Learning to Recognize Faces by Local Feature Design and Selection

Yanwei Pang, Lei Zhang, and Zhengkai Liu

Abstract—Studies in neuroscience suggest that both global and local feature information are crucial for perception and recognition of faces. It is widely believed that local feature is less sensitive to variations caused by illumination, expression and illumination. In this paper, we target at designing and learning local features for face recognition. We designed three types of local features. They are semi-global feature, local patch feature and tangent shape feature. The designing of semi-global feature aims at taking advantage of global-like feature and meanwhile avoiding suppressing AdaBoost algorithm in boosting weak classifies established from small local patches. The designing of local patch feature targets at automatically selecting discriminative features, and is thus different with traditional ways, in which local patches are usually selected manually to cover the salient facial components. Also, shape feature is considered in this paper for frontal view face recognition. These features are selected and combined under the framework of boosting algorithm and cascade structure. The experimental results demonstrate that the proposed approach outperforms the standard eigenface method and Bayesian method. Moreover, the selected local features and observations in the experiments are enlightening to researches in local feature design in face recognition.

Keywords—Face recognition, local feature, AdaBoost, subspace analysis.

I. INTRODUCTION

FACE recognition has been an active research topic in the past few decades due to its scientific challenges and potential applications. Studies in psychology and neuroscience suggest that both global and local features are crucial for the perception and recognition of faces [1-2]. These studies inspire researchers to design features and learning algorithms for machine recognition of faces.

According to feature type, face recognition methods can be categorized into holistic matching methods, local feature based matching methods and hybrid methods [1]. In hybrid methods, both local feature and global face region are used.

The current state-of-the-art of holistic matching methods is characterized by a family of subspace methods. Sirvovich *et al* [3] used Karhunen-Loeve (KL) transform to represent a

Yanwei Pang is with the Department of Electronic Engineering and Information Science, Hefei 230027, China. (phone: 86 3601342; fax: 86 3601342; e-mail: pyw@ ustc.edu.cn).
Lei Zhang is with Microsoft Research Asia, Beijing 100080, China (e-mail: leizhang@microsoft.com).
Zhengkai Liu is with the Department of Electronic Engineering and Information Science, Hefei 230027, China. e-mail: pyw@ ustc.edu.cn).

face image. Using Principal Component Analysis (PCA), which is related to KL transform, Turk *et al* [4] proposed the famous "eigenface" methods. Swets *et al* [5] proposed to use the Most Discriminant Features (MDF) for image matching. Peter [6] et al further developed the famous "fisherface" method in which class-specific information is used. Moghaddam et al [7] proposed a Bayesian face recognition method which uses Probabilistic Principal Component (PPCA) to estimate density in high dimensional space. Wang et al [8] further developed a unified subspace analysis method that exploits three subspace dimensions, i.e intrinsic difference, transform difference and noise.

Compared to holistic features, local features are less sensitive to pose variations, lighting variations, expression variations and occlusion. Brunelli *et al* [9] utilized four facial regions for template matching. Pentland *et al* [10] extended eigenface to eigen-features such as eigen-eye, eigen-mouth and eigen-nose. Without finding the exact locations of facial features, Hidden Markov Model (HMM) based method [11] uses strips of pixels that cover the forehead, eye, nose, mouth and chin. The most representative and successful local feature based methods in academia may be Elastic Bunch Graph Matching (EBGM) proposed by Wiskott *et al* [12] and Local Feature Analysis (LFA) proposed by Pensev *et al* [13]. More recently, Heisele *et al* [14] proposed a component-based method where fourteen facial regions are chosen and SVM classifiers are constructed for recognition. Lucy *et al* [15] developed a part-based face representation method where eight facial regions are used and Gaussian Mixture Model (GMM) is employed to estimate class probability.

However, many fundamental challenges are still not well addressed by most of aforementioned local feature based methods. How to select salient regions? Why four or five facial regions are used? What kind of features should be
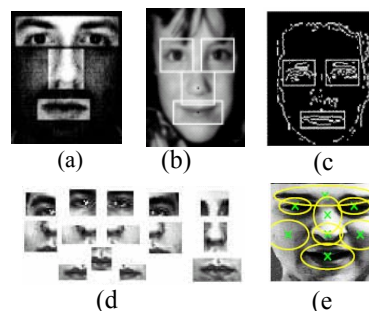

Fig. 1 Local features used by previous work [9-10,19,14-15]

World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:2, No:9, 2008

extracted from these selected regions? How to combine them optimally to get the best result? The proposed method in this paper tries to solve these problems as possible as it can in the framework of AdaBoost algorithm which can simultaneously perform both feature selection and classifier fusion.

