JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 22, 1033-1046 (2006)

Facial Expression Classification Using PCA and Hierarchical Radial Basis Function Network*

DAW-TUNG LIN
Department of Computer Science and Information Engineering
National Taipei University
Sanshia, 237 Taiwan

Intelligent human-computer interaction (HCI) integrates versatile tools such as perceptual recognition, machine learning, affective computing, and emotion cognition to enhance the ways humans interact with computers. Facial expression analysis is one of the essential medium of behavior interpretation and emotion modeling. In this paper, we modify and develop a reconstruction method utilizing Principal Component Analysis (PCA) to perform facial expression recognition. A framework of hierarchical radial basis function network (HRBFN) is further proposed to classify facial expressions based on local features extraction by PCA technique from lips and eyes images. It decomposes the acquired data into a small set of characteristic features. The objective of this research is to develop a more efficient approach to discriminate between seven prototypic facial expressions, such as neutral, smile, anger, surprise, fear, disgust, and sadness. A constructive procedure is detailed and the system performance is evaluated on a public database "Japanese Females Facial Expression (JAFFE)." We conclude that local images of lips and eyes can be treated as cues for facial expression. As anticipated, the experimental results demonstrate the potential capabilities of the proposed approach.

Keywords: intelligent human-computer interaction, facial expression classification, hierarchical radial basis function network, principal component analysis, local features

1. INTRODUCTION

The intelligent human-computer interaction (HCI) technologies play important roles in the development of advanced and ambient communication/computation. In contrast to the conventional mechanisms of passive manipulation, intelligent HCI integrates versatile tools such as perceptual recognition, machine learning, affective computing, and emotion cognition to enhance the ways humans interact with computers. Migrating from W4 (*what*, *where*, *when*, *who*) to W5+ (*what*, *where*, *when*, *who*, *why*, *how*), novel intelligent interface design has placed emphasis on both apparent and internal behavior of users [1]. Nonverbal information such as facial expression, posture, gesture, and eye gaze is suitable for behavior interpretation. Facial data analysis is one of the essential medium of perceptual processing and emotion modeling.

Received August 16, 2005; accepted January 17, 2006.
Communicated by Jhing-Fa Wang, Pau-Choo Chung and Mark Billinghurst.
* This work was supported in part by the National Science Council of Taiwan, R.O.C., under grants No. NSC 88-2213-E216-010 and No. NSC 89-2213-E216-016.
* The preliminary content of this paper has been presented in "International Conference on Neural Information Processing," Perth, Australia, November 1999. Acknowledgement also due to Mr. Der-Chen Pan at the National Taipei University for his help in performing simulations. The author would like to thank Mr. Ming-Shon Chen at Ulead System Inc., Taipei, Taiwan, for his early work and assistance in this research.

Facial expression recognition methods can be generally divided into two categories: static images vs. video sequences, based on differ use of data and feature extraction methods. Typical techniques include optical flow estimation, spatial feature analysis, and local filter analysis [2-13]. Yacoob and Davis [12] utilized optical flow method to track the dynamic movement of facial features from video sequences and classified the representation of facial feature movement into six expressions (i.e., smile, surprise, anger, fear, sadness, and disgust). The recognition rate ranges from 80% to 94%. Barlett *et al.* [2] combined optical flow and principal component analysis (PCA) for facial expression recognition. Rosenblum *et al.* [8, 9] developed a three-level radial basis function network to model the full temporal profile of facial expressions based on 2D discrete cosine transform of the entire face image sequences. Their best recognition rate is 88%. Otsuka and Ohya [7] presented a facial recognition model by combining Hidden Markov Model and optical flow, and then estimated the condition of facial muscles from image sequences. The Facial Action Coding System (FACS) derived by Ekman and Friesen has been widely used to describe the facial expression by movement of action units (AUs) [14]. The FACS is often incorporated with the above mentioned techniques to delineate the details of human expression for video sequences. Mase [6] proposed the FACS+ system for coding, analysis, and recognition of facial expressions by using an optimal estimation optical flow scheme. His approach was focused on computing the motion of facial muscles correspond to Action Units (AUs). Tian *et al.* [10] developed an automatic face analysis system based on both permanent facial features and transient features with recognition rates of 96.4% for upper face AUs and 96.7% for lower face AUs. Donato *et al.* [15] provided a more detail review of the recent techniques for facial expression recognition based on video sequences and FACS encoding.

