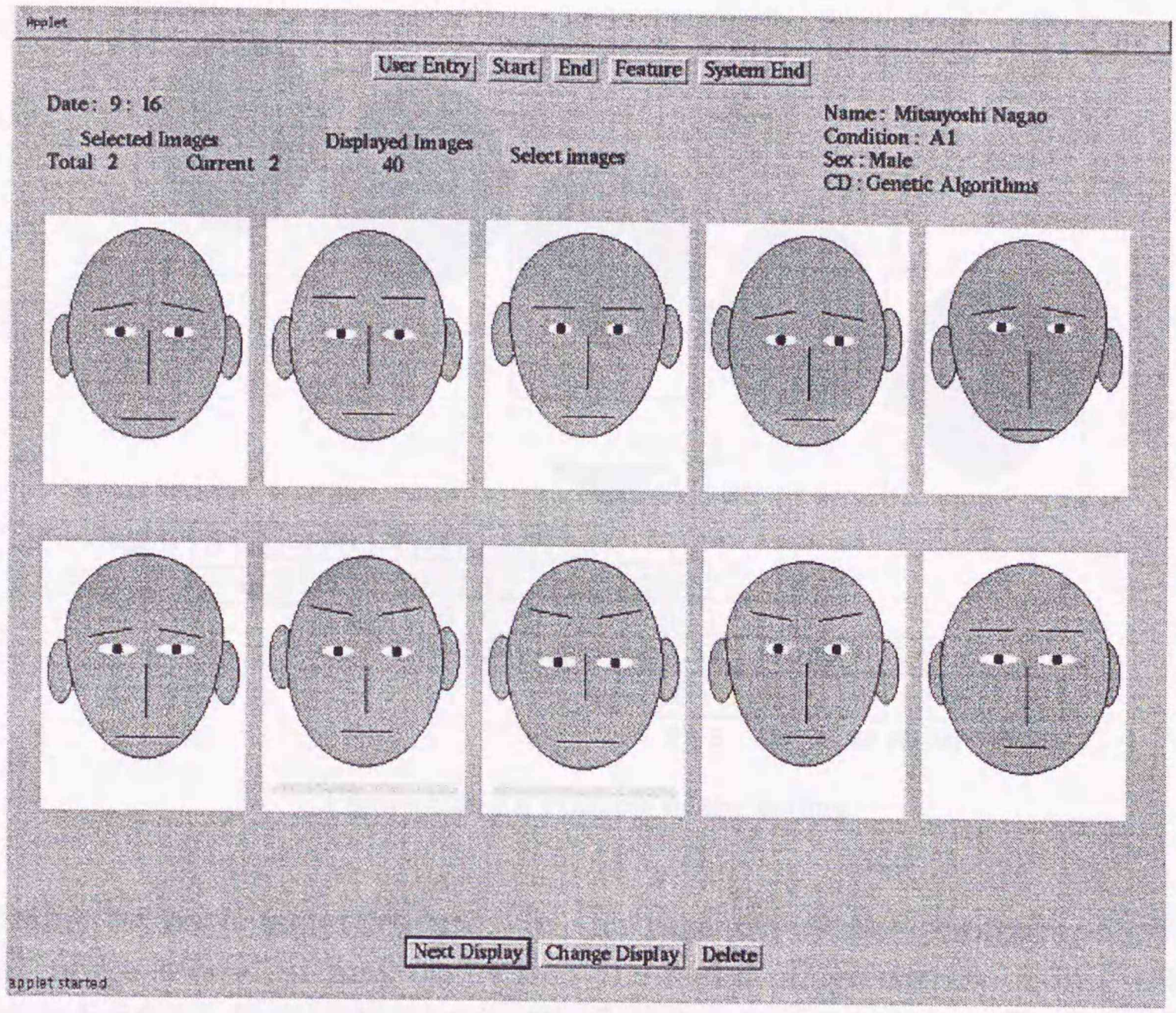


Figure 5.2: Interface for the facial image database retrieve task method

5.3 Face Cognitive Model Acquisition Method using IGA

The detail of the face cognitive model acquisition method using the IGA is described below.

5.3.1 Cording

Basically, in the proposed method, the facial images presented to the subject are retrieved from facial image database on the basis of the individuals generated by the GA. Therefore, the facial image must be encoded into the individual treated in the GA. In the coding of each individual, the values of the facial features in the facial image are encoded into the individual. The facial image used in this study consists of ten facial features, and each facial feature takes a particular value from three kinds of value. Therefore, the length of individual is ten, and each locus takes

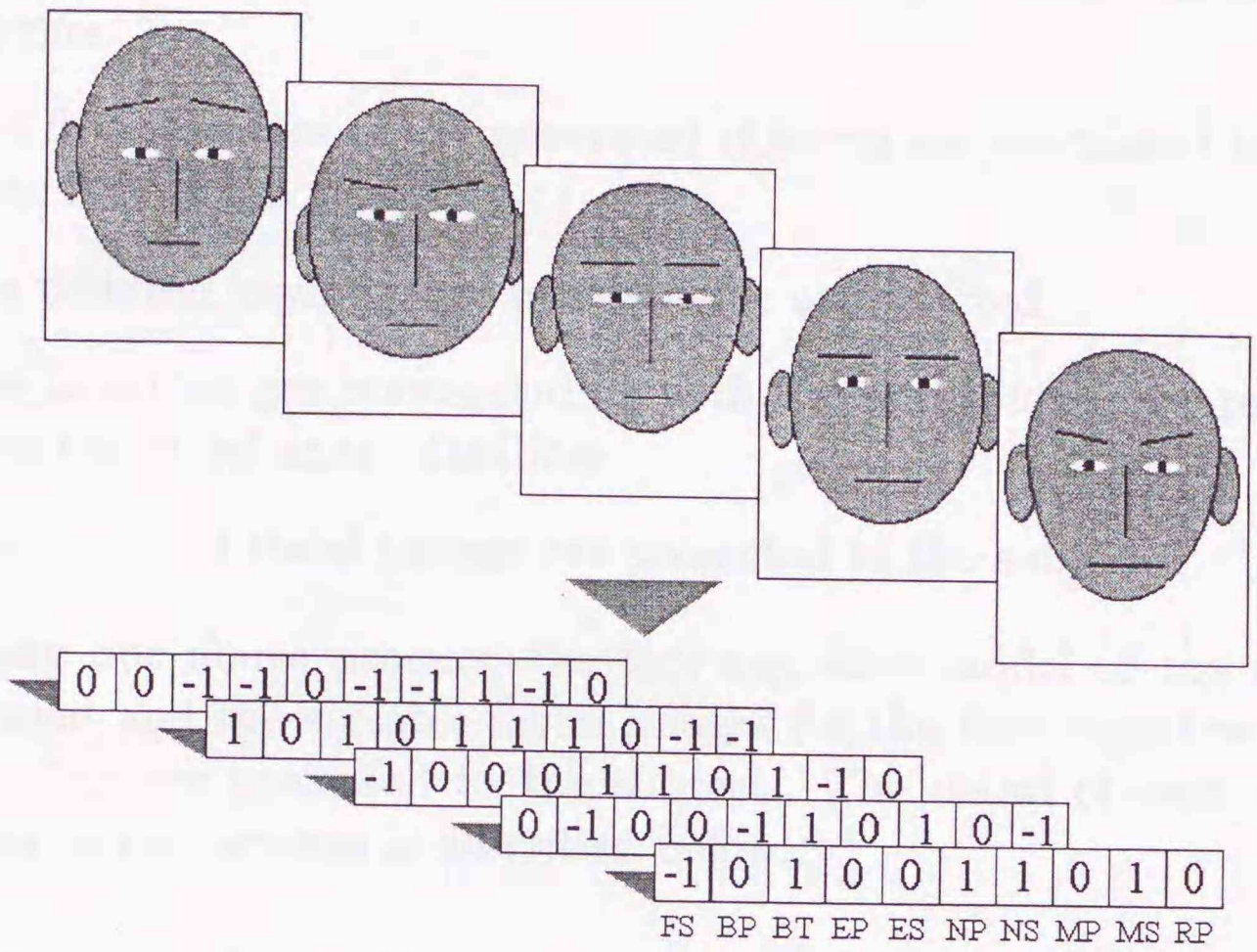


Figure 5.3: An example of the coding

a value on each facial feature. In the case where the photographic facial image is used, it may occur that the facial image corresponding with the individual does not exist in the facial image database. However, in this study, the individual perfectly corresponds with a particular facial image in the facial image database because the line-draw facial image is employed. Figure 5.3 shows an example of the coding.

