Emotion Detection Algorithm Using Frontal Face Image

Article · July 2015

Young Hoon Joo
Kunsan National University

Emotion Detection Algorithm Using Frontal Face Image

Moon Hwan Kim*, Young Hoon Joo**, and Jin Bae Park***

*Department of Electrical and Electronic Engineering, Yonsei University, Seodaemun-gu, Seoul, 120-749, Korea,
e-mail:jmacs@control.yonsei.ac.kr

**School of Electronic and Information Engineering, Kunsan National University, Kunsan, Chongbuk, 573-701, Korea,
e-mail:yhjoo@kunsan.ac.kr

***Department of Electrical and Electronic Engineering, Yonsei University, Seodaemun-gu, Seoul, 120-749, Korea,
e-mail:jbpark@yonsei.ac.kr

Abstract: An emotion detection algorithm using frontal facial image is presented in this paper. The algorithm is composed of three main stages: image processing stage and facial feature extraction stage, and emotion detection stage. In image processing stage, the face region and facial component is extracted by using fuzzy color filter, virtual face model, and histogram analysis method. The features for emotion detection are extracted from facial component in facial feature extraction stage. In emotion detection stage, the fuzzy classifier is adopted to recognize emotion from extracted features. It is shown by experiment results that the proposed algorithm can detect emotion well.

Keywords: Emotion recognition, facial analysis, feature extraction, fuzzy classifier, virtual face model.

1. Introduction

Although the technology for emotion recognition is important one which demanded in various fields, it still remains as the unsolved problem. Detecting emotion of human can be achieved by using facial image, voice, body shape, and etc.. Among them, the facial image is most frequently source to detect emotion. Especially, frontal facial image is commonly used to detect emotion. Emotion recognition procedure is not simple but complex since extracting proper feature and detecting emotion need complex steps.

There are various studies to detect human emotion [3–6]. Lien *et al.* [4] use geometric information of facial image and Joo *et al.* [3] use template vector and neural network. Black and Yacoob [8] also utilized an optical flow model of image motion for facial expression analysis. Their work explores local parameterized optical flow models for the recognition of the six basic emotional expressions (sadness, happiness, anger, disgust, fear and surprise [9]. Kobayashi and Hara [10] reported on real-time recognition, singular emotional classification and synthesis of the six basic emotional expressions. They worked on realization of an animated 3D face-robot that can recognize and reproduce the emotional expressions. They use brightness distribution data of facial image and a 3-layered back-propagation neural network for classification and synthesis of facial expressions.

In this paper, we propose a new algorithm for detecting emotion via frontal facial image. The algorithm composed of three stage: image precessing stage, facial feature extraction stage, and emotion detection stage. In image processing stage, we use the proposed image processing algorithm developed in previous study [2]. To extract more effective feature, we proposed the new feature extraction method in facial feature extraction stage. The proposed feature extraction method consists of three features regions: eye region, mouth region, and auxiliary region. In each face region, we extract feature by comparing geometric and shape information. Generally, there are only vague patterns are given as the input of system in emotion recognition problem. Therefore, it is not easy to design the emotion recognition system. To overcome this difficulty, the fuzzy classifier is adopted in emotion detection stage. When the extracted features are given, the fuzzy classifier returns the recognized emotion. The fuzzy classifier is identified by linear matrix inequality (LMI) optimization method.

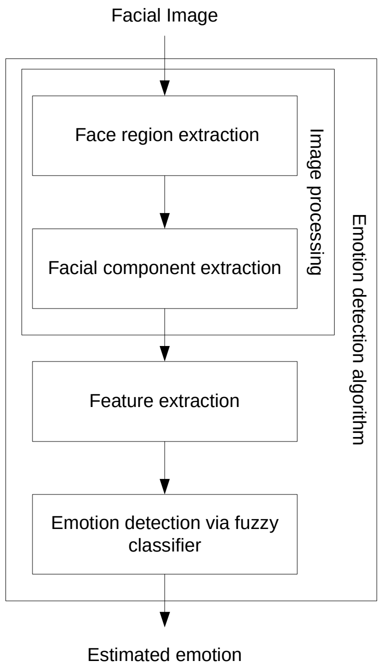

Fig. 1. Overall procedure of emotion detection algorithm.

