Facial Expression Recognition Based on Emotion Dimensions on Manifold Learning

Young-suk Shin

School of Information and telecommunication Engineering, Chosun University,
#375 Seosuk-dong, Dong-gu, Gwangju, 501-759, Korea
ysshin@chosun.ac.kr

Abstract. This paper presents a new approach method to recognize facial expressions in various internal states using manifold learning (ML). The manifold learning of facial expressions reflects the local features of facial deformations such as concavities and protrusions. We developed a representation of facial expression images based on manifold learning for feature extraction of facial expressions. First, we propose a zero-phase whitening step for illumination-invariant images. Second, facial expression representation from locally linear embedding (LLE) was developed. Finally, classification of facial expressions in emotion dimensions was generated on two dimensional structure of emotion with pleasure/displeasure dimension and arousal/sleep dimension. The proposed system maps facial expressions in various internal states into the embedding space described by LLE. We explore locally linear embedding space as a facial expression space in continuous dimension of emotion.

1 Introduction

A challenging study in automatic facial expression recognition is to detect the change of facial expressions in various internal states. Facial expressions are continuous because the expression image varies smoothly as the expression is changed. The variability of expression images can be represented as subtleties of manifolds such as concavities and protrusions in the image space. Thus automatic facial expression recognition has to be detected subtleties of manifolds in the expression image space, and it is also required continuous dimensions of emotion because the expression images consist of several other emotions and many combinations of emotions.

The dimensions of emotion can overcome the problem of discrete recognition space because the discrete emotions can be treated as regions in a continuous space. The two most common dimensions are “arousal” (calm/excited), and “valence” (negative/positive). Russell who argued that the dimensions of emotion can be applied to emotion recognition [1]. Peter Lang has assembled an international archives of imagery rated by arousal and valence with image content [2]. To recognize facial expressions in various internal states, we worked with dimensions of emotion instead of basic emotions or discrete emotion categories. The dimensions of emotion proposed are pleasure/displeasure dimension and arousal/sleep dimension.

Y. Shi et al. (Eds.): ICCS 2007, Part II, LNCS 4488, pp. 81–88, 2007.

Many studies [3, 4, 5, 6, 7] for representing facial expression images have been proposed such as Optic flow, EMG(electromyography), Geometric tracking method, Gabor representation, PCA (Principal Component Analysis) and ICA (Independent Component Analysis). At recently study, Seung and Lee [8] proposed generating image variability as low-dimensional manifolds embedded in image space. Roweis and Saul [9] showed that locally linear embedding algorithm is able to learn the global structure of nonlinear manifolds, such as the pose and expression of an individual's faces. But there have been no reports about how to contribute the intrinsic features of the manifold based on various internal states on facial expression recognition.

We explore the global structure of nonlinear manifolds on various internal states using locally linear embedding algorithm. This paper developed a representation of facial expression images on locally linear embedding for feature extraction of various internal states. This representation consists of two steps in section 3. Firstly, we present a zero-phase whitening step for illumination-invariant images. Secondly, facial expression representation from locally linear embedding was developed. A classification of facial expressions in various internal states was presented on emotion dimension having pleasure/displeasure dimension and arousal/sleep dimension using 1-nearest neighborhood. Finally, we discuss locally linear embedding space and facial expression space on dimensions of emotion.

2 Database on Dimensions of Emotion

The face expression images used for this research were a subset of the Korean facial expression database based on dimension model of emotion [10]. 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 [1]. The dimension model also has cultural universals and it was proved by Osgood, May & Morrison and Russell, Lewicka & Niit [11, 12].

The data set with dimension structure of emotion contained 498 images, 3 females and 3 males, each image using 640 by 480 pixels. Expressions were divided into two dimensions according to the study of internal states through the semantic analysis of words related with emotion by Kim et al. [13] using 83 expressive words. Two dimensions of emotion are Pleasure/Displeasure dimension and Arousal/Sleep dimension. 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. Examples of the images are shown in figure 1. Figure 2 shows a result of the dimension analysis of 44 emotion words related to internal emotion states.