5.3.2 Presumption Method of Face Cognitive Model based on IGA

In the proposed method, the face cognitive model of the subject on each step is presumed on the basis of the fitness function in which the retrieval history is considered. The presumption process in this study is as follows.

1. The fitness value for each individual is calculated on the basis of the facial images stored in the background buffer, i.e., the retrieval history.

2. The natural selection is performed by using the calculated fitness values of individuals, and the individuals employed as the parents in next generation are decided.

3. The children are generated by applying the genetic operators to the parents.

4. The fitness values of the generated children are calculated by using same fitness function.

5. Ten children having large fitness value are selected.

6. The facial images corresponding with the ten children are retrieved from the facial image database.

7. The retrieved facial images are presented to the subject.

Carrying out above process, the face cognitive model of the subject is presumed and the suitable facial images for the face cognitive model on each step are presented to the subject. The detail of each genetic operation in this process is described below.

5.3.3 Genetic Operation

5.3.3.1 Fitness Function

The subject forms a new face cognitive model which is suitable for the facial images stored in the facial image database during the retrieval task in the face cognitive model acquisition based on the facial image database retrieval task method. It has been reported that the context of stimuli in successive presentation influences the cognitive model formation process in the human brain. Thus, it is likely that more appropriate face cognitive model of the subject can be presumed by considering the context of stimuli, namely, the retrieval history, so that the effective face cognitive model acquisition can be implemented. In this fitness function, the fitness value of the individual is calculated by considering not only the information of the facial images selected on each step but also the retrieval history. The fitness function employed in this study is described below and illustrated in Figure 5.4.

$$Fitness(\boldsymbol{x}_h) = \sum_{i \in B(k)} w_i (\beta - \sum_{j=1}^{m} |x_{ij} - x_{hj}|) \tag{5.1}$$

$\boldsymbol{x}_h$: The facial image $\boldsymbol{x}_h$ in which the fitness value is calculated
x_{hj} : The value of the facial feature j on the facial image $\boldsymbol{x}_h$

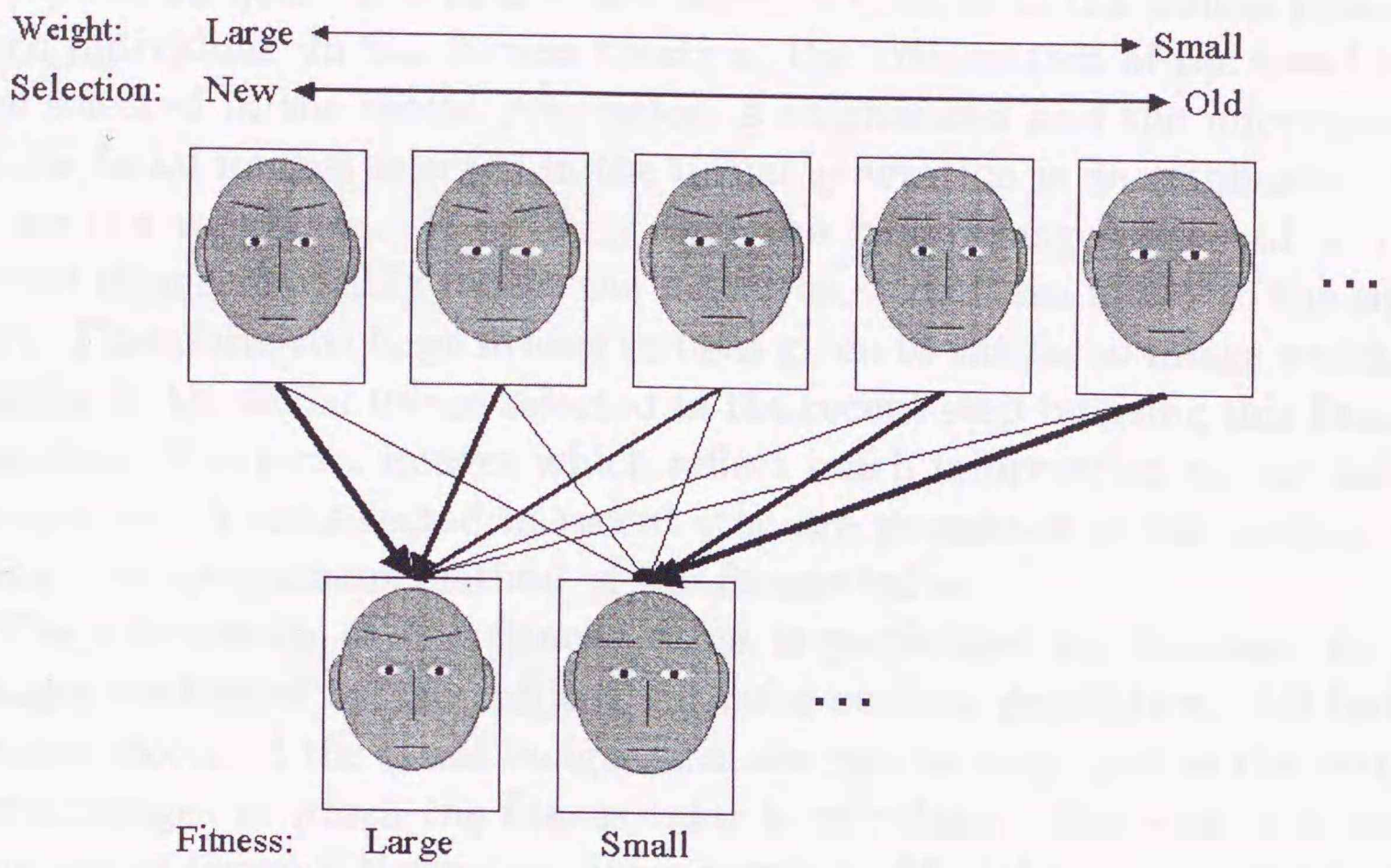


Figure 5.4: Calculation method of the fitness value

x_{ij}	:	The value of the facial feature j on the facial image $\boldsymbol{x}_i$ which is stored in the background buffer
$w_i = \alpha^{(k-T_i)}$	:	Weight for the facial image $\boldsymbol{x}_i$
α	:	Attenuation rate ($\alpha < 1$)
k	:	Current generation number
T_i	:	The generation in which the facial image $\boldsymbol{x}_i$ stored in the background buffer was selected by the subject
$B(k)$	:	A set of the facial images which are stored in the background buffer in the generation k
m	:	The number of facial features on the facial image
β	:	Constant ($\beta \geq 20$)

In this fitness function, the Euclidean distance between the facial image for calculating the fitness value and each facial image stored in the background buffer is calculated. The difference between β and the Euclidean distance is calculated in order to give the large fitness value to the facial images which are similar to the facial images stored in the background buffer. This difference is multiplied by the weight w_i based on the retrieval history. The weight is determined on the basis of the generation T_i in which the facial image $\boldsymbol{x}_i$ stored in the background buffer was select-

ed by the subject. The sum of the values is treated as the fitness value of each individual. In this fitness function, the information of the facial image selected in the recent generation is emphasized and the information of the facial images selected in the initial generation is de-emphasized by using the weight w_i. It is likely that the facial images selected in the recent step remarkably reflect the recent face cognitive model of the subject. Therefore, the large fitness value is given to the facial image which is similar to the facial image selected in the recent step by using this fitness function. The facial images which reflect much information on the facial images which are selected in recent step are presented to the subject by using this assignment method of the fitness value.

The calculation of the fitness value is performed on the past facial images evaluated by the subject until the current generation. All facial images stored in the facial image database can be employed as the target facial images in which the fitness value is calculated. However, it is time consuming because there are a large number of facial images in the facial image database. Therefore, the facial images presented to the subject in past generation are used as the target facial images in which the fitness value is calculated.

In this fitness function, the fitness value for each individual is calculated on the basis of the Euclidean distance. It has been reported that the similarity among facial images corresponds with the Euclidean distance [9]. Therefore, in this study, the Euclidean distance is employed to measure the similarity among the facial images.

Moreover, the natural selection based on the calculated fitness value is performed. The kind of facial images presented in the next generation is limited in the case where the parents in the next generation are decided only by the artificial selection. Therefore, the diversity of individuals which is the advantage in the GA is lost, and the problem of initial convergence in the search problem may occur. The diversity of the facial images can be maintained by performing the natural selection in the large population. In addition, the fitness values of the children generated by applying the genetic operators to the parents are calculated with this fitness function. The facial images corresponding to the individuals which have the large fitness value are presented to the subject. It is likely that the facial images which have the diversity and are suitable for the face cognitive model of the subject can be presented to the subject by this operation.

