Bayesian Shape Localization for Face Recognition Using Global and Local Textures

Shuicheng Yan, Xiaofei He, Yuxiao Hu, HongJiang Zhang, *Senior Member, IEEE*, Mingjing Li, and Qiansheng Cheng

***Abstract*—We present a fully automatic system for face recognition in databases with only a small number of samples (even single sample) for each individual. In this paper, the shape localization problem is formulated in the Bayesian framework. In the learning stage, the RankBoost approach is introduced to model the likelihood of local features associated with the fiducial point, while preserving the prior ranking order between the ground truth position and its neighbors; in the inferring stage, a simple efficient iterative algorithm is proposed to uncover the MAP shape by locally modeling the likelihood distribution around each fiducial point. Based on the accurately located fiducial points, two popular mutual enhancing texture features are automatically extracted and integrated for human face representation: global texture features, which are the normalized shape-free gray-level values enclosed in the mean shape, and local texture features, which are represented by the Gabor wavelets extracted at the fiducial points (eye corners and mouth, etc.). Global texture mainly encodes the low-frequency information of a face, while local texture encodes the local high-frequency components. The extensive experiments illustrate that our proposed shape localization approach significantly improved the shape location accuracy, robustness, and face recognition rate; moreover, the experiments conducted on FERET and Yale databases show that our algorithm outperformed the classical eigenfaces, fisherfaces, as well as other approaches utilizing the shape, global, and local textures.**

***Index Terms*—Bayesian shape localization, face recognition, Gabor wavelet, RankBoost.**

I. Introduction

THE face recognition problem has attracted significant interests in the last decades[3], [4], [20], [22], [26], [30]. In this paper, we mainly focus on the task for fully automatic face recognition on a database with only a small number of samples (even single sample) for each individual. Here, "automatic" means that the face images for training or to be identified do not need to be manually labeled. Difficulties of this task mainly lie in the image variations due to the differences in head pose, position, size, facial expression, and illumination. A successful face recognition methodology depends heavily on the face representation and encoding method [17]; meanwhile, the convenient approach for automatic feature extraction is expected by a practical face recognition system [28], [32]. In this work, a novel Bayesian shape localization approach is proposed for automatic feature extraction; two types of robust and mutual enhancing features are integrated for human face representation, i.e., the global texture and local texture, which characterize the low- and high-frequency information of a face, respectively.

Recently, many new algorithms on face encoding and shape localization are proposed. Liu *et al.* [17] introduced a new face encoding and recognition method, Enhanced Fisher Classifier, in which the face recognition is conducted using the Enhanced Fisher Linear Discriminant Analysis on the feature vector integrating the shape and texture, where *texture* means the shape-free gray-level values enclosed in the mean shape. Heisele [11] introduced that the component-based approaches outperform the global approaches since the local components are more robust to head pose variations. Wiskott *et al.* [28] proposed an approach named Elastic Bunch Graph Matching. In this approach, a face is encoded in the form of image graph, i.e., a set of Gabor jets at the fiducial points. The extraction of image graph is guided by maximizing the objective function evaluating the similarity between the image graph and the face graph bunch (FGB), constructed from a small number of sample images graphs. However, as it used the jets in the FGB as local experts for similarity evaluation, the evaluation may easily affected by the noise in the FGB; moreover, the term evaluating the distorting of grid relative to the FGB has the assumption that all the edges in the image graph are independent of each other and have Gaussian distributions. This may not be the case in the real world. The active shape model (ASM), developed by Cootes and his colleagues [6], and its variations [10], [11], [24], [29] are powerful tools for the shape localization problem. However, there are some limitations in conventional ASM algorithms.

1) The local features extracted at the neighborhood points of a fiducial point are often similar to that extracted at the fiducial point, which may cause ambiguity; moreover, the most representative features are not always the best discriminative features. The traditional principal components analysis (PCA) is insufficient to present discriminative likelihood to differentiate the fiducial point from its neighbors.
2) The conventional ASM searches for the optimal shape without explicit objective function, thus it can not guarantee that the searched shape has monotonously increasing posterior probability in each step in the sense of Bayesian modeling. It provides neither a robust likelihood evaluation for the searched shape nor a stop criterion for the entire searching process in a principled manner.

The accurate localization of the shape, or so-called fiducial points, of a face are essential to the face recognition problem

Manuscript received November 18, 2002; revised May 12, 2003.

S. Yan and Q. Cheng are with the Department of Information Science, School of Mathematical Sciences, Peking University, Beijing 100871, P. R. China (e-mail: scyan@math.pku.edu.cn; qcheng@pku.edu.cn).

X. He is with the Department of Computer Science, University of Chicago, IL 60637 USA (e-mail: xiaofei@cs.uchicago.edu).

Y. Hu, H. Zhang, and M. Li are with the Microsoft Research Asia, Beijing 100080, P.R. China (e-mail: i-yxhu@microsoft.com; hjzhang@microsoft.com; mjli@microsoft.com).

Digital Object Identifier 10.1109/TCSVT.2003.818359

[15], [16], [25]. In this paper, new approaches are proposed to overcome the limitations of the conventional ASM approaches. First, a semi-supervised learning algorithm, *RankBoost*, is introduced to build the local likelihood model that ensures the ground truth position will more likely have a higher likelihood than its neighbors. It aims at providing more discriminative likelihood for the fiducial points and their neighbors. Second, we present a simple effective iterative algorithm for the optimization of the objective function by dynamically locally learning the likelihood distribution around each fiducial point using the locally weighted learning method. It guarantees that the objective function increases monotonously. These two approaches are both based on accurate probability formulation, which naturally leads to a robust likelihood measure for the searched shape. It can be used as a stopping condition for the inferring process.

In general, face images vary with the change of head position, size, expression, and environment illumination. A small number of samples can not reflect all these variations, thus face recognition based on the original raw features, i.e., gray-level values, will suffer from insufficient representation ability of the training samples. A natural way to overcome this problem is to use shape free features or robust local features at the points with explicit semantic. Therefore, we propose to use two types of features, i.e., global and local texture features, for face representation. The global texture is a normalized shape-free gray-level patch enclosed by the mean face shape warped from the original images. The global texture image has the same edges and contours and is insensitive to the shape variation due to the differences in head pose and facial expressions. The local textures around the fiducial points can be efficiently represented using the Gabor-wavelet features on various scales and orientations. They are DC-free and provide robustness against varying lightness in the image and head pose variations. In this work, the global texture and local Gabor jets are integrated for face representation for their mutual enhancing in discriminating power. Note that there are some previous works using the global texture [17] or local texture [28], [32] for face representation; however, most of they used only one of them or just combined one of them with the geometric shape information. Lanitis' work [15] is most related to the framework presented in this paper. It integrated the shape, global texture and local gray-level models for face recognition. There are two differences between them: 1) we present a more robust and accurate shape localization approach and 2) the work in [15] used only the profile perpendicular to the shape contour, with single orientation and single scale, which is less representative and robust than the Gabor jets used in our framework.

The face recognition problem is a canonical pattern recognition problem. A number of face recognition methods have been proposed. These methods can be divided into two types, namely, face-based and constituent-based. The face-based approaches, such as Eigenfaces and Fisherfaces [3], utilize the original raw face as the feature vector, while the constituent-based approaches extract the features based on the detected facial features (eyes, mouth, nose, and face contour), e.g., the ASM and AAM based algorithms. Our proposed face recognition algorithm is constituent-based, and the extensive comparative experiments presented in experiment section show that our algorithm outperforms the face-based algorithms, like Eigenfaces and Fisherfaces, and the other way utilizing global and local texture.

Here, we summarize the novel contributions of our work.

