Facial Expression Recognition Based on Two Dimensions Without Neutral Expressions

Young-Suk Shin [1], Young Joon Ahn [2]

[1] Department of Information and Communication Engineering, Chosun University, Seosuk-dong, Dong-gu, Gwangju, 501-759, Republic of Korea
ysshin@mail.chosun.ac.kr

[2] Department of Mathematics Education, Chosun University, Seosuk-dong, Dong-gu, Gwangju, 501-759, Republic of Korea
ahn@ chosun.ac.kr

Abstract. We present a new approach for recognizing facial expressions based on two dimensions without detectable cues such as a neutral expression, which has essentially zero motion energy. To remove much of the variability due to lighting, a zero-phase whitening filter was applied. Principal component analysis(PCA) representation excluded the first one principal component as the features for facial expression recognition regardless of neutral expressions was developed. The result of facial expression recognition using a neural network model is compared with two-dimension values of internal states derived from ratings of facial expression pictures related to emotion by experimental subjects. The proposed algorithm demonstrated the ability to overcome the limitation of expression recognition based on a small number of discrete categories of emotional expressions, lighting sensitivity, and dependence on cues such as a neutral expression.

1 Introduction

Models for recognizing facial expressions have traditionally operated on a digitized facial image or a short digital video sequence of the facial expression being made, such as neutral, then happy, then neutral [1,2,3,4,5,6]. In general, recognition from video is more accurate than recognition from still image. Video captures well facial movements that deviate from a neutral expression. Therefore, many models for recognizing facial expressions are based on recognition from video, although there has also been work on recognition of facial expression using still images.

Most of the methods for recognizing facial expressions need reliably detectable cues such as a neutral expression, requiring it to be relatively uniform. All require the person's head to be easily found in the video. Therefore, continuous expression recognition such as a sequence of "happy, angry, surprise" was not handled well. And the expressions must either be manually separated, or interleaved with some reliably detectable cues such as a neutral expression.

Facial expression recognition models to date have treated emotions as discrete in the sense that they try to classify facial expressions into a small number of categories such as “happiness” or “surprise” [1,2,3,4,5,6,7,8]. The data in the experiments of the models are “pure” in the sense that a user willingly or naturally tried to express exactly one emotion. There is no guarantee that the facial expression recognized as sad corresponds to any genuine affective state of sadness. A feeling of sadness can occur in both “lonely” and “grief”. Categories may be fuzzy in the sense that an element can belong in more than one category at once. Therefore, discrete categories of emotions can be treated as regions in a continuous emotion space.

In this study, we present a new approach for recognizing facial expressions based on pleasure and arousal dimensions without detectable cues such as a neutral expression. In Section 2, we introduce the facial expressions in terms of two dimensions. In Section 3, firstly to remove much of the variability due to lighting, we apply a zero-phase whitening filter to the images. Secondly, we propose a principal component analysis(PCA) representation excluded the first one principle component as the features for facial expression recognition regardless of neutral expressions. In Section 4, we discuss the result of facial expression recognition using a NN model that is compared with pleasure and arousal dimension values of internal states derived from ratings of facial expression pictures related to emotion by experimental subjects.

2 Facial Expressions Based on Two Dimensions

Although emotional expression is highly varied, many theorists view its motivational basis as having a much simpler. There are two types in the previous studies of emotion model. They are the basic emotion model and the dimension model. So far, the studies of facial expression recognition have used six basic emotions developed by Paul Ekman and his colleagues [9]. The six basic emotions are happiness, surprise, fear, disgust, sadness, and anger. Their basic theory that links the facial expressions to these six categories. There is no guarantee that a user willingly or naturally tried to express exactly one emotion.

The dimension model explains that the emotion states are not independent one another and related to each other in a systematic way. This model was proposed by Russell [10], who argued that the dimension model can be applied to emotion recognition from facial expression [11]. The dimension model also has cultural universals and it was proved by Osgood, May & Morrison and Russell, Lewicka & Niit [12, 13].

In the Kim Younga et al. study [14], the dimension study about internal states through the semantic rating of emotion words which indicates two dimensions: pleasure(P)-displeasure(D), arousal(A)-sleep(S). The result of the dimension analysis of emotion word related internal states is shown in Figure1. The face images used for this research were a subset of the Korean facial expression database[15]. The data set contained 500 images, 3 females and 3 males, each image using 640 by 480 pixels. Examples of the original images are shown in figure 2.