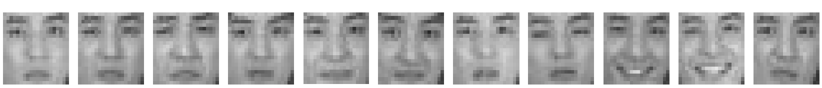


Fig. 1. Examples from the facial expression database in various internal states

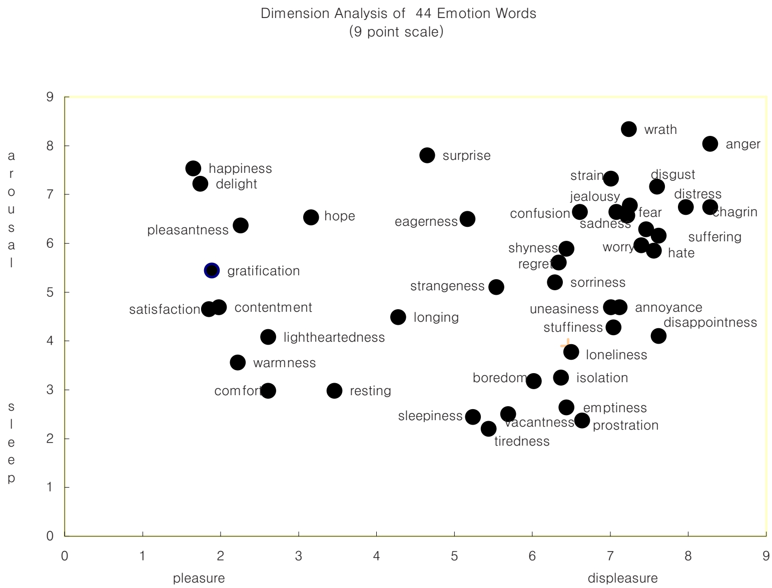


Fig. 2. The dimension analysis of 44 emotion words related to internal emotion states

3 Facial Expression Representation from Manifold Learning

This section develops a representation of facial expression images based on locally linear embedding for feature extraction. This representation consists of two steps. In the first step, we perform a zero-phase whitening step for illumination-invariant images. Second step, facial expression representation from locally linear embedding was developed.

3.1 Preprocessing

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, the rows of the images were concatenated to produce 1×400 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:

$$V = E\{XX^T\}^{-\frac{1}{2}}, \quad Z = XV \tag{1}$$

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, Z removes much of the variability due to lightening. Fig. 3(a) shows original images before preprocessing and Fig. 3(b) shows images after preprocessing.

(a) (b)

Fig. 3. (a) original images before preprocessing (b) images after preprocessing

3.2 Locally Linear Embedding Representation

Locally linear embedding algorithm[9] is to preserve local neighbor structure of data in both the embedding space and the observation space and is to map a given set of high-dimensional data points into a surrogate low-dimensional space.

Similar expressions on continuous dimension of emotion can be existed in the local neighborhood on the manifold. And the mapping from the high-dimensional data points to the low dimensional points on the manifold is very important for dimensionality reduction. LLE can overcome the problem of nonlinear dimensionality reduction, and its algorithm does not involve local minima [9]. Therefore, we applied the locally linear embedding algorithm to feature extraction of facial expressions.

LLE algorithm is used to obtain the corresponding low-dimensional data ***Y*** of the training set ***X***. D by N matrix, ***X*** consists of N data item in D dimensions. ***Y*,** d by N matrix, consists of d < D dimensional embedding data for ***X***. LLE algorithm can be described as follow.

Step 1: compute the neighbors of each data point, ***X***

Step 2: compute the weights ***W*** that best reconstruct each data point from its neighbors, minimizing the cost in eq. (2) by two constraints.

$$\varepsilon(W) = \left\| x_i - \sum_{j=1}^{K} W_{ij} x_{ij} \right\|^2 \tag{2}$$

First, each data point x_i is reconstructed only from its neighbors, enforcing $W_{ij} = 0$ if x_i and x_j are not in the same neighbor. Second, the rows of the weight matrix have sum-to-one constraint $\sum_{j=1} W_{ij} = 1$. These constraints compute the optimal weights W_{ij} according to the least square. *K* means nearest neighbors per data point.

Step 3: compute the vectors ***Y*** best reconstructed by the weights ***W***, minimizing the quadratic form in eq.(3) by its bottom nonzero eigenvectors.

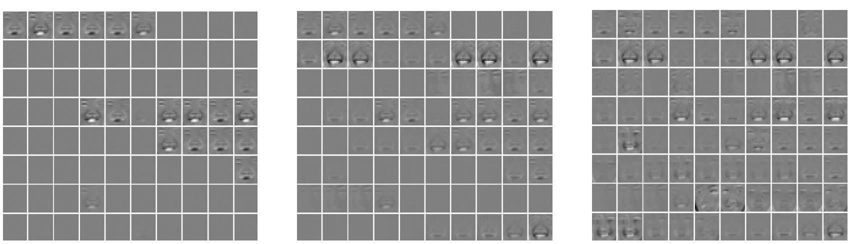
$$\Phi(Y) = \left\| y_i - \sum_{j=1}^{k} W_{ij} y_{ij} \right\|^2 \tag{3}$$