Previous local feature based methods [9-10,19,14-15] intuitively extract features from facial components such as eyes, nose and mouth (See Fig. 1). Different with them, we design three types of local features: semi-global features (Fig. 4.4(b)), local patches (Fig. 5) and tangent shapes [24]. Studies in neurolscience suggest that both global and local feature information are crucial for the perception and recognition of faces [1-2]. Intuitively, it seems that combining of global and local features could enhance recognition rate. However, usually the classifier established from the global face region is so strong that it might suppress the powerfulness of the AdaBoost algorithm. To utilize AdaBoost algorithm to select and combine multiple features, as a tradeoff, semi-global features, instead of global features, and local features are combined. The second type feature is local patch. By learning (feature selection), one can find which regions are salient for recognizing faces. The third type feature is tangent shape [24]. Shape information is of mid-level semantic and may be more discriminative for recognition than low level features.

The remainder of this paper is organized as follows. In section II the details of feature design are described. In section III issues related to the learning algorithm are presented. Experimental results are shown in section IV and conclusions are drawn in section V.

II. FEATURE DESIGN

A. Semi-Global Features

Studies suggest that both global and local feature information are crucial for the perception and recognition of faces. It seems that global and local features should be combined to enhance recognition rate. However, usually the classifier established from the global face region is so strong that it might suppress the powerfulness of the AdaBoost algorithm. To utilize AdaBoost algorithm, as a tradeoff, semi-global feature, instead of global feature is proposed. By "semi", we mean that semi-global feature can be regarded as either global or local feature. Semi-global feature has the property of local feature in that only part of the facial region is used. Meanwhile, it resembles global feature because it covers several facial components (see Fig. 3.4(b)).

Inspired by the work [18], we adopt the discriminant vectors (i.e fisherfaces) obtained by LDA to design semi-global features .

In subspace analysis method, each pixel in an image corresponds to a dimension (or axis) in the high dimensional space. If dimension i carries discriminant information, it also means that pixel i and features extracted from pixel i carry discriminant information. In LDA subspace method, a linear transform matrix $\mathbf{W}=[\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_m]$ is obtained by maximize

the ratio between within-class scatter and between-class scatter. As illustrated in Fig. 2 for the two dimensional case, $\mathbf{w}_2=(-0.34,0.93)$ is the discriminant vector. Projected onto the direction $\mathbf{w}_2$, data can be separated completely. Because the angle between e_2 axis and vector $\mathbf{w}_2$ is smaller than that between e_1 axis and $\mathbf{w}_2$, so it is safe to say that $\mathbf{e}_2$ axis carries more discriminant information than e_1 axis. In another way, because $0.93=|w_{22}|>|w_{21}|=0.34$, so axis $\mathbf{e}_2$ carries more discriminant information than axis $\mathbf{e}_1$.
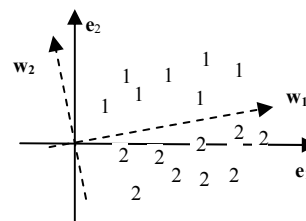

Fig. 2 Find the most discriminant axis in 2D
space. There are two classes

We use thçok kee m discriminant vectors (fisherfaces) to statistically determine the discriminating ability of each dimension. By summing the absolute value of $\mathbf{w}_i$ over i, we get a new vector:

$$\mathbf{w}' = \sum_{i=1}^{m} | \mathbf{w}_i | \qquad (1)$$

Because $\mathbf{w}'$ reflects the saliency of each pixel (dimension), we denote $\mathbf{w}'$ as saliency map. Discarding those small value elements in the saliency map, we can get the semi-global region. Semi-global features are features that are extracted from pixels covered by semi-global region. In section , we will show the semi-global region and analyze its property.

From the semi-global region, we extract Gabor features. As in [12], 40 Gabor kernels are employed. Let $\mathbf{I}_{sg}$ denotes an image masked by semi-global region and $\mathbf{g}_k$ ($k=1,\ldots,40$) represent Gabor kernels. Then the extracted features are

$$\mathbf{f}_k = \mathbf{I}_{sg} \otimes \mathbf{g}_k \quad k=1,\ldots,40 \qquad (2)$$

where $\otimes$ represents convolution operator.

Classifier established from each Gabor-filtered semi-global feature $\mathbf{f}_k$ is expected to be not too strong to suppress the powerfulness of AdaBoost algorithm.

B. Local Features

We claim that both semi-global features and local features are crucial for face recognition. In this section, local feature extraction will be discussed.

Local Patches

To extract local features, overlapped local patches are used. For a given scale (20×20 in a 115×115 image for example), a window slides from left to right, from top to bottom with r (for example, 90%) overlapping area between two adjacent patches. Note that the size of a patch should not be too large, otherwise the resultant classifier might not be *weak* [16]. The size is empirically determined by experiments. The saliency map in Fig 1.1(a) might also provide a clue to choose local patch.

World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:2, No:9, 2008

We can take original pixel intensity of local patches as features and input them to the AdaBoost algorithm. Because in this work the weak classifier is chosen to be Bayesian classifier and probabilistic principal component analysis [7] is adopted, the extracted features are PCA features in essence.

More Features

As stated, PCA features are extracted from local patches. We can further extract other features as well. As long as the weak classifier established from these features is somewhat better than randomly guessing, the performance of the final classifier can be boosted.

Thus LDA features are considered. For two class problem in Fig. 3, LDA features cannot always be superior to PCA features though the former uses class-specific information. In Fig. 3(a), both PCA and LDA can separate the two classes correctly because there is no overlap between the projections of the two classes. But in Fig. 3(b), PCA projections overlapped completely and there is no overlapping for LDA projections. However, in Fig. 3(c), PCA is superior to LDA. This is because that LDA is prone to overfit when the *small sample size problem* occurs [20]. For different situations, PCA and LDA exhibit different functions. Therefore, in addition to PCA features, we expect to add LDA features to improve recognition performance. They are selected and combined by the AdaBoost algorithm.

However, belonging to subspace methods, both PCA and LDA are sensitive to geometric transformation such as translation and rotation. In subspace method, each pixel corresponds to an axis (dimension) in high dimensional space.
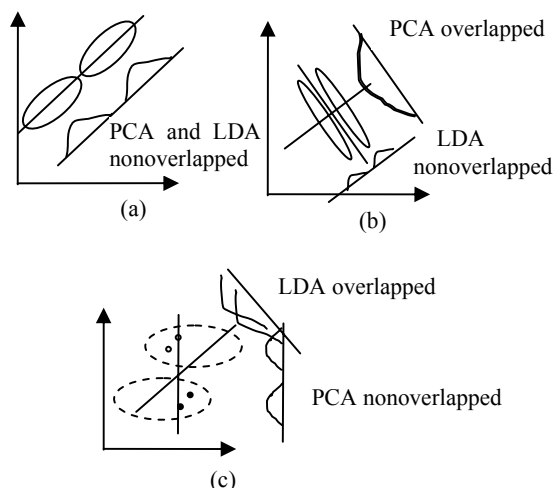

Fig. 3 Comparing PCA and LDA in 2D space

Even translation in two dimensional images is nonlinear in high dimensional space. The performance of the subspace method degrades drastically due to geometric transformation. To address such drawback, we decide to add singular values as additional features. Singular Value Decomposition (SVD) has been an important pattern recognition method [21-22]. If

A is an image patch, the SVD theory states that A can be decomposed as:

$$\mathbf{A} = \mathbf{U\Sigma V}$$

The diagonal elements of ∑ form the singular value features.

It has been proven that singular values are invariant to translation and rotation and are robust to noise and perturbation [21]. In practice, it is difficult to locate facial components (such as two eyes) accurately to pixel level. So geometric transformation such as translation and rotation can not be avoided. It is expected that a feature vector formed by singular values can contribute to improve face recognition performance.

Shape information is usually considered not stable when pose variations occur. Few works has been done to use shape feature [23]. But shape information might be helpful if we could constrain our task to frontal view only. Using Bayesian Tangent Shape Model (BTSM) proposed by Zhou *et al* [24], we obtain six facial components: left eyebrow, right eyebrow, left eye, right eye, nose, mouth and the whole face contour. Tangent shapes are used because its coordinates are independent of the absolute coordinates (refer to [24] for details). Tangent shapes of these components and their combinations are extracted. As long as the weak classifier established from these features is somewhat better than randomly guessing, the performance of the final classifier can be boosted.

III. LEARNING ALGORITHM

In this Section, some issues related to learning algorithm will be presented. Both semi-global features and local features are selected by the AdaBoost learning algorithm.

We adopted Viola's AdaBoost algorithm [16] which is a variant of the original AdaBoost algorithm [17]. The training error drops exponentially fast with respect to iteration number. Moreover, AdaBoost has good generalization ability and is not prone to overfitting. So, we can input a lot of features to the AdaBoost learning algorithm. These features include local patches, LDA features, SVD features and tangent shapes.

The AdaBoost algorithm deals well with two class problem. In most cases, the task of face recognition is multiple-class problem. Just as [7], we convert multi-class problem to binary class problem by defining intra-class Ω_I and extra-class Ω_E. Let $\mathbf{x}_i$ and $\mathbf{x}_j$ are two feature vectors, and $\Delta=\mathbf{x}_i\text{-}\mathbf{x}_j$. The class label of $\mathbf{x}$ is $l(\mathbf{x})$. Then the intra-class and extra-class are defined respectively by

$$\Omega_I = \{\Delta \mid \ell(\mathbf{x}_i) = \ell(\mathbf{x}_j)\} \text{ and } \Omega_E = \{\Delta \mid \ell(\mathbf{x}_i) \neq \ell(\mathbf{x}_j)\} .$$

Note that in our method, $\mathbf{x}$ are not necessary the image difference. They might be the difference between two LDA feature vectors or shape vectors.

A. Weak Classifier Learning with Weighted Samples

In this subsection, the form of the weak classifier is given. Meanwhile, we discuss how to train the weak classifier with weighted samples

World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:2, No:9, 2008

Given intra-class Ω_I and extra-class Ω_E, we employee Bayesian classifier [7] as the weak classifier. The following two factors guarantee the weakness the Bayesian classifier. 1) The Bayesian classifier is learned from local or semi-global facial region. Obviously, it is hard to get a strong classifier from small local facial region. 2) The complexity of the Bayesian classifier can be controlled by adjusting the parameter M in equation (5). The bigger the M is, the more complex the weak classifier is.

Two feature vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ are determined to belong the same individual if

$$P(\Omega_I \mid \Delta) > P(\Omega_E \mid \Delta) \qquad (3)$$

The posteriori probability is given by

$$P(\Omega_I \mid \Delta) = \frac{P(\Delta \mid \Omega_I)P(\Omega_I)}{P(\Delta \mid \Omega_I)P(\Omega_I) + P(\Delta \mid \Omega_E)P(\Omega_E)} \qquad (4)$$

Assume both Ω_I and Ω_I are Gaussian distributed, we have

$$P(\Delta \mid \Omega) \approx \frac{\exp(-0.5\sum_{i=1}^{M} \frac{\mathbf{y}_i^2}{\lambda_i})}{(2\pi)^{M/2} \prod_{i=1}^{M} \lambda_i^{0.5}} \qquad (5)$$

where principal component $\mathbf{y}_i$ are obtained by solving the following eigenproblem

$$\mathbf{C}\lambda_i = \lambda_i \mathbf{u}_i \qquad (6)$$

$$\mathbf{C} = \sum_i^N p_i \Delta_i \Delta_i^T \qquad (7)$$

where p_i is the priori probability of Δ_i. See [7] for details.

Parameter M is determined by the energy ratio r:

$$r = \sum_{i=1}^{M} \lambda_i \Big/ \sum_{i=1}^{N} \lambda_i \qquad (8)$$

To train the above Bayesian classifier with sample weights $\mathbf{w}$ in the AdaBoost algorithm, we can simply substitute $w(i)$ for p_i :

$$\mathbf{C} = \sum_i^N p_i \Delta_i \Delta_i^T = \sum_i^N w(i) \Delta_i \Delta_i^T \qquad (9)$$

Furthermore, one can directly weight each sample Δ_i

$$\tilde{\Delta}_i = \sqrt{w(i)}\Delta_i \qquad (11)$$

Then it holds

$$\mathbf{C} = \sum_i^N w(i) \Delta_i \Delta_i^T = \sum_i^N \tilde{\Delta}_i \tilde{\Delta}_i^T \qquad (12)$$

B. Two Classes to Multi Classes

By defining intra-class Ω_I and extra-class Ω_E , face recognition problem is converted to two-classe problem. However, face recognition is usually multi-class problem (more than two persons). One must convert the two-class classification results back to multi-class classification results.

It is worth noting that cascade learning structure must be adopted. If there are C=500 persons and each person has m=2 images, then the number of positive samples is p=500 while the number of negative samples is $n = p \times [(m/(m-1) \times (c-1)] = 499000$. AdaBoost learning using the cascade structure can deals well with this unbalance classification problem. The AdaBoost classifier in each stage has the following form:

$$H(\mathbf{x}) = sign(\sum_{t=1}^{T} \alpha_t h_t(\mathbf{x}) - \tau)$$

where

$$\tau = \begin{cases} 0, & if \quad Y = \{+1, -1\} \\ 0.5, & if \quad Y = \{0, 1\} \end{cases}$$

$$\sum_{t=1}^{T} \alpha_t = 1$$

Y is the class label set. h_t denotes weak hypothesis and a_t is its coefficient. τ is a threshold for this stage.