2. Preliminaries

In this section, we present image processing stage to extract fundamental information from facial image. Figure 1 shows the overall procedure of emotion detection algorithm. In im-

This work is a result of the URC Project supported by MIC(Ministry of Information and Communication), KOREA.

age processing stage, the facial region is extracted and then facial components are extracted. In face region extraction algorithm, the fuzzy color filter and histogram analysis method are used.
It is not easy to recognize the skin color in give image because skin color changes depend on personality and illumination condition. However the skin color should be detected by using some type filter to extract the face region. To solve this difficulty, fuzzy color filter is used. The fuzzy color filter is based on fuzzy inference system. The vague skin colors are represented as fuzzy sets and memorized in fuzzy rules. The structure of fuzzy rule is represented as,

$$\begin{aligned} R_i \ &: \text{IF } x_1 \text{ is } M_{i1} \text{ and } \dots \text{ and } x_m \text{ is } M_{im} \\ &\text{THEN } y_i(x) \ = \ a_i, i = 1, \dots, l \end{aligned} \tag{1}$$

where $x_i \in \mathbb{R}$ is the ith color input, $M_{i1}, \dots, M_{im}$ are the antecedent fuzzy sets, $y_i(x)$ is the consequent output of the ith rule, $x = [x_1, \dots, x_m]^T \in F \subset \mathbb{R}^{\gtrdot}$ is the input feature vector, F is the feature vector set, and a_i is the consequent parameter and mean the weight for rule i. The output of the fuzzy rule system is inferred by following equations:

$$Y(x) = \frac{\sum_{i=1}^{l} \left\{ \prod_{j=1}^{m} \mathrm{e}^{-\frac{(c_j^i - x_j)^2}{v_j^i}} \right\} a_i}{\sum_{i=1}^{l} \prod_{j=1}^{m} \mathrm{e}^{-\frac{(c_j^i - x_j)^2}{v_j^i}}} \tag{2}$$

where c_j^i and v_j^i are the center the width of membership function of the jth feature in the ith rule.
Finally, the final output of fuzzy color filter $\hat{Y}(x)$ is calculated as

$$\hat{Y}(x) = \alpha u(Y(x) - Y_{min}) \tag{3}$$

where α is the offset value for gray image, $u(x)$ is the unit step function, and Y_{min} is minimum value of $Y(x)$. Therefore, if $Y(x)$ is greater than Y_{min}, the final output has α. Adjusting Y_{min}, we can change robustness of skin color filter. The identification method for fuzzy skin filter is discussed in [2] specifically.
When the given image is filtered via fuzzy skin filter, we can get filtered gray image $I_g \subset \mathbf{R^{m \times n}}$. Let $z_{ij} \in I_g$ is the one pixel include the face region information. By adding the gray value of each pixel, we obtain the following horizontal histogram vector $o = [o_1, \dots, o_m]$ and vertical histogram vector$p = [p_1, \dots, o_n]$,

$$o_i = \sum_{j=1}^{n} z_{ij} \tag{4}$$

$$p_j = \sum_{i=1}^{m} z_{ij}. \tag{5}$$

To extract more specific face region, histogram analysis method is used. When the gray image is converted into histogram vector o and p, the largest segment can be detected via Algorithm 1. The largest segment in o and p becomes the width and height of final face region.

Algorithm 1: Histogram Segmentation Algorithm

Data: $W = [w_1, \dots, w_c]$: histogram input
d: minimum depth of segment
r: minimum width of segment
Result: The largest segment in given histogram W
$egde_{begin}$=false
for $i = 1$ *to* c **do**
$\hat{w}_i = \prod_{k=-r}^{-1} \delta(w_{i+k} - e_{i+k}^1) \prod_{k=1}^{r} u(w_{i+k} - e_{i+k}^1) - \prod_{k=-r}^{-1} u(w_{i+k} - e_{i+k}^2) \prod_{k=1}^{r} \delta(w_{i+k} - e_{i+k}^2)$
where
$e_i^1 = \begin{cases} 0 & , -r \le i \le -1 \\ d & , 0 \le i \le r \end{cases}, e_i^2 = \begin{cases} d & , -r \le i \le -1 \\ 0 & , 0 \le i \le r \end{cases}$
if $\hat{w}_i = 1$ *and egde$_{begin}$=false* **then**
$temp = i$
$egde_{begin}$=true
endif
if $\hat{w}_i = -1$ *and egde$_{begin}$=true* **then**
$(temp, i)$ is new segment.
$egde_{begin}$=false
endif
end

