

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Generation of Emotional Feature Space for Facial Expression Recognition Using Self-Mapping

Masaki Ishii<sup>1</sup> and Makoto Nishida<sup>2</sup>

<sup>1</sup>Akita Prefectural University, <sup>2</sup>Akita University  
<sup>1,2</sup>Japan

## 1. Introduction

The study of facial expression recognition for the purpose of man-machine emotional communication is attracting attention lately (Akamatsu, 2002a; Akamatsu, 2002b; Akamatsu, 2002c; Akamatsu, 2003; Fasel & Luetlin, 2003; Pantic & Rothkrantz, 2000; Tian et al., 2001).

The shape (static diversity) and motion (dynamic diversity) of facial components such as the eyebrows, eyes, nose, and mouth manifest expression. Considering facial expression from a viewpoint of static diversity, because facial configurations differ for every person, it is presumed that a facial expression pattern appearing on a face when facial expression is manifested includes subject-specific features. In addition, from the viewpoint of dynamic diversity, because the dynamic change of facial expression originates in a subject-specific facial expression pattern, it is presumed that the displacement vector of facial components has subject-specific features.

On the other hand, although a facial expression pattern appearing on a face by emotion is peculiar to an individual, the internal emotion that humans express on the face and the emotion that humans recognize from the facial expression are considered to be person-independent and universal. For example, one expresses the common emotion of happiness using various facial expressions, whereas another person might recognize the common emotion of happiness from these various facial expressions. Pantic et al. argue that a natural facial expression always includes various emotions, and that a pure facial expression rarely appears (Pantic & Rothkrantz, 2000). Furthermore, they suggest that it is not realistic to classify all facial expressions appearing on a face into one of the six basic emotion categories: anger, sadness, disgust, happiness, surprise and fear, and that quantitative classification into many emotion categories must be performed instead.

Precedent works on the quantification of emotion recognized from facial expression are found in the field of psychology. Especially, the mental space model of Russell et al. is well known: each facial expression is arranged on a space centering on "pleasantness" and "arousal", particularly addressing the semantic antithetical nature of emotion (Russell & Bullock, 1985). Russell et al. discovered that facial expression stimuli can be conceptualized as arranged in a circle in the mental space described above (The Circumplex

Model). Yamada found a significant correlation between "slantedness" and "curvedness/openness" of facial components and "pleasantness" and "arousal" on the mental space (Yamada, 2000). This observation underscores the importance of clarifying a relationship of correspondence between the change of facial components accompanying emotional expression (physical parameters) and recognized emotion (psychological parameters).

We propose the following two subjects for recognizing emotion from facial expression.

The first subject concerns a facial expression pattern as a physical parameter. Expression carries personality, and facial expression patterns, as physical parameters, differ among people. For those reasons, the classification problem of facial expression is fundamentally a problem with an unknown number of categories. Therefore, it is an important subject how subject-specific facial expression categories are extracted using a common and person-independent technique.

The second subject relates to emotion as a psychological parameter. Although a facial expression pattern is peculiar to an individual, emotion as a psychological parameter is person-independent and universal. Moreover, the grade of the emotion that is recognized changes according to that of physical change of a facial expression pattern. It is therefore considered a key subject to match the amount of physical change of a subject-specific facial expression pattern with that of psychological change corresponding to that extent, to estimate the grade of emotion.

This chapter presents a generation method of a subject-specific emotional feature space using the Self-Organizing Maps (SOM) (Kohonen, 1995) and the Counter Propagation Networks (CPN) (Nielsen, 1987). The feature space expresses the correspondence relationship between the change of facial expression pattern and the strength of emotion on the two-dimensional space centering on "pleasantness" and "arousal".

Specifically, the proposed method first hierarchically classified facial expression images using a SOM, and extracted subject-specific facial expression categories. Next, data expansion of facial expression patterns based on the similarity and continuity of each facial expression category was performed using CPN. Then a subject-specific emotion feature space was generated; the space matches a physical and psychological parameter by inputting the coordinate value on the Circumplex model of Russell as a teaching signal of CPN.

Experimental results suggested that our method was useful to estimate strength and mixture level of six basic emotions.

## 2. Algorithms of SOM and CPN

### 2.1 Self-Organizing Maps (SOM)

The SOM is a learning algorithm that models the self-organizing and adaptive learning capabilities of a human brain (Kohonen, 1995). A SOM comprises two layers: an input layer, to which training data are supplied; and a Kohonen layer, in which self-mapping is performed by competitive learning. The learning algorithm of a SOM is described below.

1) Let  $w_{i,j}(t)$  be a weight from an input layer unit  $i$  to a Kohonen layer unit  $j$  at time  $t$ . Actually,  $w_{i,j}$  is initialized using random numbers.

2) Let  $x_i(t)$  be input data to the input layer unit  $i$  at time  $t$ ; calculate the Euclidean distance  $d_j$  between  $x_i(t)$  and  $w_{i,j}(t)$  using (1).

$$d_j = \sqrt{\sum_{i=1}^I (x_i(t) - w_{i,j}(t))^2} \quad (1)$$

3) Search for a Kohonen layer unit to minimize  $d_j$ , which is designated as a winner unit.  
4) Update the weight  $w_{i,j}(t)$  of a Kohonen layer unit contained in the neighborhood region of the winner unit  $N_c(t)$  using (2), where  $a(t)$  is a learning coefficient.

$$w_{i,j}(t+1) = w_{i,j}(t) + a(t)(x_i(t) - w_{i,j}(t)) \quad (2)$$

5) Repeat processes 2)–4) up to the maximum iteration of learning.

## 2.2 Counter Propagation Networks (CPN)

The CPN is a learning algorithm that combines the Grossberg learning rule with the SOM (Nielsen, 1987). A CPN comprises three layers: an input layer to which training data are supplied; a Kohonen layer in which self-mapping is performed by competitive learning; and a Grossberg layer, which labels the Kohonen layer by the counter propagation of teaching signals. A CPN is useful for automatically determining the label of a Kohonen layer when a category in which training data will belong is predetermined. This labeled Kohonen layer is designated as a category map. The learning algorithm of a CPN is described below.

1) Let  $w_{i,n,m}^i(t)$  and  $w_{i,n,m}^j(t)$  respectively indicate weights to a Kohonen layer unit  $(n, m)$  at time  $t$  from an input layer unit  $i$  and from a Grossberg layer unit  $j$ . In fact,  $w_{i,n,m}^i$  and  $w_{i,n,m}^j$  are initialized using random numbers.

2) Let  $x_i(t)$  be input data to the input layer unit  $i$  at time  $t$ , and calculate the Euclidean distance  $d_{n,m}$  between  $x_i(t)$  and  $w_{i,n,m}^i(t)$  using (3).

$$d_{n,m} = \sqrt{\sum_{i=1}^I (x_i(t) - w_{i,n,m}^i(t))^2} \quad (3)$$

3) Search for a Kohonen layer unit to minimize  $d_{n,m}$ , which is designated as a winner unit.  
4) Update weights  $w_{i,n,m}^i(t)$  and  $w_{i,n,m}^j(t)$  of a Kohonen layer unit contained in the neighborhood region of the winner unit  $N_c(t)$  using (4) and (5), where  $a(t)$ ,  $\beta(t)$  are learning coefficients, and  $t_j(t)$  is a teaching signal to the Grossberg layer unit  $j$ .

$$w_{i,n,m}^i(t+1) = w_{i,n,m}^i(t) + a(t)(x_i(t) - w_{i,n,m}^i(t)) \quad (4)$$

$$w_{i,n,m}^j(t+1) = w_{i,n,m}^j(t) + \beta(t)(t_j(t) - w_{i,n,m}^j(t)) \quad (5)$$

5) Repeat processes 2)–4) up to the maximum iteration of learning.