1) A general Bayesian shape localization approach is introduced for accurately localizing the fiducial points of a face. It makes use of the prior ranking order between the ground truth position of the fiducial point and its neighbors in the construction of the local likelihood model. In addition, a local model based iterative algorithm is presented to search for the optimal positions of the fiducial points. The experimental results show that it outperforms the conventional ASM significantly in position accuracy, robustness, stability and the recognition accuracy.
2) Two popular mutual enhancing texture features: i.e., global texture and local texture (Gabor jets) are integrated, which results in a robust face encoding method for face recognition.

The rest of this paper is organized as follows. Section II introduces our framework for face recognition. In Section III, the Bayesian shape localization approach is introduced in detail. Face recognition using global and local textures is discussed in Section IV. The evaluations and experiments are presented in Section V. Finally, the concluding remarks and future work are given in Section VI.

II. System Overview

In this paper, we mainly focus on the task for automatic face recognition on database with a small number of samples (even single sample) for each individual. There are three major problems for this task: 1) what kind of features used to encode the human face are insensitive to the image variations due to the differences in head pose, position, size, and illumination; 2) how to automatically extract the desired features from a face image without any user interaction; and 3) which kind of similarity measure is optimal to match the gallery and probe face image. In Sections II-A–C, we present our solutions to the above three questions and a unified framework is presented for face recognition.

A. Global and Local Textures for Face Representation

In real world applications, there are probably significant differences between the geometrical structures of the training samples and those of the probing faces for an individual due to the variations in face pose, position, and expression. A natural way to collapse these variances is to use shape-free features or local features extracted at the fiducial points with explicit semantic, like the eye and mouth corners, nose tip, and face contour. Consequently, we present a novel face encoding method by integrating the global and local textures, which characterize the low- and high-frequency components of a face, respectively.

Global texture is a normalized shape-free gray-level patch enclosed in the mean face shape. It is a shape normalized image transformed from the original face image. The warping operation uses the triangular meshes, as shown in Fig. 1, to build the correspondence between the original image and warped image.

Fig. 1. The triangular mesh constructed from the mean shape and four warped global texture samples from the FERET database.

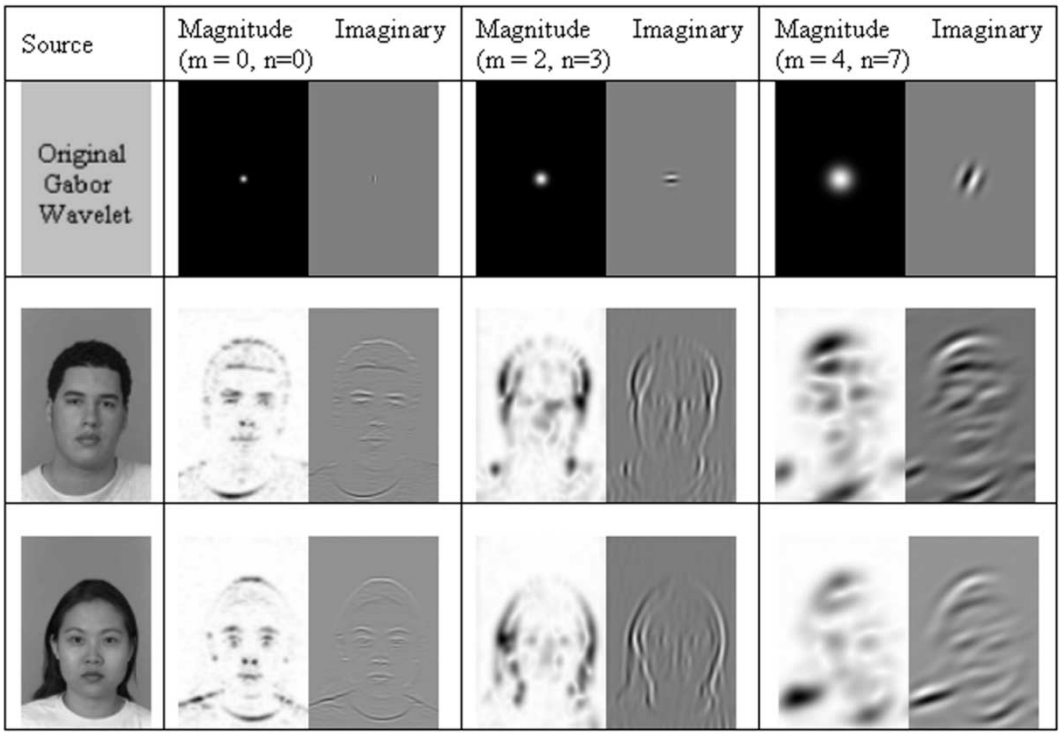

Fig. 2. Gabor wavelet representation: examples of three kernels from the images in the FERET database.

An affine transformation is performed for each triangular region. Some warped texture samples from FERET [23] database are listed in Fig. 1. After the warping operation, all the images have the same shape and contour, and the local details are weakened in most cases; thus the global texture mainly characterizes the low frequency components of the original face image.

The local texture is defined as a small patch of gray-level values in a face image I around a fiducial point $\vec{x} = (x, y)$. A typical representation of the local texture is the Gabor wavelet, i.e., Gabor jet [14], [19], [21], which is defined as a convolution based on the Gabor-wavelet transformation function

$$J_j(\vec{x}) = \int I(\vec{x}')\varphi_j(\vec{x}'-\vec{x})d^2x' \tag{1}$$

where Gabor kernels

$$\varphi_j(\vec{x}) = \frac{\vec{k}_j^2}{\delta^2}\exp\left(-\frac{\vec{k}_j^2\vec{x}^2}{2\delta^2}\right)\left[\exp(i\vec{k}_j\cdot\vec{x}) - \exp\left(-\frac{\delta^2}{2}\right)\right] \tag{2}$$

is the two-dimensional (2-D) plane wave with wavelet vector $\vec{k}_j$ restricted by a Gaussian envelope function with relative width δ. We set $\delta = \pi$ in all our experiments and use a discrete set of five different frequencies and eight orientations as follows:

$$\vec{k}_j = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_m \cos\varphi_n \\ k_m \sin\varphi_n \end{pmatrix},$$
$$k_m = \frac{0.5\pi}{(\sqrt{2})^m}, \quad \varphi_n = n\frac{\pi}{8} \tag{3}$$

where $m = 0, \ldots, 4$, $n = 0, \ldots, 7$, $j = 8m + n$. Thus, there are 40 features for each Gabor jet. Two examples with three of the Gabor kernels are showed in Fig. 2. In our study, only the magnitudes are used, since they are insensitive to the position while the phrases are very sensitive to position. The Gabor jet directly represents the local properties and mainly characterizes the high frequency information of a face.

The global texture and local texture characterize two different parts of the information in a face, and they can mutually enhance the discriminating power of each other. In our work, these two kinds of textures are integrated to encode a face.

B. Automatic Feature Extraction by Bayesian Shape Localization

A practical system should be able to automatically extract the features for face encoding, especially for a real time face recognition system. Elastic Bunch Graph Matching (EBG) and ASM are two popular algorithms for automatic localization of the fiducial points. They have been widely used in the tasks such as shape localization, expression analysis and face recognition. Compared with EBG, ASM has its advantages at building the global shape model and local likelihood/appearance models in a statistical framework. All the similarity measures efficiently make use of the prior distribution of the features. A typical ASM algorithm has two steps: single point localization and shape parameter estimation. Different methods are proposed to improve the performance of these two steps. However, there are some limitations in the ASM framework. First, the local likelihood model characterizes only the local features at the ground truth positions, whereas the feature that best *characterizes* the ground truth positions is not always able to best *distinguish* the ground truth position from its neighbors. Thus, the ground truth position may have a lower likelihood than those of its neighbors. This may happens even on the training samples. Second, there is no explicit objective function to be optimized for the fiducial point localization problem. Consequently, it provides neither a robust evaluation for the searched shape nor a stop criterion for the entire searching process in a principled manner.