Facial expression recognition from still images is a more difficult problem than from video sequences due to the fact that less information during expression actions is available [16]. Cottrell and Metcalfe [17] applied PCA and backpropagation neural network to recognize facial expression, gender, and identity from static images. Similar approach was proposed by Padgett and Cottrell [18] combining PCA and feed-forward network to recognize six emotions. The hit ratio is improved to 84% by only feeding the features of eyes and mouth. Matsuno *et al.* [19] proposed an approach which recognized facial expressions from static images based on pre-computed facial expressions parameters. They computed the amount of image gradient magnitude warping vector to learned vectors of four facial expressions (happiness, sadness, anger, and surprise). Chen and Huang [16] modified linear discriminate analysis (LDA) algorithm and presented a new clustering based feature extraction method for facial expression recognition. The performances are 86.7%, 98.2%, and 89.1% for neutral vs. non-neutral, smile vs. non-smile, and anger vs. non anger, respectively. A constructive feed-forward neural network was further proposed for facial expression recognition with pruning technique by Ma and Khorasani [20]. Four facial expressions were studied: surprise, anger, happiness, and disgust. The best recognition rates are 100% and 93.75% (without rejection). Other approaches such as multiple discriminate analysis, 2D Gabor wavelet representation, linear discriminate analysis, feature point position comparison, and potential field analysis can also been found in the literatures [5, 21-23].

In this paper we are concerned with automatic classification of facial expressions from still images. The psychologists have indicated that as least six emotions are universally associated with distinct facial expressions, including smile, sadness, surprise, fear, anger, and disgust. Examples of Japanese Females Facial Expression (JAFFE) databases are shown in Fig. 1 [21]. We will also focus on these expressions. The organization of the remainder of this paper is as follows. In section 3, the main components of the proposed system are illustrated. We utilized PCA in the pre-processing stage to extract features from face imagery. We further proposed a hierarchical radial basis function network (HRBFN) to fulfill the facial expression differentiation task. In section 4, the experimental results are presented to demonstrate the potential capabilities of the proposed approach. Finally, conclusions are stated in section 5.

Fig. 1. Examples of seven principal facial expressions in JAFFE [21]: smile, disgust, anger, surprise, fear, neutral, and sadness (from left to right).

2. FACIAL EXPRESSIONS RECONSTRUCTION WITH PCA

Most approaches in computer recognition of faces and expressions have been focused on detecting individual features such as eyes, head outline, mouth, or defining a face model by position, size, and relationships among these features [24]. Features extraction plays an essential role in the pre-processing stage. Principal Component Analysis (PCA) has been commonly used to faces recognition problems [5, 22, 25, 26]. Typical PCA algorithm (Eigenface/Fisherface) is one of the main streams of research on face feature processing [26, 27]. PCA has advantage over other face recognition schemes in its speed and simplicity. We utilized PCA in the pre-processing stage to extract features from face imagery.

The basis of the ordinary image space is composed of all single pixel vectors. However, the image space is not a optimal space for face representation and categorization. The aim of applying PCA is to build a face space which better describes the face images. The basis vectors of this face space are called the principal components. These components will be uncorrelated and will maximize the variance accounted in the original basis. It can also reduce the dimension of the feature space. The computation complexity is thus reduced [28].

By concatenating each row of the image by row, a face image can be transformed to a column vector. Assume the width and height of the image are n pixels and m pixels respectively, the size of the transformed vector of this image will be $n \times m$ by 1. Given M facial expression images as training data, we convert these images to corresponding column image vectors Γ_i, where $i = 1, 2, \ldots, M$. Compute the mean of training data $\Psi =$

$\frac{1}{M}\sum_{n=1}^{M}\Gamma_n$ and let the normalized vectors be $\Phi_i = \Gamma_i - \Psi$. We want to seek a set of M orthonormal vectors, u_n, that best represents the distribution of the data. The kth vector, u_k, is chosen such that

$$\lambda_k = \frac{1}{M}\sum_{n=1}^{M}(u_k^T\Phi_n)^2 \tag{1}$$

is maximum, where $u_l^T u_k = \begin{cases} 1 & \text{if } l = k \\ 0 & \text{otherwise} \end{cases}$. Extend Eq. (1) and multiply u_k^T to both sides of the equation, we obtain $\lambda_k u_k^T = u_k^T\left(\frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^T\right)u_k u_k^T$. Assume the covariance matrix $C = AA^T$, where $A = [\Phi_1, \Phi_2, \ldots, \Phi_M]$, we obtain

$$\lambda_k u_k^T = u_k^T C. \tag{2}$$

Transpose both side of Eq. (2), we derive $\lambda_k(u_k^T)^T = C^T(u_k^T)^T = C^T u_k$ and $C^T = AA^T = C$. Then we can conclude that $\lambda_k u_k = Cu_k$. Thus the vectors u_k and scalars λ_k are the eigenvectors and eigenvalues, respectively, of the covariance matrix $C = AA^T$.