6) After learning is completed, compare weights  $w_{j,n,m}^i$  observed from each unit of the Kohonen layer; and let the teaching signal of the Grossberg layer with the maximum value be the label of the unit.

3. Proposed Method

Figure 1 depicts the procedure used for the proposed method. The proposed method consists of following three steps.

- Step1: Extraction of subject-specific facial expression categories using the SOM.
- Step2: Generation of a Facial Expression Map using the CPN.
- Step3: Generation of a Emotion Map using the CPN.

The proposed method is explained in detail below.

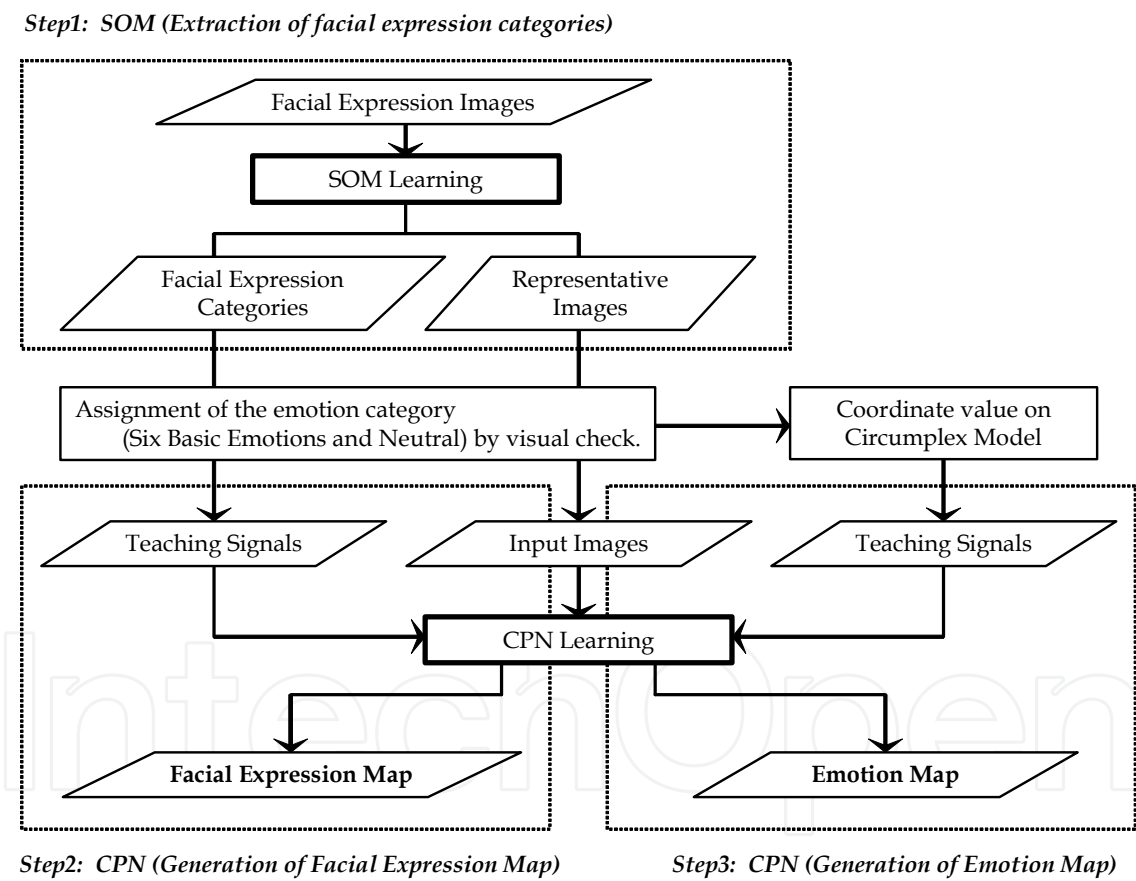


Fig. 1. Flow chart of proposal method.

### 3.1 Extraction of Facial Expression Category

The proposed method was used in an attempt to extract a subject-specific facial expression category hierarchically using a SOM with a narrow mapping space. A SOM is an unsupervised learning algorithm and classifies given facial expression images self-organizedly based on their topological characteristics. For that reason, it is suitable for a classification problem with an unknown number of categories. Moreover, a SOM compresses the topological information of facial expression images using a narrow mapping space and performs classification based on features that roughly divide the training data. We speculate that repeating these hierarchically renders the classified amount of change of facial expression patterns comparable; thereby, a subjectspecific facial expression category can be extracted. Figure 2 depicts the extraction procedure of a facial expression category. Details of the process are explained below.

- 1) Expression images described in Section 4 were used as training data. The following processing was performed for each facial expression. The number of training data is assumed as  $N$  frames.
- 2) Learning was conducted using a SOM with a Kohonen layer of five units and an input layer of  $40 \times 48$  units (Fig. 2(a)), where the number of learning sessions  $t$  was set as 10,000 times. The initial value of neighborhood  $N_c(t)$  was set as the forth neighbor of a winner unit. Then  $N_c(t)$  was linearly reduced to the first neighbor at the maximum learning number. The initial value of learning coefficient  $a(t)$  was set to 0.5, and was  $a(t)$  linearly reduced to zero similarly.
- 3) The weight of the Kohonen layer  $W_{ij}$  ( $0 \leq W_{ij} \leq 1$ ) was converted to a value of 0-255 after the end of learning, and a visualized images were generated (Fig. 2(b)), where  $n_1 - n_5$  are the number of training data classified into each unit.
- 4) Five visualized images can be considered as representative vectors of the training data classified into each unit ( $n_1 - n_5$ ). Therefore, whether a visualized image was suitable as a representative vector was judged using a threshold process. Specifically, for the upper and lower faces presented in Fig. 2(c), a correlation coefficient between a visualized image and classified training data was determined for each unit. The standard deviation of those values was computed. When the standard deviation of both regions was 0.005 or less in all five units, the visualized image was considered to represent training data and the subsequent hierarchization processing was cancelled. It was continued when any unit over 0.005 remained.
- 5) The correlation coefficient of weight  $W_{ij}$  between each adjacent unit in the Kohonen layer was computed. The Kohonen layer was divided into two bordering on between the units of the minimum (Fig. 2(b)).
- 6) The training data ( $N_1$  and  $N_2$ ) classified into both sides of the border were used as new training data; processing described above was repeated recursively. Consequently, the hierarchic structure of a SOM was generated (Fig. 2(b), 2(d)).
- 7) The lowermost hierarchy of the hierarchic structure was defined as a facial expression category (Fig. 2(e)). Five visualized images were defined as representative images of each category after learning completion. Then the photographer of the facial expression images performed visual confirmation to each facial expression category and conducted implication in emotion categories.