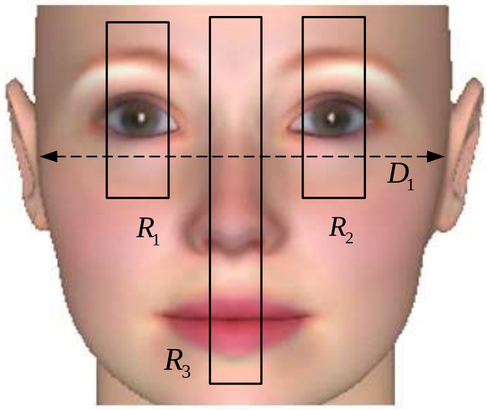

Fig. 2. Three regions in VFM.

The detailed description for Algorithm 1 presented in [2]. When face region is extracted, the facial components are extracted by using virtual face model (VFM) based histogram analysis. VFM is proposed to reduce searching space of histogram analysis method. VFM contains position and length information of each facial component. In this paper, we used VFM proposed in [2]. There are three face region to extract the facial component. By investigate each face region, we can extract accurate facial component in fast time. The detailed algorithm is presented in [2].

3. Feature Extraction for Emotion Recognition

The feature vector extraction method is most important key point in emotion recognition problem. Especially, it is necessary to get good feature vector to make better recognition accuracy. In the facial feature extraction stage, we propose a new feature vector extraction method. The proposed method divide whole image into three feature region: eye region, mouth region, and auxiliary region. Several information are extracted from each region: geometric and shape informa-

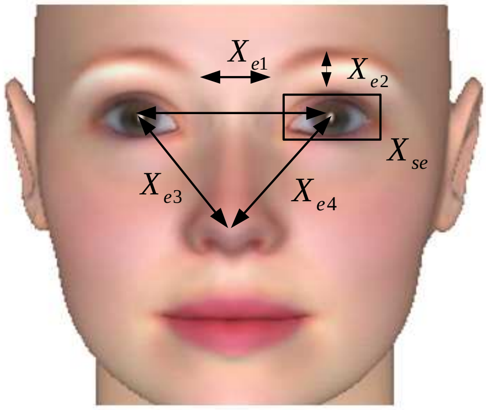

Fig. 3. Position of features in eye region

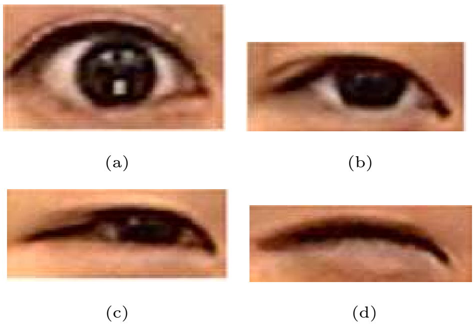

Fig. 4. Eye template for comparison.

Table 1. Features in eye region.

Features	Description	Size
X_{e1}	Distance between two eye brow	1×1
X_{e2}	Distance between eye and eye brow	1×1
X_{e3}	Distance between nose and eye(left side)	1×1
X_{e4}	Distance between nose and eye(right side)	1×1
X_{se}	Error between eye and template	4×1

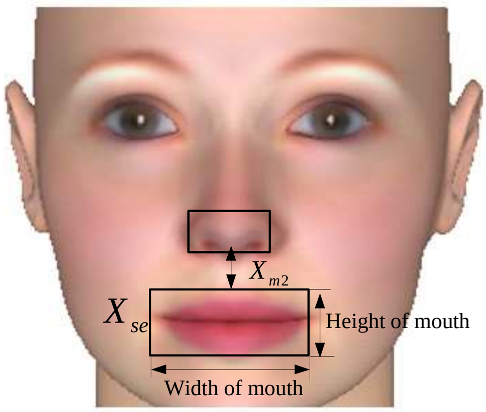

Fig. 5. Position of features in mouth region.

Table 2. Features in mouth region.