In this paper, we formulate the shape localization problem in a Bayesian framework and propose a set of new approaches to overcome the limitations of the conventional ASM algorithms. First, a new *ranking prior likelihood model* is introduced to ensure that ground truth position will more likely have a higher likelihood than its neighbors via the *RankBoost* algorithm [5], [9]. It aims at providing correctly ordered likelihoods for the candidate points by efficiently combining the "weak" evaluation functions, which are built using different combinations of the principal components of the local features. Second, since the proposed objective function is complex and has no close form solution, we propose an *adaptive local likelihood distribution model* motivated by a lazy learning algorithm called locally weighted learning [1], [2], to optimize the objective function iteratively.

C. Face Similarity Measure Methods

As discussed previously, we mainly focus on the face recognition on the database with a small number of training samples, even a single face, for each person. The similarity can be evaluated in the original feature space and the dimensionality reduced space. In the original feature space, we systematically compared the performance by using L_1 distance, Euclidean distance, and cosine distance for similarity measure. Moreover, we evaluate the performance in the reduced representation subspace obtained by PCA. It is important to note that LDA and its variations are not suitable to the case that only a single face per individual is available. This is because the intra-class scatter matrix is a zero matrix in this case.

In Sections III–VI, we first introduce our proposed Bayesian shape localization algorithm in detail, and then the face recognition algorithm using global and local textures are presented.

III. Bayesian Shape Localization

The task of shape localization is to infer the optimal face shape with the maximal posterior probability from an image I, i.e.,

$$S^* = \arg\max_{s\in\mathbb{S}_s} p(S|I) \tag{4}$$

where $\mathbb{S}_s$ is the low-dimensional shape space [7], [8], [13] learned by principal component analysis (PCA) from the training shapes [6]. A shape $S = ((x_1, y_1), \ldots, (x_K, y_K)) \in \mathbb{R}^{2K}$ is a sequence of K fiducial points of a face. In (4), S is reconstructed from the shape parameter $s \in \mathbb{S}_s$ as: $S = Ts_c(\bar{S} + Us)$, where $\bar{S}$ is the average shape and U is composed of the first k leading eigenvectors; $Ts_c(\cdot)$ is the 2-D geometry transformation function based on four parameters: scale (r), rotation (θ), and translations T_xT_y.

From the Bayesian rule and the assumption that the local features from different fiducial points are independent to each other, the problem can be reformulated as

$$\begin{aligned} S^* &= \arg\max_{s\in\mathbb{S}_s} P(I|S)P(S) \\ &= \arg\max_{s\in\mathbb{S}_s} \prod_{i=1}^{K} P_i^{n_i}(I|S_i) \times P(S) \end{aligned} \tag{5}$$

where $P_i^{n_i}(I|S_i)$ is the likelihood of the ith fiducial point with n_i being the normal direction to the shape contour at S_i.

The local likelihood model plays an important role in the shape localization problem. It presents the fundamental likelihood evaluation for the candidate positions of fiducial points. In order to accurately locate the fiducial points, it is desirable that the local likelihood model should present higher likelihood for the ground truth position than its neighbors. However, the conventional local likelihood models in ASMs characterize only the local features at the ground truth positions and ignore the requirement for accurate shape localization. Thus the accuracy is affected by the incorrectly ranked likelihoods provided by these models in many cases.

Given an image, it is difficult to obtain the optimal shape of (5) because of the difficulties in exploring the global structure of $P_i^{n_i}(I|S_i)$. In a conventional ASM, the solution shape is approximated using a two-step iterative framework. First, each point is relocated to the new position with the maximal likelihood $P_i^{n_i}(I|S_i)$ in the neighborhood of the original position using the local likelihood models; then the relocated shape is constrained by the statistical shape space $\mathbb{S}_s$, which implicitly maximizes $P(S)$. Due to the lack of the constraints of the term $P(S)$, the first step is easily affected by noise, which undermines the accuracy and stability of the previous algorithms.

A. Ranking Prior Local Likelihood Model

It is often the case that, in fiducial point localization, the points around a ground truth position have similar local features especially in the low-resolution images. The ambiguity between the ground truth position and its neighbors demands that the

local likelihood model should be able to correctly rank the likelihoods of these ambiguous positions. A natural way is to take into account the ranking priors of these positions in the construction of the local likelihood model. Consequently, we propose a novel ranking prior local likelihood model to not only characterize the local features of a ground truth position as conventional ASM algorithms do, but also preserve the likelihood ranking order between the ground truth position and its neighbors.

Our proposed ranking prior local likelihood model is constructed using a newly emerged boosting algorithm *RankBoost* [5], [9]. Boosting is a method to produce highly accurate prediction rule by combining many "weak" rules, which may be only moderately accurate. *RankBoost* is a variation of the boosting algorithm. It aims at providing high accurate ranking evaluation and was originally used in the applications like web page ranking.

For the sake of simplicity, we shall omit the point sequence number in the deduction of the ranking prior local likelihood model. For a fiducial point, let $\chi = \{x$: local feature vector at ground truth position or its neighbors $\}$ be the sample space. The crucial pairs with prior ranking orders are presented in set $\Omega = \{(x, x') : R(x, x') > 0\}R(x, x')$ where $R(x, x')$ is the likelihood difference evaluation function; It is normalized to satisfy $\sum_{(x,x')\in\Omega} R(x, x') = 1$. In the shape localization problem, x and x' are the local features extracted at the ground truth position and its neighbors, respectively. Let $F = \{f_i, i = 1, \ldots, L\}$ denote the "weak" ranking evaluation function set, where L is the number of the functions; f_i satisfies $0 \leq f_i \leq 1$ for all $x \in \chi$. *RankBoost* is designed to find a likelihood evaluation function $H : \chi \to \mathbb{R}$ with "minimal" weighted number of the incorrectly ranked crucial pairs, i.e.,

$$H^* = \arg\min_{H\in\mathbb{S}_F} \sum_{(x,x')\in\Omega} R(x, x') \left[\tilde{H}(x, x')\right] \tag{6}$$

where $\tilde{H}(x, x') = H(x) - H(x')$; $\left[\tilde{H}(x, x')\right] = 1$ if $\tilde{H}(x, x') < 0$ which means that x and x' are misranked, else 0; $\mathbb{S}_F$ is the likelihood evaluation function space spanned by the functions in F. The whole process of *RankBoost* is referred to in [9]. In our local likelihood models, we set the bounds for the coefficient at each step for better stability and faster convergence.

As the projection to each principal component of the local features at the ground truth position presents fundamental ranking evaluation for all the samples, we construct the "weak" ranking evaluation function set using different combinations of these principal components. Denoting L' as the number of the principal components, there are $(2^{L'} - 1)$ kinds of different combinations

$$f_i(x) = \exp\left\{-\frac{1}{2}\sum_{k=1}^{l}\frac{\tilde{x}_{jk}^2}{\lambda_{jk}}\right\} \tag{7}$$

where $\tilde{x}_{jk}$ is the projection of x to the jkth principal component, λ_{jk} is the jkth largest eigenvalue, and $1 \leq j1 < j2 < c\cdots < jl \leq L'$. Since the ranking order of candidate pairs is only determined by the sign of their likelihood difference, we normalize $\oint H(x)dx = 1$.

By taking into account the prior information in the modeling process, the ranking prior local likelihood model is able to present more accurate likelihoods for the candidate positions than the traditional model does. We have conducted experiments to compare single point localization accuracy of the ranking prior model and the traditional model. It is observed that only 76% of the ground truth positions have higher likelihoods than their neighbors by using traditional local likelihood models, while the number is increased to over 89% by using our proposed ranking prior local likelihood models.