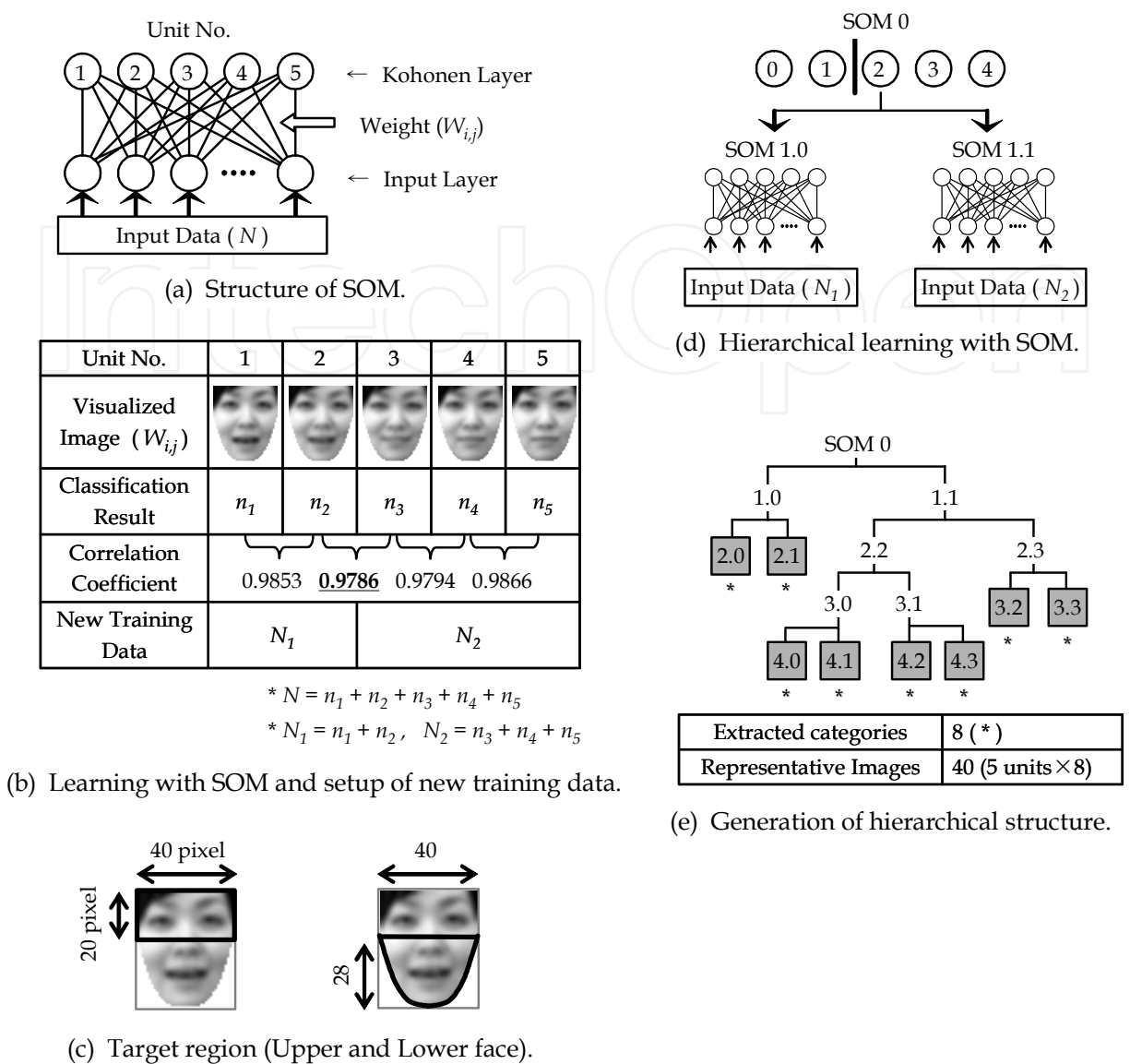


Fig. 2. Extraction procedure of a facial expression category.

3.2 Generation of Facial Expression Map

It is considered that recognition to a natural facial expression requires generation of a facial expression pattern (mixed facial expression) that interpolates each emotion category. The proposed method used the representative image obtained in Section 3.1 as training data and carried out data expansion of facial expression patterns between each emotion category using CPN with a large mapping space. The reason for adopting CPN, a supervised learning algorithm, is that the teaching signal of training data is known by processing in Section 3.1. The mapping space of CPN has a greater number of units than the number of training data, and has a torus structure because it is presumed that a large mapping space allows CPN to perform data expansion based on the similarity and continuity of training data. Figure 3 depicts the generation procedure of facial expression map and emotion map described in Section 3.3. The details of processing are described below.

- 1) In fact, CPN has a structure comprising an input layer of  $40 \times 48$  units and a two-dimensional Kohonen layer of  $30 \times 30$  units. In addition, the Grossberg layer 1 of seven units was prepared, to which the teaching signal of six basic facial expressions and a neutral facial expression were input (Fig. 3(a)).
- 2) Representative images obtained in Section 3.1 were used as training data, and learning was carried out for each subject. As the teaching signal to the Grossberg layer 1, 1 was input into units that mean emotion categories of representative images, otherwise 0. The number of learning was set to 20,000 times. The initial value of the radius of neighborhood  $N_c(t)$  was set as  $1/2$  of that of Kohonen layer. Then the radius was reduced linearly to the first neighbor at the maximum learning number. The initial values of training rate coefficients  $\alpha(t)$  and  $\beta(t)$  were both set to 0.5. Then they were linearly reduced to zero similarly.
- 3) The processing described above was repeated to the maximum learning number.
- 4) The weights ( $W_{g1}$ ) of the Grossberg layer 1 were compared for each unit of the Kohonen layer after learning completion; an emotion category of the greatest value was used as the label of the unit. A category map generated by the processing described above was defined as a subject-specific facial expression map (FEMap).

### 3.3 Generation of Emotion Map

Even if the facial expression pattern appearing on a face is peculiar to an individual, the internal emotion that humans express on the face and the emotion that humans recognize from the facial expression are considered to be person-independent and universal. Therefore, it is presumed necessary to match the grade of emotion based on a common index for each subject to the grade of change of facial expression patterns extended in Section 3.2. The proposed method is centered upon the Circumplex model of Russell (Russell & Bullock, 1985) as a common index. Specifically, the coordinate values based on the Circumplex model were input as teaching signals of CPN, in parallel to processing in Section 3.2. Then generation of an emotion feature space was tried, which matches the grade of change of facial expression patterns and the grade of emotion. The details of processing are described as follows.

- 1) The Grossberg layer 2 of one unit that inputs the coordinate values of the Circumplex model was added to the CPN structure (Fig. 3(a)).
- 2) Each facial expression stimulus is arranged in a circle on a space centering on "pleasantness" and "arousal" in the Circumplex model (Fig. 3(b)). The proposed method expresses this circular space as the complex plane depicted in Fig. 3(c), and complex number based on the figure were input to the Grossberg layer 2 as teaching signals. For example, when a inputted training data is an emotion category of happiness, a teaching signal for Grossberg layer 2 is  $\cos(\pi/4) + i \sin(\pi/4)$ .
- 3) This processing was repeated to the maximum learning number.
- 4) Each unit of the Kohonen layer was plotted onto the complex plane after learning completion based on the values of the real and imaginary parts of the weight ( $W_{g2}$ ) on Grossberg layer 2. Then this complex plane was defined as a subject-specific emotion map (EMap).



