FACIAL EXPRESSION ANALYSIS BY KERNEL EIGENSPACE METHOD BASED ON CLASS FEATURES (KEMC) USING NON-LINEAR BASIS FOR SEPARATION OF EXPRESSION-CLASSES

Y. Kosaka, K. Kotani

School of Information Science, Japan Advanced Institute of Science and Technology
1-1 Asahidai Tatsunokuchi Ishikawa 923-1292 Japan
{y-kosaka, ikko}@jaist.ac.jp

ABSTRACT

In the facial expression recognition by analyzing feature-vectors with linear transformation, an accuracy of recognition is depending on expression-classes. The accuracy falls remarkably when feature vectors of expression-classes are linearly non-separable in a feature space. This paper describes a new method of facial expression analysis and recognition by using non-linear transformation for separating each expression-classes. Our new method, namely KEMC, consists of the non-linear transformation defined by kernel functions for transforming higher dimensional space and EMC (Eigenspace Method based on Class features). This paper also shows experimental results of facial expression classification by KEMC.

1. INTRODUCTION

Nonverbal information plays an inportant part in a human communication; especially facial expression is indispensable information in order to carry about 70% of our feeling and emotion. For example, when a recognition of facial expression becomes possible from face image, we consider that man-machine communication can be facilitated.

Although Principal Component Analysis (PCA) is a fundamental method of pattern analysis, PCA cannot extract the feature of expression with sufficient accuracy. Kurozumi proposed Eigen-space Method based on Class-feature (EMC), EMC chooses bases to maximizes differences of the withn-class scatters and between-class scatters, in order to acquire bases which expresses the feature of expression for facial patterns. EMC has separated the feature of expression with sufficient accuracy compared with PCA [1]. However recognition rates by EMC for the expression of "Anger", "Disgust", "Sadness" and "Fear" were lower than "Happiness" and "Surprise". The accuracy falls remarkably when feature vectors of exprssion-classes are linearly non-separable in a feature space. This paper describes a new method of facial expression analysis and recognition by using non-linear transformation for separating each expression-classes. For the separation of a linearly non-separable expression-class, each class is solved on two or more linearly separable classes, the nonlinear mapping of facial patterns are carried out for linearly separable, etc. can be considered. Our new method, namely Kernel EMC (KEMC), consists of the non-linear transformation defined by kernel functions for transforming higher dimensional space and EMC.

2. ANALYSIS METHOD

2.1. Kernel EMC(KEMC)

Let F be a set of expression-class. Assume that M_f facial-patterns are given for each $f \in F$, and let $\mathbf{x}_{fm}$ be an N-dimensional vector of the m-th facial-pattern where $m = 1, 2, \ldots, M_f$. The idea of KEMC is to yield a nonlinear discriminant in the input space through the kernel trick and EMC. The nonlinear map Φ of the $\mathbf{x}_{fm}$ is carried out at $\Phi(\mathbf{x}_{fm})$. We approximate $\mathbf{x}_{fm}$ by using $\Phi(\mathbf{x}_{fm})$ and an orthogonal D ($D \leq M-1$) vectors $\mathbf{v}_d = (v_{1d}, v_{2d}, \ldots, v_{Nd})^T$, $d = 1, 2, \ldots, D$ the approximated pattern $\tilde{\mathbf{x}}_{fm}$ is obtained as

$$\tilde{\mathbf{x}}_{fm} = \sum_{d=1}^{D} z_{dfm}\mathbf{v}_d \quad (1)$$

KEMC considers $\tilde{\mathbf{x}}_{fm}$ to be an ideal approximation pattern for all patterns in each $f \in F$, and finds a basis $\mathbf{v}_d$ to minimize the mean-square error $\varepsilon^2(\mathbf{v}_d)$ so that

$$\varepsilon^2(\mathbf{v}_d) = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f} ||\mathbf{u}_f - \tilde{\mathbf{x}}_{fm}||^2 \quad (2)$$

$$\mathbf{u}_f = \frac{1}{M_f}\sum_{m=1}^{M_f} \Phi(\mathbf{x}_{fm}) \quad (3)$$

The D-dimensional vector $\mathbf{z}_{fm}$ obtained from $\mathbf{v}_d, d = 1, 2, \ldots, D$ turns into the feature vector used for analysis. The dth ele-

0-7803-8554-3/04/$20.00 ©2004 IEEE.

ment of $\mathbf{z}_{fm}$ can be expressed as follows.

$$z_{dfm} = (\mathbf{v}_d \cdot \Phi(\mathbf{x}_{fm})) \tag{4}$$

Equation(2) is substituted for equation(1).