Features	Description	Size
X_{m1}	$\frac{\text{Width of mouth}}{\text{Height of mouth}}$	1×1
X_{m2}	Distance between nose and mouth	1×1
X_{se}	Error between mouth and template	6×1

Fig. 6. Mouth template for comparison.

tion. Table 1 shows the specific features of eye region. Eight features are extracted from eye region. Figure 3 shows the location for features in eye region. First four features represent geometric information of eye and eye brow. Remained four features represent shape information of eye. This shape information is acquired from comparing with template. Figure 4 shows the template for comparison. Table 2 shows the features in mouth region. Figure 5 shows the position of features in mouth region. There are two features for geometric information and six features for shape information. The template for comparison is shown in Figure 6. Table 3 shows the features in auxiliary region. If winkles exist, features have one. If winkles do not exist, features have zero. Figure 7 shows the auxiliary region and corresponding features. Since size of facial image is not static value, we need to normalize the feature vector. In this paper, all features are normalized by width of facial image. Comparing images is not easy and spends much time to compute. To overcome this difficulty, new calculated method is used to compare facial component image with template. Let X_w, X_h, and X_p are width, height, and the number of pixel in image. The similarity S can be calculated as

$$S = |X_w - T_w| + |X_h - T_h| + \left| \frac{X_w}{X_h} - \frac{T_w}{T_h} \right| + |X_p - T_p| \quad (6)$$

where T_w, T_h, and T_p are width, height, and the number of pixel in template. Table 4 shows the specific features of eye region in template. Three features for four templates are represented. Table 5 shows the feature values of mouth region in template. Features in templates are normalized by width of facial image.

4. Fuzzy Classifier Based Emotion Detection

Generally, pattern recognition problem has specific desired pattern to recognize. For example, the finger print recog-

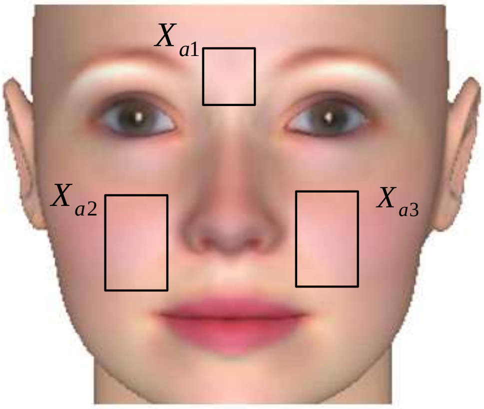

Fig. 7. Feature in auxiliary region.

Table 3. Features in auxiliary region.

Features	Description	Size
X_{a1}	Existence of winkles between eyes	1×1
X_{a2}	Existence of winkles in left cheek	1×1
X_{a3}	Existence of winkles in left cheek	1×1

nition problem, definite finger print is given to recognize. However, there are only vague patterns are given in emotion recognition problem. We can't clearly guarantee that the extracted feature vector represents only one specific emotion. Therefore, it is not easy to design the emotion recognition system. To overcome this difficulty, a fuzzy classifier is adopted in emotion detection stage. The fuzzy classifier is one of the most powerful classifier to solve classification problem with vague input. The fuzzy classifier can be described as set of fuzzy rules. The structure of fuzzy rule is represented as,

Table 4. Feature values of eye template

Template	X_w	X_h	$\frac{X_w}{X_h}$	X_p
Fig. 4(a)	0.22	0.31	0.66	13.1
Fig. 4(b)	0.23	0.24	0.46	9.6
Fig. 4(c)	0.23	0.25	0.49	10.5
Fig. 4(d)	0.23	0.25	0.48	11.4

Table 5. Feature values of mouth template

Template	X_w	X_h	$\frac{X_w}{X_h}$	X_p
Fig. 6(a)	1.79	0.37	0.40	20.12
Fig. 6(b)	1.83	0.34	0.35	20.10
Fig. 6(c)	1.55	0.49	0.61	18.89
Fig. 6(d)	1.81	0.54	0.63	20.51
Fig. 6(e)	1.76	0.34	0.39	15.98
Fig. 6(f)	2.14	0.41	0.37	28.0

$$R_i \ : \text{IF } x_1 \text{ is } M_{i1} \text{ and } \ldots \text{ and } x_m \text{ is } M_{im} \tag{7}$$

$$\text{THEN } y_i(x) \ = \ a_{i1}x_1 + \ \ldots + \ a_{im}x_m + \ b_i L$$

$$i = 1, \ldots, l \tag{8}$$

where $x_i \in \mathbb{R}$ is the ith color input, $M_{i1}, \ldots, M_{im}$ are the antecedent fuzzy sets, $y_i(x)$ is the consequent output of the ith rule, $x = [x_1, \ldots, x_m]^T \in F \subset$ is the input feature vector, F is the feature vector set, and a_{ij} and b_i is the consequent parameters.