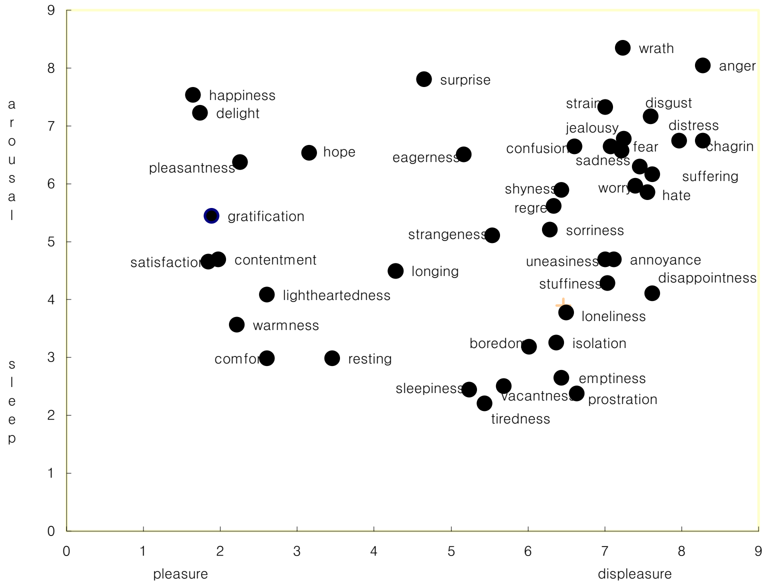


Fig. 1. The two dimensions of emotion

Expressions were divided into two dimensions according to the study of internal states through the semantic analysis of words related with emotion by Kim Younga et al. using 83 expressive words. Each expressor of females and males posed 83 internal emotional state expressions when 83 words of emotion are presented. 51 experimental subjects rated pictures on the degrees of expression in each of the two dimensions on a nine-point scale. The images were labeled with a rating averaged over all subjects.

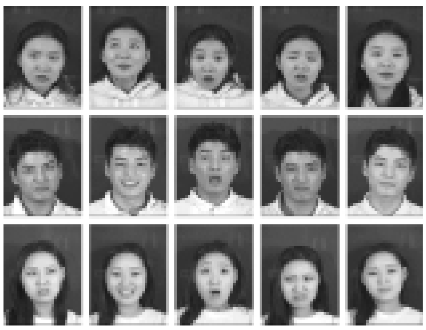


Fig. 2. Examples from the facial expression database containing 83 posed internal emotional state expressions

3 PCA Representations for Facial Expression Recognition

3.1 Preprocessing for Illumination-Invariance

The face images used for this research were centered the face images with coordinates for eye and mouth locations, and then cropped and scaled to 20x20 pixels. The luminance was normalized in two steps. First, a "sphering" step prior to principal component analysis is performed. The rows of the images were concatenated to produce 1 $\times$ 400 dimensional vectors. The row means are subtracted from the dataset, X. Then X is passed through the zero-phase whitening filter , V, which is the inverse square root of the covariance matrix:

$$V = E\{XX^T\}^{-\frac{1}{2}} \quad (1)$$
$$W = XV$$

This indicates that the mean is set to zero and the variances are equalized as unit variances. Secondly, we subtract the local mean gray-scale value from the sphered each patch. From this process, W removes much of the variability due to lightening. Figure 3(a) shows the cropped images before normalizing. Figure 3(b) shows the cropped images after normalizing.

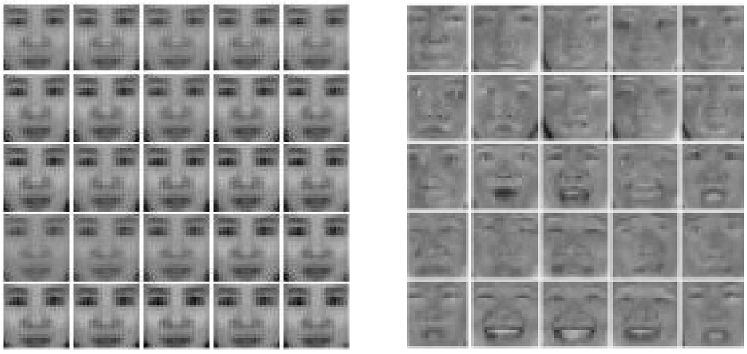
(a) (b)

Fig. 3. (a) The cropped images before normalizing. (b) The cropped images after normalizing

3.2 Principal Component Analysis Representation

Some of the most successful algorithms for face recognition applied PCA representation are "eigen faces[16]" and "holons[17]". These methods are based on learning mechanisms that are sensitive to the correlations in the face images. PCA provides a dimensionality-reduced code that separates the correlations in the input.