5.3.3.2 Natural Selection

I use the roulette strategy for the natural selection. In the natural selection based on the roulette strategy, the individual is stochastically chosen as the parent in the next generation in accordance with its fitness value. In this natural selection, the number of individuals selected as the parents is p, and the duplication is allowed.

5.3.3.3 Crossover

The children which inherit the features of the parents are generated by the crossover operator. In this crossover, I use the uniform crossover operator. The uniform crossover operator can generate the children which are not influenced by the sequence of the features of parents. The generation method of the children based on the uniform crossover operator is illustrated in Figure 5.5. As shown in Figure 5.5, a mask is prepared in order to generate two children from two parents. The mask takes either 1 or 0 on each feature. If the value of mask is 1 on each feature, the feature of parent $\boldsymbol{x}_i$ is inherited to the child $\boldsymbol{x}'_i$ and the feature of parent $\boldsymbol{x}_j$ is inherited to the child $\boldsymbol{x}'_j$. Conversely, if the value of mask is 0, the feature of parent $\boldsymbol{x}_i$ is inherited to the child $\boldsymbol{x}'_j$ and the feature of the parent $\boldsymbol{x}_j$ is inherited to the child $\boldsymbol{x}'_i$. In the GA, various crossover operators have been proposed, e.g., one point crossover operator, two point crossover operator and so on. The one point crossover operator is superior to the uniform crossover operator in terms of the convergence speed. In the case where the uniform crossover operator is employed, the performance in the convergence speed will decay [10]. However, the uniform crossover operator can present the facial images having the diversity to the subject.

5.3.3.4 Mutation

The mutation operator stochastically converts a value on each locus of the individual. In this study, $1/m$ is employed as the mutation rate.

5.4 Experimental Method

I designed the experiments with some subjects in order to confirm the effectiveness of the proposed method. The experimental method is described below.

In this experiment, six subjects were recruited from Hokkaido University. The subjects received the instruction of the facial image database

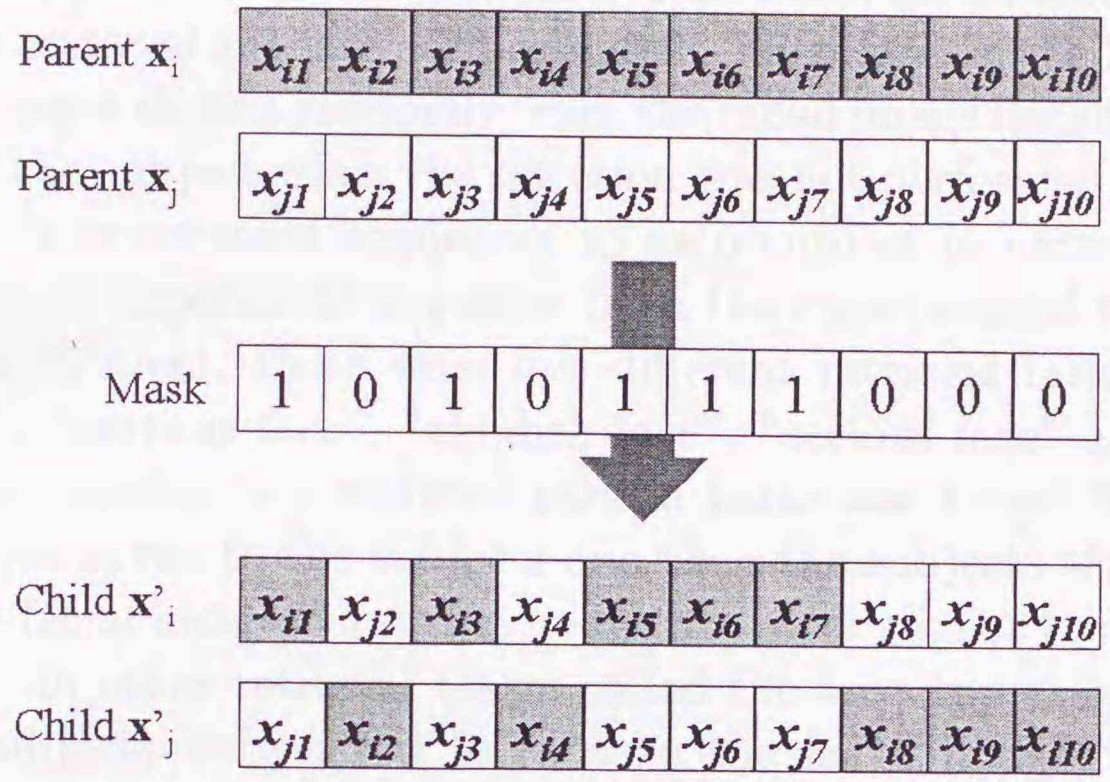


Figure 5.5: An example of uniform crossover

retrieval and then the retrieval practice was given to the subjects in order to familiarize themselves with the facial image database retrieval and the mouse operation. After the instruction and the practice retrieval, the facial image retrieval experiment was started.

In this experiment, two different methods were employed in order to compare with the proposed method (IGA method). One is the RN method which presents the facial images chosen randomly from the facial image database to the subject. Another is the AS method which decides the parents in next generation by the artificial selection, which is used in the conventional IGA. In the AS method, the facial images selected on each step by the subject were employed as the parents in next generation, and the children were generated by applying the genetic operator or operators to these parents. That is, the children were generated only by the mutation operator when the subject selected one facial image, and they were generated by the crossover and mutation operators when the subject selected some facial images. Moreover, the application of the crossover and mutation operators to the facial image or facial images which were selected by the subject was continued until the facial images which were different from the facial images presented already were generated. I employed the uniform crossover as the crossover operator in the AS method. In the conventional IGA, the user has to select one individual at least on each step. However, in the face cognitive model acquisition based on the facial image database retrieval task method, there may be the case where

the selection by the subject is not performed when the suitable facial images for the retrieval goal are not presented. Therefore, in the AS method, the facial images chosen randomly from the facial image database were presented to the subject when the selection was not performed. Moreover, I assigned six experiment sequences to each subject in order to remove the influence of experiment sequence from the experimental result.

In this experiment, there were five different retrieval tasks; i.e., "favorite face", "nervous face", "childish face", "serious face" and "similar face". In the similar face retrieval task, a particular target face such as Figure 5.1 was given to the subjects and then the subjects were asked to retrieve the facial images that they judge are most similar to a particular target face. In other retrieval tasks, called the face impression retrieval task, the subjects were asked to retrieve the facial images in which a particular target impression was most received. The face impressions retrieval task was the experiments to confirm the basic performance for the face cognitive model acquisition. On the other hand, the similar face retrieval task was the experiment to evaluate the proposed method in the convergence performance point of view. In addition, in these retrieval tasks, the subjects were asked to continue the retrieval task until they were satisfied with ten facial images in the background buffer.

The parameter setting used in the IGA method is as follows.

- Population size: 200
- Crossover rate: 0.3
- Mutation rate: 0.1
- α: 0.9
- β: 20

These parameters were employed in order to prevent the initial convergence and to maintain the diversity of facial images.

5.5 Experimental Result

The experimental results of the face impression retrieval task and the similar face retrieval task are described below.

5.5.1 Experimental Result of Face Impression Retrieval Task

The experimental results are shown in Table 5.1, Figure 5.6 and Figure 5.7. In Table 5.1, Step and Time are the number of steps and the time which the subjects took to complete the face impression retrieval task, respectively. These results are the average of 24 experiments; i.e., 4 face impression retrieval tasks $\times$ 6 subjects. In the face cognitive model acquisition based on the facial image database retrieval task method, it is employed as a measurement criterion to evaluate its effectiveness that the subjects can retrieve appropriate facial images in a small number of steps and in a short time.

In Table 5.1, the number of steps was 27.5 for IGA, 30.5 for AS and 34.0 for RN. In addition, to complete the face impression retrieval task, the subjects took 5 minutes 35 seconds for IGA, 7 minutes 37 seconds for AS and 5 minutes 35 seconds for RN. I found from these results that the subjects could complete the retrieval task with IGA in the smallest number of steps. However, the subject could complete the retrieval task with RN in the shortest time. I calculated the evaluation time for the one step in order to investigate this cause, which is calculated by dividing the retrieval time by the number of steps. The result was 10 seconds for RN, 13 seconds for IGA and 15 seconds for AS. This result suggests that in RN, the total time taken to complete the retrieval task was shortest in three methods because the subjects encountered many faces that they were able to quickly discount as unsuitable for their face cognitive model. In AS, the subjects took long evaluation time for the one step. I could obtain from a subject the verbal report that it is difficult to recognize the difference among the facial images presented on the display because the similar faces to each other were presented. Therefore, it is likely that the total time taken to complete the retrieval task was long in AS. On the other hand, it seems to be that in IGA, the result of total time which was shorter than AS and was longer than RN was indicated because the facial images which reflected the face cognitive model of the subjects and which had the diversity were presented to the subject. Moreover, the analysis of variance was applied to these results of Step and Time. However, the statistically significant result was not indicated. In the result of the face impression retrieval task, its difference among the subjects was large. In addition, the number of subjects was small. Therefore, it is likely that this result is caused by the number of the subjects and the kind of the retrieval task.