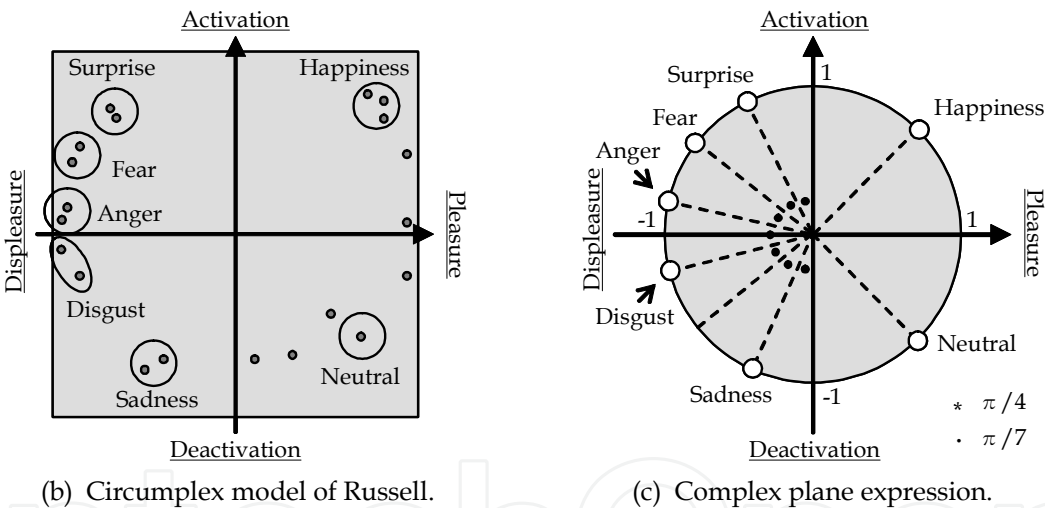
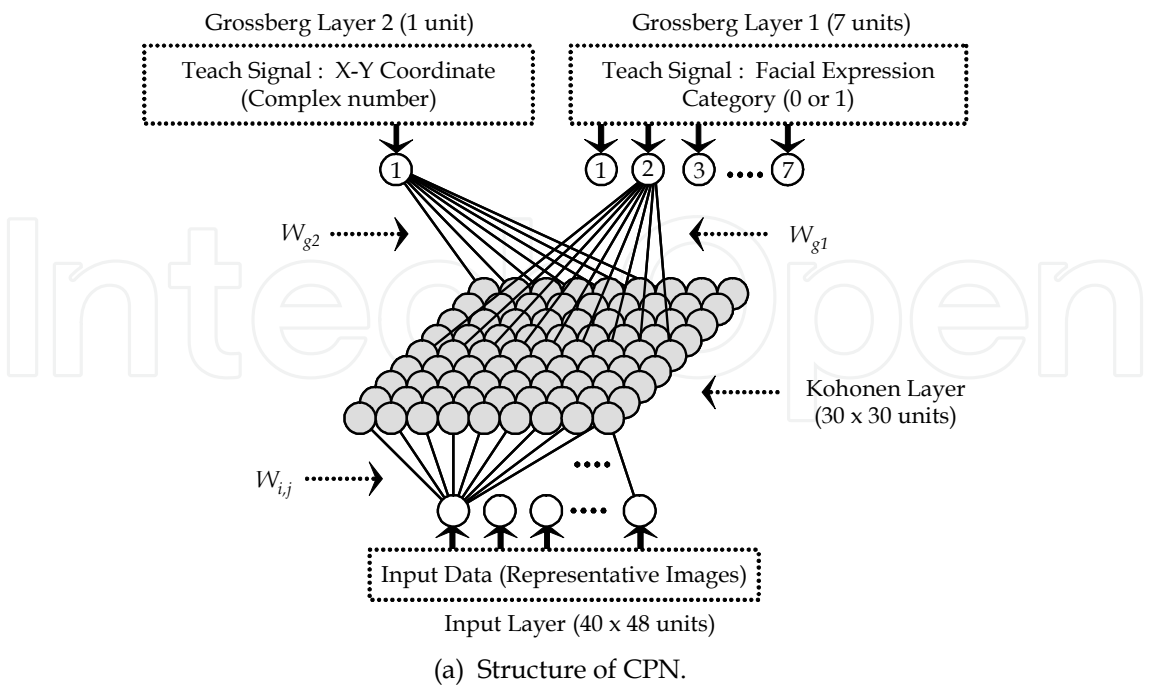


Fig. 3. Generation procedure of facial expression map (FEMap) and emotion map (EMap).

4. Expression Images

Open facial expression databases are generally used in conventional studies (Pantic et al., 2005; Gross, 2005). These databases contain a few images per expression and subject. For this study, we obtained facial expression images of ourselves because the proposed method extracts person-specific facial expression categories and the representative images of each category from large quantities of data.

This chapter presents a discussion of six basic facial expressions and a neutral facial expression that four subjects manifested intentionally. Basic facial expressions were obtained as motion videos including a process in which a neutral facial expression and facial expressions were manifested five times respectively by turns for each facial expression.

Neutral facial expressions were obtained as a motion video for about 20 s. The motion videos were converted into static images (30 frame/s, 8 bit gray, 320 × 240 pixels) and used as training data. Table 1 presents the number of frames of all the subjects’ training data. A region containing facial components was processed in this chapter; extraction and normalization of a face region image were performed according to the following procedures. Figure 4 shows an example of face region images after extraction and normalization.

- 1) A face was detected using Haar-like features (Lienhart & Maydt, 2002); a face region image normalized into a size of 80 × 96 pixels was extracted.
- 2) The image was processed using a median filter for noise removal. Then smoothing processing was performed after dimension reduction of the image using coarse grain processing (40 × 48 pixels).
- 3) A pseudo outline that is common to all the subjects was generated; the face region containing facial components was extracted.
- 4) Histogram linear transformation was performed for brightness value correction.

ID	An.	Sa.	Di.	Ha.	Su.	Fe.	Ne.	Total
A	454	448	533	530	473	479	605	3,522
B	284	337	323	335	325	323	599	2,526
C	493	556	592	475	455	471	599	3,641
D	574	518	457	473	459	468	605	3,554

Table 1. Number of frames of all subjects' training data (ID, Subject; An., Anger; Sa., Sadness; Di., Disgust; Ha., Happiness; Su., Surprise; Fe., Fear; Ne., Neutral).

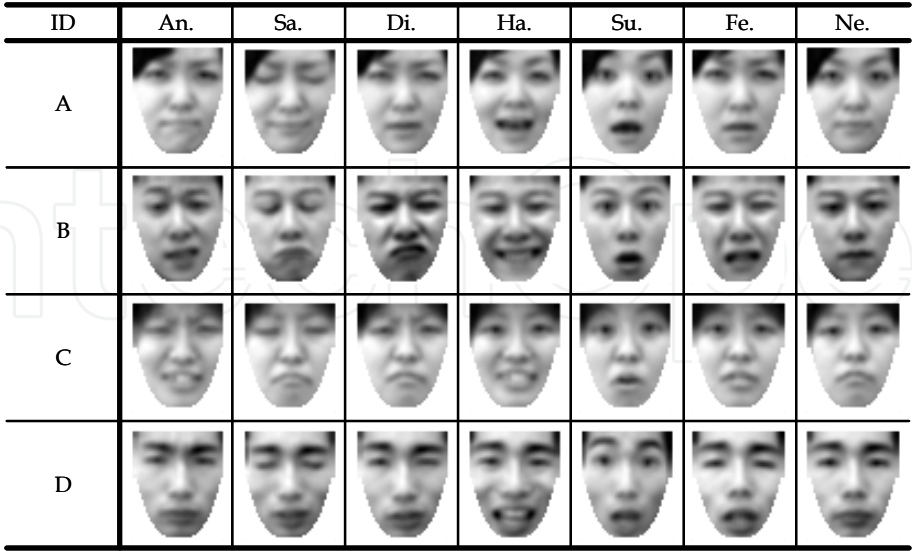


Fig. 4. Examples of facial expression images (ID, Subject; An., Anger; Sa., Sadness; Di., Disgust; Ha., Happiness; Su., Surprise; Fe., Fear; Ne., Neutral).

5. Results and Discussion

5.1 Extraction of Facial Expression Categories

Tables 2 and 3 present the number of facial expression categories and representative images for each subject extracted using the proposed method, which demonstrates that the number of facial expression categories differ for each subject. For example, for anger, 6 and 8 categories were extracted respectively from Subjects A and C, whereas only 3 and 2 categories were extracted respectively from Subjects B and D. This fact suggests that Subjects A and C have more facial expression patterns to express anger compared with Subjects B and D.

Figure 5 presents the hierarchical structure of a SOM for the happiness of Subject A. Expression categories extracted from the figure comprise eight categories of Happiness and 11 categories of neutral facial expressions, from which 95 representative images were generated. Figure 6 shows the facial expression category of happiness, in which categories of different grades of change of facial expression patterns are extracted, such as mouth wide open, mouth narrowly open, eye closed and mouth shut, eye closed and mouth open, and smile. The topological characteristic of facial expression images differs for every subject with the personality of facial expression. Because the proposed method performs hierarchical classification based on topological characteristics that are peculiar to a subject, it is presumed that the extraction result of a different facial expression category was obtained for each subject.