B. Locally Weighted Learning for Optimal Shape Inferring

As analyzed previously, it is difficult to directly undertake the optimization process for the complex global structure of distribution $P_i^{n_i}(I|S_i)$. In the control theory literature, locally weighted learning [1], [2] is a widely used lazy learning algorithm. It dynamically models the complex function using simple local models, with no necessity to find an appropriate structure for the global distribution. We extend this idea to locally model the complex function $P_i^{n_i}(I|S_i)$ using semi-Gauss functions. Consequently, the optimum of the shape localization problem in the neighborhood of the original shape can be obtained using these local models, namely adaptive local likelihood distribution models. In Sections III-B1 and 2, we will first introduce this new model and then present a new optimization approach for the shape localization problem.

1) Adaptive Local Likelihood Distribution Model: There is no closed-form solution for the shape localization problem. A natural way is to search for the solution in an iterative approach. Let S^k be the initial shape for the kth iteration, the task of each step is to find the optimal shape in the neighborhood of S^k using local optimization approach. Following the ideas of the locally weighted learning approach, $P_i^{n_i}(I|S_i)$ can be locally approximated around S^k using the likelihoods presented by the ith ranking prior local likelihood model. Note that n_i can be approximated by n_i^k in the neighborhood of S_i^k. This local model dynamically approximates the likelihood distribution around each fiducial point, so called *adaptive local likelihood distribution model*. Its construction process has two steps: 1) local neighbor selection and 2) local distribution model construction.

In the first step, the local neighbors of S_i^k are sampled according to the following distribution:

$$p(x, y) = \begin{cases} \frac{1}{\pi\theta^2}, & \text{when } \left\|(x, y) - S_i^k\right\| \leq \theta, \\ 0 & \text{else.} \end{cases} \tag{8}$$

where θ is the coefficient which determines the sampling range around the point S_i^k. Denoting $\{A_m^{(k,i)}\}$ as the sample set, the likelihood for each sample can be obtained as

$$Cf_i\left(A_m^{(k,i)}\right) = H_i\left(x\left(A_m^{(k,i)}, n_i^k\right)\right) \tag{9}$$

where $x(A_m^{(k,i)}, n_i^k)$ is the local features extracted at $A_m^{(k,i)}$ and $H_i(\cdot)$ is the learned ranking prior local likelihood evaluation function for the ith fiducial point.

In the second step, the likelihood distribution in the neighborhood of S_i^k is approximately modeled from the samples $\{A_m^{(k,i)}\}$ and their likelihoods $\{Cf_i(A_m^{(k,i)})\}$

$$p_{N(S_i^k)}^{n_i^k}(I|S_i) = C \cdot N\left(\mu_i^k, \sum_i^k\right) \tag{10}$$

where parameters C, μ_i^k, and $\sum_i^k$ can be easily learned by using least square method.

2) Shape Parameter Inferring: From (5), we can write the object function for shape localization problem as

$$F(S,I) = \prod_{i=1}^{K} P_i^{n_i}(I|S_i) \times P(S). \tag{11}$$

In the $(k+1)$th iteration, $F(S,1)$ can be optimized in the neighborhood of S^k in term of the adaptive local likelihood distribution model, then the object function is changed to

$$F_{N(S^k)}(S,I) = \prod_{i=1}^{K} P_{N(S_i^k)}^{n_i^k}(I|S_i) \times P(S). \tag{12}$$

Thus, the local optimum of the shape localization problem can be derived using the energy function

$$\begin{aligned} S^* &= \arg\min_{S\in\mathbb{S}_s} \sum_{i=1}^{K} En_{N(S_i^k)}^{n_i^k}(I;S_i) + En(S) \\ &= \arg\min_{S\in\mathbb{S}_s} \sum_{i=1}^{K} \left\|Ts_c(M_i^s) - \mu_i^k\right\|_{\sum_i^k}^2 + \sum_{j=1}^{K'} \frac{s_j^2}{\lambda_j} \end{aligned} \tag{13}$$

where $En(\cdot)$ is the corresponding energy function of the distribution $p(\cdot)$, namely $p(x) \propto \exp\{-En(x)\}$, K' is the dimension of $\mathbb{S}_s$, λ_j is the jth largest eigenvalue of the covariance matrix of the training shapes, s is the corresponding shape parameter, $M^s = \bar{S} + Us$, and Ts_c is the geometry transformation function based on the transformation parameter $c = (r, \theta, T_x, T_y)$

$$\begin{aligned} Ts_c(x,y) &= Tr\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \end{pmatrix} \\ &= \begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \end{pmatrix} \end{aligned} \tag{14}$$

where $a = r\cos\theta, b = r\sin\theta$.

In (13), the optimization function is multinomial and has no close form solution. As discussed in [12], the solution can be approximated iteratively using a two-step optimization method as following.

Transformation parameter estimation given s: Given the shape parameters s, M^s, and $En(S)$ are constant. In this case, we only need to minimize the following energy function:

$$En(c) = \sum_{i=1}^{K} \left\| \begin{pmatrix} a & -b \\ b & a \end{pmatrix} M_i^s + \begin{pmatrix} T_x \\ T_y \end{pmatrix} - \mu_i^k \right\|_{\sum_i^n}^2. \tag{15}$$

In order to obtain the optimal transformation parameter, we set the partial derives of $En(c)$ to zero. That is, the optimal parameters are obtained by solving the following linear functions:

$$\begin{aligned} \sum_{k=1}^{4}\left[\sum_{i=1}^{K}\left(\frac{\partial Ts_c(M_i^s)}{\partial c_l}\right)^T \sum_i^k \left(\frac{\partial Ts_c(M_i^s)}{\partial c_l}\right)\right] c_k \\ = \sum_{i=1}^{K}\left(\frac{\partial Ts_c(M_i^s)}{\partial c_l}\right)^T \sum_i^k \mu_i^k \quad l = 1,\ldots,4. \end{aligned} \tag{16}$$

Shape parameter estimation given c: Given the transformation parameter c, the energy function is changed to

$$\sum_{i=1}^{K} \left\| Tr(\bar{S}_i + U_i s) + \begin{pmatrix} T_x \\ T_y \end{pmatrix} - \mu_i^k \right\|_{\sum_i^n}^2 + \sum_{j=1}^{K'} \frac{s_j^2}{\lambda_j} \tag{17}$$

where U_i is a matrix consisting of the $(2i-1)$ and $2i$ row of U. Using the same approach as above, s can be obtained by solving the following linear functions:

$$\begin{aligned} \left[\sum_{i=1}^{K}(TrU_i)^T \sum_i^k (TrU_i) + \frac{1}{\lambda_j}\right] s \\ = \sum_{i=1}^{K}(TrU_i)^T \sum_i^k (\mu_i^k - Tr\bar{S}_i - T) \end{aligned} \tag{18}$$

where $T = (T_x, T_y)^T$.

For a given image, we start with the mean shape using proper transformation parameter c_0, i.e., $S^0 = Ts_{c_0}(\bar{S})$. The whole iterative searching process of the shape localization problem can be outlined as follows:

Set the iteration number $K = 0$.

Construct the adaptive local likelihood distribution model for each fiducial point using the points sampled from the distribution function (8) and their corresponding likelihoods provided by the learnt ranking prior local likelihood model.

The shape S^{k+1} with optimal parameter s and the transformation parameter c in the neighborhood of S^k are derived using the above iterative parameter inferring method.

If $F(S^{k+1},I) - F(S^k,I) \le \tilde{\theta}$ or $k \ge K_{\max}(\tilde{\theta})$ is the least decreased value that $F(S^k,1)$ must achieve in each iteration, $K_{\max}$ is the manually defined maximal iteration number), exit; otherwise, let $k = k+1$, go step 2.