The most important criterion to evaluate the performance of the face

	Step	Time
RN	34.0	5 min. 35 sec.
AS	30.5	7 min. 37 sec.
IGA	27.5	5 min. 53 sec.

Table 5.1: Experimental results of face impression retrieval task

cognitive model acquisition based on the facial image database retrieval task method is the quality of the facial images retrieved by the subjects. The quality evaluation of the facial images was performed in the nervous face retrieval task by using the Analytic Hierarchy Process (AHP) [11]. The AHP is the method to decide the priority for the evaluation objects by performing pairwise comparison among the evaluation objects. In order to evaluate the quality of the facial images, there is a method that the subjects select most appropriate facial image from the facial images which were retrieved by using each method. However, more accurate quality evaluation can be implemented by using the AHP.

In the quality evaluation of the facial images based on the AHP, at first, the subjects were asked to select two facial images which were most suitable for the retrieval goal in the facial images retrieved by using each method. The pairwise comparison among these six facial images chosen was performed. In the AHP, the pairwise comparison is difficult when the number of evaluation objects exceeds seven. Therefore, two most appropriate facial images retrieved by each method were chosen, and the pairwise comparison was performed by using these six facial images. The pairwise comparison among the facial images was performed by using the table such as Figure 5.6 (the name of method was not described in the table). In this pairwise comparison, nine grade evaluation criteria including "Same" was employed. For example, as shown in Figure 5.6, the subjects inserted 9 into the corresponding position if the facial image on the right hand side was absolutely more suitable for the retrieval goal than the facial image on the left hand side, or the subjects inserted 1/3 into the corresponding position if the facial image on the left hand side was weakly more suitable for the retrieval goal than the facial image on the right hand side. The priorities of the facial images were determined by performing this pairwise comparison. In the AHP, it is necessary to check the consistency in the pairwise comparison performed by the subject. The consistency is around 0.1 when the correct pairwise comparison was