The output of the fuzzy rule system is inferred by following equations:

$$Y(x) = \frac{\sum_{i=1}^{l} h_i(x) a_i}{\sum_{i=1}^{l} h_i(x)} \tag{9}$$

$$h_i(x) = \prod_{j=1}^{m} \mu_{M_{ij}}(x_j) \tag{10}$$

where $h_i(x)$ is the firing strength of the ith rule and $\mu_{M_{ij}}(x_j)$ is the membership degree of the jth feature of the ith rule. The membership function is defined as,

$$\mu_{M_{ij}} = \mathrm{e}^{-\frac{(c_j^i - x_j)^2}{v_j^i}} \tag{11}$$

where c_j^i is the center and v_j^i is the width of the jth feature of the ith rule. For computational convenience, the output $Y(x)$ can be represented as following matrix equation:

$$H = \begin{bmatrix} d_1(x) \\ \vdots \\ d_i(x) \\ \vdots \\ d_l(x) \end{bmatrix}, A = \begin{bmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{i1} & \ldots & a_{im} \\ \vdots & & \vdots \\ a_{l1} & \ldots & a_{lm} \end{bmatrix}, B = \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_l \end{bmatrix} \tag{12}$$

$$d_i(x) = \frac{h_i(x)}{\sum_{j=1}^{l} h_j(x)}. \tag{13}$$

Fuzzy classifier design almost means arriving at a hard classifier because most of the pattern recognition systems require hard labels for objects being classified. In order to convert the soft label $Y(x)$ to the hard label $Y_c(x)$, we use the following mapping equation:

$$Y_c(x) = \arg_g \min\{|g - Y(x)|\}, \ g \in \{1, \ldots, n\} \tag{14}$$

where n is the number of classed and g is the index of the class and n denotes the number of classes.

Finally, to design fuzzy classifier for emotion recognition, we should identify V_i and c_i in the antecedent part and A and B in the consequent part. It is not easy to identify system parameters in classifier because the designed classifier has 19 inputs and 5 rules at least. In this paper, we used LMI optimization method developed in [1], [2]. Algorithm 2 shows the LMI optimization procedure. After membership functions in antecedent part are identified, the parameters in consequent part are identified. The specific descriptions are presented in previous studies [1], [2].

Algorithm 2: Fuzzy classifier identification algorithm.

for $i \in \{1, 2, \ldots, n\}$ **do**
$W \leftarrow$ Calculate variance of the data set
Add following LMI constraints to LMI system,
$V_i > 0$
$V_i < \gamma W$
for $x \in C_i$ **do**
Add following LMI constrain to LMI system,
$$\begin{bmatrix} \gamma & \star \\ V_i x - Q_i & \gamma \end{bmatrix} > 0$$
end
$q_i, V_i \leftarrow$ Solving the LMI optimization problem
$c_i \leftarrow V_i^{-1} q_i$
end
for $i \in \{1, 2, \ldots, n\}$ **do**
for $x \in \mathbf{F}$ **do**
$h_i(x) \leftarrow \mathrm{e}^{-(x-c_i)^T V_i^T V_i (x-c_i)}$
end
end
for $x \in \mathbf{F}$ **do**
Determine the desired output $Y_d \leftarrow i$
Add following LMI constraint to LMI system,
$$\begin{bmatrix} \gamma & \star \\ Y_d - H^T(Ax+B) & I \end{bmatrix} > 0$$
end
$A, B \leftarrow$ Solving LMI optimization problem by using interior point method.