These results indicate that hierarchical use of SOM with a narrow mapping space is useful as a category extraction technique for facial expressions that are peculiar to an individual. It is particularly applicable when the classification problem of facial expression is considered to be a classification problem with an unknown number of categories.

( ) : Neutral

ID	An.	Sa.	Di.	Ha.	Su.	Fe.	Ne.	Total
A	6 ( 7 )	6 ( 10 )	8 ( 6 )	8 ( 11 )	4 ( 13 )	6 ( 12 )	0 ( 9 )	106
B	3 ( 5 )	4 ( 1 )	6 ( 6 )	6 ( 6 )	6 ( 6 )	9 ( 6 )	0 ( 6 )	70
C	8 ( 7 )	8 ( 4 )	1 ( 5 )	8 ( 8 )	5 ( 8 )	7 ( 10 )	0 ( 2 )	81
D	2 ( 1 )	4 ( 8 )	9 ( 6 )	8 ( 10 )	5 ( 9 )	8 ( 6 )	0 ( 1 )	77

Table 2. Number of facial expression categories extracted with proposed method.

ID	An.	Sa.	Di.	Ha.	Su.	Fe.	Ne.	Total
A	30	30	40	40	20	30	340	530
B	15	20	30	30	30	45	180	350
C	40	40	5	40	25	35	220	405
D	10	20	45	40	25	40	205	385

Table 3. Number of representative images of each category.

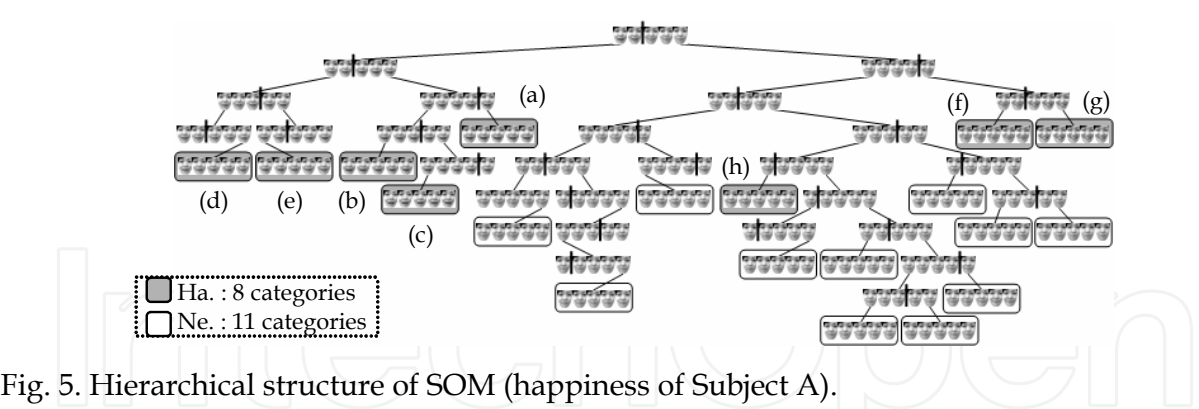


Fig. 5. Hierarchical structure of SOM (happiness of Subject A).

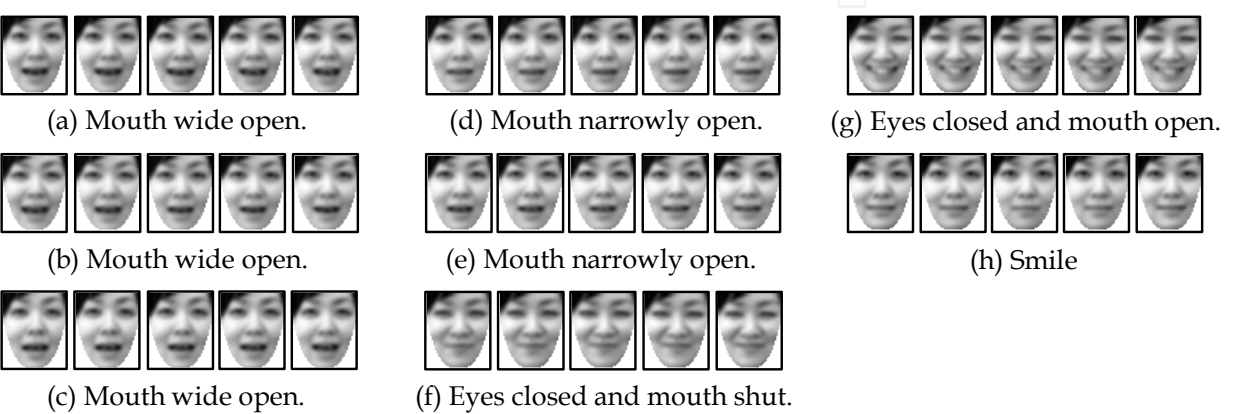


Fig. 6. Representative images (Happiness of Subject A).

5.2 Generation of FEMap and EMap

Figures 7 and 8 respectively present the FEMaps and EMaps generated using the proposed method.

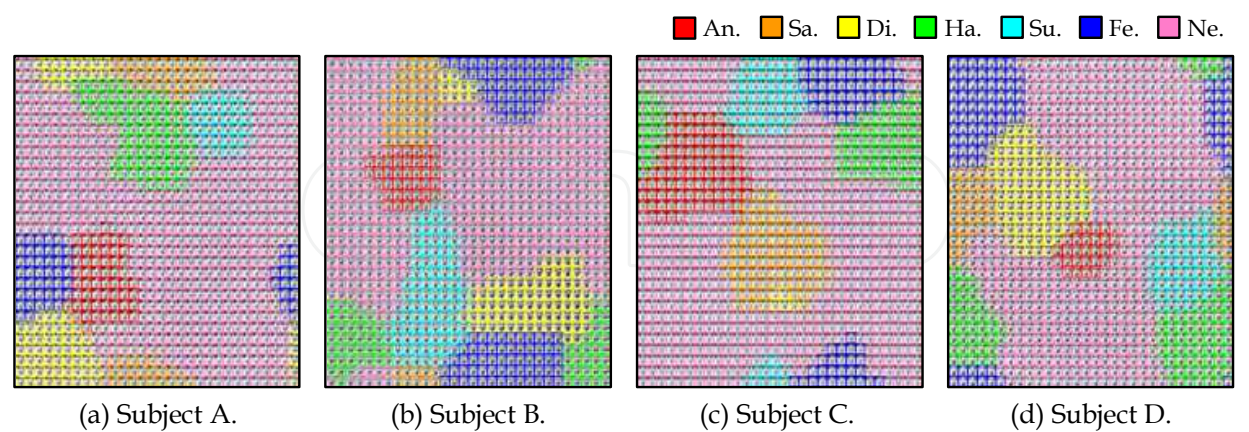


Fig. 7. Generation results of Facial Expression Map (FEMap).

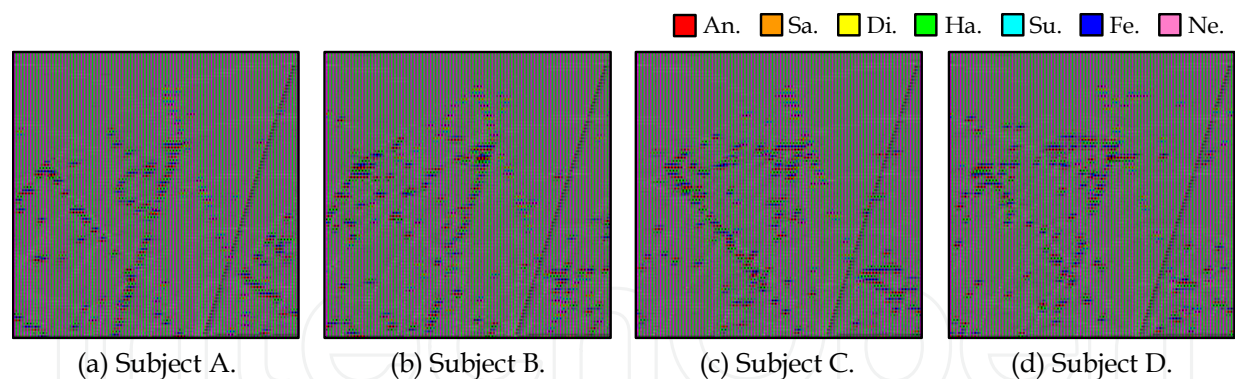


Fig. 8. Generation results of Emotion Map (EMap).