$$\varepsilon^2(\mathbf{v}_d) = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}\left|\left|\mathbf{u}_f - \sum_{d=1}^{D} z_{dfm}\mathbf{v}_d\right|\right|^2 \tag{5}$$

$$\bar{z}_{df} = (\mathbf{v}_d \cdot \mathbf{u}_f) \tag{6}$$

Equation(7) substitutes equation(3) and equation(4) for equation(5).

$$\varepsilon^2(\mathbf{v}_d) = \frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}||\bar{x}_f||^2 - \sum_{d=1}^{D}\mathbf{v}_d^T\mathbf{S}'\mathbf{v}_d \tag{7}$$

where

$$\begin{aligned}\mathbf{S}' &= \frac{1}{M}\sum_{f\in F}M_f\mathbf{u}_f\mathbf{u}_f^T \\ &-\frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}(\Phi(\mathbf{x}_{fm})-\mathbf{u}_f)(\Phi(\mathbf{x}_{fm})-\mathbf{u}_f)^T\end{aligned} \tag{8}$$

The value of $\sum_{d=1}^{D}\mathbf{v}_d^T\mathbf{S}'\mathbf{v}_d$ is maximized in order to minimize the value of $\varepsilon^2(\mathbf{v}_d)$. At that time , $\mathbf{v}_d$ is the eigenvector $\mathbf{v}_d$ corresponding in descending order of eigenvalue of $\mathbf{S}'$. However it is difficult to calculate, since $\Phi(\mathbf{x})$ which is a nonlinear map function exists in $\mathbf{S}'$. Then, it changes so that it may be easy to calculate $\mathbf{v}_d^T\mathbf{S}'\mathbf{v}_d$. $\mathbf{v}_d$ is expressed with linear combination of $\Phi(\mathbf{x}_{fm})$ in the pattern space mapped by Φ.

$$\mathbf{v}_d = \sum_{f\in F}\sum_{m=1}^{M_f}a_{fm}\Phi(\mathbf{x}_{fm}) \tag{9}$$

$\mathbf{u}_f$ and $\mathbf{v}_d^T$ change into the form where a kernel is used, by equation (9).

$$\mathbf{v}_d^T\mathbf{u}_f = \mathbf{a}^T\begin{pmatrix}\frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{1\ 1},\mathbf{x}_m)\\ \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{1\ 2},\mathbf{x}_m)\\ \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{1\ 3},\mathbf{x}_m)\\ \cdots\\ \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{2\ 1},\mathbf{x}_m)\\ \cdots\\ \frac{1}{M_f}\sum_{m=1}^{M_f}k(\mathbf{x}_{f\ m},\mathbf{x}_m)\end{pmatrix} = \mathbf{a}^T\mathbf{m}_f \tag{10}$$

This reads

$$\mathbf{v}_d^T\mathbf{S}'\mathbf{v}_d = \mathbf{a}_d^T\mathbf{R}\mathbf{a}_d. \tag{11}$$

$$\begin{aligned}\mathbf{R} &= \frac{1}{M}\sum_{f\in F}M_f\mathbf{m}_f\mathbf{m}_f^T \\ &-\frac{1}{M}\sum_{f\in F}\sum_{m=1}^{M_f}(\zeta_{fm}-\mathbf{m}_f)(\zeta_{fm}-\mathbf{m}_f)^T\end{aligned} \tag{12}$$

$$\zeta_{fm} = (k(\mathbf{x}_{1\ 1},\mathbf{x}_m), k(\mathbf{x}_{1\ 2},\mathbf{x}_m),\ldots,k(\mathbf{x}_{f\ m},\mathbf{x}_m))^T \tag{13}$$

$\mathbf{a}_k$is the eigenvector of $\mathbf{R}$.

$$\mathbf{R}\mathbf{a}_d = \lambda_d\mathbf{a}_d \tag{14}$$

z_{dfm} can be drawn with $\mathbf{a}_d$, equation(4) and (9) to

$$z_{dfm} = (\mathbf{v}_d \cdot \Phi(\mathbf{x}_{fm})) = \sum_{f\in F}\sum_{m=1}^{M_f}a_{dfm}k(\mathbf{x}_{fm},\mathbf{x}). \tag{15}$$