In a task such as facial expression recognition, the first 1 or 2 principal components of PCA do not address the high-order dependencies of the facial expression images, that is to say, it just displays the neutral face. Figure 4(a) shows PCA representation that included the first 1 principle component. But selecting intermediate ranges of components that excluded the first 1 or 2 principle components of PCA did address well the changes in facial expression (Figure 4(b)).



(a) (b)

Fig. 4. (a) PCA representation only included the first 1 principle component (b) PCA representation excluded the first 1 principle component

Therefore, to extract information of facial expression regardless of neutral expression, we employed the 200 PCA coefficients, P_n, excluded the first 1 principle component of PCA of the face images. The principal component representation of the set of images in W in Equation(1) based on P_n is defined as $Y_n = W * P_n$. The approximation of W is obtained as:

$$\overline{W} = Y_n * P_n^T . \qquad (2)$$

The columns of $\overline{W}$ contains the representational codes for the training images (Figure 4(b)). The representational code for the test images was found by $\overline{W}_{test} = Y_{test} * P_n^T$. Best performance for facial expression recognition was obtained using 200 principal components excluded the first 1 principle component.

4 Results

The system for facial expression recognition uses a three-layer neural network. The first layer contained the representational codes derived in Equation (2). The second layer was 30 hidden units and the third layer was two output nodes to recognize the two dimensions: Pleasure-Displeasure and Arousal-Sleep.

Training applies an error back propagation algorithm. The activation function of hidden units uses the sigmoid function. 500 images for training and 66 images excluded from the training set for testing are used. The 66 images for test include 11 expression images of each six people. The first test verifies with the 500 images trained already. Recognition result produced by 500 images trained previously showed

100% recognition rates. The rating result of facial expressions derived from 9 point scale on two dimension for degrees of expression by subjects was compared with experimental results of a neural network(NN). The dimension values of human and NN in each of the two dimensions are given as vectors of $\overline{H}$ and $\overline{N}$. The similarity of recognition result between human and NN was obtained as:

$$S(\vec{H},\vec{N}) = \frac{\vec{H}\cdot\vec{N}}{\|\vec{H}\|\|\vec{N}\|}\min(\frac{\|\vec{H}\|}{\|\vec{N}\|},\frac{\|\vec{N}\|}{\|\vec{H}\|}) \quad (3)$$

Table 1 describes a degree of similarity of expression recognition between human and NN on the continuous two-dimensions of emotion and indicates a part of all. The result of expression recognition of NN appears very similar to the result of expression recognition of human. In Table 1, the result of expression recognition of NN was matched to the nearest emotion word within 83 emotion words related to internal emotion states. Figure 5 and 6 show the correlation of the expression recognition between human and NN in each of the two dimensions.

The statistical significance of the similarity for expression recognition between human and NN on each of the two dimensions was tested by Person correlation analysis. The correlation in the Pleasure-Displeasure dimension between human and NN showed 0.77 at the 0.01 level and 0.51 at the 0.01 level in the Arousal-Sleep dimension.

Table 1. The result data of expression recognition between human and NN derived from two people (Abbreviation: P-D,pleasure-displeasure;A-S,arousal-sleep;)

Named emotional word of Pictures(person)	Human		Neural Network		Recognition on Neural Network	Similarity
	P – D	A – S	P –D	A – S		
depression(a)	6.23	4.43	5.22	4.41	boredom	0.89
crying(a)	6.47	4.10	6.16	5.19	sorry	0.94
gloomy(a)	7.37	5.53	7.53	6.84	strain	0.90
strange(a)	6.17	5.17	5.72	4.44	envy	0.89
proud(a)	3.07	4.47	1.69	4.54	satisfaction	0.86
confident(a)	3.47	4.57	2.90	5.35	grateful	0.93
despair(a)	6.23	5.97	5.35	5.08	strangeness	0.85
sleepiness(a)	5.00	1.80	3.13	2.96	resting	0.74
likable(a)	1.97	4.23	1.42	3.96	warmness	0.89
delight(a)	1.17	4.20	3.41	5.87	pleasantness	0.62
boredom(a)	6.77	5.50	5.05	5.65	strangeness	0.85
pleasantness (b)	1.40	5.47	3.12	4.35	contentment	0.88
depression (b)	6.00	4.23	7.10	4.28	stuffiness	0.88
crying(b)	7.13	6.17	7.46	7.07	displeasure	0.91
gloomy(b)	5.90	3.67	6.93	5.73	sadness	0.76
strangeness(b)	6.13	6.47	5.70	3.18	boredom	0.69
proud(b)	2.97	5.17	4.56	2.31	sleepiness	0.71
confident(b)	2.90	4.07	2.63	3.60	satisfaction	0.89
despair(b)	7.80	5.67	7.19	5.61	sadness	0.94
sleepiness(b)	6.00	1.93	6.34	3.07	emptiness	0.88
likable(b)	2.07	4.27	3.52	5.12	longing	0.75
delight(b)	1.70	5.70	1.79	4.92	contentment	0.87