A FEMap was generated as depicted in Fig. 7, in which the adjacency and area of emotion categories differs for each subject. Table 4 shows the number of units of each emotion category on the FEMap. Table 5 provides the number of winner units when inputting the training data of Table 3 into the FEMap. For example, 252 out of 900 were winner units for Subject A. This implies that the remaining 648 units were generated as units that interpolate between training data by learning of CPN.

ID	An.	Sa.	Di.	Ha.	Su.	Fe.	Ne.	Total
A	54	56	78	82	36	57	537	900
B	40	51	84	69	88	112	456	900
C	90	87	5	81	62	78	497	900
D	27	47	89	109	55	102	471	900

Table 4. Number of units of each emotion category on FEMap.

ID	A	B	C	D
Number of Kohonen layer units	900	900	900	900
Number of winner units (Inputs are training data)	252	183	244	231
Number of extended units	648	717	656	669

Table 5. Number of winner units when inputting training data.

On the other hand, the EMap was generated as a feature space based on a common index for each subject, as portrayed in Fig. 8. Figure 9 presents details of the EMap of Subject A, where (a) presents the plot of winner units when inputting training data, and (b) shows the plot of all units (900 units). Figure 9(a) portrays that winner units, when inputting training data, concentrate near the coordinate values of the teaching signal of CPN, based on the Circumplex model of Russell (Fig. 3(c)). On the other hand, Fig. 9(b) reveals that the extended units interpolate the result of Fig. 9(a). The region of happiness shown in Fig. 9(b) is magnified into Fig. 9(c). Because region A in Fig. 9(c) is near the coordinate value of the



teaching signal of happiness, it is suggested that facial expressions representing pure happiness are plotted in this region. Meanwhile, because Region C shows less “arousal” than Region A, facial expressions of a transition process from happiness to neutral is plotted in this region. Similarly, facial expressions of the emotion of surprise are mixed in Region B; facial expressions of sadness are mixed in Region D.

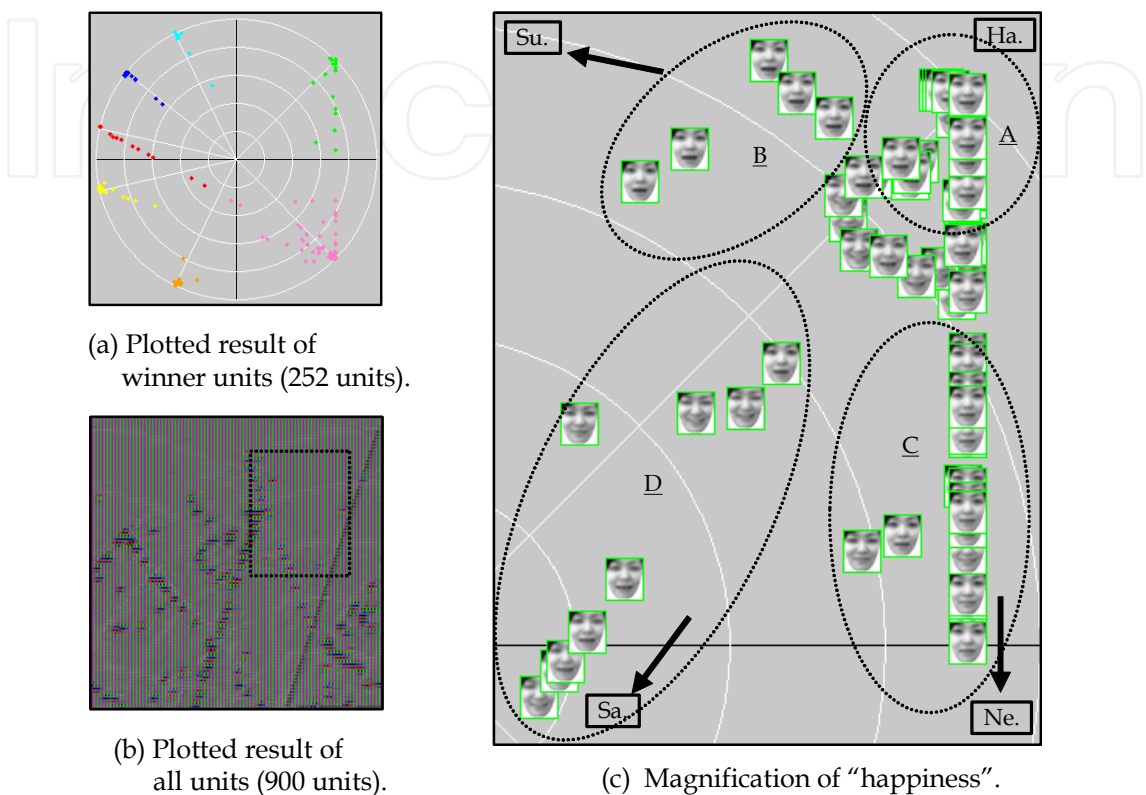


Fig. 9. Detailed analysis of EMap (Happiness of Subject A).

These results suggest that data expansion is performed based on the similarity and continuity of training data, and that more facial expression patterns such as mixed facial expressions between emotion categories can be generated in the mapping space of CPN. It is also inferred that the grade of emotion with “pleasantness” and “arousal” as indices can be matched to the grade of change of facial expression patterns on the EMap by assigning coordinate values based on the Circumplex model of Russell as a teaching signal of CPN.

6. Evaluation Experiments

An experiment was conducted that was expected to estimate the grade of emotion using time series data of newly obtained facial expressions to verify the usefulness of the proposed method. Test data for six basic facial expressions were obtained in the same conditions as those described in Section 4.

Figure 10 shows the recognition result for “fear” and “surprise” of Subject A and B, which reveals pleasantness value and arousal value gradually change with the change of facial expression pattern. Moreover, the change of pleasantness value and arousal value is similar, although facial expression patterns of two subjects are different.