IV. Face Recognition Using Global and Local Textures

Three types of features: shape, global, and local textures, can be extracted from a shape localization approach. In this work, we integrate the global and local textures for face encoding due to the following observations and analysis.

1) The shape is unstable for the head pose and facial expression variations; it may be mislocated due to the ambiguities especially for the contour points, while the local textures are more robust comparatively.

2) The global and local textures characterize the low and high frequency components of a face, respectively; and they can enhance each other in discriminating power.

With our proposed Bayesian face localization approach, the manually defined fiducial points in a face can be located in an iterative way. Let S_f denote the final searched fiducial points. The global texture is extracted by warping the gray levels enclosed by the shape S_f to the mean shape via the triangular mesh [6]. Let $x^{\text{glo}} \in \mathbb{R}^{N_1}$ denote the extracted normalized global texture, i.e., $\left\|x^{\text{glo}}\right\|_2 = 1$, where N_1 is the number of the fiducial points. The local Gabor jets are computed as (1) at the searched fiducial points of S_f. Let $x^{\text{loc}} \in \mathbb{R}^{K \times N_2}$ denote the normalized local Gabor features in a matrix form, *i.e.,* $\|x^{\text{loc}}\|_2 = 1$. Now, the human face can be represented as a high-dimensional vector, i.e., $x = [(x^{\text{glo}})^T, (x_1^{\text{loc}})^T, \ldots, (x_k^{\text{loc}})^T]^T$. The face recognition may be conducted in the original space characterized by the global and local textures and the dimensionality-reduced space. PCA and LDA are two of the most popular subspace learning algorithms for face recognition. In this paper, PCA is applied for dimensionality reduction, since LDA cannot work in the case that only a single sample is available for each individual, like our experiments on the FERET database.

A. Recognition in the Original Space

In the original space, we first conduct the down sampling at the global textures for robust representation. We systematically investigated three similarity measures.

1) The Cosine distance:

$$S_{\cos}(x,y) = \frac{1}{K}\sum_{i=1}^{K} \frac{\sum_{j=1}^{N_2} x_{i,j}^{loc} y_{i,j}^{loc}}{\sqrt{\sum_{j=1}^{N_2} (x_{i,j}^{loc})^2 \sum_{j=1}^{N_2} (y_{i,j}^{loc})^2}} + \lambda_1 \frac{\sum_{j=1}^{N_1} x_j^{glo} y_j^{glo}}{\sqrt{\sum_{j=1}^{N_1} (x_j^{glo})^2 \sum_{j=1}^{N_1} (y_j^{glo})^2}} \quad (19)$$

where λ_1 is a suitable constant. The first term, defined as in [28], measures the similarity of the local Gabor wavelet features and the second term measures the similarity of the global textures.

2) The distance:

$$S_{L_1}(x,y) = \frac{1}{K \times N_2}\sum_{i=1}^{K}\sum_{j=1}^{N_2} |x_{i,j}^{loc} - y_{i,j}^{loc}| + \lambda_2 \frac{1}{N_1}\sum_{j=1}^{N_1} |x_j^{glo} - y_j^{glo}| \quad (20)$$

where λ_2 is a suitable constant.

3) The L_1/Euclidean distance:

$$S_{L_2}(x,y) = \frac{1}{K \times N_2}\sum_{i=1}^{K}\sum_{j=1}^{N_2} ||x_{i,j}^{loc} - y_{i,j}^{loc}||_2 + \lambda_3 \frac{1}{N_1}\sum_{j=1}^{N_1} ||x_j^{glo} - y_j^{glo}||_2 \quad (21)$$

where λ_3 is a suitable constant.

Our experiments show that the recognition using L_1 distance outperforms that by the other two methods.

B. Recognition in the Dimensionality-Reduced Space

Sometimes it may be beneficial to perform dimensionality reduction before we utilized any classification techniques. High dimensionality creates several problems for face recognition. First, learning from examples is computationally infeasible if it has to rely on high-dimensional representations. The reason for this is known asthe *curse of dimensionality*: the number of examples necessary for reliable generalization grows exponentially with the number of dimensions. Learnability thus necessitates dimensionality reduction. Second, for storage and efficiency concern, dimensionality reduction is needed.

PCA is optimal as to reducing redundancy. The leading components obtained by PCA maximize the data variance in those directions, and hence capture the most informative features of the data set. Let P denote the projection matrix whose column vectors are the leading eigenvectors of the data covariance. Thus, the face image can be represented in the reduced subspace as follows:

$$\tilde{x} = Px. \quad (22)$$

In the dimensionality-reduced subspace, we investigated the face recognition performance by using different similarity measures, i.e., cosine distance, L_1 distances, and Euclidean distance.

V. Experiments

In this section, two types of experiments were systematically conducted. We first conducted experiments to evaluate the accuracy, robustness and stability of our proposed Bayesian shape localization algorithm; then three types of face recognition experiments on FERET and Yale databases were presented: 1) the performance of our algorithm using different similarity measure; 2) comparison with the classical Eigenfaces and fisherfaces on the Yale database; and 3) comparison of the different approach utilizing the shape, global and local texture.

A. Bayesian Shape Localization

The experiments have been conducted on a data set consisting of 500 frontal face images, in which each face area contains about $150 * 150$ pixels and many faces have ambiguous fiducial points such as those shown in Fig. 6. All faces were manually labeled with 83 fiducial points. Four hundred of them were randomly selected for model construction and the remaining 100 for testing.

For comparison, ASM and our proposed Bayesian shape localization approach, referred as RPBF (**R**anking **P**rior likelihoods for **B**ayesian shape localization **F**ramework) were trained on the same data set, in a four-level image pyramid (Resolution is reduced 1/2 level by level). All results were obtained by searching maximally five times per layer (exit each layer as described in Section III-B2). RPBF is a fast algorithm. It costs only 80 ms per iteration (on P4 1.8G computer with 512M memory) although it is slower than the classical ASM for the consumption at locally learning of the local likelihood distribution. It takes about eight iterations to converge in average. As we discuss in

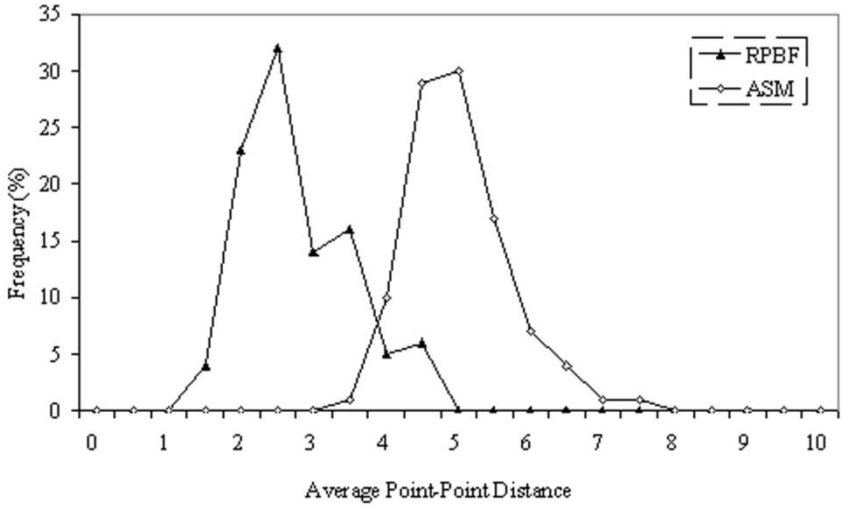

Fig. 3. Shape localization accuracy comparison: ASM versus RPBF. Note that most results of RPBF have smaller point–point distances than those of ASM.