Let $\mathbf{x}$ be a probe feature vector, $\{\mathbf{y}_i\}$ be the gallery set and $\mathbf{z}_i=\mathbf{x}-\mathbf{y}_i$. If $\mathbf{z}_i$ can pass through s stages of the cascade structure, it holds

$$\sum_{t=1}^{T^{(1)}} \alpha_t h_t^{(1)}(\mathbf{z}) - \tau^{(1)} > 0$$

$$\sum_{t=1}^{T^{(2)}} \alpha_t h_t^{(2)}(\mathbf{z}) - \tau^{(2)} > 0$$

…

$$\sum_{t=1}^{T^{(s)}} \alpha_t h_t^{(s)}(\mathbf{z}) - \tau^{(s)} > 0$$

where the superscript (i) represents the stage order of cascade structure. Then

$$\left[\sum_{t=1}^{T^{(1)}} \alpha_t h_t^{(1)}(\mathbf{x}) + \sum_{t=1}^{T^{(2)}} \alpha_t h_t^{(2)}(\mathbf{x}) + \cdots + \sum_{t=1}^{T^{(s)}} \alpha_t h_t^{(s)}(\mathbf{x}) \right]$$
$$- \left[\tau^{(1)} + \tau^{(2)} + \cdots + \tau^{(s)} \right] > 0$$
$$\tilde{H}(x) - \Gamma^{(s)} > 0 \qquad (13)$$

$$\tilde{H}(x) = \sum_{t=1}^{T^{(1)}} \alpha_t h_t^{(1)}(\mathbf{x}) + \sum_{t=1}^{T^{(2)}} \alpha_t h_t^{(2)}(\mathbf{x}) + \cdots + \sum_{t=1}^{T^{(s)}} \alpha_t h_t^{(s)}(\mathbf{x})$$

$$\Gamma^{(s)} = \tau^{(1)} + \tau^{(2)} + \cdots + \tau^{(s)}$$

All $\mathbf{z}$ which satisfy equation (13) form a set

$$Z = \{z : \tilde{H}(z) - \Gamma^{(s)} > 0\} \qquad (14)$$

Take s as variable. Select s^* such that

$$s* = \arg\min_s |Z| = |\{z : \tilde{H}(z) - \Gamma^{(s)} > 0\}| \qquad (15)$$

s.t $|Z| \geq 1$

where $|Z|$ represents the number of elements of the set Z. Then a new set is formed:

$$Z* = \{z : \tilde{H}(z) - \Gamma^{(s*)} > 0\} \qquad (16)$$

The class label of the probe $\mathbf{x}$ is considered to be the same as that of the sample i^* in the gallery set:

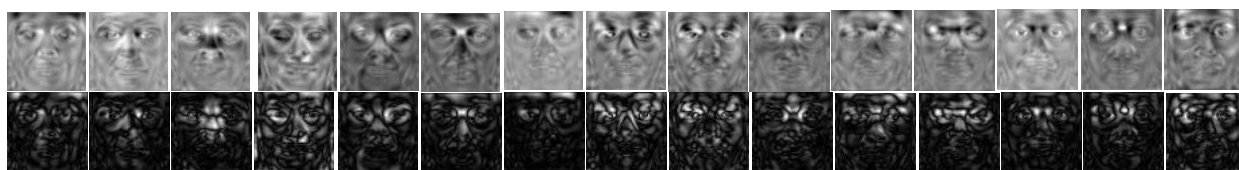$$i^* = \arg\max_i \tilde{H}(\mathbf{z}_i) \qquad (17)$$

where $\mathbf{z}_i \in Z*$.

IV. EXPERIMENTAL RESULTS

In this section, after describing the experiment configuration, we will illustrate the designed semi-global features and the selected local features. Finally, the performance comparison with two baseline methods is reported in terms of recognition accuracy

A. Configuration

We use FA/FB of FERET [25] database for training and testing. There are 1195 individuals in both FA and FB in our