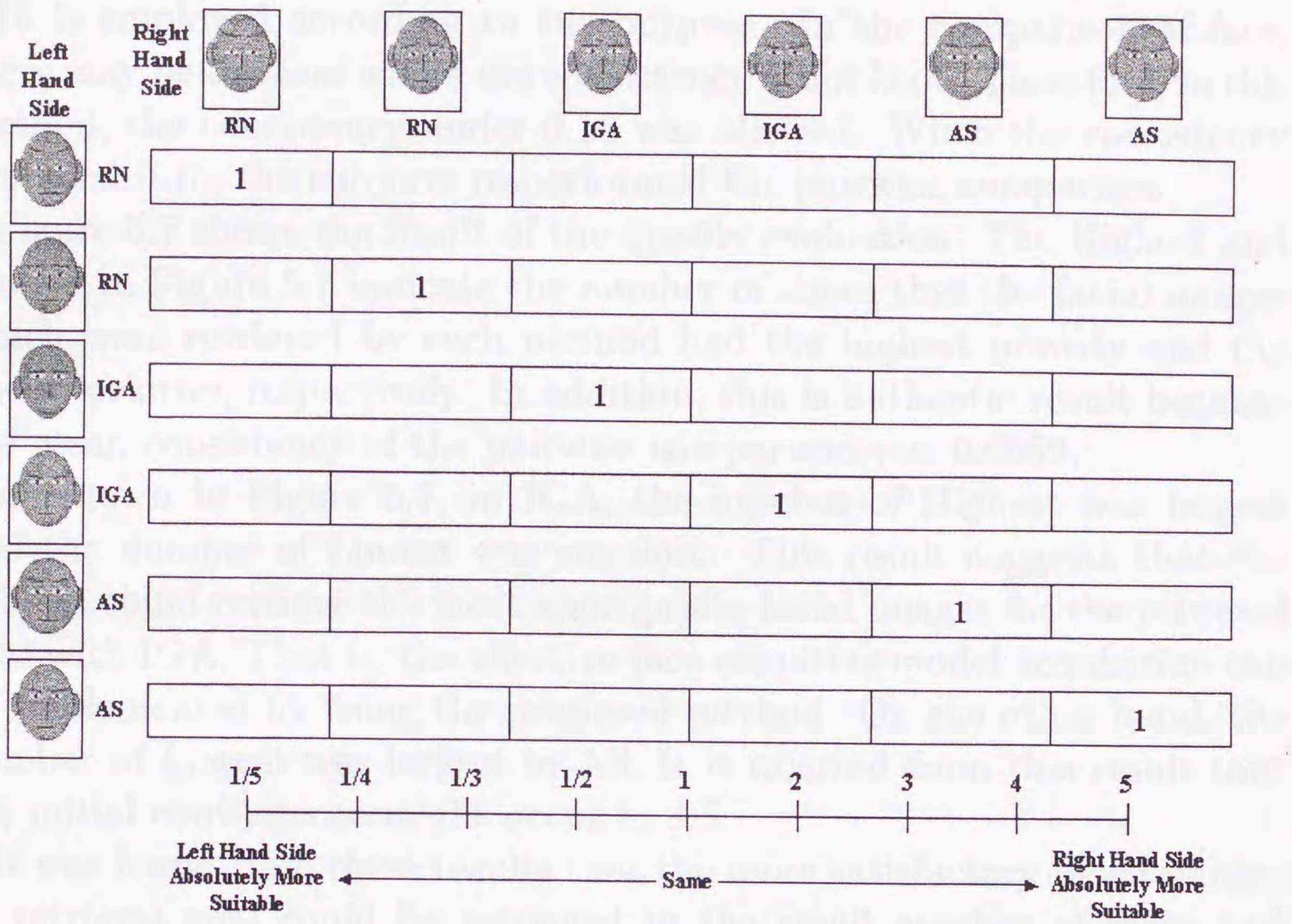


Figure 5.6: Table for pairwise comparison

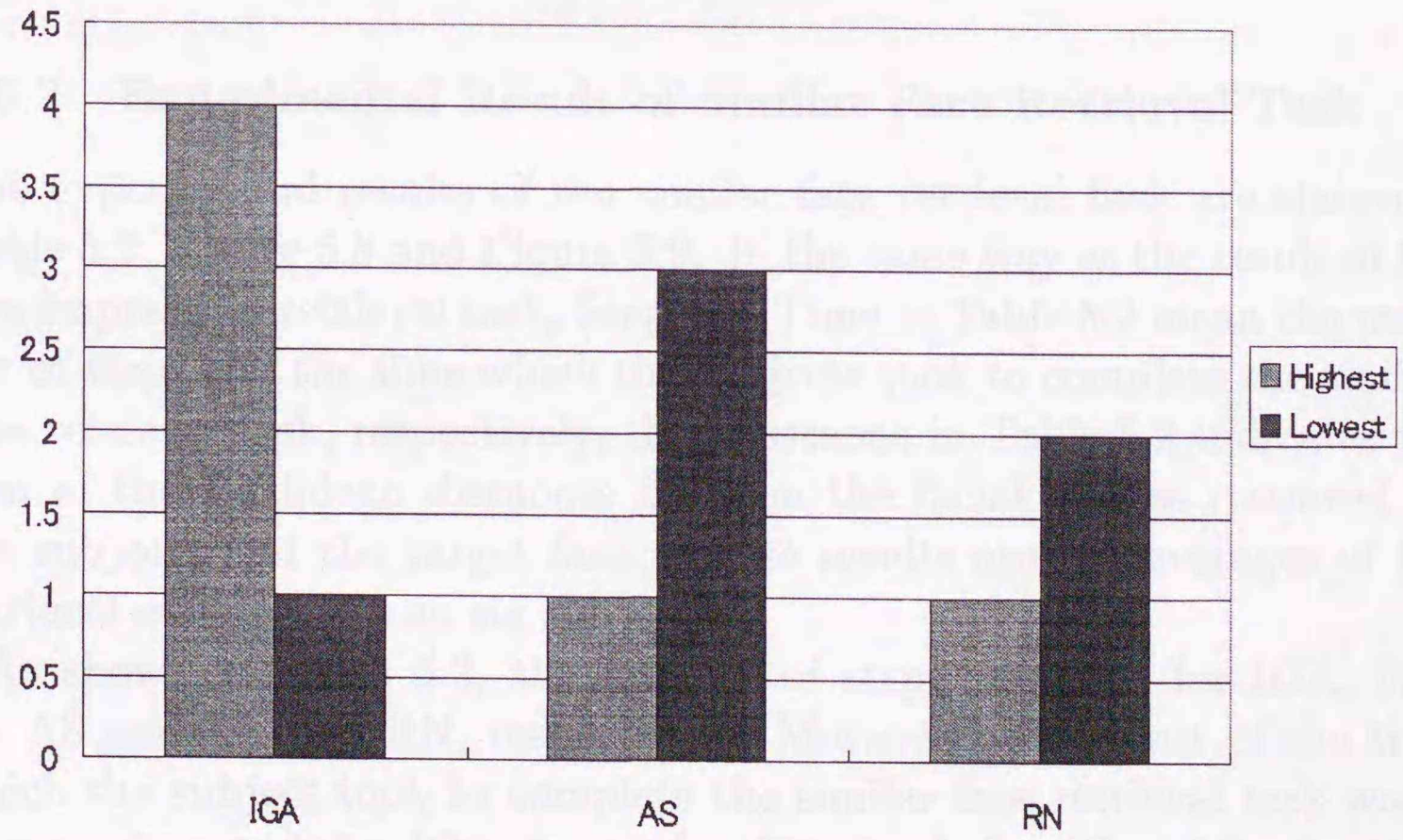


Figure 5.7: Quality of the retrieved facial images

performed. However, there is the case where the consistency of around 0.15 is employed according to the purpose. In the comparison of face, there may be the case where the consistency is not kept. Therefore, in this method, the consistency under 0.15 was allowed. When the consistency exceeded 0.15, the subjects re-performed the pairwise comparison.

Figure 5.7 shows the result of the quality evaluation. The Highest and Lowest in Figure 5.7 indicate the number of times that the facial images which were retrieved by each method had the highest priority and the lowest priority, respectively. In addition, this is authentic result because the mean consistency of the pairwise comparison was 0.0869.

As shown in Figure 5.7, in IGA, the number of Highest was largest and the number of Lowest was smallest. This result suggests that the subject could retrieve the most appropriate facial images for the retrieval goal with IGA. That is, the effective face cognitive model acquisition can be implemented by using the proposed method. On the other hand, the number of Lowest was largest in AS. It is inferred from this result that the initial convergence might occur in AS.

It was found from these results that the more satisfactory facial images as retrieval goal could be retrieved in the small number of steps and in the short time by using the proposed method than the conventional method although the statistically significant result could not be obtained. Namely, these results revealed that the proposed method can implement effective face cognitive model acquisition.

5.5.2 Experimental Result of Similar Face Retrieval Task

The experimental results of the similar face retrieval task are shown in Table 5.2, Figure 5.8 and Figure 5.9. In the same way as the result of the face impression retrieval task, Step and Time in Table 5.2 mean the number of steps and the time which the subjects took to complete the similar face retrieval task, respectively. The Distance in Table 5.2 indicates the sum of the Euclidean distances between the facial images retrieved by the subjects and the target face. These results are the averages of the retrieval experiments on six subjects.

As shown in Table 5.2, the number of steps was 36.3 for IGA, 63.3 for AS and 77.1 for RN, respectively. Moreover, the result of the time which the subject took to complete the similar face retrieval task was 6 minutes 9 seconds for IGA, 8 minutes 49 seconds for AS and 8 minutes 9 seconds for RN. I found from these results that in IGA, the subject could complete the similar face retrieval task in the smallest number of steps

	Step	Time	Distance
RN	77.1	8 min. 9 sec.	65.7
AS	63.3	8 min. 49 sec.	54.6
IGA	36.3	6 min. 9 sec.	62.6

Table 5.2: Experimental results of similar face retrieval task

and the shortest time. In this experiment, the analysis of variance was also applied to these results of Step and Time. Although the statistically significant result could not be obtained in these results, there is a high possibility that the statistically significant result can be obtained in the similar face retrieval task by increasing the number of subjects because the difference among the subjects was not so large.

The result of the sum of the Euclidean distances between the target face and the facial images retrieved was 62.6 for IGA, 54.6 for AS and 65.7 for RN. Moreover, the sum of the Euclidean distances between the target face and the facial images presented on each step is shown in Figure 5.8. These results suggest that in the case of RN, the result of the Euclidean distance to the target face was large because the similar facial images to the target face were not frequently presented to the subjects. In AS, the facial images which had the shortest Euclidean distance to the target face in three methods were retrieved. The AS has superior performance to retrieve similar facial images because the parents in next generation are decided only on the basis of the facial images selected by the subjects. However, as shown in Figure 5.9, the variance of the facial images presented on each step rapidly decreased in the initial step. That is, it is likely that the initial convergence problem arose in AS. Therefore, it is reasonable to suppose that the number of steps and the time were large and long, although the Euclidean distance between the target face and the facial images retrieved was small. On the other hand, in IGA, the sum of the Euclidean distances between the target face and the facial images retrieved was larger than AS. As shown in Figure 5.8, in IGA, the convergence speed was slow. Therefore, it is likely that in IGA, the subject could not retrieve similar facial images to the target face in the small number of steps and the short time. The convergence speed can be improved by adjusting the parameter settings. However, too fast convergence speed may prevent the acquisition of accurate face cognitive model. Thus, the appropriate parameter settings must be performed so

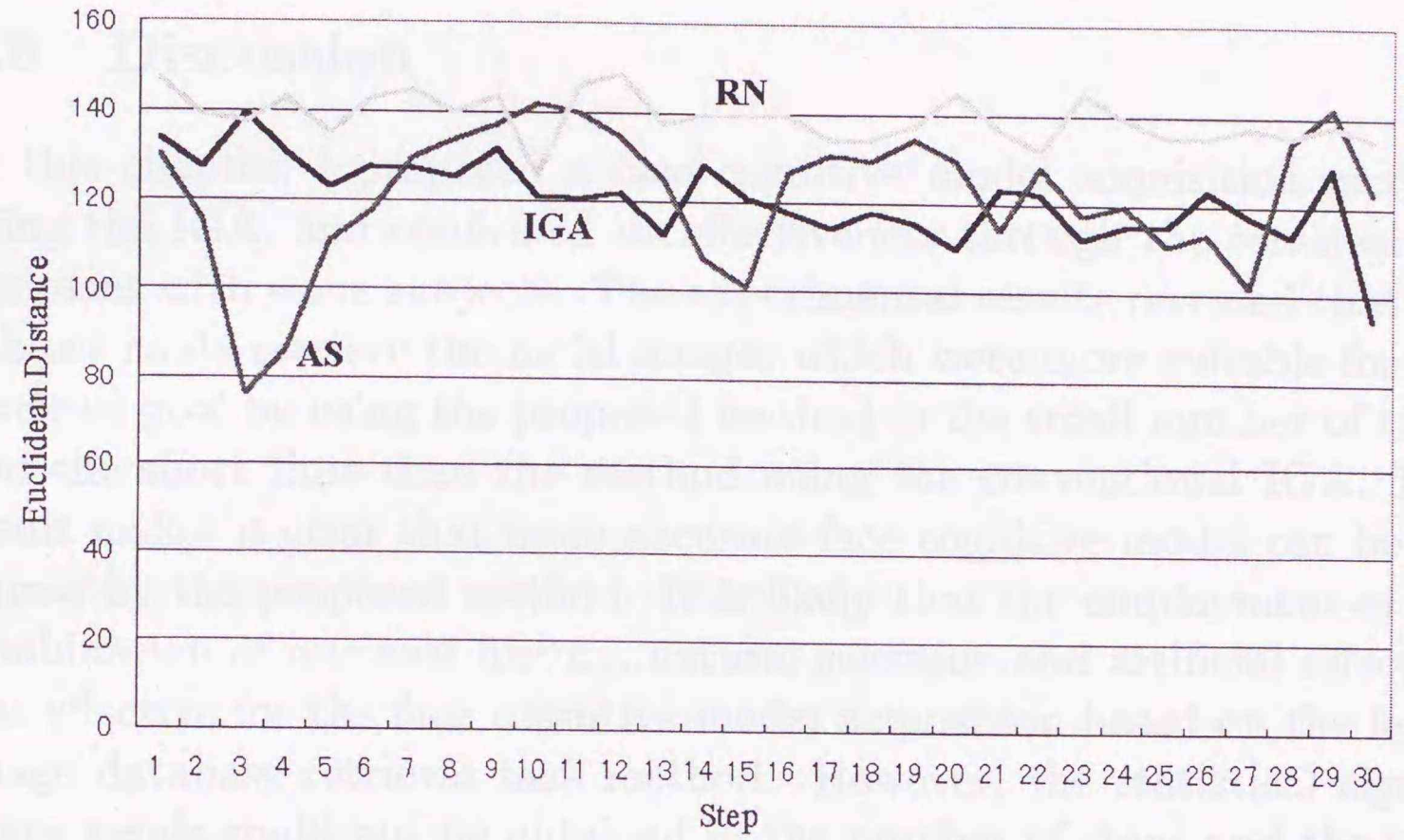


Figure 5.8: Sum of Euclidean distances between the facial images presented on each step and the target face