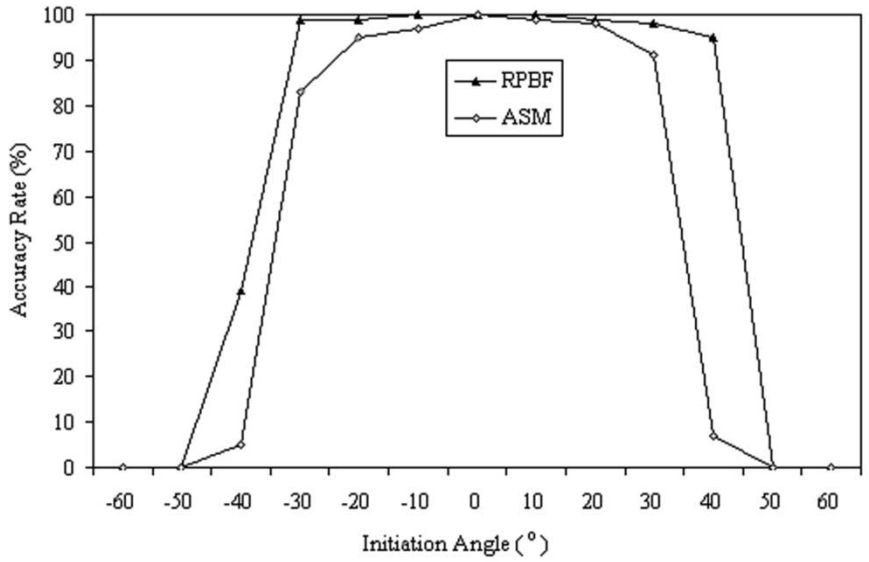

Fig. 4. The capture range comparison of in-plane rotation angle: ASM versus RPBF.

the following, it greatly outperforms ASM in accuracy, stability, and robustness.

The most commonly used criterion to evaluate the searched shape is the average point–point or point–curve distance between the searched shape and the ground truth. In all our experiments, the results were evaluated using the average point–point distance.

1) Accuracy and Robustness of Shape Localization: The statistics of the average point–point distance between the searched shape and the manually labeled shape of ASM and RPBF are presented in Fig. 3. The vertical axis represents the distribution of point–point distances. It shows that most results of RPBF have smaller point–point distances than those of ASM.

The capture range of in-plane rotation angle is an important criterion to evaluate the shape localization algorithms. The results illustrated in Fig. 4 demonstrate that RPBF can capture larger percentage of cases in range of the in-plane rotation $[-40°, 40°]$.

2) Algorithmic Stability: The algorithm stability was measured by the standard deviation of the results from different initializations. The results of RPBF and ASM are compared in

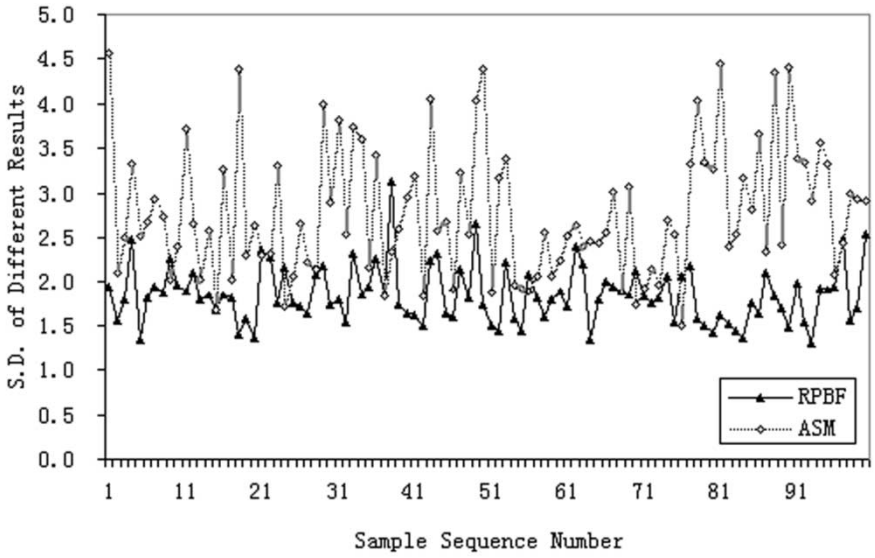

Fig. 5. Standard deviation of results derived from different initializations for each example compared between ASM and RPBF.

Fig. 5, where the vertical axis describes the average standard deviation of the results obtained from nine different initializations. The result shows that RPBF is more stable for initializations compared to the conventional ASM.

The comparison between RPBF and ASM is presented in Fig. 6. As can be seen, ASM failed to accurately locate the contour points due to the ambiguity caused by the furrow on the face, while RPBF accurately located the contour. Fig. 7 shows some results of the images from the FERET database and they were directly used for global and local textures extraction in the following face recognition experiments.

(a) (b)

Fig. 6. (a) A case ASM fails for the ambiguous furrow (b) PRSSL performs well.

B. Face Recognition Experiments

In this subsection, we tested our face algorithm using global and local textures based on RPBF on the FERET and YALE databases. The algorithm performance was evaluated based on the automatically extracted global and local textures using our proposed automatic Bayesian shape localization approach. We systematically compared the performance of different combinations of the features; and we evaluated the performances of our algorithm using the following similarity measures: L_1 distance, Euclidean distance and cosine distance; moreover, experiments comparing our proposed algorithm with the classical Eigenfaces and Fisherfaces on the Yale database were also conducted.

1) Performance on the Feret Database: The FERET database was provided by the U.S. Army Research Laboratory. We used 450 face images corresponding to 150 individuals such that each person has three images in gallery "*ba*" (frontal face), "*be*" (face with pose $+15^{\circ}$) and "*bf*" (face with pose -15°) of size $256 * 384$ with 256 gray-level values. We systematically investigated the performance by using one image per person for training. The fiducial points were automatically located using our proposed Bayesian shape localization approach as described in Fig. 7. The results from using different way utilizing the shape, global and local texture based on the similarity measure criterion S_{L_1} at the original feature space are listed in Table I. It is observed that the recognition rate obtained by using integrated global and local textures is always higher than that obtained by using only one of the features; meanwhile, it also shows that our proposed Bayesian shape localization algorithm improves the recognition rate comparing to the classical ASM algorithm due to its better accuracy at shape localization. Meanwhile, it shows that mostly the face recognition performance using integrated shape, global texture, and local gray levels as [15] (in the experiments, we used the same weights for the three types of parameters) is worse than that using local texture and is always worse than that using integrated global and local texture (Gabor jets).

To evaluate the performance of our algorithm with different similarity measures, we conducted the experiments using gallery "*bf*" as training samples and gallery "*be*" as testing samples. These experiments ware performed in the dimensionality reduced subspace by PCA. Fig. 8 shows that face recognition using L_1 distance outperforms that using the