Because the size of covariance matrix C is $nm \times nm$, it is time-consuming to determine nm eigenvalues and eigenvectors. It is necessary to reduce the complexity of the computation. Define a matrix $C' = A^TA$, then the size of C' is M by M and let v_i denote the eigenvectors of the matrix C', we obtain $A^TAv_i = \mu_i v_i$. Multiply A to both sides, we have

$$AA^TAv_i = \mu_i Av_i. \tag{3}$$

We can observe from Eq. (3) that Av_i is the eigenvector of $C = AA^T$ [29]. So we can now reduce the method by derive p ($p < nm$) eigenvectors (v_i) of the matrix C' and derive the p eigenvectors of covariance matrix C by multiplying A to v_i. The face images can be represented in the way by projecting the data in the image space onto the face space (also called Eigenfaces). The dimension of the projected data in the feature space is much smaller than that in the original image space. After the Eigenfaces are obtained from the training set, we may transform face image into feature space by a simple projection operation $x_i = u_i^T(\Gamma - \Psi)$, $i = 1, 2, \ldots, p$, where x_i denotes the projection of the face image Γ projected onto the i-th Eigenface component u_i, where Ψ is the average face image of the training set. The projections constitute a vector $X^T = [x_1, x_2, \ldots, x_p]$ called projection vector. We will treat the projection vector X as the features for further training process.

To illustrate the feasibility of using eigen feature to fulfill expression classification task, we modify the PCA reconstruction method for preliminary evaluation [30]. Notice that if the input image is much similar to some expression training set, the reconstructed image will has less distortion than the image reconstructed from other eigen vectors of training expressions. Based on this episode, we divide the training set into seven classes according to different expression and compute the eigen space of each class. For a test face image, we first project it onto the eigen space of each class independently and then derive reconstructed image from each eigen space. By measuring the similarity (mean-square error) between input image and the reconstructed image of each class, we can identify the class of input image whose reconstructed image is most similar to the input one. The procedure of the developed PCA reconstruction method is delineate in Fig. 2.

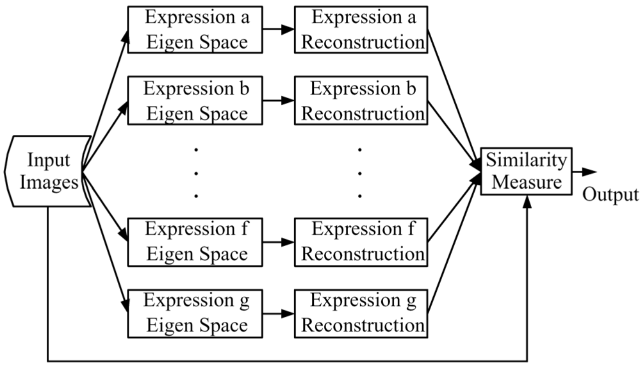


Fig. 2. Classification procedures of PCA reconstruction.

3. THE COMPONENTS OF HRBFN ARCHITECTURE

3.1 Basic RBFN Description

Neural network technology offers a number of tools such as learning and adaptation, generalization and robustness, feature extraction and distributed representation. The neural net approach has been shown fruitful in solving face recognition problems [31-36]. The radial basis function neural network (RBFN) theoretically provides a sufficient large network structure such that any continuous function can be approximated to within an arbitrary degree of accuracy by appropriately choosing radial basis function centers [28]. The RBFN is trained using sample data to approximate a function in multidimensional space. A basic topology of RBFN is depicted in Fig. 3. The RBFN is a three-layered network. The first layer constitutes input layer in which the number of nodes is equal to the dimension of input vector. In the hidden layer, the input vector is transformed by radial basis function as activation function: $\varphi(\vec{x};\ \vec{c}_j) = \exp\left(-\frac{1}{2\sigma^2}\|\vec{x}-\vec{c}_j\|^2\right)$, where $\|\cdot\|$ denotes a norm (usually Euclidean distance) of the input data sample vector $\vec{x}$ and the center $\vec{c}_j$ of radial basis function. The kth output is computed by equation

$$F_k(x) = \sum_{j=1}^{m} w_{kj} \cdot \varphi(\vec{x};\ \vec{c}_j), \tag{4}$$

where w_{kj} represents a weight synapse associates with the jth hidden unit and the kth output unit with m hidden units. Given a set of N different points $\{\vec{x}_i \in R^p \mid i = 1, 2, \ldots, N\}$ as input pattern and the corresponding set of desired target values $\{\vec{d}_i \in R^k \mid i = 1, 2, \ldots, N\}$, the goal is to find a function $F: R^p \to R^k$ that satisfies the condition: $F(\vec{x}_i) = \vec{d}_i$, $i = 1, 2, \ldots, N$. Thus, we obtain the following result derived from Eq. (4):