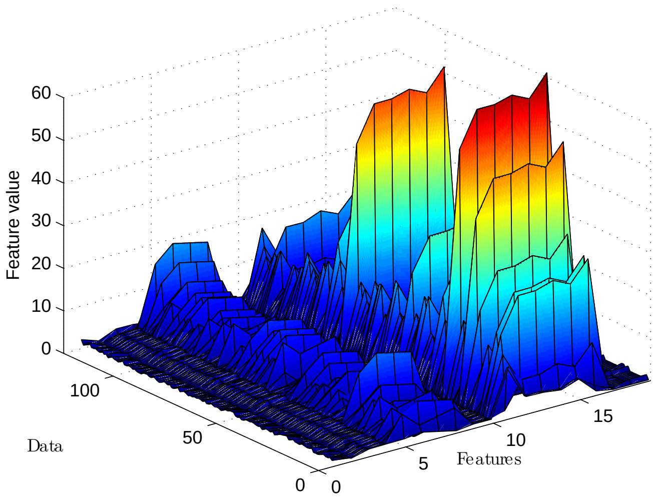

Fig. 8. Facial features extracted from 124 facial image

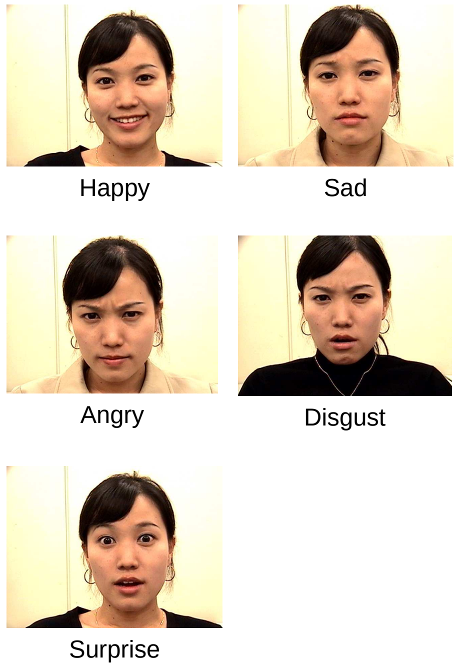

Fig. 9. Five sample image of facial image database

5. Experimental Results

In the experiment, the performance of the proposed emotion detection algorithm is evaluated. To evaluate the performance, 150 facial images are acquired from 20 men and 20 women. To evaluate the performance, we make facial image database consist of 20 men and 10 women. Under same position and slightly different illumination condition, five facial images representing five emotions are acquired from each person. Figure 9 shows the five sample images of database. The used five emotions are happy, sad, disgust, surprise, and angry. To detect human emotion, following algorithm is used. Notice that the proposed database is not easy to detect emotion because there are no limitation for hair style and wearing glasses. Moreover, we can't guarantee that the acquired facial image express accurate corresponding emotion since all people in the database is normal people. However, most of all facial images represent corresponding emotion surely. The whole emotion detection procedure can be described as following steps:

Step 1 Apply fuzzy color filter and extract face region by the method of histogram analysis
Step 2 Apply color filter and extract facial component by using VFM and histogram analysis method.
Step 3 Extract feature vector from facial component.
Step 4 Identify fuzzy classifier with the extracted feature vector.
Step 5 Test fuzzy classifier

From Step 1 and Step 2, the facial components of 124 facial images are correctly extracted. We failed to extract facial component from another 36 facial images. Therefore, the performance of the facial image analysis algorithm amounts 82.7 %. Feature vectors are then extracted from 124 facial images. Figure 8 shows the surface of extracted feature vector. In Step 4, the fuzzy classifiers with five rules are trained

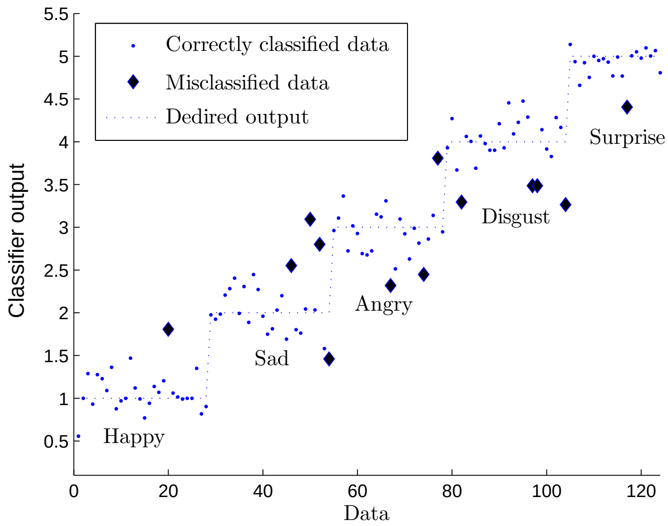

Fig. 10. Result of emotion detection algorithm

Table 6. Emotion detection accuracy

Facial component extraction sucess	82.7%
Fuzzy classification accuracy	89.5%
Final emotion detection accuracy	74.0%