2.2. Expression recognition using the eigenspace method

Let $\mathbf{x}'$ be an input pattern which is an N-dimensional vector and each element of $\mathbf{x}'$ corresponds to the value of each pixel. By the projection of equation (14) , we obtain

$$\begin{aligned}\mathbf{z}' = &\left[\sum_{f\in F}\sum_{m=1}^{M_f}a_{1fm}k(\mathbf{x}_{fm},\mathbf{x}'), \sum_{f\in F}\sum_{m=1}^{M_f}a_{2fm}k(\mathbf{x}_{fm},\mathbf{x}')\right. \\ &\left.,\ldots,\sum_{f\in F}\sum_{m=1}^{M_f}a_{Dfm}k(\mathbf{x}_{fm},\mathbf{x}')\right]^T\end{aligned} \tag{16}$$

The dictionary vector $\bar{\mathbf{z}}_f$ of the class $f\in F$ in the expression feature space is defined as

$$\bar{\mathbf{z}}_f = \sum_{m=1}^{M_f}\mathbf{z}_{fm}. \tag{17}$$

The input vector $\mathbf{z}'$ is compared with the dictionary vector $\bar{\mathbf{z}}_f$, and the expression class of an input pattern is determined. R_f is used for comparison. The input pattern $\mathbf{x}'$ belongs to the expression class f in case R_f is the minimum.

$$R_f = 1 - \frac{\mathbf{z}'\cdot\bar{\mathbf{z}}_f}{|\mathbf{z}'||\bar{\mathbf{z}}_f|} \tag{18}$$

3. EXPRESSION RECOGNITION EXPERIMENT

This chapter shows experimental results of expression analysis by KEMC. The recognition rates for every expression class are measured, and the characteristics of recognition for the different kernel function are also shown. Polynomial kernels (poly) for the kernel function is as shown,

$$k(\mathbf{x}_1,\mathbf{x}_2) = (\mathbf{x}_1^T\mathbf{x}_2)^p \tag{19}$$

and gauss kernels (rbf)

$$k(\mathbf{x}_1,\mathbf{x}_2) = \exp\left(\frac{-||\mathbf{x}_1-\mathbf{x}_2||^2}{2\sigma^2}\right). \tag{20}$$

3.1. The experimental subjects

- Recognition for open-data and closed-data by EMC , KEMC and KPCA[2].
- Recognition for closed-data with the parameter of each kernel function experimentally.
- Recognition for open-data of KEMC and KPCA with the parameter of the kernel function learned by using closed-data.
- closed-data : Input patterns is contained in dictionary patterns.
- open-data : Input patterns is not contained in dictionary patterns.

3.2. Facial image data

For an experiment, we prepared facial images of 60 person with 6 facial expressions. These images were obtained by MINOLTA VIVID 700 simultaneously under the same illuminating condition. All images are nomalized by using positions of eyes and a nose as the preprocessing such as the size and position changes. By this processes, each image normalized to an image of 100×100 pixels. Figure 1 shows examples of facial image for the experiment. (Anger;An. Disgust;Di. Fear;Fe. Happiness;Ha. Sadness;Sa. Surprise;Su.)

Fig. 1. Examples of facial image $\mathbf{Upper}$: An. Di. Fe. $\mathbf{Upper}$: Ha. Sa. Su.

Each facial image is selected by subjective assessment test that is obeerved by viewers. The number of feature vector in the dictionary and input patterns are decided from selected images. In each facial expression, 60% of all facial images are used for dictionaries, and 40% is used for input patterns. Table 1 shows numbers of image for dictionary and input pattern for the experiment.

Table 1. Number of image in a dictionary and input patterns

	An.	Di.	Fe.	Ha.	Sa.	Su.	Total
Dictionary	22	22	15	30	23	33	146
Input	14	15	11	19	16	22	96
Total	36	37	26	49	39	55	242

3.3. Experimental result of the facial expression recognition

Table 2 and 3 show the recognition rate for the closed-data and the open-data. Comparison of each rates are as below,
(1)Recognition rate of KPCA and KEMC
The average recognition rate of KEMC is about 20 points higher than KPCA. For example, the highest rate is 77.2% by KEMC (poly), and second one is 65.2% by KEMC (rbf). Since the recognition rates of EMC, KPCA (poly) and KPCA (rbf) is 61.9%, 44.9% and 42.3% respectively, the nonlinear mapping of facial pattern by KEMC is more effective for expression recognition than PCA.
(2)Recognition rate of EMC and KEMC
For the open-data,

- The average of recognition rate by KEMC (poly) is 15 point higher than EMC.
- In the case of EMC, the recognition rate on “Anger”, “Disgust”, “Sadness” and “Fear” is lower than “Happiness” and “Surprise”.
- The recognition rate of KEMC (poly) of “Anger”, “Disgust” , “Sadness” and “Fear” is higher than EMC, especially “Anger” and “Sadness” is 18 to 25 point higher.