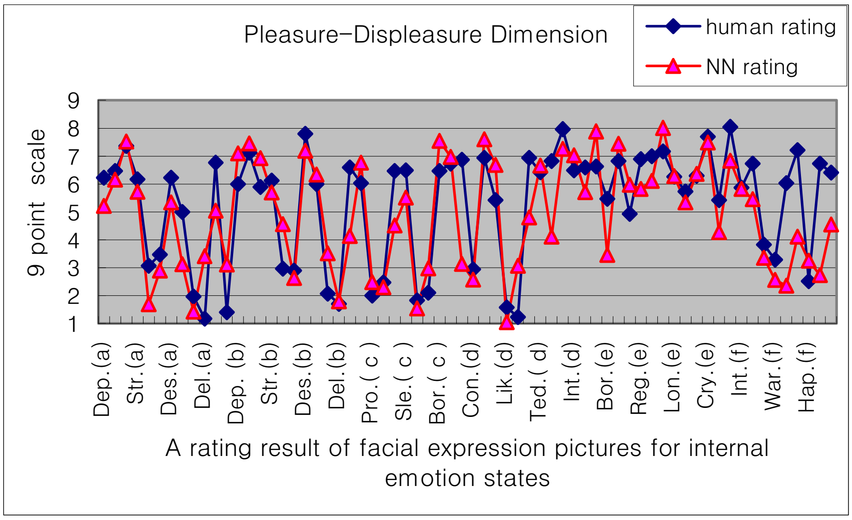


Fig. 5. A rating result of facial expression recognition in Pleasure-Displeasure dimension (Abbreviation: Dep.,depression; Str.,strangeness;Des.,despair;Del.,delight;Pro.,proud; Sle.,sleepiness;Bor.,boredom; Con.,confusion;Lik.,likable;Ted.,tedious;Int.,intricacy; Reg.,regret; Lon.,loneliness; Cry., crying; War., warmness; Hap.,happiness.)

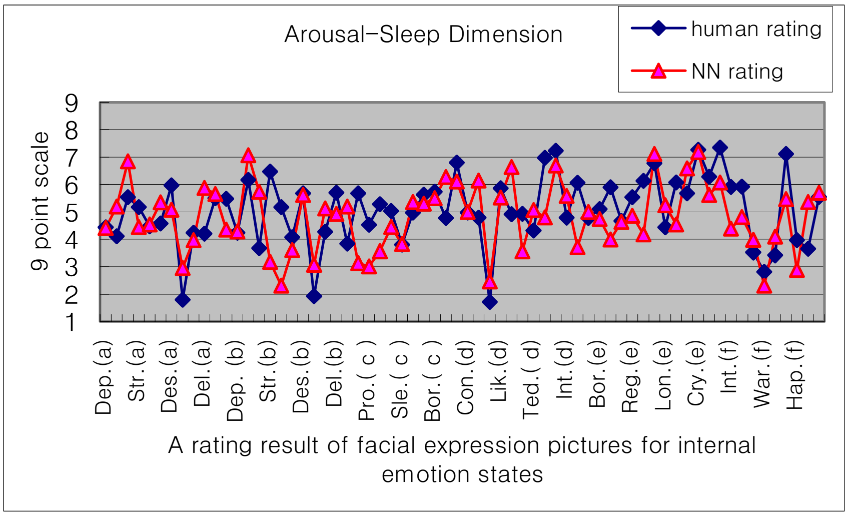


Fig. 6. A rating result of facial expression recognition in Arousal-Sleep dimension

Our results allowed us to extend the range of emotion recognition and to recognize on the continuous two dimensions of emotion with illumination-invariance without detectable cues such as a neutral expression. The result of expression recognition between human and NN on the continuous two-dimensional structure of emotion showed four significant conclusions.