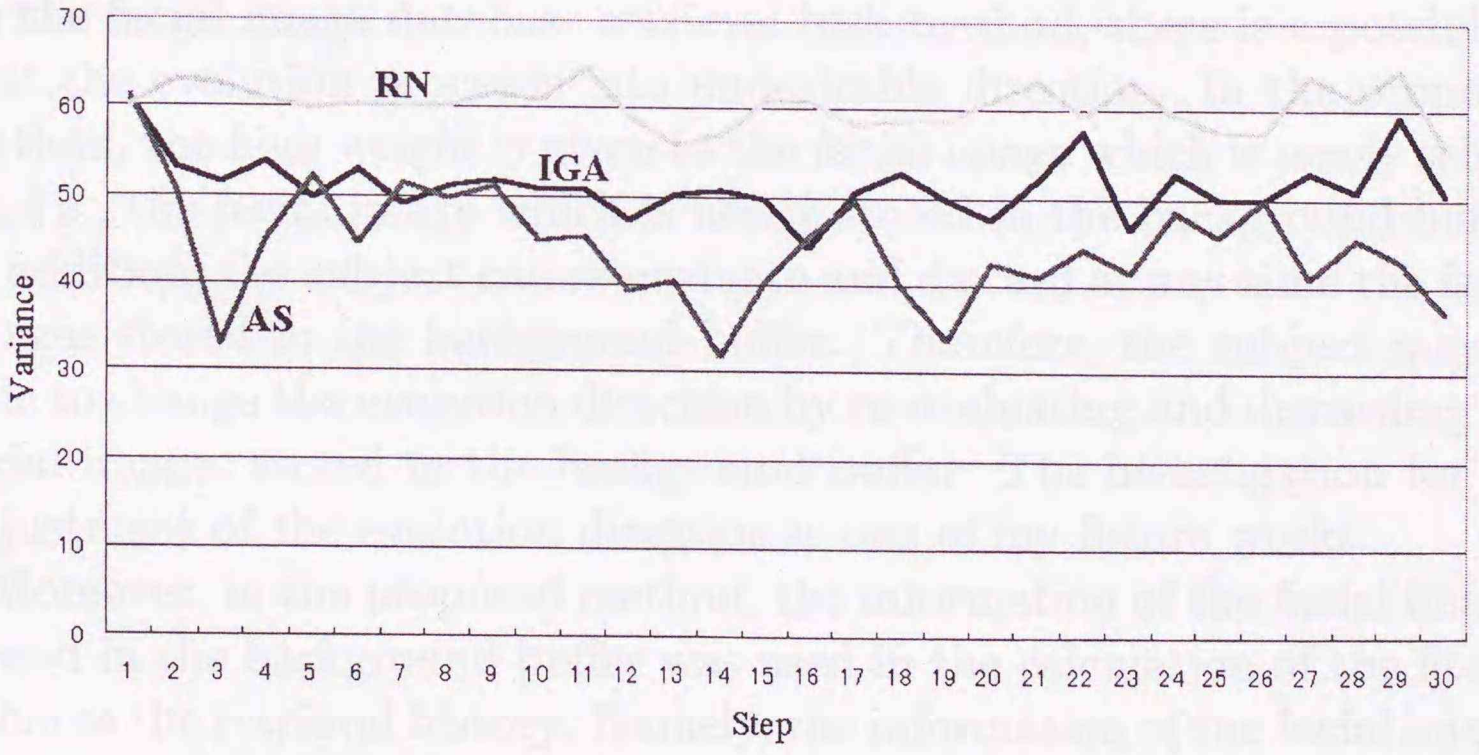


Figure 5.9: Variance of the facial images presented on each step

that the performance of the acquisition of face cognitive model does not decreased.

5.6 Discussion

In this chapter, I proposed a face cognitive model acquisition method using the IGA, and confirmed its effectiveness through the retrieval experiment with some subjects. The experimental results revealed that the subject could retrieve the facial images which were more suitable for the retrieval goal by using the proposed method in the small number of steps and the short time than the method using the conventional IGA. This result makes it clear that more accurate face cognitive model can be acquired by the proposed method. It is likely that the employment of the combination of retrieval history, natural selection and artificial selection was effective for the face cognitive model acquisition based on the facial image database retrieval task method. However, the statistical significance result could not be obtained in the number of steps and the time that the subjects took to complete the retrieval task. To demonstrate the statistically significance result, it is necessary to design the retrieval experiments with many subjects because the number of subjects was small in this experiment.

In the proposed method, the retrieval history was utilized in the calculation of the fitness value for the individual. In the case where the conventional IGA is used for the face cognitive model acquisition based on the facial image database retrieval task method, there is a possibility that the evolution proceeds into undesirable direction. In the proposed method, the high weight is given to the facial image which is newly selected, i.e., the facial image which is newly stored in the background buffer. In addition, the subject can re-evaluate and discard at any time the facial images stored in the background buffer. Therefore, the subject may be able to change the evolution direction by re-evaluating and discarding the facial images stored in the background buffer. The investigation for the adjustment of the evolution direction is one of my future works.

Moreover, in the proposed method, the information of the facial images stored in the background buffer was used in the calculation of the fitness value as the retrieval history. Namely, the information of the facial images selected by the subject only was considered in order to presume the face cognitive model. However, the facial images which are not selected by the subject may also contain the information for the face cognitive model

of the subject. Therefore, it is likely that more effective method can be implemented by considering both.

In this method, the facial images evaluated once were not presented to the subject again. However, in the face cognitive model acquisition based on the facial image database retrieval task method, the face cognitive model of the subject may change. That is, the facial images judged to be not suitable for the face cognitive model once may change to the suitable facial images for the face cognitive model in the future step. Therefore, it is necessary to examine whether the facial images evaluated once by the subject should be presented to the subjects or not.

Finally, the facial image used in this study was the line-draw facial image and the number of attributes was ten. It is difficult to obtain the large difference among the methods in the case where the number of attributes is small. Therefore, the effectiveness of the proposed method should be also confirmed in the photographic facial image database.

5.7 Conclusion

In this chapter, I proposed a face cognitive model acquisition method using the IGA in order to implement the acquisition of accurate face cognitive model in the facial image database retrieval task method. Moreover, I confirmed the effectiveness of the proposed method through the retrieval experiment with some subjects. The experimental results revealed that the effective face cognitive model acquisition based on the facial image database retrieval task method could be implemented by using the proposed method. However, it is likely that the efficiency of the proposed method can be promoted by finding more appropriate fitness function and parameter settings. Furthermore, in this study, I used the line-draw facial image in order to simplify the construction of the facial image database and the evaluation of the proposed method. However, the face cognitive acquisition with photographic facial image should be implemented to acquire the face cognitive model which is utilized for the face recognition methods. These are my future works.

Bibliography

[1] S. Akamatsu, "Computational Studies on KANSEI Information Conveyed by Human Face", Journal of Japan Society for Fuzzy Theory and Systems, Vol.10, No.2, pp.225-235, April 1998. (in Japanese)

[2] O. Hasegawa, S. Morishima and M. Kaneko, "Processing of Facial Information by Computer", Journal of Institute of Electronics, Information and Communication Engineers, Vol.80-A, No.8, pp.1231-1249, Aug. 1997. (in Japanese)

[3] M. Oda, "An Image Retrieval System that Uses Human Cognitive Properties on Image Concept Formation", Journal of Information Processing Society of Japan, Vol.35, No.7, pp.1449-1456, July 1994. (in Japanese)

[4] M. Oda and T. Kato, "AN EVALUATION OF FEATURE SALIENCY OF FACES FOR IMPROVING THE AMBIGUOUS-FACE RETRIEVAL SYSTEM", Proceedings of 15th International Symposium on Human Factors in Telecommunications, pp.415-417, March 1995.