World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:2, No:9, 2008

(1) $\mathbf{w}_1, \mathbf{w}_2,\ldots\mathbf{w}_{15}$ (1st line) and $|\mathbf{w}_1| \, |\mathbf{w}_2|,\ldots|\mathbf{w}_{15}|$ (2nd line). Discrimnant features lie mostly around eyes.
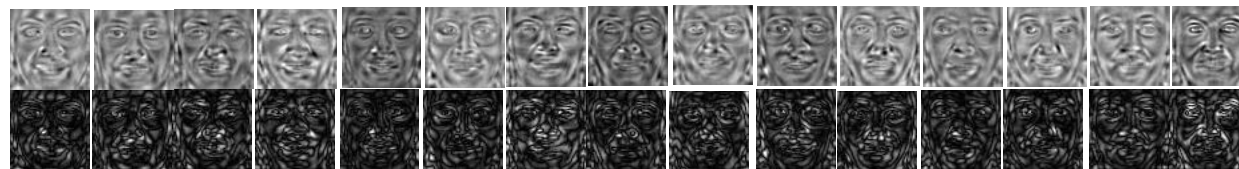
(2) $\mathbf{w}_{103}, \mathbf{w}_{104},\ldots\mathbf{w}_{117}$ (1st line) and $|\mathbf{w}_{103}| \, |\mathbf{w}_{104}|,\ldots|\mathbf{w}_{117}|$ (2nd line). There are discriminant features around the two sides of a nose.
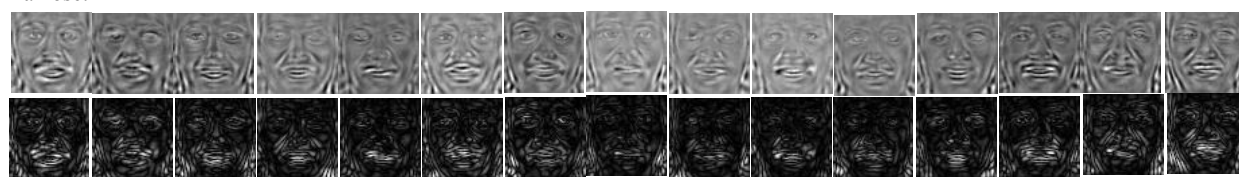
(3) $\mathbf{w}_{185}, \mathbf{w}_{186},\ldots\mathbf{w}_{199}$ (1st line) and $|\mathbf{w}_{185}| \, |\mathbf{w}_{186}|,\ldots|\mathbf{w}_{199}|$ (2nd line). Discriminant features lie mostly around mouth corners.

(a)　(b)　(c)　(d)　(e)　(f)　(g)　(h)　(i)　(j)
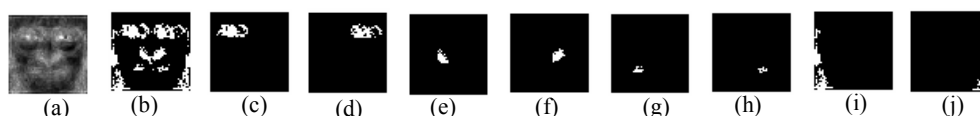
(4) (a) is saliency map $\mathbf{w}' = \sum_{i=1}^{199}|\mathbf{w}_i|$, (b) is binarized semi-global region . (b) can be segmented into (c)-(j)

Fig. 4 Discriminant vectors, saliency map and the semi-global region

experiments with two images of each individual. We use 495 pairs of images for training set. The remained 700 pairs of images are used for testing with 700 images in gallery set and another 700 images in probe set. In the first experiment, 700 images in FA are used as gallery set and 700 images in FB as probe set (we denote this probe set as "subFB"). Similarly, in the second experiment, "subFA" is used as probe set.

All images are normalized by the eye center positions. The size of the final images is 115×115. To alleviate the influence of lighting variations, histogram equalization is applied for each image. Then, each image is normalized to be zero mean and unit variance. Compared to conventional method, no mask (such as ellipse mask) is used because many features will be extracted from nearby a face boundary.

In the cascade learning structure, 3 stages are used with recall rates 0.994, 0.989 and 0.970 respectively. To decrease the complexity of the weak classifier, the energy ratio r in equation (8) is set to be 0.4. There are 9,536 local features in total.

In the following two subsections, we will discuss and analyze the designed and selected features in accordance with the iteration order of the AdaBoost algorithm.

B. Designed Semi-Global Feature

Semi-global region is derived from the saliency map $\mathbf{w}'$ (see Equation (1)). Let the eigenvalue corresponding to $\mathbf{w}_i$ be λ_{I} which is the solution of the generalized eigenproblem:

$$\mathbf{S}_b\mathbf{w}_i = \lambda_i\mathbf{S}_w\mathbf{w}_i , \ \lambda_1 > \lambda_2 > \cdots > \lambda_m$$

where the value of m is determined by the ratio

$$\sum_{i=1}^{C-1} \lambda_i \Big/ \sum_{i=1}^{C-1} \lambda_i \approx q$$

In our experiment, the class number C equals to 495 and q is set to be 0.95. Consequently, 199 (m=199 discriminant vectors are computed. The resultant $\mathbf{w}_i$, $|\mathbf{w}_i|$ and the saliency map $\mathbf{w}'$ re shown in Fig. 4.

Fig. 5 The first 14 selected local patches

World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:2, No:9, 2008

It can be seen that the first few $|\mathbf{w}_i|$ (i=1,...,15) highlight discriminant features around two eyes. The last few $|\mathbf{w}_i|$ (i=185,...,199) highlight discriminant features around two mouth corners. Fig. 4.4 (a) visualizes $\mathbf{w}'$. To demonstrate the characteristics of the semi-global region, we segment Fig. 4.4 (b) into several local parts which are shown in Fig. 4.4 (c-j). One can find that two mouth corners, instead of the whole mouth, are retained. Intuitively, when the status of a mouth alters from open to close or from close to open, its appearance changes drastically. But its two corners are more stable comparing to the whole mouth. Thus the extracted features form two mouth corners are more robust than those from the whole mouth. Fig 1(a)-(e) show the local features used in [9],[10],[19],[14] and [15] respectively. Little attention was paid on mouth corners.