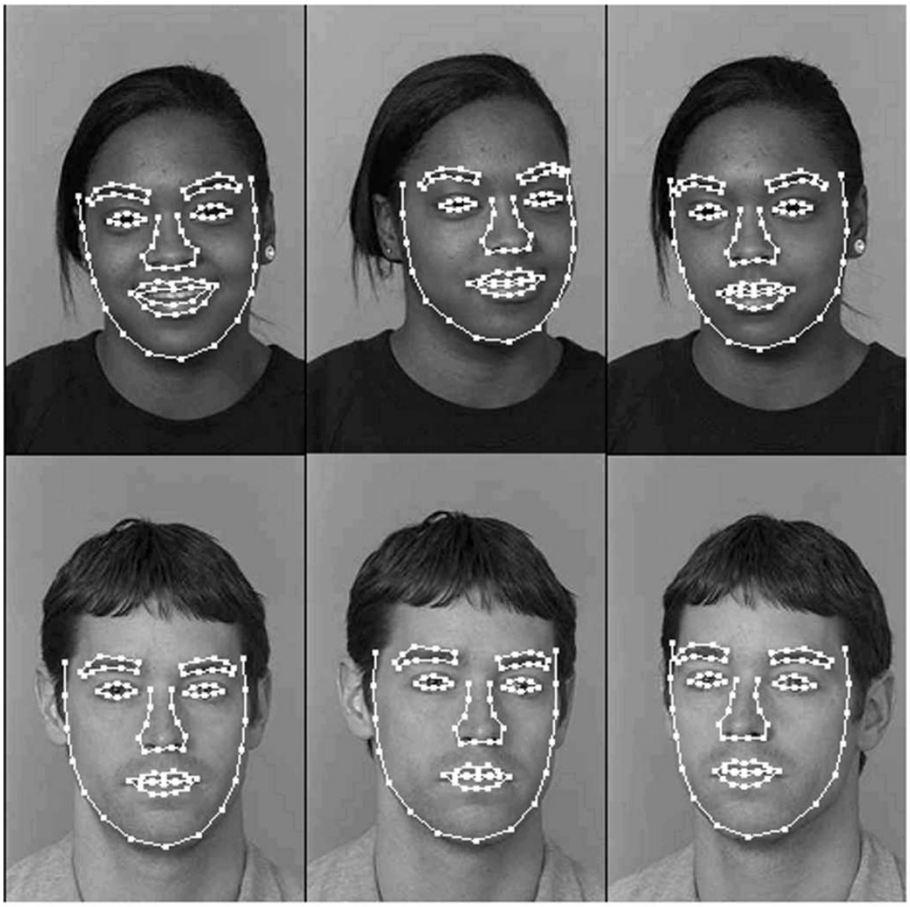

Fig. 7. Some shape localization results of RPBF from the images of the FERET database.

TABLE I
RECOGNITION RATES IN THE ORIGINAL FEATURE SPACE USING L_1 DISTANCE ON THE FERET DATABASE WITH ONLY ONE SAMPLE FOR TRAINING PER PERSON. NOTE THAT THE FIRST FIVE EXPERIMENTS WERE CONDUCTED BASED ON THE SHAPE LOCALIZATION RESULTS FROM RPBF; AND THE LAST ONE WAS CONDUCTED BASED ON THE SHAPE LOCALIZATION RESULTS FROM ASM AND USED THE GLOBAL AND LOCAL TEXTURES

Training Gallery	Testing Gallery	Results (Global)	Results (Local)	Results Ref [15]	Results (G+L)	ASM Results (G+L)
ba	*be*	91.33%	98.00%	92.67%	99.33%	94.67%
ba	*bf*	95.33%	95.33%	95.33%	98.00%	93.33%
be	*ba*	92.67%	97.33%	95.33%	98.67%	93.33%
be	*bf*	71.33%	93.33%	89.33%	94.67%	90.00%
bf	*ba*	92.00%	94.00%	93.33%	94.67%	89.33%
bf	*be*	70.00%	93.33%	94.00%	96.00%	91.33%

other two similarity measures; and face recognition using the Euclidean distance performs worst.

2) Performance on the Yale Database: The Yale face database [27] was constructed at the Yale Center for Computational Vision and Control. It contains 165 grayscale images of 15 individuals. These images demonstrate variations in lighting condition (left-light, center-light, right-light), facial expression (normal, happy, sad, sleepy, surprised, and wink), and with/without glasses. We systematically compared the performance of four methods using the shape, global and local textures. All these experiments were conducted in the original feature space using Cosine distance for similarity measure and we tested their performances when using different number of training samples. As can be seen in Fig. 9, the recognition method integrating the global and local textures outperforms the other three methods in all the experiments with different number of training samples; the recognition method using global textures and shape outperforms that using only global textures, but is worse than that using local textures and much worse than our algorithm.

Eigenfaces and Fisherfaces are both face-based algorithms. The face-based algorithm directly utilizes the gray-level values of the original raw image which are sensitive to changes of head pose and expressions. Meanwhile, the Fisherfaces algorithm can only produce a small number of discriminating features, which limits its recognition performance. Table II system-

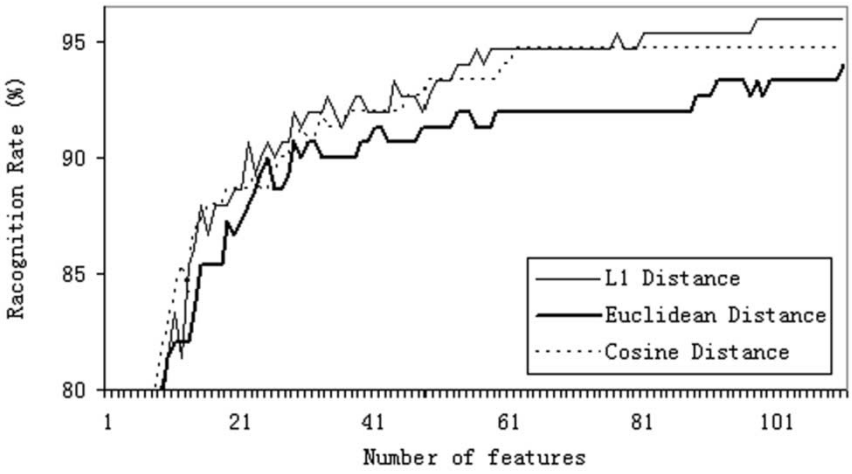

Fig. 8. Face recognition rate versus feature number in the dimensionality-reduced space by PCA using L_1, Euclidean and Cosine distance for similarity measure on the FERET database with gallery "*bf*" for training and "*be*" for testing.

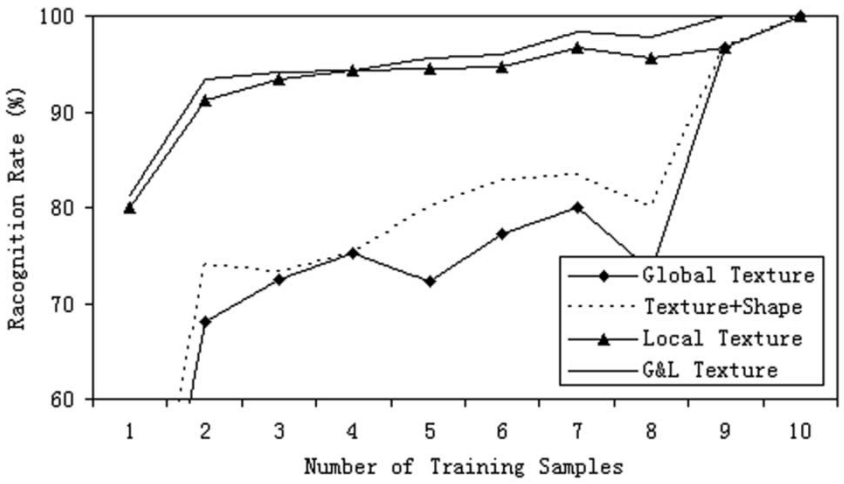

Fig. 9. The recognition rate versus the number of the training samples in original feature space using the cosine distance on the Yale database.

TABLE II
RECOGNITION ACCURACY COMPARISON ON THE YALE DATABASE USING SIX SAMPLES FOR TRAINING AND THE OTHER FIVE FOR TESTING EACH INDIVIDUAL: EIGENFACES, FISHERFACES, FACE RECOGNITION USING GLOBAL TEXTURE, LOCAL TEXTURE, AND INTEGRATED GLOBAL AND LOCAL TEXTURES. NOTE THAT THE BEST RESULT FOR EIGENFACES IS ACHIEVED IN THE SUBSPACE WITH DIMENSION 33

Eigenfaces/Dim	Fisherfaces	Global Texture	Local Texture	G&L Textures
74.67%/33	81.33%	77.33%	94.66%	96.00%

atically compares the following face recognition approaches: Eigenfaces, Fisherfaces, and three face recognition algorithms utilizing global and local textures on the Yale database. For each individual, six images were randomly selected for training, and the remaining five for testing. We used cosine distance for similarity measure for all these experiments. The results show that face recognition in the original feature space using the warped global texture outperforms Eigenfaces approach; and Fisherfaces outperforms Eigenfaces; moreover, our algorithm using integrated global and local textures performs better than all the other four algorithms.