[5] H. Kitano (eds.), "Genetic Algorithm 4", Sangyou Tosho Publishers, Tokyo, Japan, 2000. (in Japanese)

[6] H. Takagi, T. Unemi and T. Terano, "Perspective on Interactive Evolutionary Computing", Journal of Japanese Society for Artificial Intelligence, Vol.13, No.5, pp.693-703, Sept. 1998. (in Japanese)

[7] C. Caldwell and V. S. Johnston, "Tracking a Criminal Suspect through Face-Space with a Genetic Algorithm", Proceedings of International Conference on Genetic Algorithms, pp.416-421, 1991.

[8] M. Oda, S. Akamatsu and H. Fukamachi, "Adaptability of K-L Expansion Technique for Ambiguous Face Image Retrieval", Journal of Institute of Electronics, Information and Communication Engineers, Vol. J79-A, No.2, pp.288-297, Feb. 1996. (in Japanese)

[9] V. Bruce, S. Yoshikawa(trans.), "RECOGNIZING FACES", SAIENSU-SHA Publishers, Tokyo, Japan, 1990. (in Japanese)

[10] D. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison-Wesley Publishing Co., 1989.

[11] K. Tone, "Decision Make Method like Game", JUSE Press, Tokyo, Japan, 1986. (in Japanese)

Chapter 6

Prospect for Development of Comprehensive Face Recognition System

This chapter is concerned with the future works in my study. The prospect for the development of comprehensive face recognition system which can be constructed by using the face recognition methods proposed in this dissertation is described, and several application fields in which such face recognition system can be utilized are discussed.

In this doctoral dissertation, a facial expression recognition method based on the MTS, a robust face identification method against the lighting and face position fluctuations and a robust face identification method against the appearance fluctuations were proposed. One of further directions of this study will be to combine these proposed methods. As a beginning, in the facial expression recognition, the facial expression recognition method which is robust against the photographic fluctuations and appearance fluctuations will be implemented by combining these methods. Moreover, the performance of facial expression recognition will be improved by introducing the function of face identification into the facial expression recognition. That is, the personalization in the facial expression recognition can be performed. On the other hand, needless to say, the robust face identification against the photographic and the appearance fluctuations can be realized. The robust face identification against facial expressions will be also carried out.

Furthermore, in this dissertation, the face cognitive model acquisition method using the IGA was proposed in order to implement the face recognition methods which have the consistency with the face cognitive structure of human. Therefore, it is obvious that the facial expression recognition and face identification which have the consistency with the face cognitive structure of human will be implemented. The comprehensive

face recognition system which is robust against various fluctuations and has the consistency with the face cognitive structure of human will be realized by combining the advantages of the proposed face recognition methods.

In addition, the knowledge about the face cognitive structure of human which is obtained by using the face cognitive model acquisition method can be not only used for the face recognition but also be utilized for the face composition. Recently, the research on the face which is composed by using computer graphics and animation technologies has been carried out in order to implement user-friendly man-machine interfaces. This subject is called the face composition. The considerably true-to-life composition of virtual human has come to be able to be accomplished by the development of various techniques in computer graphics and animation field. However, the purpose of the face composition is the implementation of the interface to support the communication based on face between human and computer or among humans. Therefore, the face composed by computer must be able to give to the user the information which human face can impart. The face composition which can give such information to the user can be fulfilled by the knowledge about the face cognitive structure of human which is obtained by utilizing the face cognitive model acquisition method. What is more, the system which can perform the communication based on face between human and computer can be carried out by introducing the face composition function into the face recognition system. If such system can be realized, more effective and interactive computer environment can be constructed. The application fields in which the comprehensive face recognition system or the face recognition system with face composition function can be utilized are described below.

Recently, the research on the communication among humans utilizing the computer network environment has been carried out. The comprehensive face recognition system or the face recognition system with face composition function can be utilized for the communication through the computer network. The computer supported cooperative work and the virtual reality are taken up as the representative studies of the communication using the computer network. These studies are briefly introduced below.

1. **Computer Supported Cooperative Work (CSCW):** With the spread of computer network, the research on the system to support the cooperative work in a group through the computer network en-

vironment is carried out. This subject is called Groupware or Computer Supported Cooperative Work (CSCW). In the CSCW, some users perform the interaction with each other and cooperatively work through the computer network environment. Each user has a common display to do work. The figure of each user is shown in this display. The group accomplishes an arbitrary task through the dialogue concerning the task. The information such as facial expression, line of sight and so on is very important element in the communication based on the CSCW. That is, in the communication based on the CSCW, it is required to recognize other user's facial expressions, to assume the direction of the line of sight and to join line of sight together. The face recognition system can be used to solve such problems

2. **Virtual Reality (VR):** In the Virtual Reality (VR) environment, the figure of user which is composed by computer graphics technique is shown in the display. In the communication among the users on the VR environment, the facial expressions are important element. Moreover, it is required to identify the person who communicates with oneself. Therefore, the face recognition system can be utilized in order to satisfy these requirements.

It is clear that the face recognition system can not only assist the communication among humans through the computer network but can also support the communication between human and computer. The application fields on the support of the communication between human and computer are briefly introduced below.

1. **Multi-modal Interface:** Recently, various computers and electric appliances have been developed. They have many buttons to operate them. The user has to read a thick manual in order to master its operation. It is difficult to learn the operation of computers or electric appliances, and it is burden for the user. From this background, the research on the user interface which can operate the computers or electric appliances by using facial expression, speech, gesture and line of sight and so on is carried out. This interface is called multi-modal interface. In this multi-modal interface, the functions to recognize the facial expression and to identify the user are needed. Therefore, the face recognition system is also effective for this multi-modal interface.

2. **Interface Agent:** The paradigm called agent has been proposed in terms of the development of more user-friendly and useful computer environment. In this paradigm, the software is regarded as a faithful and dependable servant, and the software performs the task for the user instead of him/her. This software is called agent. It is assumed that agent has various kinds of function. The agent having the function for the intermediation between human and computer is called interface agent. Recently, the development of the interface agent with face called the humanized agent has been attempted. The humanized agent has to obtain the information about the user from face, e.g., facial expression, personal data and so on. Moreover, the agent must give various kinds of information to the user by using own face composed by computer graphics. Therefore, it is obvious that the agent must have the face recognition function which can flexibly recognize the user's face and the face composition function which can give the various kinds of the information to the user. The face recognition system with face composition function is useful for this agent paradigm.

As mentioned above, the comprehensive face recognition system is effective for the support of the communication among humans or between human and computer. However, each technology in the face recognition can be utilized in various fields. Therefore, it is very significant studies in the engineering to develop the effective face recognition methods.

Chapter 7

Summary

In this doctoral dissertation, face recognition and recognition methods were focused on for the implementation of flexible and adaptable human interface. Human face plays very important roles in the face-to-face communication among humans because humans can obtain various kinds of the information such as intention, emotion, state and so on from face. It is clear that the effective human interface can be implemented if computer can recognize or understand the information which face represents. Therefore, it is required to develop the effective recognition methods for the information which face represents, i.e., face recognition methods. The face recognition contains facial expression recognition, facial identification, facial image retrieval, face cognitive model acquisition and so on. In this dissertation, especially, the facial expression recognition, face identification and face cognitive model acquisition were taken up as the typical problems in the face recognition, and the face recognition methods for these problems were proposed.

Chapter 1 was concerned with the face recognition and recognition methods. The typical problems in the face recognition were introduced. In particular, the facial expression recognition, face identification and face cognitive model acquisition which were treated in this dissertation were surveyed. Moreover, the MTS and IGA which were used for the face recognition methods proposed in this dissertation were explained, and the recent works and the related works in these methods were surveyed.

In chapter 2, a facial expression recognition based on the MTS was proposed. In the facial expression recognition, it is important to find a common facial expression pattern on each facial expression and then perform the recognition based on this pattern. The MTS is one of statistical pattern recognition methods, which has been proposed by Taguchi. In the pattern recognition based on the MTS, the correlations among the attributes which are common in a set of multi-dimensional training data

are automatically extracted and then the pattern recognition is performed on the basis of these correlations. Therefore, the MTS can allow the implementation of the facial expression recognition based on the common facial expression pattern on each facial expression. The experiment with some subjects was designed in order to confirm the effectiveness of the proposed facial expression recognition. The experimental results revealed that effective facial expression recognition could be implemented by using the MTS.