C. Selected Local Features

Fig. 5 shows the first few selected local patches. Consistent with previous works (see Fig. 1), eye regions are the most important regions. However, different with previous work, we find that patches around eyes play more important role than the patches which cover the whole eyes. Intuitively, when the status of an eye alters from open to close or from close to open, its appearance changes greatly. However, the patches around it (esp. the eye corners) are more stable relative to the whole eye. So, the extracted features from those patches are more robust than those from the patch which covers the whole eye for face recognition.

As to shape feature, only tangent shape of the face contour is selected in the experiments by the AdaBoost algorithm. The interpretation might be that face contour, in frontal view, is more stable than the shape of the internal facial components such as the eyes and the mouth. Face contour can measure the aspect ratio of a face which is an intrinsic property of a face. In Fig. 6, the face contour of the left face image is quit distinct from the face contour of the right one. Compared to the left one, the contour of the right one is more narrower/thinner.
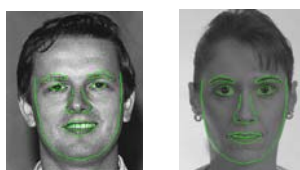

Fig. 6 Shape (face contour) contributes for recognition

D. Performance Comparison

We compare our method with two baseline methods, eigenface method [4] and Moghaddam's method [7]. For these two methods and the proposed method, their parameters are well adjusted so that they can achieve their best results. For eigenface method, the eigenfaces are computed from the training set with 495 persons rather than the gallery set with 700 persons. Fig. 7 shows their performances respectively. It

can be seen that the proposed method outperforms both eigenface method and Moghaddam's method.

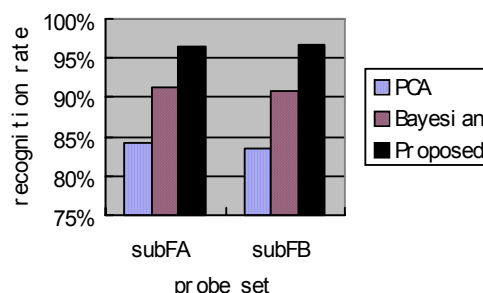Both eigenface method and Moghaddam's method are

Fig. 7 Performance comparison

global-based whose performances can be degraded by global illumination variations and expression changes. Making full use of semi-global features and local features of a face image, the proposed method can eliminate the unfavorable influence of such variations to some extent.

V. CONCLUSIONS

We designed three types of local features. They are semi-global features, local patches and tangent shape which are selected and combined by AdaBoost learning algorithm. Semi-global region is obtained from the saliency map which is derived from LDA. Improved recognition rates in our experiments confirm the following statement: 1) Both semi-global and local features are crucial for face recognition; 2) Mouth corners are more robust than whole mouth and contribute more for recognition; 3) Local patches around an eye are more discriminative than the patches which cover the whole eye; 4) For frontal view, it is helpful to use shape information for recognition. Face contour can measure the aspect ratio of a face which is an intrinsic property of a face. Experiment results on FERET database show that the proposed method is promising. In the future, more experiments will be conducted to test the generalization ability of the proposed features and other learning algorithms.

REFERENCES

[1] W. Zhao, R. Chellappa, A. Rosenfeld, J. Phillips, "Face recognition: A Literature Survey," ACM Computing Surveys, vol. 12, 399-458, 2003.
[2] B. Bruce, Recognizing Faces, London, Lawrence Erlbaum Associates, 1998.
[3] L. Sirovich, M. Kirby, "Low –dimensional Procedure for the Characterization of Human Face," Journal of the Optical Society of America, vol. 4, pp. 519-524, 1987.
[4] M. Turk, A. Pentland, "Face Recognition Using Eigenfaces," IEEE CVPR, 586-591, 1991.
[5] D.L Swets, J. Weng, "Using Discriminant Eigenfeatures for Image Retrieval," IEEE PAMI, vol. 18, pp. 831-836, 1996.
[6] N.B. Peter, P.H. Joao, J.K. David, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE PAMI, vol. 19, no. 7, 711-720, 1997.
[7] B. Moghaddam, B. Jebara, A. Pentland, "Bayesian Face Recognition," Pattern Recognition, vol. 33, 1771-1782, 2000.
[8] X. Wang, X. Tang, "Unified Subspace Analysis for Face Recognition," IEEE ICCV, pp. 679-686, 2003.