(1) the two-dimensional structure of emotion in the facial expression recognition appears as a stabled structure for the facial expression recognition. The correlation results of each dimension through Person correlation analysis were significant over 0.5 at the 0.01 level.
(2) Pleasure-Displeasure dimension is analyzed as a more stable dimension than Arousal-Sleep dimension. Pleasure-Displeasure dimension was significant 0.77 at the 0.01 level, while Arousal-Sleep dimension was significant 0.51 at the 0.01 level. This result corresponds to a research for validating the stability of two-dimensional structure of emotion about emotion word[18].
(3) When the whole face was presented, facial expressions were successfully recognized. This fact was reflected by PCA representation excluded the first 1 principle component. This finding suggests that holistic analysis is important for facial expression recognition.
(4) We propose that the inference of emotional states within a subject from facial expressions may depends more on the Pleasure-Displeasure dimension than Arousal-Sleep dimension. It may be analyzed that the perception of Pleasure-Displeasure dimension may be needed for the survival of the species and the immediate and appropriate response to emotionally salient, while the Arousal-Sleep dimension may be needed for relatively detailed cognitive ability for the personal internal states.

References

1. Mase, K.: Recognition of facial expression from optical flow. IEICE Transactions, E 74, **10** (1991) 3473-3483
2. Yacoob, Y., Davis, L.S.: Recognizing human facial expression from long image sequences using optical flow. IEEE Trans. Pattern Anal. Machine Intell. **18**(6) (1996) 636-642
3. Lien, J.: Automatic recognition of facial expressions using hidden Markov models and estimation of expression intensity. Ph.D. Thesis, Carnegie Mellon University, (1998)
4. Oliver, N. Pentland, A., Berard, F.: LAFTER:a real-time face and lips tracker with facial expression recognition. Pattern Recognition **33** (2000) 1369-1382
5. Cohen, I., Sebe, N., Garg, A., Chen, L. S., Huang, T. S.: Facial expression recognition from video sequence. Proc. Int'l Conf. Multimedia and Exp(ICME) (2002) 121-124
6. Cohen, I. :Semisupervised learning of classifiers with application to human-computer interaction. PhD thesis, Univ. of Illinois at Urbana-Champaign (2003)
7. Bartlett, M., Viola, P., Sejnowski, T., Larsen, J., Hager, J., Ekman, P.: Classfying Facial Action. In: Advances in Neural Information Processing Systems 8. D. Touretzky et al. editors, MIT Press, Cambridge, MA (1996)
8. Essa, I., Pentland, A.:Coding, analysis, interpretation, and recognition of facial expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence, **19** (1997) 757-763
9. Ekman, P., Friesen, W.V.: Facial action coding system. Consulting Psychologists, Palo Alto, CA., (1977)
10. Russell, J. A.: Evidence of convergent validity on the dimension of affect. Journal of Personality and Social Psychology, **30**, (1978) 1152-1168
11. Russell, J. A. :Culture and categorization of emotion. Psychological Bulletin, **110**, (1991) 426-450

12. Osgood, C. E., May, W.H. and Miron, M.S.: Cross-curtral universals of affective meaning. Urbana:University of Illinoise Press, (1975)
13. Russell, J. A., Lewicka, M. and Nitt, T.: A cross-cultural study of a circumplex model of affect. Journal of Personality and Social Psychology, 57, (1989) 848-856
14. Younga, K., Jinkwan, K., Sukyung, P., Kyungja, O., Chansub, C.: The study of dimension of internal states through word analysis about emotion. Korean Journal of the Science of Emotion and Sensibility, **1** (1998) 145-152
15. Saebum, B., Jaehyun, H., Chansub, C.: Facial expression database for mapping facial expression onto internal state. '97 Emotion Conference of Korea, (1997) 215-219
16. Turk, M, Pentland, A. : Eigenfaces for recognition. Journal of Cognitive Neuroscience **3**(1) (1991) 71-86
17. Cottrell, G., Metcalfe, J.: Face, gender and emotion recognition using holons. In Touretzky, D., editor, Advances in Neural information processing systems (3) San Maleo, CA. Morgan aufmann (1991) 564-571
18. Jinkwan, K., Hyesshin, M., Kyungja, O.: Validating the stability of two-dimensional structure of emotion. Korean Journal of the Science of Emotion and Sensibility, **2**(1) (1999) 43-52