Therefore, KEMC (poly) is able to separate each classes with high accuracy ,especially for“Anger” and “Sadness”, compared with EMC.
(3)The feature vector of EMC and KEMC
Figure 2 and 3 show the feature vectors of “Anger”, “Disgust”, “Sadness” and “Fear” for open-data that is mapped to eigen space at higher rank 2 base of EMC and KEMC. An x-axis is the value of the 1st base and a y-axis is the value of the 2nd base. Fig. 2 and 3 show the following things.

- In EMC, since “Anger” and “Disgust”, “Sadness” and “Fear”became the distribution near as for the feature vector, the rate of discernment of each expression falls.

- In KEMC, although “Disgust” and “Fear” of the feature vector are near distributions, the separation accuracy of “Anger” and “Sadness” is high. Therefore, the high recognition rate was obtained.

KEMC is separating an expression-classes in the space which mapped face patterns using the nonlinear mapping function, and it was able to create the eigen space which was suitable for separation of each expression-classes compared with EMC.

Table 2. The recognition rate on the closed-data[%]

	An.	Di.	Fe.	Ha.	Sa.	Su.	Av.
EMC	82	100	80	100	100	100	94
KEMC (poly, p=13)	100	91	87	100	100	100	96
KEMC (rbf, $\sigma = 3.0$)	100	100	100	100	100	100	100
KPCA (poly, p=10)	59	14	53	100	57	100	65
KPCA (rbf, $\sigma = 5.0$)	100	100	100	100	100	100	100

Table 3. The recognition rate on the open-data[%]

	An.	Di.	Fe.	Ha.	Sa.	Su.	Av.
EMC	57	47	46	84	56	82	62
KEMC (poly, p=13)	79	60	64	84	81	96	77
KEMC (rbf, $\sigma = 3.0$)	64	53	46	81	56	91	65
KPCA (poly, p=10)	36	7	36	63	50	77	45
KPCA (rbf, $\sigma = 5.0$)	0	7	46	68	38	96	42

4. CONCLUSION

This paper described a new method of facial expression analysis “KEMC” based on a nonlinear mapping of feature vectors by kernel method and eigen-space method fitting on class-feature “EMC”. Since KEMC was able to separate the expression classes that were linearly non-separable, KEMC could achieve high accuracy for recognition of wide expression classes. As a result of experimental test, KEMC obtained about 18 to 25 points higher recognition rate on “Sadness” and “Anger” than EMC and PCA.

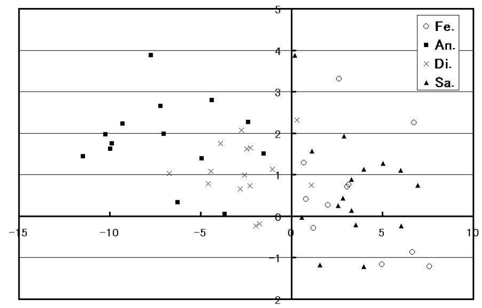

Fig. 2. The distribution of higher rank 2 basees to open-data; EMC

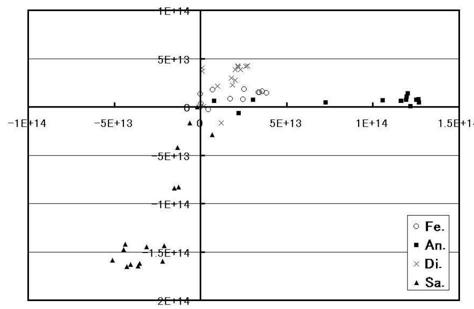

Fig. 3. The distribution of higher rank 2 basees to open-data; KEMC

5. REFERENCES

[1] T. Kurozumi, Y. Shinza, Y. Kenmochi, K. Kotani, “Facial individuality and expression analysis by eigenspace method based on class features or multiple discriminant analysis”, Proc. of 1999 IEEE ICIP, Vol.1, pp.648-652, 1999.

[2] B. Schölkopf, A. Smola and K. R. Müller, “Nonlinera Component Analysis as a Kernel Eigenvalue Problem”, Neural Computation, vol. 10, pp. 1299-1319, 1998.

[3] M. -H. Yang, “Kernel eigenfaces vs. kernel fisherfaces: face recognition using kernel methods”, Proc. of Int’l Conf. on Automatic Face and Gesture Recognition, pp. 215-220(2002)

[4] Q. Liu, R. Huang, H. Lu, and S. Ma, “Face recognition using kernel based Fisher discriminant analysis”, Proc. of Int’l Conf. on Automatic Face and Gesture Recognition, pp. 197-201(2001)