VI. DISCUSSIONS AND FUTURE WORK

In this paper, a unified framework is introduced for fully automatic face recognition. A novel Bayesian shape localization approach is proposed for automatic fiducial point localization and the automatically extracted global and local textures are integrated for face encoding.

The shape localization problem is formulated in a Bayesian framework and two novel methods are proposed for highly confident inference of the optimal shape. The likelihood of the local features associated with the fiducial point is modeled using the RankBoost method which is introduced to ensure that the ground truth position has higher likelihood than its neighbors. This model is learned in a semi-supervised manner and presents discriminative likelihood output for the local features. On the other hand, the optimal shape is inferred in an iterative method that locally models the likelihood distribution around each fiducial point via the semi Locally Weighted Learning method and simplifies the task into a general multinomial optimization problem in each step.

The global and local textures (Gabor jets) characterize the low- and high-frequency information of a face, respectively. They are both robust to head position, which results in a robust face representation for face recognition, especially for the database with only a small number of samples per individual. The experiments have shown that our algorithm has better accuracy, robustness, and stability in shape localization and outperforms the classical face recognition approaches, such as Eigenfaces and Fisherfaces. Moreover, we systematically investigated the performance of face recognition using different combinations of the features and using different types of similarity measures.

In this paper, our proposed framework has shown good performance for near frontal face recognition. Our framework can be easily extended to automatic multiview shape localization by using three-dimensional face modeling and face recognition may be conducted using stronger classifier like using Fisher discriminant analysis. We are currently exploring these extensions in theory and practice.

REFERENCES

[1] C. G. Atkeson, A. W. Moore, and S. Schaal, "Locally weighted learning," *Artif. Intell. Rev.*, vol. 11, pp. 11–73, 1997.

[2] ——, "Locally weighted learning for control," *Artif. Intell. Rev.*, vol. 11, pp. 75–113, 1997.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 711–720, July 1997.

[4] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1042–1052, Oct. 1993.

[5] W. W. Cohen, R. E. Schapire, and Y. Singer, "Learning to order things," *J. AIR*, vol. 10, pp. 243–270, 1999.

[6] T. F. Cootes and C. J. Taylor. (2001) Tech. Rep.: Statistical Models of Appearance for Computer Vision. [Online]. Available: http://www.isbe.man.ac.uk~bim/refs.html

[7] I. L. Dryden, "General shape and registration analysis," in *Stochastic Geometry: Likelihood and Computation*, 1997.

[8] ——, "Statistical shape analysis in high-level vision," presented at the IMA Workshop on Image Analysis and High Level Vision Modeling, 2001.

[9] Y. Freund, R. Iyer, R. E. Schapire, and Y. singer, "An efficient boosting algorithm for combining preference," in *Proc. 15th Int. Conf. Machine Learning*, 1998.

[10] B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever, "A nonlinear gray-level appearance model improves active shape model segmentation," in *Proc. IEEE Workshop on Mathematical Models in Biomedical Image Analysis*, 2001, pp. 205–212.

[11] B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: Global versus component-based approach," in *Proc. IEEE Int. Conf. Computer Vision*, Vancouver, BC, Canada, July 2001.

[12] A. Hill, T. F. Cootes, and C. J. Taylor, "Active shape models and the shape approximation problem," *Image Vis. Comput.*, vol. 14, no. 8, pp. 601–608, Aug. 1996.

[13] J. T. Kent and K. V. M. Shape, "Procrustes tangent projections and bilateral symmetry," *Biometrica*, 2001.

[14] V. Krüger and G. Sommer, *Proc. Gabor Wavelet Networks for Object Representation. DAGM-Symp*, 2000, pp. 309–316.

[15] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 743–756, July 1997.

[16] Y. Li, S. Gong, and H. Liddell, "Modeling faces dynamically across views and over time," in *Proc. IEEE Int. Conf. Computer Vision*, Vancouver, BC, Canada, July 2001.

[17] C. Liu and H. Wechsler, "A shape and texture based enhanced fisher classifier for face recognition," *IEEE Trans. Image Processing*, vol. 10, pp. 598–608, Apr. 2001.

[18] Q. Liu, R. Huang, H. Lu, and S. Ma, "Face recognition using kernel based fisher discriminant analysis," in *Proc. 5th Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, May 2002.

[19] M. J. Lyons, J. BudyNek, A. Plante, and S. Akamatsu, "Classifying facial attributes using a 2-D gabor wavelet representation and discriminant analysis," in *Proc. IEEE Conf. AUTOMATIC and Gesture Recogniton, FG'2K*, Grenoble, France, Apr. 2000, pp. 202–207.

[20] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 7, pp. 696–710, July 1997.

[21] K. Okada and M. J. Lyons, "On Gabor Wavelet-Based Image Processing for Nissl-Stained Rat Brain Slices," Comput. Sci. Dept., Univ. of Southern California, Tech. Rep. 02-756, 2002.

[22] P. J. Phillips, "Support vector machines applied to face recognition," *Adv. Neural Inform. Proces. Syst.*, vol. 11, pp. 803–809, 1998.

[23] I. J. Philips, H. Wechsler, J. Huang, and P. Rauss, "The FERET dabase and evaluation procedure for face recognition algorithms," *Image Vis. Comput.*, vol. 16, pp. 295–306, 1998.

[24] M. Rogers and J. Graham, "Roboust active shape model search," in *Proc. 7th Eur. Conf. Computer Vision*, vol. 4, Copenhagen, Denmark, May 2002, pp. 517–530.

[25] S. Romdhani, V. Blanz, and T. Vetter, "Face identification by matching a 3D morphable model using linear shape and texture error functions," in *Proc. 7th Eur. Conf. Computer Vision*, vol. 4, 2002, pp. 3–19.

[26] M. Turk and Pentalend, "Eigenfaces for recognition," *J. Cognitive Neurosci.*, Mar. 1991.

[27] (2002) Yale Univ. Face Database. [Online]. Available: http://cvc.Yale.edu/projects/Yalefaces/Yalefaces.html

[28] L. Wiskott, J. M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 775–779, July 1997.

[29] S. C. Yan, C. Liu, S. Z. Li, L. Zhu, Z. H. J. Zhang, H. Shum, and Q. S. Cheng, "Texture-constrained active shape models," in *Proc. 1st Int. Workshop Generative-Model-Based Vision*, Copenhagen, Denmark, May 2002.

[30] M.-H. Yang, "Kernel eigenfaces vs. Kernel fisherfaces: Face recognition using kernel methods," in *Proc. 5th Int. Conf. Automatic Face and Gesture Recognition*, Washington, DC, May 2002.

[31] J. Yang, Y. Yu, and W. Kunz, "An efficient LDA algorithm for face recognition," in *Proc. 6th Int. Conf. Control, Automation, Robotics and Vision*, Singapore, 2000.

[32] Z. Zhang, M. Lyons, M. Schuster, and S. Akamatsu, "Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron," in *Proc. 3rd Int. Conf. Automatic Face and Gesture Recognition (FG '98)*, 1998, pp. 454–459.

[33] W. Zhao, R. Chellappa, and P. J. Phillips, "Subspace Linear Discriminant Analysis for Face Recognition," Center for Automation Research, University of Maryland, Tech. Rep. CAR-TR-914, 1999.