by 124 data. The 95 antecedent membership function and consequent parameters are identified. The identified parameters are omitted by page limitation. Then the classification performance is evaluated. The experimental results on 124 facial images are shown in Table 6 and 7. The final emotion recognition rate is calculated by multiplying facial image analysis accuracy and fuzzy classification accuracy. Table 6 summarize accuracy of emotion recognition system. Figure 10 shows the experiment result of fuzzy classifier for 124 patterns. There are 13 facial images we can't classify. The emotion recognition accuracies on each emotion are evaluated in Table 7. In Table 7, we could know that most of recognition accuracy is similar.

6. Conclusion

We have presented an emotion detection algorithm by using frontal facial image. The algorithm composed of three stages. In image processing stage, the face region and facial components are extracted. The fuzzy color filter and histogram analysis methods are adopted to extract face region from facial image. The facial components then are extracted by using VFM and histogram analysis method. To extract feature for emotion detection, the new feature extraction methods are proposed in feature extraction stage. The emotion of facial image is detected by using fuzzy classifier with extracted features. In the experimental results, it is shown that the proposed algorithm detect emotion well.

Table 7. Recognition accuracy of five emotion

Happy	79.7%
Sad	69.9%
Angry	72.3%
Disgust	69.9%
Surprise	78.5%

References

[1] M. H. Kim, Y. H. Joo, and J, "TS fuzzy classifier using a linear matrix inequality," *Journal of Korea Fuzzy and Intelligent System*, vol. 14, no.1, pp. 46-51, 2004.

[2] Y. H. Joo, K. H. Jeong, M. H. Kim, J. B. Park, J. LEE and Y. J. CHO, "Facial image analysis algorithm for emotion recognition," *Journal of Korea Fuzzy and Intelligent System*, vol. 14, no. 7, pp. 801-806, 2004.

[3] Y. H. Joo and J. H. Oh, "Emotion Recognition Using Template Vector and Neural-Network," *Journal of Korea Fuzzy and Intelligent System*, Vol. 13, No.6, pp. 710-715, 2003.

[4] J.J. Lien, T. Kanade, J.F. Cohn, and C.C. Li, "Automated facial expression recognition based on FACS action units," *Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp.390-395, 1998.

[5] S. Morishima and H. Harashima, "Emotion space for analysis and synthesis of facial expression," *IEEE International Workshop on Robot and Human Communication*, pp. 188-193, 1993.

[6] J. Zhao and G. Kearney, "Classifying facial emotions by back propagation neural networks with fuzzy inputs," *International Conference on Neural Information Processing*, vol. 1, pp. 454-457, 1996.

[7] S. Morishima, F. Kawakami, H. Yamada, and H. Harashima, "A Modelling of Facial Expression and Emotion for Recognition and Synthesis," *Symbiosis of Human and Artifact, Elsevier*, pp. 251-256, 1995.

[8] M.J. Black and Y. Yacoob, "Recognising Facial Expressions in Image Sequences using Local Parameterised Models of Image Motion," *International Journal on Computer Vision*, Vol. 25, No. 1, pp. 23–48, 1998.

[9] P. Ekman and W.V. Friesen, *Unmasking the Face*, Prentice Hall, New Jersey, 1975.

[10] H. Kobayashi and F. Hara, "Facial interaction between animated 3D face robot and human beings," *IEEE International Conference on System, Man and Cybernetics*, pp. 3732–3737, 1997.

[11] N.M. Thalmann, P. Kalra and M. Escher, "Face to virtual face," *Proceedings of the IEEE*, Vol. 86, No.5, pp. 870–883, 1998.

[12] I.A. Essa and A.P. Pentland, "Coding analysis interpretation and recognition of facial expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 757–763, 1997.