In chapter 3, a robust face identification method against lighting and face position fluctuations was proposed. The major problem in the face identification is to implement the robustness against various fluctuations. In this chapter, the lighting fluctuation and face position fluctuation which inevitably occur in taking the facial image were taken up, and the implementation of the robust face identification against these fluctuations was tried. One of methods to implement the robustness against these fluctuations is the identification based on many facial images taken in various situations in which these fluctuations occur. However, it burdens the user to take many facial images in various situations. Therefore, in the proposed method, one original facial image of the user was taken and the artificial lighting and face position fluctuations were inserted into it. The robust face identification against lighting and face position fluctuations in the real environments can be implemented by performing the identification based on the facial images involving the artificial fluctuations. The experiment with some subjects was designed, and the effectiveness and the robustness of the proposed face identification method were revealed from the experimental results.

In chapter 4, a face identification method based on the division of region on facial image was proposed. The implementation of the robustness against appearance fluctuations such as hairstyle or glasses are required in order to develop practical face identification system. In the conventional face identification, the attributes used for the identification are extracted from a whole of facial image and then the pattern recognition is performed by using them. The face identification is performed on the basis of the one evaluation criterion for the similarity calculated by the pattern recognition. However, it is difficult to implement the robustness against the appearance fluctuations by using the conventional method because the extracted attributes include the large amount of noise which is caused by hairstyle or glasses. In the proposed method, the facial image was divided into some simple regions. Therefore, the facial image

is divided into the regions which include the large amount of the noise and the regions which do not include so much noise. On each region, the attributes were extracted and the pattern recognition was performed by using them. The face identification was carried out by integrating the evaluation criteria for the similarity which were obtained on these regions. The robust face identification against the appearance fluctuations can be implemented by dividing the facial image into some regions and performing the identification based on the integration of some evaluation criteria on these regions. Through the experiment with some subjects, the effectiveness of the proposed method was confirmed. The experimental results revealed that the proposed method could implement more robust face identification against the appearance fluctuations than the conventional method.

In chapter 5, a face cognitive model acquisition method using the IGA was proposed. The face cognitive model of human must be clarified in order to develop the face recognition methods which have the consistency with the face cognitive structure of human. Oda et al. has proposed a facial image database retrieval task method to acquire the face cognitive model of human. In my study, the acquisition of the accurate face cognitive model was attempted by introducing the IGA in which human cognitive characteristic is considered into the facial image database retrieval task method. The experiments with some subjects were designed in order to confirm the effectiveness of the proposed face cognitive model acquisition method. It was found from the experimental results that the accurate face cognitive model of human could be acquired by using the proposed method.

Chapter 6 was related to the prospect for the development of comprehensive face recognition system. In this dissertation, the face recognition methods for the facial expression and the face identification were proposed. Moreover, the face cognitive model acquisition method to implement the face recognition methods which have the consistency with the face cognitive structure of human was proposed. The facial expression recognition and face identification which have the consistency with the face cognitive structure of human can be implemented by introducing to these methods the knowledge about the face cognitive model which is obtained by using the face cognitive model acquisition method. Moreover, this face cognitive model can be utilized for the face composition. That is, the face composition which can represent various kinds of the information can be implemented on the basis of this face cognitive model. The system

which supports the nonverbal communication using face between human and computer will be able to be constructed by combining the proposed face recognition methods and the face composition. In this chapter, the prospect for the development of the comprehensive face recognition system and the application fields in which such face recognition system can be utilized was discussed.

Acknowledgments

The author would like to express his heartfelt thanks to Professor Azuma Ohuchi of Hokkaido University for constant instruction and guidance from the inception to the completion of the present research.

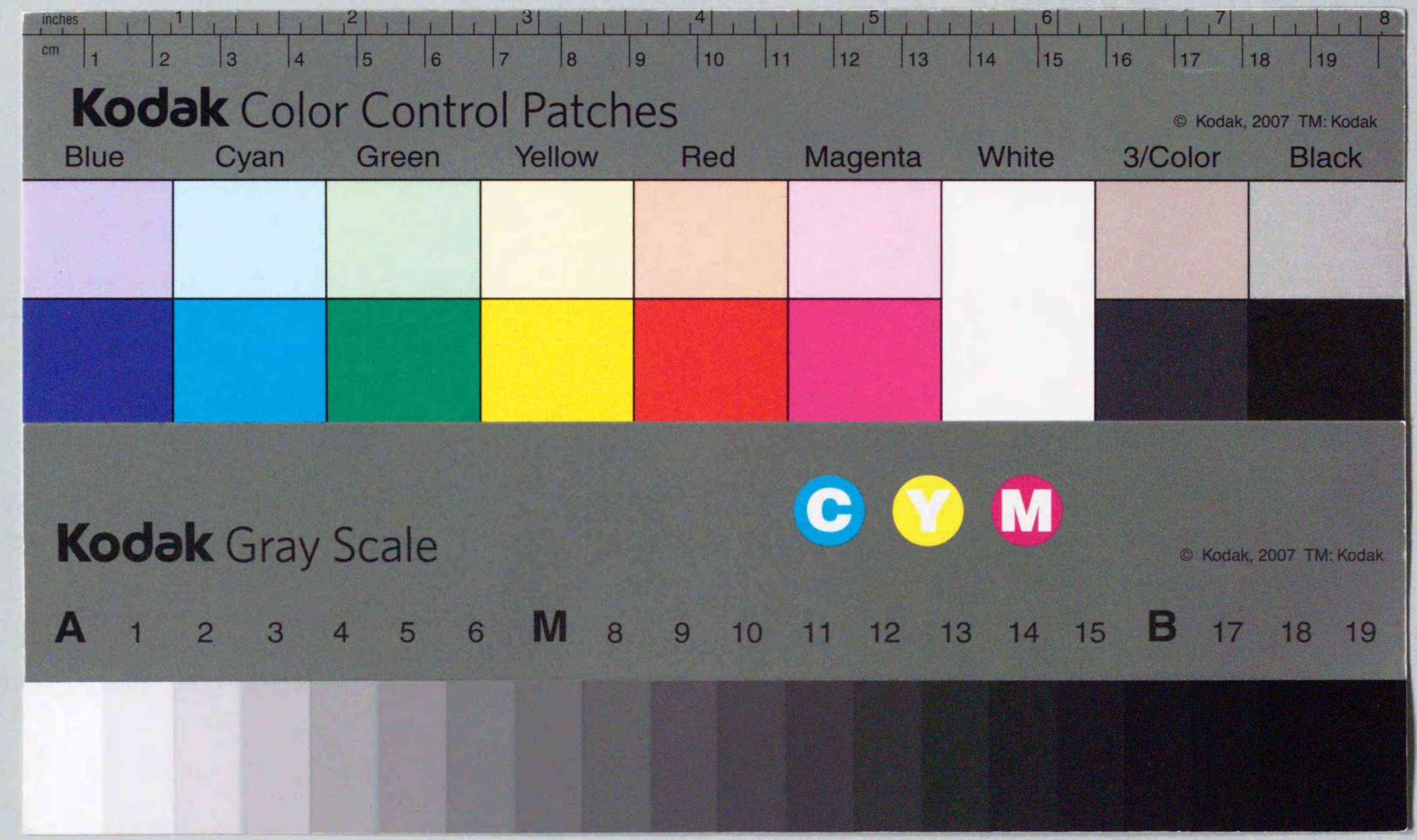
Deep thanks are also due to Professor Takashi Ohmori, Professor Yukinori Kakazu and Professor Mitsuo Wada of Hokkaido University.

He wishes to thank to Associate Professor Masahito Yamamoto and Assistant Hidenori Kawamura of Hokkaido University for profitable discussions and suggestions.

He wishes to extend deep appreciation to Associate Professor Keiji Suzuki of Future University Hakodate and Associate Professor Tamotsu Mitamura of Asahikawa Medical College for their valuable discussions and comments.

In addition, the author's thanks go to many students of the office, Laboratory of Harmonious Systems Engineering, Division of Complex Systems Engineering, especially, to Mr. Hajime Saito.

Finally, and most important, he expresses his gratitude to his family, Katsusada Nagao, Yoko Nagao and Yumie Nagao, for their support and encouragement.


inches
cm
Kodak Color Control Patches
© Kodak, 2007 TM: Kodak
Blue
Cyan
Green
Yellow
Red
Magenta
White
3/Color
Black
Kodak Gray Scale
C
Y
M
© Kodak, 2007 TM: Kodak
A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19