This optimization is performed subjected to constraints. Considering that the cost $\Phi(Y)$ is invariant to translation in ***Y***, $\sum_i y_i = 0$ is to remove the degree of freedom by requiring the coordinates to be centered on the origin. Also, $\frac{1}{N}\sum_i y_i y_i^T = I$ is to avoid degenerate solutions of ***Y***=0. Therefore, eq.(3) can be described to an eigenvector decomposition problem as follow.

$$\begin{aligned} \Phi(Y) = \left\| y_i - \sum_{j=1}^{k} W_{ij} y_{ij} \right\|^2 &= \arg\min_Y \|(I-W)Y\|^2 \\ &= \arg\min_Y Y^T (I-W)^T (I-W) Y \end{aligned} \tag{4}$$

The optimal solution of eq.(3) is the smallest eigenvectors of matrix $(I-W)^T(I-W)$. The eigenvalues which are zero is discarded because discarding eigenvectors with eigenvalue zero enforces the constraint term. Thus we need to compute the bottom (d+1) eigenvectors of the matrix.

Therefore we obtain the corresponding low-dimensional data set ***Y*** in embedding space from the training set ***X***. Figure 4 shows facial expression images reconstructed from bottom (d+1) eigenvectors corresponding to the d+1 smallest eigenvalues discovered by LLE, with K=3 neighbors per data point. Especially, the first eight components d=8 discovered by LLE represent well features of facial expressions. Facial expression images of various internal states mapped into the embedding space described by the first two components of LLE (See Fig. 5). From figure 5, we can explore the structural nature of facial expressions in various internal states on embedding space modeled by LLE.



(a) (b) (c)

Fig. 4. Facial expression images reconstructed from bottom (d+1) eigenvectors (a) d=1, (b) d=3, and (c) d=8

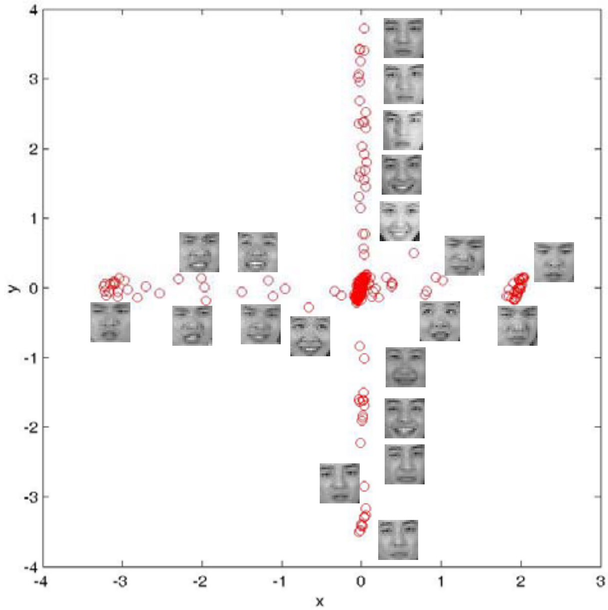


Fig. 5. 318 facial expression images of various internal states mapped into the embedding space described by the first two components of LLE

The further a point is away from the center point, the higher is the intensity of displeasure and arousal dimensions. The center points coexists facial expression images of various internal states.

4 Result and Discussion

Facial expression recognition in various internal states with features extracted by LLE algorithm was evaluated by 1-nearest neighborhood on two dimensional structure of emotion having pleasure/displeasure dimension and arousal/sleep dimension. 252 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 class label which is recognized consists of four sections on two dimensional structure of emotion. Fig. 6 shows the sections of each class label.

Table 1 gives a result of facial expression recognition recognized by proposed algorithm on two dimensions of emotion and indicates a part of all. The recognition result in the Pleasure/Displeasure dimension of test set showed 90.9% and 56.1% in the Arousal/Sleep dimension. In Table 1, the first column indicates the emotion words of 11 expression images used for testing, the second and third columns include each dimension value on bipolar dimensions of test data. The fourth column in Table 1 indicates the class label(C1,C2,C3,C4) of test data and the classification results recognized by proposed algorithm are shown in the fifth column.

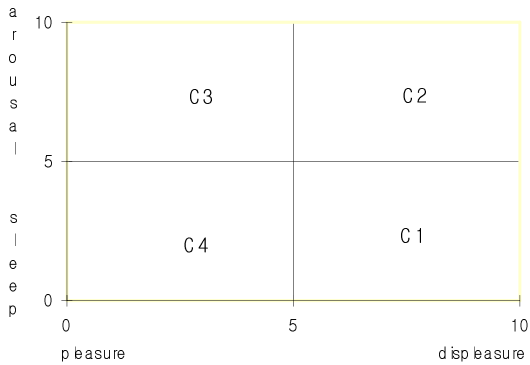


Fig. 6. The class region on two dimensional structure of emotion

Table 1. A result data of facial expression recognition recognized by proposed algorithm (Abbreviation: P-D, pleasure/displeasure; A-S, arousal/sleep;)

Emotion word (person)	Test data		Class label of test data	Recognized class label on proposed algorithm
	P – D	A – S		
pleasantness (a)	1.40	5.47	3	3
depression (a)	6.00	4.23	1	1
crying(a)	7.13	6.17	2	2
gloomy(a)	5.90	3.67	1	1
strangeness(a)	6.13	6.47	2	1
proud(a)	2.97	5.17	3	1
confident(a)	2.90	4.07	4	3
despair(a)	7.80	5.67	1	1
sleepiness(a)	6.00	1.93	4	1
likable(a)	2.07	4.27	4	3
delight(a)	1.70	5.70	3	3
gloomy(b)	6.60	3.83	1	2
strangeness(b)	6.03	5.67	2	4
proud(b)	2.00	4.53	4	3
confident(b)	2.47	5.27	4	1
despair (b)	6.47	5.03	2	2
sleepiness(b)	6.50	3.80	1	1
likable(b)	1.83	4.97	4	4
delight(b)	2.10	5.63	3	4
boredom(b)	6.47	5.73	2	3
tedious(b)	6.73	4.77	1	1
Jealousy(b)	6.87	6.80	2	2

This paper explores two problems. One is to explore a new approach method to recognize facial expressions in various internal states using locally linear embedding algorithm. The other is to explore the structural nature of facial expressions in various internal states on embedding space modeled by LLE.

As a result of the first problem, the recognition results of each dimension through 1-nearest neighborhood were significant 90.9% in Pleasure/Displeasure dimension and 56.1% in the Arousal/Sleep dimension. The two dimensional structure of emotion in the facial expression recognition appears as a stabled structure for the facial expression recognition. Pleasure-Displeasure dimension is analyzed as a more stable dimension than Arousal-Sleep dimension. In second case, facial expressions in continuous dimension of emotion was showed a cross structure on locally linear embedding space. The further a point is away from the center point, the higher is the intensity of displeasure and arousal dimensions. From these results, we can know that facial expression structure on continuous dimension of emotion is very similar to structure represented by the manifold model.

Thus our result may be analyzed that the relationship of facial expressions in various internal states can be facilitated on the manifold model. In the future work, we will consider learning invariant manifolds of facial expressions.

Acknowledgements. This work was supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2005-042-D00285).

References

1. Russell, J. A.: Evidence of convergent validity on the dimension of affect. Journal of Personality and Social Psychology, **30**, (1978) 1152-1168
2. Peter J. L.: The emotion probe: Studies of motivation and attention. American Psychologist, 50(5) (1995) 372-385
3. Donato, G., Bartlett, M., Hager, J., Ekman, P. and Sejnowski, T.: Classifying facial actions, IEEE PAMI, 21(10) (1999) 974-989
4. Schmidt, K., Cohn, J. :Dynamics of facial expression:Normative characteristics and individual difference, Intl. Conf. On Multimedia and Expo, 2001
5. Pantic, M., Rothkrantz, L.J.M.: Towards an Affect-Sensitive Multimodal Human Computer Interaction, Proc. Of IEEE. 91 1370-1390
6. Shin, Y., An, Y.: Facial expression recognition based on two dimensions without neutral expressions, LNCS(3711) (2005) 215-222
7. Bartlett, M.: Face Image analysis by unsupervised learning, Kluwer Academic Publishers (2001)
8. Seung, H. S., Lee, D.D.:The manifold ways of perception, Science (290), (2000) 2268-2269
9. Roweis, S.T., Saul, L.K..:Nonlinear Dimensionality reduction by locally linear embedding, Science (290), (2000) 2323-2326
10. Bahn, S., Han, J., Chung, C.: Facial expression database for mapping facial expression onto internal state. '97 Emotion Conference of Korea, (1997) 215-219
11. Osgood, C. E., May, W.H. and Miron, M.S.: Cross-curtral universals of affective meaning. Urbana:University of Illinoise Press, (1975)
12. 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
13. Kim, Y., Kim, J., Park, S., Oh, K., Chung, 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