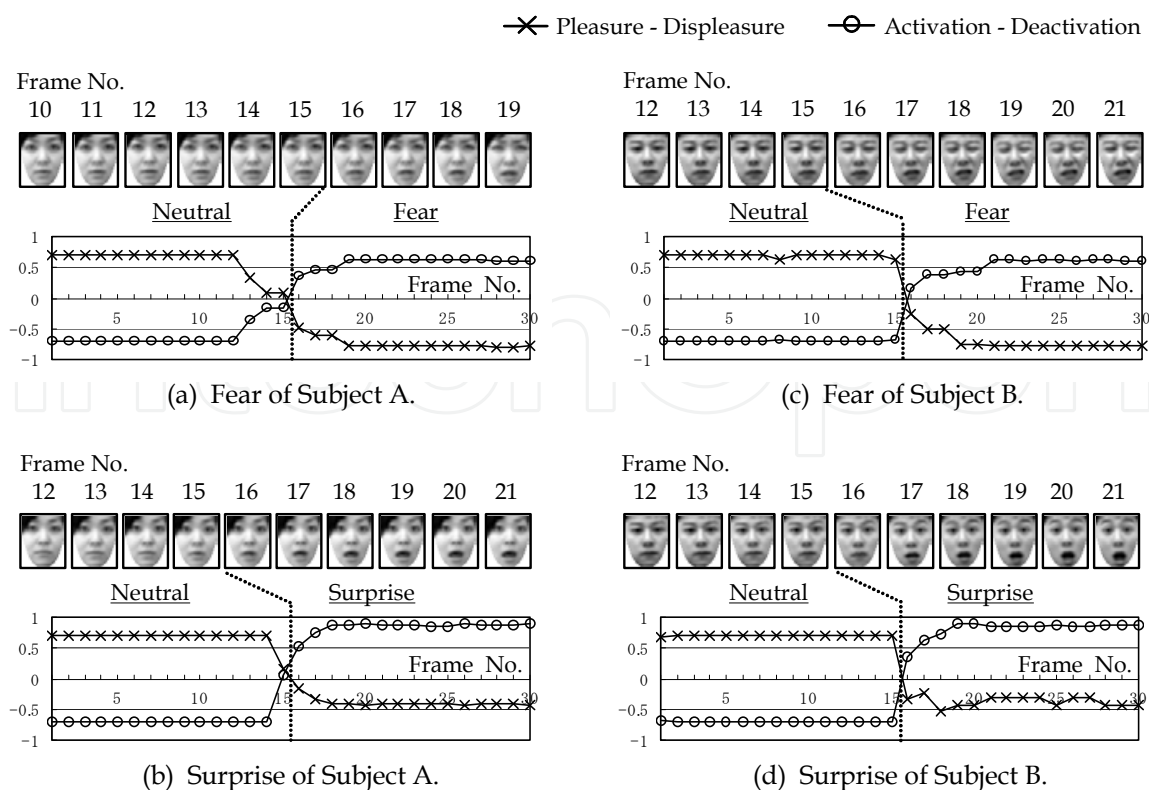


Fig. 10. Recognition result for “fear” and “surprise” of Subject A and B.

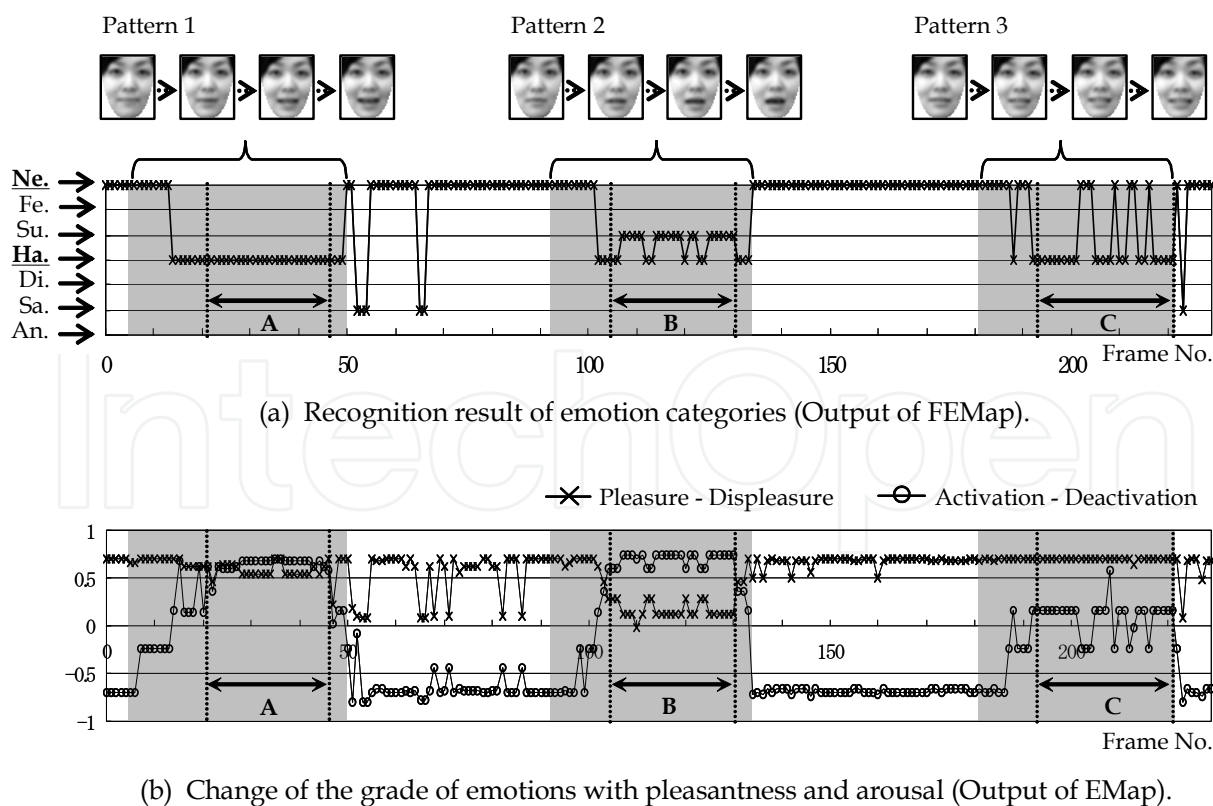


Fig. 11. Recognition result of happiness which is different facial pattern.

Figure 11 depicts the recognition result for the happiness of Subject A. The test data contain facial expressions of three patterns of happiness for which the way of manifestation differ, as shown in Fig. 11(a). That figure portrays that all the frames in Section A are recognized as happiness, where "pleasantness" and "arousal" remain high, as shown in Fig. 11(b). Section B is recognized as happiness or surprise, and "pleasantness" remains low compared to Section A; Section C is recognized as happiness or neutral, and "arousal" remains low.

Figure 12 expresses the state transition of Patterns 1, 2, and 3 on an EMap. Regions A, B, and C in the figure mark regions in which the frames of Sections A, B, and C in Fig. 11 are plotted. Although "arousal" increases from near the coordinate value of neutral expression gradually in Patterns 1, 2 and 3, they deviate from one another when the values turn positive, as presented in Fig. 12. The state of Pattern 1 is stabilized near the coordinate value of happiness (region A). Therefore, it is surmised that these time series data manifest the typical emotion of happiness. In addition, because Pattern 2 is stable near the middle of happiness and surprise (region B), it is conjectured to represent the state of excitement and upsurge in which the emotion of happiness and surprise is mixed. Furthermore, because Pattern 3 is stable near the middle of happiness and neutral expression (region C), it is inferred that a state of satisfaction and relief is expressed, i.e. low-level manifestation of happiness.

These results demonstrate that an EMap can quantify the grade of emotion based on common indices, such as "pleasantness" and "arousal", to the grade of change of facial expression patterns, and that emotion estimation can be performed to facial expressions for which two or more emotions are mixed.