World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:2, No:9, 2008

[9] R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates", IEEE PAMI, vol. 15, no. 10, 1993, pp. 1042-1052.

[10] A. Pentland, B. Moghaddam, T. Starner, "View-Based and modular eigenspaces for face reocognition,", IEEE CVPR, pp. 84-91, 1994.

[11] F.S. Samaria, "Face Recognition Using Hidden Markov Models". PhD thesis, University of Cambridge, 1994.

[12] L. Wiskott, J. Fellous, N. Kruger, C. Malsburg, "Face Recognition by Elastic Bunch Graph Matching," IEEE PAMI, vol. 19, no. 7, pp. 775-779, 1997.

[13] P.S. Penev, "Local Feature Analysis: a General Statistical Theory for Object Representation," Network: Computation in Neural Systems, vol. 7, pp. 477-500, 1996.

[14] B. Heisele, T. Serre, M. Pontil, T. Vetter, T. Poggio, "Categorization by Learning and Combining Object Parts", Advances in Neural Information Processing Systems, vol. 2, pp. 1239-1245, 2002.

[15] S. Lucey, T. Chen, "A GMM Parts Based Face Representation for Improved Verification Through Relevance adaptation," IEEE CVPR, vol. 2, 855-861, 2004.

[16] R.E. Schapire, "A Brief Introduction to Boosting," Proc. ICAI, pp. 1401-1406, 1999.

[17] P. Viola, J. Michael, "Rapid Object Detection Using a Boosted Cascade of Simple Features," CVPR, pp. 511-518, 2001.

[18] J. Peng, D.R. Heisterkamp, H.K. Dai, "LDA/SVM Driven Nearest Neighbor Classification," IEEE Trans. on Neural Networks, vol. 14, no. 4, pp. 940-942, 2003.

[19] B. Guo, K.M. Lam, K.Ho Lin, "Human Face Recognition Based on Spatially Weighted Hausdorff Distance," Pattern   Recognition Letters, vol 24, no. 1-3, pp. 99–507, 2003.

[20] A.M. Martinez, A.C. Kak, "PCA versus LDA," IEEE PAMI, vol. 23, no. 2, pp. 228-233, 2001.

[21] Z. Hong, "Algebraic Feature Extraction of Image For Recognition," Pattern Recognition, vol. 24, pp. 211-219, 1991.

[22] Y. Tian, T. Tan, Y. Wang, Y. Fang, "Do Singular Values Contain Adequate Information for Face Recognition?" , Pattern recognition, vol. 36, pp. 649-655, 2003.

[23] A. Lanitis, C.J. Taylor, T.F. Cootes, "Automatic Face Identification System Using Flexible Appearance Models," Image and Vision Computing, vol. 13, pp. 393-401, 1995.

[24] Y. Zhou, L. Gu, H.J. Zhang, "Bayesian Tangent Shape Model," IEEE CVPR, vol. 1, pp. 109-116, 2003.

[25] P.J. Phillips, H. Moon, S. Rivzi, P. Rauses, "The FERET Evaluation Methodology for Face Recognition Algorithm," IEEE PAMI, vol. 22, pp. 1090-1104, 2000.

Yanwei Pang was born in 1976 in Hebei, China. He received his B.E. Degree from the Hebei Normal University, in 1998, M.E. Degree from the North China Institute of Technology, in 2001, and PhD Degree from the University of Science and Technology of China (USTC), in 2004, respectively, all in Electronic Engineering. He also spent ten months at Microsoft Research Asia as a visiting student. He is currently a postdoctor at USTC. His research interests include face recognition, manifold learning and computer vision.

Lei Zhang is a researcher in Microsoft Research Asia. He received his B.S. and M.S. degrees in Computer Science from Tsinghua University in 1993 and 1995 respectively. After two years of working in the industry, he returned to Tsinghua University and received his Ph.D. degree in Computer Science in 2001. Then he joined the media computing group in Microsoft Research Asia, and moved to the web search and mining group since 2004. His current research interests are content-based image retrieval and classification, image processing, face detection and annotation, computer vision, photo management and sharing. He is a member of IEEE and a member of ACM.

Zhengkai Liu received his B.E. Degree in Electronic Engineering from the University of Science and Technology of China, Hefei, China, in 1964. He is currently a Professor at the Department of Electronic Engineering and Information Science (EEIS), University of Science and Technology of China (USTC). He has been the Chairman of the Research Committee in the Department of EEIS of the USTC, and the head of the Multimedia Communication Lab and the Information Processing Center (USTC). His research interests include artificial neural networks, image processing and computer vision.