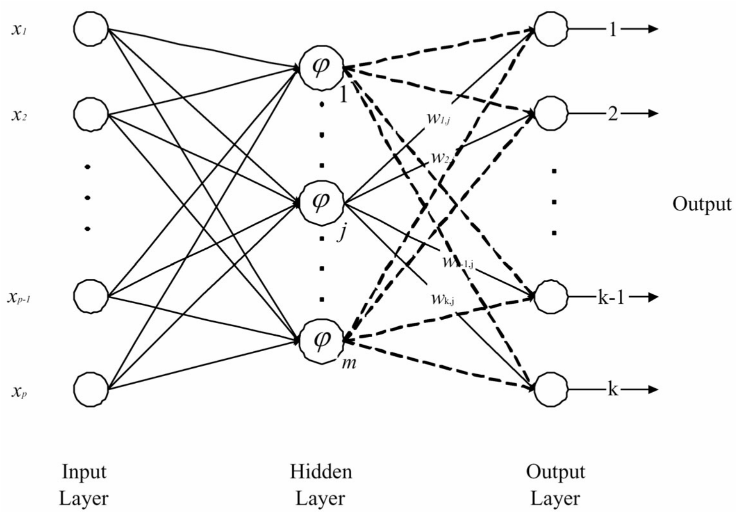


Fig. 3. Basic topology of RBFN.

$$\begin{bmatrix} \varphi_{11} & \varphi_{12} & \dots & \varphi_{1m} \\ \varphi_{21} & \varphi_{22} & \dots & \varphi_{2m} \\ \vdots & & & \\ \varphi_{N1} & \varphi_{N2} & \dots & \varphi_{Nm} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1k} \\ w_{21} & w_{22} & \dots & w_{2k} \\ \vdots & & & \\ w_{m1} & w_{m2} & \dots & w_{mk} \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1k} \\ d_{21} & d_{22} & \dots & d_{2k} \\ \vdots & & & \\ d_{N1} & d_{N2} & \dots & d_{Nk} \end{bmatrix} \tag{5}$$

where $\varphi_{ij} = \varphi(\vec{x}_i;\ \vec{c}_j)$, and $i = 1, 2, \dots, N$, $j = 1, 2, \dots, m$. We can then rewrite Eq. (5) to the form $\varphi \cdot W = D$. Thus, the weight matrix can be obtained by the least square approximation algorithm $W = (\varphi^T\varphi)^{-1}\varphi^T D$.

We employed the RBFN to classify the facial expressions images in the Eigen-space domain extracted via PCA as described in the previous section. The architecture is depicted in Fig. 3. The major advantages of RBFN over other models such as feed-forward neural network and backpropagation are its fast training speed and local feature convergence [28]. We first divided the training set images into seven classes (six principal expressions plus a neutral expression). Then the Eigen-vectors of the correlation matrix of the training data were computed. Next, the principal components of the projection of each image were collected as the training set for RBFN. For a test image, we project it onto the Eigen-space obtained previously. Then, the projection vector was fed into RBFN. Thus, the output of the RBFN will indicate which class this input image belongs to.

3.2 Hierarchical RBFN (HRBFN) Model

The cognitive and emotional states of a person can be correlated with visual features derived from images of the mouth and eye regions [1, 37]. Examining the cues for facial expression suggested by Ekman and Friesen [38], the actions of eyes, brows, and lips can significantly discriminate the changes of various expressions. Many researches on facial expression representation have focused on the specific feature motion of upper face (i.e., eyes and brows) and lower face (lips) [10, 18]. To improve the performance of classification, we further proposed a hierarchical RBFN model (HRBFN) and replaced RBFN indicated in the forth block in Fig. 4. The schematic diagram of the proposed HRBFN model is presented in Fig. 5. We divided the classification process into two stages: eyes phase and lips phase. In the first layer (eyes phase), we intend to categorize the expressions into k ($2 \leq k \leq 6$) coarse classifiers according to Eigen-features of eyes. Each classifier is aimed to differentiate candidates of a subset of expressions. The number of expressions to be recognition is denoted as n_k for classifier k. The notation is shown in Fig. 5. The philosophy of the first layer classifier design is illustrated in section 4. It is a coarse to fine hierarchical classification model.

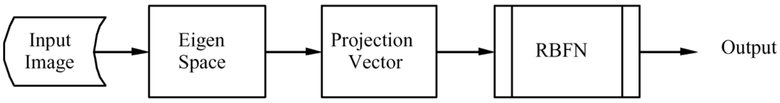


Fig. 4. The recognition flow chart of single stage RBFN.

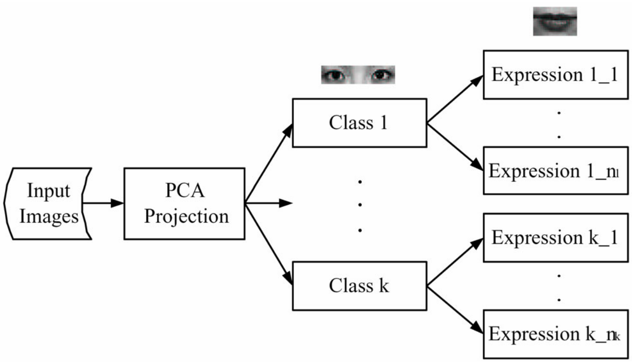


Fig. 5. Schematic block diagram of the proposed HRBFN model.

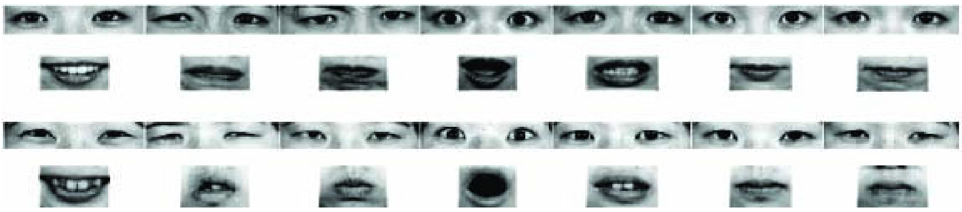
4. EXPERIMENTAL RESULTS

In this section, we demonstrate the capabilities of the proposed HRBFN approach in classifying seven facial expressions. The JAFFE database used in the experiments consists of 212 frontal pose images of ten Japanese females. Each person posed some examples of each of the seven fundamental facial expressions, such as neutral, smile, sadness, surprise, fear, anger, and disgust. The database are partitioned into two sets, namely, 73 training images and 139 test images without overlapping. Images of different expressions are randomly selected from ten persons to make training set and the remaining images are used as the test set. To investigate the local effect of the source images, two types of images were further acquired for the experiments (examples were illustrated in Fig. 4):

Type A Face images without hair, shoulders. Image size: 80×80.

Type B Images of eyes and mouth region with size of 80×20 and 45×30, respectively.

(a) Type A – face images



(b) Type B – eyes and lips images.

Fig. 6. Simulation sample images extracted from JAFFE database.

We used the images acquired from the database and transformed them into type A and type B images to perform the experiments according to the procedure introduced in the previous sections with PCA reconstruction, single stage RBFN, and HRBFN approaches. Table 1 lists the test results. For comparison, the classification results of single stage RBFN with PCA pre-processing is illustrated in Table 2. The detail classification performance comparison of each expression for type B image data are shown in Tables 3 and 4.

Table 1. Classification rate (%) of PCA reconstruction method.

Data type	Training set	Test set	
		Top one match	Top two match
Type A – face	100	78.42	89.21
Type B – lips	100	69.06	83.45
Type B – eyes	100	73.38	87.77

Table 2. Classification rate (%) of single stage RBFN with PCA pre-processing method.

Data type	Training set	Test set		
		Top one match	Top two match	Top three match
Tyep A – face	100	54.68	79.14	86.33
Type B – face	100	69.78	84.17	92.09
Tyep B – eyes	100	44.60	74.82	82.73

Table 3. Comparison of the classification rate (%) of different approaches on type B eyes test set.

Method	Sadness	Smile	Disgust	Neutral	Surprise	Fear	Anger
PCA Reconstruction	42.86	68.42	70.0	80.0	85.0	89.47	80.0
PCA + RBFN	28.57	52.63	25.0	50.0	30.0	100	30.0

Table 4. Comparison of the classification rate (%) of different approaches on type B lips test set.

Method	Sadness	Smile	Disgust	Neutral	Surprise	Fear	Anger
PCA Reconstruction	61.90	78.95	50.0	75.0	85.0	78.9	55.0
PCA + RBFN	61.90	68.42	70.0	60.0	75.0	78.95	75.0

As we can observe from Tables 1 and 2, the classification rate of using PCA reconstruction method based on face images is up to 89.21% (top two matches). The correct classification can be further extended to 92.09% and 82.73% based on local lips and eyes features and single stage RBFN, respectively, by counting the top three matched candidates. Thus, it is possible to re-organize the training set and construct a hierarchical RBFN as proposed in this paper.

To obtain the design guideline of the hierarchical classifiers of the multiple layers delineation structure, we want to observe whether local image (lips or eyes) eigen-feature can help to do the preliminary clustering task. The confusion matrices of the single stage RBFN top one match test results are recorded in Tables 5 and 6. Not all expressions were

Table 5. Confusion matrix of facial expression classification measured by single stage RBFN method on type B eyes test images from JAFFE database.

I\O	Sadness	Smile	Disgust	Neutral	Surprise	Fear	Anger
Sadness	6	1	1	0	0	13	0
Smile	0	10	1	0	0	8	0
Disgust	1	1	15	0	0	13	0
Neutral	0	3	0	10	0	7	0
Surprise	0	0	0	3	6	14	0
Fear	0	0	0	0	0	19	0
Anger	0	0	0	0	0	14	6

Table 6. Confusion matrix of facial expression classification measured by single stage RBFN method on type B lips test images from JAFFE database.

I\O	Sadness	Smile	Disgust	Neutral	Surprise	Fear	Anger
Sadness	13	0	1	1	0	0	6
Smile	4	13	1	1	0	0	0
Disgust	4	2	14	0	0	1	1
Neutral	3	0	0	12	0	4	1
Surprise	3	1	1	0	15	0	0
Fear	1	0	1	1	1	15	0
Anger	3	1	1	0	0	0	15

equally well recognized by the system. As we can see from the confusion matrices, expression "fear" was often mis-classified by others. Similarly, expressions "sadness" and "disgust" were also mis-classified. Thus, we intend to divide the mid-level classification into various categories and train the lips image individually for each classifier to perform less intensive identification task in the second stage. Various scenarios can be created. We have tested all cases. Three of the well trained cases are: scenario A is formed by using expression "surprise" as one subset and the reset expressions are in one subset, scenario B contains two subsets 0 ("surprise, neutral") and the reset of expressions, scenario C also divided into two parts with ("smile", surprise", neutral") and (others). Classification results of the proposed HRBFN are illustrated in Table 7. The best classification ratios of the first phase, eyes layer, is 100% for training set and 95.68% for test set. The performance of lips phase is around 72.66%. Lyons reported 75% generalization rate over expression identification on the same database but with six classes [21]. Both simulation results are compatible. Our results for the eyes phase are outperforming. Besides, the proposed HRBFN is simple and with low complexity. From these results, we conclude the local face features is useful for differentiating expressions when a more appropriate modular classifiers are provided.

Table 7. Classification rate (%) of the proposed HRBFN model.

	Scenario A		Scenario B		Scenario C	
	eye phase	lips phase	eye phase	lips phase	eye phase	lips phase
Training set	100	100	100	100	100	100
Test set	95.68	71.94	93.52	72.66	91.37	69.78

Table 8. Comparison of the classification rate (%) of different HRBFN arrangements on test set.

Scenario	Phase	Sadness	Smile	Disgust	Neutral	Surprise	Fear	Anger
A	eyes	100	100	100	95.0	80.0	94.73	100
	lips	66.67	73.68	70.0	65.0	80.0	73.68	75.0
B	eyes	100	100	100	70.0	90.0	100	95.0
	lips	57.14	68.42	80.0	70.0	85.0	78.95	70.0
C	eyes	95.24	84.21	90.0	95.0	90.0	84.21	100
	lips	57.14	57.89	60.0	95.0	65.0	73.68	80.0

5. CONCLUSIONS

In this paper, we proposed a hierarchical RBFN model (HRBFN) and combined PCA feature extraction method to tackle facial expressions classification problems. The related methods were also investigated. From our experimental results, full face images provide confusing information for distinguishing the facial expressions. We proposed a new construction of HRBFN model. As expected, it could strain the difficult expression classification efficiently. This cascade strategy works properly. We conclude that local images of lips and eyes can be treated as cues for facial expression. They are good enough to provide feature for discrimination. Moreover, the images of eyebrows can also be adopted to test the recognition performance.

Nevertheless, the proposed work is constrained to a proper extraction of lips and eyes images. This limitation can be resolved by advanced development of computer vision technology. In fact, the recognition of facial expressions is a challenging task. The appearance of one expression may vary from person to person. Variations of each expression may also large in different databases. Furthermore, some expressions are ambiguous, such as surprise and fear. The difference between such expressions is hard to discriminate. We also found that most of the misjudgment was classified into expressions that looks similarly to the expression it should be. This is an interesting phenomenon and may be the important clues for the future research to improve the classification performance. More work should be conducted to validate the technique on other larger databases such as CMU AMP Face Expression Database [39] and POFA image set of Ekman and Friesen [40]. In addition, we can apply learning mechanism to determinate the optimal combination of first layer classifiers to yield good performance. Meanwhile, the PCA reconstruction scheme can be also re-arranged hierarchically to seek changes of improving hit ratio.

REFERENCES

1. Z. Duric, W. Gray, R. Heishman, F. Li, A. Rosenfeld, M. Scholles, C. Schunn, and H. Wechsler, "Integrating perceptual and cognitive modeling for adaptive and intelligent human computer interaction," in *Proceedings of the IEEE*, Vol. 90, 2002, pp. 1272- 1289.
2. M. Bartlett, P. Viola, T. Sejowski, L. Larsen, J. Hager, and P. Ekman, "Classifying facial action," in M. Mozer, D. S. Touretzky, and M. Hasselmo, (eds.), *Advances in Neural Information Processing Systems*, Vol. 8, The MIP Press, 1996, pp. 823-829.
3. A. Colmenarez, B. Frey, and T. S. Huang, "Embedded face and facial expression recognition," in *Proceedings of the International Conference on Image Processing*, Vol. 1, 1999, pp. 633-637.
4. I. Essa and A. Pentland, "Coding, analysis, interpretation, and recognition for facial expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, 1997, pp. 757-763.
5. T. Kurozumi, Y. Shinza, Y. Kenmochi, and K. Kotani, "Facial individuality and expression analysis by eigenspace method based on class features or multiple discriminant analysis," in *Proceedings of the International Conference on Image Processing*, Vol. 1, 1999, pp. 648-652.
6. K. Mase, "Recognition of facial expression from optical flow," *IEICE Transactions*, Vol. E74, 1991, pp. 3474-3483.
7. K. Mase, "Recognizing multiple persons' facial expressions using HMM based on automatic extraction of significant frames from image sequences," in *Proceedings of the International Conference on Image Processing*, Vol. 2, 1997, pp. 546-549.
8. M. Rosenblum, Y. Yacoob, and L. Davis, "Human emotion recognition from motion using a radial basis function network architecture," in *Proceedings of the IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, 1994, pp. 43-49.

9. M. Rosenblum, Y. Yacoob, and L. Davis, "Human expression recognition from motion using a radial basis function network architecture," *IEEE Transactions on Neural Networks*, Vol. 7, 1996, pp. 1121-1138.
10. Y. L. Tian, T. Kanade, and J. Cohn, "Recognizing action units for facial expression analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, 2001, pp. 97-115.
11. Y. Yacoob and L. Davis, "Recognizing human facial expression from long image sequences using optical flow," *IEEE Transactions on Annual Machine Intelligence*, Vol. 16, 1994, pp. 636-642.
12. Y. Yacoob, H. M. Lam, and L. S. Davis, "Recognizing faces showing expressions," in *Proceedings of the International Workshop on Automatic Face and Gesture Recognition*, 1995, pp. 278-283.
13. D. Yang, T. Kunkhiro, H. Shimoda, and H. Yoshikawa, "A study of real-time image processing method for treating human emotion by facial expression," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Vol. 2, 1999, pp. 360-364.
14. P. Ekman and W. Friesen, *The Facial Action Coding System, Consulting Psychologists Press*, Palo Alto, CA, 1965.
15. G. Donato, M. Bartlett, J. Hager, P. Ekman, and T. Sejnowski, "Classifying facial actions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, 1999, pp. 974-989.
16. X. W. Chen and T. Huang, "Facial expression recognition: a clustering-based approach," *Pattern Recognition Letter*, Vol. 24, 2003, pp. 1295-1302.
17. G. Cottrell and J. Metcalfe, "EMPATH: face, gender, and emotion recognition using hoons," in J. Moddy, R. P. Lippman, and D. S. Touretzky, (eds.), *Advances in Neural Information Processing Systems*, Vol. 3, 1991, pp. 564-571.
18. C. Padgett and G. Cottrel, "Identifying emotion in static face images," in *The Second Joint Symposium of Neural Computation*, Vol. 5, 1995, pp. 91-101.
19. K. Matsuno, C. Lee, and S. Tsuji, "Recognition of human facial expressions without feature extraction," in *Proceedings of the European Conference on Computer Vision*, 1994, pp. 513-520.
20. L. Ma and K. Khorasani, "Facial expression recognition using constructive feedforward neural network," *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, Vol. 34, 2004, pp. 1588-1559.
21. M. Lyons, J. Budynek, and S. Akamatsu, "Automatic classification of single facial images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, 1999, pp. 1357-1362.
22. K. Matsuno and S. Tsuji, "Recognizing human facial expressions in a potential field," in *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, Vol. 2, 1994, pp. 44-49.
23. Y. Shinza, Y. Saito, Y. Kenmochi, and K. Kotani, "Facial expression analysis by integrating information of feature-point positions and gray levels of facial images," in *Proceedings of the IEEE International Conference on Image Processing*, Vol. 1, 2000, pp. 466-469.
24. S. Carey and R. Dimond, "From piecemeal to configurational representation of faces," *Science*, Vol. 195, 1997, pp. 312-313.

25. A. Lanitis, C. Taylor, and T. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, 1997, pp. 743-756.
26. M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, Vol. 19, 1997, pp. 743-756.
27. P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherface: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, 1997, pp. 711-720.
28. S. Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing Company, New York, 1994.
29. H. Abdi, D. Valentin, and A. J. O'Toole, "A generalized auto-associator model for face processing and sex categorization: from principal components to multivariate analysis," in D. Levine, (ed.), *Optimality in Biological and Artificial Networks*, 1996, pp. 317-337.
30. M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, Vol. 3, 1991, pp. 71-86.
31. M. Fleming and G. Cottrell, "Categorization of faces using unsupervised feature extraction," in *Proceedings of the International Joint Conference on Neural Networks*, Vol. 2, 1990, pp. 65-70.
32. T. Kohonen, E. Oja, and P. Lehtio, "Storage and processing of information in distributed associative memory systems," in G. E. Hinton and J. A. Anderson, (eds.), *Parallel Models of Associative Memory*, 1981, pp. 105-143.
33. S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: a hybrid neural network approach," Technical Report, No. UMIACS-TR-96-16, Institute for Advanced Computer Studies, University of Maryland, College Park, 1996.
34. M. Er, S. Wu, J. Lu, and H. Toh, "Face recognition with radial basis function neural networks," *IEEE Transactions on Neural Networks*, Vol. 13, 2002, pp. 697-710.
35. T. J. Stonham, "Practical face recognition and verification with WIZARD," in H. Ellis, M. Jeevs, F. Newcombe, and A. Young, (eds.), *Aspects of Face Processing*, 1986, pp. 426-441.
36. M. Yoneyama, A. Ohtake, Y. Iwano, and K. Shirai, "Facial expressions recognition using discrete Hopfield neural network," in *Proceedings of the International Conference on Image Processing*, Vol. 1, 1997, pp. 117-120.
37. M. Knapp and J. Hall, *Nonverbal Communication in Human Interaction*, Harcourt Brace Jovanovich, New York, 1997.
38. P. Ekman and W. Friesen, *Unmaking the Face*, Prentice-Hall, Englewood Clifs, New Jersey, 1975.
39. Advanced Multimedia Processing Lab., Carnegie Mellon University, http://amp.ece.cmu.edu/projcets/faceauthentication/download.htm.
40. P. Ekman and W. V. Friesen, "Pictures of facial affect," Human Interaction Laboratory, University of California Medical Center, San Francisco, 1976.

Daw-Tung Lin (林道通) received the B.S. degree in Control Engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1985, and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Maryland at College Park, MD, U.S.A., in 1990 and 1994, respectively. From February 1995 to August 2005, he was as an Associate Professor of the Department of Computer Science and Information Engineering at the Chung Hua University, Hsinchu, Taiwan. He is currently an Associate Professor of the Department of Computer Science and Information Engineering at the National Taipei University, Taipei County, Taiwan. His current research interests are pattern recognition, intelligent surveillance, neural networks, fuzzy systems, and medical image processing.