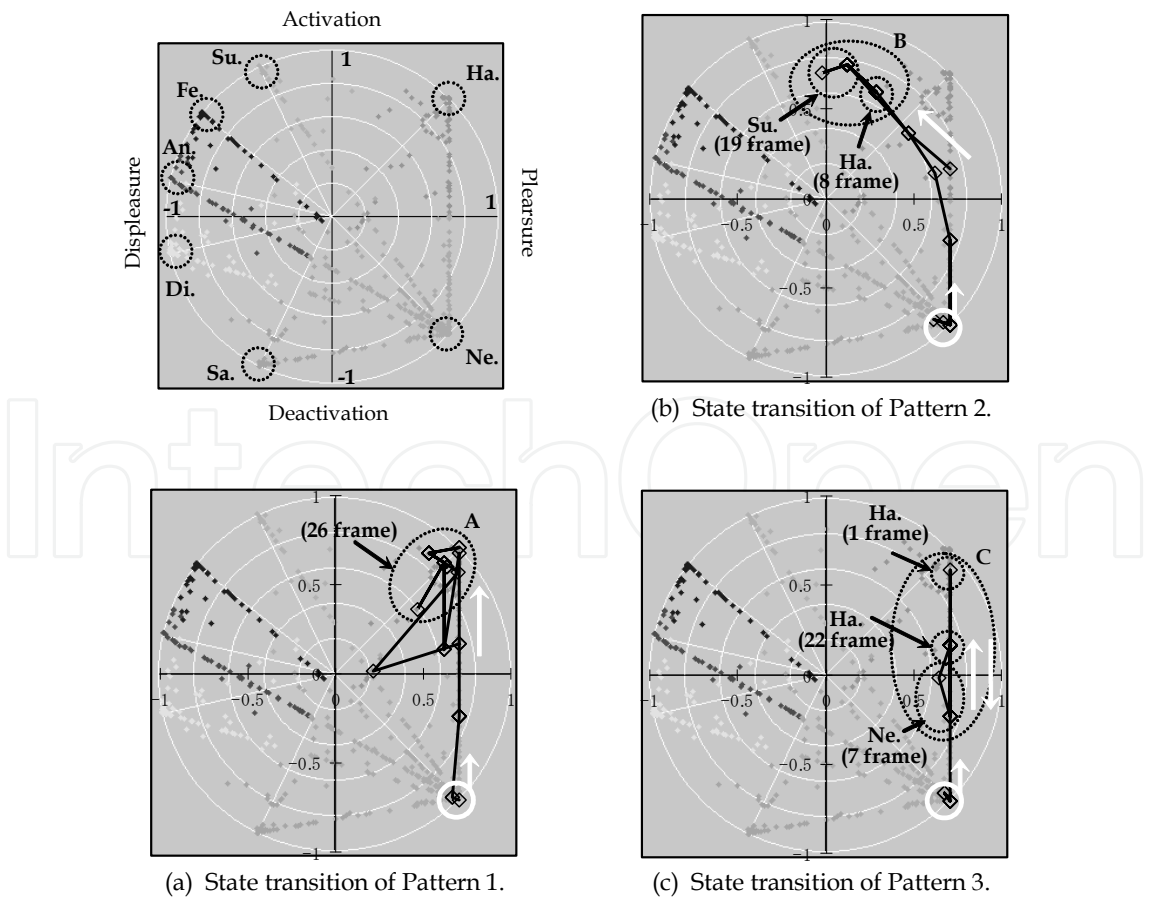


Fig. 12. State transition on EMap (Pattern 1, 2 and 3 in Fig.11).

## 7. Conclusion

The study described in this chapter examined the method of generating a subject-specific emotion feature space for estimating the grade of emotion. The essential results obtained in this chapter are the following.

- 1) Hierarchical use of SOM with a narrow mapping space enables extraction of subject-specific expression categories.
- 2) The grade of emotions with "pleasantness" and "arousal" as indices can be matched to the grade of change of facial expression patterns on an EMap that is generated using the proposed method.
- 3) An EMap enables quantification of the grade of emotion to the grade of change of facial expression patterns, and to conduct emotion estimation to mixed facial expressions.

Objective evaluation on the relationship between the temporal change of facial expression patterns and state transition on an EMap accompanied by the context of a scene will be performed in a future study using natural facial expressions during conversation.

## 8. Acknowledgments

This work was supported in part by Ministry of Education, Culture, Sports, Science and Technology (MEXT) Grants-in-Aid for Young Scientists (B): No. 18700192 and No. 20700174. This chapter is based on "Generation of Emotional Feature Space based on Topological Characteristics of Facial Expression Images," by M. Ishii, K. Sato, H. Madokoro and M. Nishida, which appeared in the Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, (FG 2008), Amsterdam, The Netherlands, September 2008. © 2008 IEEE.

## 9. References

- Akamatsu, S. (2002a). Recognition of Facial Expressions by Human and Computer [I]: Facial Expressions in Communications and Their Automatic Analysis by Computer, The Journal of the Institute of Electronics, Information, and Communication Engineers, Vol.85, No.9, pp.680-685. (in Japanese)
- Akamatsu, S. (2002b). Recognition of Facial Expressions by Human and Computer [II]: The State of the Art in Facial Expression Analysis-1; Automatic Classification of Facial Expressions, The Journal of the Institute of Electronics, Information, and Communication Engineers, Vol.85, No.10, pp.766-771. (in Japanese)
- Akamatsu, S. (2002c). Recognition of Facial Expressions by Human and Computer [III]: The State of the Art in Facial Expression Analysis-2; Recognition of Facial Actions, The Journal of the Institute of Electronics, Information, and Communication Engineers, Vol.85, No.12, pp.936-941. (in Japanese)
- Akamatsu, S. (2003). Recognition of Facial Expressions by Human and Computer [IV: Finish]: Toward Computer Recognition of Facial Expressions Consistent with the Perception by Human, The Journal of the Institute of Electronics, Information, and Communication Engineers, Vol.86, No.1, pp.54-61. (in Japanese)
- Fasel, B., & Luetttin, J. (2003). Automatic Facial Expression Analysis: A Survey, Pattern Recognition, Vol.36, pp.259-275.

- Gross, R. (2005). Face Databases, Handbook of Face Recognition, S.Li and A.Jain, ed., Springer-Verlag.
- Kohonen, T. (1995). Self-Organizing Maps, Springer Series in Information Sciences.
- Lienhart, R. & Maydt, J. (2002). An Extended Set of Haar-like Features for Rapid Object Detection, Proc. IEEE Int. Conf. Image Processing, Vol.1, pp.900-903.
- Nielsen, R.H. (1987). Counterpropagation Networks, Applied Optics, vol.26, No.23, pp.4979-4984.
- Pantic, M. & Rothkrantz, L.J.M. (2000). Automatic Analysis of Facial Expressions: The State of the Art, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.22, No.12, pp.1424-1445.
- Pantic, M.; Valstar, M.F., Rademaker, R. & Maat, L. (2005). Webbased Database for Facial Expression Analysis, Proc. IEEE Int. Conf. Multimedia and Expo, pp.317-321.
- Russell, J.A. & Bullock, M. (1985). Multidimensional Scaling of Emotional Facial Expressions: Similarity from Preschoolers to Adults, J. Personality and Social Psychology, Vol.48, pp.1290-1298.
- Tian, Y.L.; Kanade, T. & Cohn, J.F. (2001). Recognizing Action Units for Facial Expression Analysis, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.23, No.2, pp.97-116.
- Yamada, H. (2000). Models of Perceptual Judgments of Emotion from Facial Expressions, Japanese Psychological Review, Vol.43, No.2, pp.245-255.

IntechOpen

IntechOpen

IntechOpen



## **Self-Organizing Maps**

Edited by George K Matsopoulos

ISBN 978-953-307-074-2

Hard cover, 430 pages

**Publisher** InTech

**Published online** 01, April, 2010

**Published in print edition** April, 2010

The Self-Organizing Map (SOM) is a neural network algorithm, which uses a competitive learning technique to train itself in an unsupervised manner. SOMs are different from other artificial neural networks in the sense that they use a neighborhood function to preserve the topological properties of the input space and they have been used to create an ordered representation of multi-dimensional data which simplifies complexity and reveals meaningful relationships. Prof. T. Kohonen in the early 1980s first established the relevant theory and explored possible applications of SOMs. Since then, a number of theoretical and practical applications of SOMs have been reported including clustering, prediction, data representation, classification, visualization, etc. This book was prompted by the desire to bring together some of the more recent theoretical and practical developments on SOMs and to provide the background for future developments in promising directions. The book comprises of 25 Chapters which can be categorized into three broad areas: methodology, visualization and practical applications.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Masaki Ishii and Makoto Nishida (2010). Generation of Emotional Feature Space for Facial Expression Recognition Using Self-Mapping, Self-Organizing Maps, George K Matsopoulos (Ed.), ISBN: 978-953-307-074-2, InTech, Available from: <http://www.intechopen.com/books/self-organizing-maps/generation-of-emotional-feature-space-for-facial-expression-recognition-using-self-mapping>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821



© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen