社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2006－ＡＶＭ－55 (10)
2006／12／14

Facial Feature Land-marking with Optimized Gabor Parameters based on Maximization of Separation between Features

Fan CHEN[1] Kazunori KOTANI[1]

School of Information Science,
Japan Advanced Institute of Science and Technology
Nomi, Ishikawa, 923-1292, Japan[1]

Abstract

Elastic Bunch Graph Matching(EBGM) is a popular method in automatic localization of facial feature points, where the selection of optimal Gabor parameters plays a key role in the extraction of Gabor jets with a high degree of discrimination. We propose a method for the selection of parameters by minimizing the energy function consisting of within-class and between-class scatters using the gradient descendent method. We formulate the learning rule and design the algorithm for the learning of parameters. Numerical experiments have been performed to investigate the performance of estimated parameters in improving the precision of automatic localization.

1 Introduction

Localization of automatically selected or prespecified fiducial points plays a key role in both appearance based approaches and model based ones as prior pose normalization and feature extraction, respectively.[1] In general, the performance of facial feature localization is evaluated from the following aspects: accuracy, robustness against illumination, pose and scale variations, and computation expense, etc. EBGM in Ref.[2] provides a robust method of localization against the variations of brightness by using a bank of Gabor responses (Gabor jet) as a discriminant feature, where shape variations are modeled by an elastic graph. To further improve the accuracy, statistical approaches, such as SVM, have been applied to EBGM to explore the discrimination of extracted graphs.[3] Elastic graph matching with morphology features has also been proposed. [4] We still adopt the Gabor filters, a biologically relevant model for the receptive field[5,6], for robust feature extraction against illumination and varying contrast by making the Gabor kernel DC-free and normalizing the Gabor jet. We consider improving the robustness against scaling by extracting and scaling the face region to a specified face size. The accuracy of EGBM in land-marking is increasable by improving the discriminant of jets among fiducial points nearby, which is implemented by the optimization of Gabor parameters in the present work. The robustness against distortion and rotation is also improvable by learning the Gabor parameters based on training samples.

Various approaches have been proposed for the selection of Gabor parameters, e.g., post-selection of Gabor response according to the output of SVM classifiers [7], Neural Network (NN)-based selection of parameters for texture segmentation [8] and information diagram based search [9], etc. The selection by NN introduces extra hyper-parameters for the network, whose convergence is difficult to control when dealing with a large number of high-dimensional Gabor responses. Due to the large size of the Gabor family used in landmarking, the information diagram-based search in the candidate space of parameters is highly computationally expensive. We perform the optimization of Gabor parameters by the minimization of an energy function which maximizes the discrimination of Gabor jets.

In Section 2, we give a full view of the whole localization system, and formulate the energy-minimization based approach for the optimization of Gabor parameters. We also give the learning algorithm. Numerical experiments have been performed in Section 3 to investigate the performance under optimized Gabor parameters. Discussions on the advantages and shortcomings of the proposed approach will also be given. Finally, we conclude the present work and explain our future work.

2 Feature Localization by EBGM with Optimized Gabor Parameters

A two-stage localization, i.e., one stage for coarse localization and one for fine localization, is used in our system to improve the computational speed. The Adaboost face detector, proposed by Viola and Jones [10], provides a fast and accurate detection of facial region with reasonable robustness to both illumination and scale variations in the coarse localization stage, whose detailed explanation can be found in Ref.[10]. A default graph is then applied to the extracted face image as the initial positioning for the following finer localization. The refinement of the positions for those fiducial points is executed by EBGM, the major topic of the present paper, starting with an initial point which has

been put in a neighboring region of the ground truth after the first stage.

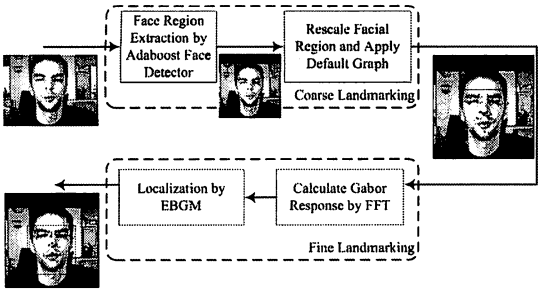


Fig. 1 Block diagram of the land-marking system used in the present work.

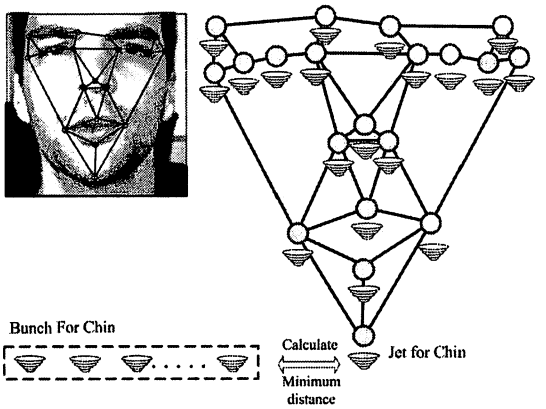


Fig. 2 Schematic diagram of EBGM land-marking.

2.1 Fine Land-marking by EBGM

A schematic diagram for the idea of EBGM is given in Fig.2. The structural information of selected fiducial points is described by a graph. The points are processed sequentially in a descending order of their reliability, which is decided subjectively or by experiments, e.g., the extraction of an eye might be more reliable than that of the chin tip. After the coarse localization, most of the points have been located within the neighborhood of the ground truth. For each point, an initial position is guessed from the interpolation of others that have already been estimated. A discriminant feature is calculated from that position and compared to a template bunch to measure the similarity. A search, full or selective, is then executed inside a specified range for the detection of the optimal position with highest overall similarity. A flow chart is given in Fig.3 for EBGM land- marking.

Discriminant features are extracted by a family of K Gabor kernels in Ref.[2]. In a common form, the k-th

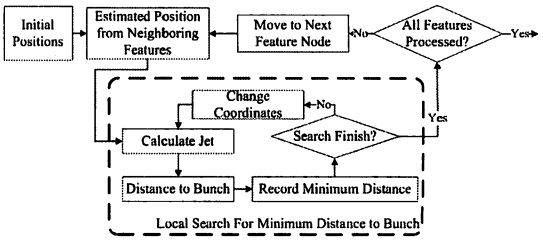
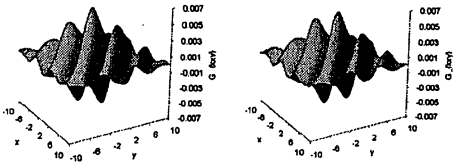


Fig. 3 Flow chart for EBGM land-marking.

Gabor kernel is defined as

$$G_{kxy} = \frac{\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_{kx}^2}+\frac{y^2}{\sigma_{ky}^2}\right)\right]}{2\pi\sigma_{kx}\sigma_{ky}}\left\{\exp\left[i\left(\omega_{kx}x+\omega_{ky}y\right)\right]\right.$$
$$\left.-\exp\left[-\frac{\left(\omega_{kx}\sigma_{kx}\right)^2+\left(\omega_{ky}\sigma_{ky}\right)^2}{2}\right]\right\}, \quad (1)$$

which is a complex sinusoid centered at frequency $(\omega_{kx}, \omega_{ky})$ and modulated by a Gaussian envelope, as shown in Fig.4. σ_{kx} and σ_{ky} are the standard deviations of the elliptical Gaussian along x and y. The second term inside the bracket makes the Gabor kernel DC-free, so that robustness against brightness variation can be improved.



(a.) Real part (b.) Imaginary part

Fig. 4 A Gabor kernel.

In Ref.[2], $\sigma_{kx} = \sigma_{ky}$ is adopted to simplify the Gabor kernel. The jet J for one point is defined as a K-dimensional vector $[J_k|k \in \{1, \cdots, K\}]$ formed by responses of Gabor filters, where $J_k = \alpha_k \exp(i\phi_k)$, with α_k as the magnitude and ϕ_k as the phase information. Wiskott et al. [2] further defined similarity between two jets using their phase information, i.e.,

$$S(J, J') = \frac{\sum_k \alpha_k \alpha_k' \cos(\phi_k - \phi_k' - \vec{d}\vec{\omega}_k)}{\sqrt{\sum_k \alpha_k^2}\sqrt{\sum_k (\alpha_k')^2}}, \quad (2)$$

where $\vec{\omega}_k = [\omega_{kx} \quad \omega_{ky}]$ and $\vec{d}\vec{\omega}_k$ compensates for the phase rotation. $\vec{d}$ stands for the displacement of two points. Further exploration of discriminant from the graph by classifiers has been done to improve the localization accuracy[3], utilization of discriminant features other than Gabor was also proposed.[4] We consider the derivation of optimal Gabor parameters to improve the accuracy of localization in the present work.

Tabel 1 Matrices for calculating the differential

$\mathbf{B}_\omega : K \times K,$	$(\mathbf{B}_\omega)_{i,j} = \sum_k \delta_{i,2k}\delta_{i+1,j} - \delta_{i,2k+1}\delta_{i-1,j};$
$\mathbf{C}_\bullet : XY \times XY,$	$(\mathbf{C}_x)_{i,j} = \delta_{i,j} i_x, (\mathbf{C}_y)_{i,j} = \delta_{i,j} i_y;$
$\mathbf{B}_{\sigma_\bullet} : K \times K,$	$(\mathbf{B}_{\sigma_\bullet})_{i,j} = \sum_k \delta_{i,j}\delta_{\lfloor i/2 \rfloor,k}\sigma_{k\bullet}^{-3};$
$\mathbf{D}_{\sigma_\bullet} : K \times K,$	$(\mathbf{D}_{\sigma_\bullet})_{i,j} = \sum_k \delta_{i,j}\delta_{\lfloor i/2 \rfloor,k}\sigma_{k\bullet}^{-1};$
$\mathbf{C}_{\sigma_\bullet} : XY \times XY,$	$(\mathbf{C}_{\sigma_x})_{i,j} = \delta_{i,j} i_x^2, (\mathbf{C}_{\sigma_y})_{i,j} = \delta_{i,j} i_y^2;$
$\mathbf{E}_{\omega_\bullet} : XY \times XY,$	$(\mathbf{E}_{\omega_x})_{i,j} = \sum_k \delta_{i,j}\delta_{\lfloor i/2 \rfloor,k}\omega_{kx}\sigma_{kx}^2,$
	$(\mathbf{E}_{\omega_y})_{i,j} = \sum_k \delta_{i,j}\delta_{\lfloor i/2 \rfloor,k}\omega_{ky}\sigma_{ky}^2;$
$\mathbf{E}_{\sigma_\bullet} : XY \times XY,$	$(\mathbf{E}_{\sigma_x})_{i,j} = \sum_k \delta_{i,j}\delta_{\lfloor i/2 \rfloor,k}\omega_{kx}^2\sigma_{kx},$
	$(\mathbf{E}_{\sigma_y})_{i,j} = \sum_k \delta_{i,j}\delta_{\lfloor i/2 \rfloor,k}\omega_{ky}^2\sigma_{ky}.$

2.2 Optimization of Parameters for Gabor Kernels

We collect a training set of P patterns, i.e., $Q = \{q_p | p \in \{1, \cdots, P\}\}$, by cutting the neighboring regions of all M fiducial points from N training images, where $P = N \times M$. For the p-th pattern, we use p_c and p_s for the label of the fiducial point and that of the subject to which this pattern belongs. We take a reversed-order raster scan for each pattern, i.e., $\mathbf{q}^p = [q_{xy}^p | x \in \{X/2, \cdots, -X/2\}, y \in \{Y/2, \cdots, -Y/2\}]$, so that the convolution of Gabor kernel on a certain feature point is identical to a matrix multiplication in a window centered on this point. X and Y are the width and height of the window around the feature. We use XY to denote the window size. Dividing the jets into different clusters according to the fiducial points that they represent, we search for Gabor parameters that increase the similarity among jets within the same cluster and decrease the similarity between different clusters by minimizing the following energy function:

$$F = \sum_{p,r,r>p} \left(\frac{\lambda\delta_{p_s,r_s}(1-\delta_{p_c,r_c})}{N_s} - \frac{\delta_{p_c,r_c}}{N_c}\right) S(J^p, J^r) \quad (3)$$

with $N_s = \sum_{p,r,r>p} \delta_{p_s,r_s}(1-\delta_{p_c,r_c})$ and $N_c = \sum_{p,r,r>p} \delta_{p_c,r_c}$. $S(J^p, J^r)$ is the metric of similarity between two jets J^p and J^r. $\delta_{a,b}$ is the Kronecker delta function which takes 1 for $a = b$ and 0 for $a \neq b$. The first term tends to increase the degree of separation between different fiducial points from the same person, while the second term is a regularization term that improves the generalization ability of the extracted bunch from training samples. λ is a hyper- parameter introduced to control the relative strength of between-cluster scatter to that of within cluster scatter. Detailed formulations are given in the following part of this section. For convenience of formulation, we separate the direct current component from the Gabor kernel in Eq.(1). A common Gabor kernel is

$$G_{kxy} = C_{kxy} + iS_{kxy}, k \in \{1, \cdots, K\}, \quad (4)$$

Tabel 2 Algorithm for learning Gabor parameters

a) Initialize the parameters by taking certain divisions in angle and wave length. Set T as the max iteration time and ε as the threshold for convergence. Set $t = 0$;

b) Construct q^p by raster scanning the cropped face image in a reverse order;

c) Calculate $\mathbf{u}^p, \mathbf{u}_{\omega x}^p, \mathbf{u}_{\omega y}^p, \mathbf{u}_{\sigma x}^p, \mathbf{u}_{\sigma y}^p$ for all patterns;

d) Calculate $\partial F^{(t)}/\partial\theta$ for both $\omega_x, \omega_y, \sigma_x, \sigma_y$;

e) Update $\omega_x^{(t+1)}, \omega_y^{(t+1)}, \sigma_x^{(t+1)}, \sigma_y^{(t+1)}$;

f) Calculate energy function $F^{(t+1)}$;

g) Increase t by 1;

h) If $\left|F^{(t+1)} - F^{(t)}\right| < \varepsilon$ or $t > T$, exit. Otherwise, repeats steps (c) $\sim$ (h).

$$C_{kxy} = \frac{\cos(\omega_{kx}x + \omega_{ky}y)}{2\pi\sigma_{kx}\sigma_{ky}} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_{kx}^2} + \frac{y^2}{\sigma_{ky}^2}\right)\right], \quad (5)$$

$$S_{kxy} = \frac{\sin(\omega_{kx}x + \omega_{ky}y)}{2\pi\sigma_{kx}\sigma_{ky}} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_{kx}^2} + \frac{y^2}{\sigma_{ky}^2}\right)\right]. \quad (6)$$

Especially, all filters in a Gabor family are put into rows of $2K \times XY$ matrix $\mathbf{G}$, whose element reads

$$\mathbf{G}_{l,xy} = \sum_k \delta_{l,2k} C_{kxy} + \delta_{l,2k+1} S_{kxy}. \quad (7)$$

while $\Theta_k = \{\sigma_{kx}, \sigma_{ky}, \omega_{kx}, \omega_{ky}\}$ is the set of parameters. The DC components of filters are written into another matrix $\mathbf{G^{DC}}$, which is defined as $\mathbf{G}_{l,xy}^{\mathbf{DC}} = \sum_k \delta_{l,2k} D_{kxy}$, with

$$D_{kxy} = \frac{\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_{kx}^2} + \frac{y^2}{\sigma_{ky}^2} + \omega_{kx}^2\sigma_{kx}^2 + \omega_{ky}^2\sigma_{ky}^2\right)\right]}{2\pi\sigma_{kx}\sigma_{ky}} \quad (8)$$

Therefore, the response of the Gabor bank reads

$$\mathbf{u}^p = [u_0^p, \cdots, u_{2K-1}^p]^T = \{\mathbf{G} - \mathbf{G^{DC}}\}\mathbf{q}^p \quad (9)$$

Because jets from each point are extracted independently during the learning process, and all samples are centered on the ground truth position, we let $\vec{d} = [0 \quad 0]$ for similarity defined in Eq.(2). This similarity is equal to the cosine of the angle between $\mathbf{u}^p$ and $\mathbf{u}^r$, i.e.,

$$S(J^p, J^r, \vec{d} = [0 \quad 0]) = \cos(\mathbf{u}^p, \mathbf{u}^r) = \frac{[\mathbf{u}^p]^T\mathbf{u}^r}{|\mathbf{u}^p||\mathbf{u}^r|}. \quad (10)$$

Accordingly, the energy function is rewritten as

$$F = \sum_{p,r,r>p} \left\{\frac{\lambda(1-\delta_{p_c,r_c})\delta_{p_s,r_s}}{N_s} - \frac{\delta_{p_c,r_c}}{N_c}\right\} \cos(\mathbf{u}^p, \mathbf{u}^r).$$

The differential of energy function to parameter θ is

$$\frac{\partial F}{\partial \theta}=\sum_{p,r,r>p}\{\frac{\lambda(1-\delta_{p_c,r_c})\delta_{p_s,r_s}}{N_s}-\frac{\delta_{p_c,r_c}}{N_c}\}\frac{\partial \cos(\mathbf{u}^p,\mathbf{u}^r)}{\partial\theta} \quad (11)$$

For a parameter $\theta_k \in \Theta_k$, the differential reads

$$\frac{\partial \cos(\mathbf{u}^p,\mathbf{u}^r)}{\partial\theta_k}=\frac{1}{|\mathbf{u}^p||\mathbf{u}^r|}\frac{\partial([\mathbf{u}^p]^T\mathbf{u}^r)}{\partial\theta_k} -\frac{[\mathbf{u}^p]^T\mathbf{u}^r}{2|\mathbf{u}^p||\mathbf{u}^r|}\left\{\frac{\partial([\mathbf{u}^p]^T\mathbf{u}^p)}{|\mathbf{u}^p|^2\partial\theta_k}+\frac{\partial([\mathbf{u}^r]^T\mathbf{u}^r)}{|\mathbf{u}^r|^2\partial\theta_k}\right\}. \quad (12)$$

Accordingly, we have

$$\begin{aligned}\frac{\partial([\mathbf{u}^p]^T\mathbf{u}^r)}{\partial\theta_k}&=(\sum_{xy}\frac{\partial(C_{kxy}-D_{kxy})}{\partial\theta_k}q^p_{xy})u^r_{2k}\\ &+(\sum_{xy}\frac{\partial S_{kxy}}{\partial\theta_k}q^p_{xy})u^r_{2k+1}+(\sum_{xy}\frac{\partial(C_{kxy}-D_{kxy})}{\partial\theta_k}q^r_{xy})u^p_{2k}\\ &+(\sum_{xy}\frac{\partial S_{kxy}}{\partial\theta_k}q^r_{xy})u^p_{2k+1} \quad (13)\end{aligned}$$

If we let $\theta = [\theta_0,\cdots,\theta_{K-1}]^T$ be the parameter vector for the Gabor family, and define $\hat{\mathbf{u}}^p_\theta = [\hat{u}^p_{\theta_{0,0}},\cdots,\hat{u}^p_{\theta_{K-1,2K-1}}]^T$ which satisfies

$$\hat{u}^p_{\theta_{\lfloor l/2\rfloor,l}}=\sum_{k,xy}\frac{q^p_{xy}(\delta_{l,2k}\partial(C_{kxy}-D_{kxy})+\delta_{l,2k+1}\partial S_{kxy})}{\partial\theta_k} \quad (14)$$

we will have a simplified notation as

$$\frac{\partial([\mathbf{u}^p]^T\mathbf{u}^r)}{\partial\theta}=\mathbf{A}[\hat{\mathbf{u}}^p_\theta\circ\mathbf{u}^r+\hat{\mathbf{u}}^r_\theta\circ\mathbf{u}^p] \quad (15)$$

where $\mathbf{A}$ is a $K\times 2K$ matrix and $\mathbf{A}_{k,l}=\delta_{\lfloor l/2\rfloor,k}$ for $\forall k,l,k\in\{0,\cdots,K-1\},l\in\{0,\cdots,2K-1\}$. The symbol $\circ$ defines the component-wise multiplication of two matrices. We summarize the equations for $\hat{\mathbf{u}}^p_\theta$ as

$$\hat{\mathbf{u}}^p_{\omega_\bullet}=\mathbf{G}_{\omega_\bullet}\mathbf{q}^p,\ \hat{\mathbf{u}}^p_{\sigma_\bullet}=\mathbf{G}_{\sigma_\bullet}\mathbf{q}^p, \quad (16)$$

with

$$\begin{cases}\mathbf{G}_{\omega_\bullet}=\mathbf{B}_\omega\mathbf{G}\mathbf{C}_\bullet+\mathbf{E}_{\omega_\bullet}\mathbf{G}^{\mathbf{DC}}\\ \mathbf{G}_{\sigma_\bullet}=\mathbf{B}_{\sigma_\bullet}(\mathbf{G}-\mathbf{G}^{\mathbf{DC}})\mathbf{C}_{\sigma_\bullet}-\mathbf{D}_{\sigma_\bullet}(\mathbf{G}-\mathbf{G}^{\mathbf{DC}})\\ \quad +\mathbf{E}_{\sigma_\bullet}\mathbf{G}^{\mathbf{DC}}\end{cases} \quad (17)$$

where the symbol $\bullet$ could be replaced by x or y for the horizontal or vertical direction of 2D Gabor, respectively. Each matrix in Eq. (17) is summarized in Table 1. Note that we use $i_y=\lfloor i/X\rfloor, i_x=(i-i_yX)$ to denote the x and y coordinates of pixel i , respectively. Finally, we have

$$\frac{\partial F}{\partial\theta}=\mathbf{A}\sum_p\sum_{r,r>p}\{(1-\delta_{p_c,r_c})\delta_{p_s,r_s}\frac{\lambda}{N_s}-\frac{\delta_{p_c,r_c}}{N_c}\} \left\{\frac{\hat{\mathbf{u}}^p_\theta\circ\mathbf{u}^r+\hat{\mathbf{u}}^r_\theta\circ\mathbf{u}^p}{|\mathbf{u}^p||\mathbf{u}^r|}-\frac{[\mathbf{u}^p]^T\mathbf{u}^r}{|\mathbf{u}^p||\mathbf{u}^r|}\left[\frac{\hat{\mathbf{u}}^p_\theta\circ\mathbf{u}^p}{|\mathbf{u}^p|^2}+\frac{\hat{\mathbf{u}}^r_\theta\circ\mathbf{u}^r}{|\mathbf{u}^r|^2}\right]\right\} \quad (18)$$

The learning of optimal Gabor parameters is performed using a form of Gradient descendent learning, while the learning rule is

$$\theta^{(t+1)}=\theta^{(t)}-\eta\Delta\theta^{(t)}=\theta^{(t)}-\eta\frac{\partial F^{(t)}}{\partial\theta}. \quad (19)$$

The algorithm is summarized in Table 2.

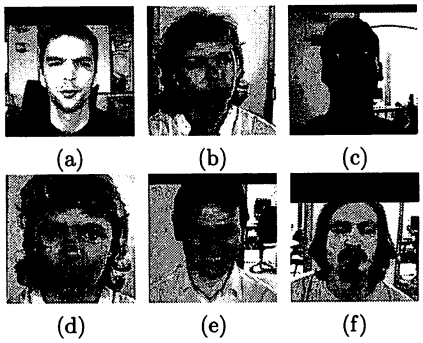


Fig. 5 Samples of extracted images for a finer localization. (a) Frontal face, (b) Pose variation, (c) Brightness variation, (d) Scale variation (e) Expression, and (f) Occlusion.

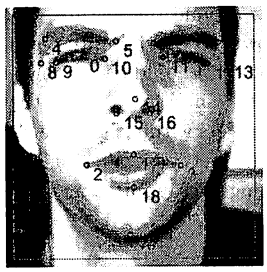


Fig. 6 Selected fiducial features and their positions.

3 Numerical Experiments and Discussions

In our numerical experiments, we focus on the investigation of performance under the optimized parameters achieved by the proposed approach. The Bio ID Face Database [11] has been used as the testing database. We randomly picked two mutually-exclusive sets from the whole database. One set, consisting of 200 images, is used as the universe of training samples, from which samples are randomly selected to form the training set. Another set with 100 images is prepared for testing. When dealing with the finer land- marking by EBGM, we confront the following problems, i.e., variations in pose, illumination, and scale, morphing caused by facial expressions, and occlusions of fiducial points by eye closing or by glasses and beard. Some

examples of extracted face images before finer localization was performed are shown in Fig.5.

For each image, 20 points are selected as fiducial features and their true positions are manually marked, as shown in Fig.6. We use an Adaboost face detector to extract the face region for each image in the training set, and then rescale it to a size of 200 pixels x 200 pixels. All images that failed to be detected by the Adaboost face detector were not counted in our experiments. Around the ground truth of each fiducial point, a window of size 51 x 51 pixels is extracted and reorganized in the reversed order of raster scanning to produce Q for the learning of Gabor parameters.

To evaluate the accuracy of localization, a normalized distance is defined by taking the ratio of bias in the present point to the inter-eye distance, where the inter-eye distance is calculated from the ground truth for each image. In our experiments, 40 Gabor kernels are used, which include 5 wavelength and 8 angle divisions. The learning rate is 0.3 and the threshold ε is 0.0001. Starting from the initial position, a full search is performed in a neighborhood of size 16 x 16 pixels.

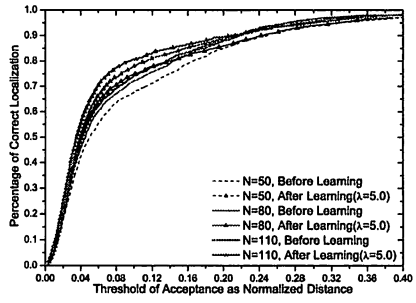


Fig. 7 Comparison of localization accuracy from parameters before and after learning. A smaller bunch size also achieves higher accuracy with optimized parameters. N is the depth of bunch (also the number of images patterns) and N=P/20.

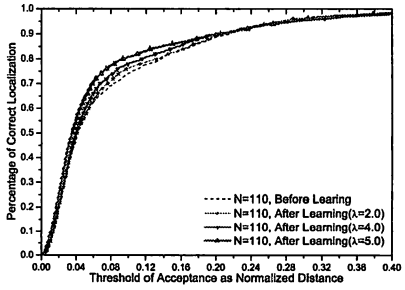


Fig. 8 Comparison between localization accuracy under different λ.

Major results of experiments are listed as follows:

1) In Fig.7, comparisons are made between results from parameters before and after learning. As depicted in Fig.7, a bunch with less depth achieves higher accuracy with learned optimal Gabor parameters.

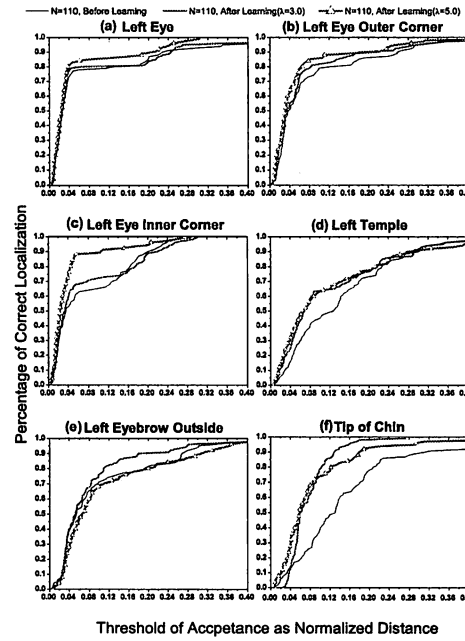


Fig. 9 Comparison of localization accuracy with Gabor parameters before and after learning for different fiducial points.

2) The transition of localization accuracy under different hyper- parameter λ is shown in Fig.8. Because the parameter is optimized on the majority of samples by the learning on the training samples, the overall accuracy has been improved. On the other hand, even larger bias might occur for outliers in the testing set, e.g., images with beard or glasses in eye area, etc.

3) We also made detailed comparisons on each point. All fiducial points are divided into two categories: first-tier points include eyes, corners of eyes and mouth, nose tip and others, and second-tier points include the left and right temples, the tip of chin, eye brows, the bridge of nose, etc. Results on three first-tier points, i.e., left eye, left eye outer corner, left eye inner corner, are given in Fig.9(a) to (c), while results for three second tier points, namely, the left temple, the left eyebrow outer corner and the tip of chin, are shown in Fig.9(d) to (f), respectively. Three cases under N=110, i.e., before learning, after learning with $\lambda = 3.0$, after learning with $\lambda = 5.0$, are compared. We found that EBGM with optimized Gabor parameters improve the accuracy on both the first- tier and second-tier points.

4) The transition of Jet similarity around the ground

truth of left eye is plotted to check the influence caused by different hyper-parameter in Fig.10. The similarity of the jet in left eye to all jets extracted in the same row is plotted for three cases, initial parameter, parameter from N=110 and $\lambda = 3.0$, and N=110 and $\lambda = 5.0$ in Fig.10. After learning, the distance between the left eye to its neighbors (e.g., inner corner of left eye or outer corner of left eye, etc) is enlarged, which reflects the influence of the minimization of the energy function.

From the above results, we found that the learned Gabor parameters increase the degree of discrimination of Gabor jets and accordingly the localization accuracy has been improved. Comparing with the conventional EBGM, optimized Gabor parameters could provide improved accuracy of localization with a smaller depth of bunch, which also speeds up the processing of localization. Although we performed a full search, a technique, i.e., using FFT for fast computation of Gabor jets, helps reduce the burden of computation. In a PC with 1.6GHz CPU, the whole process with both coarse and fine land-marking costs about 10 seconds for each image in our experiments. Faster speed is possible if the search is executed in a selective way.

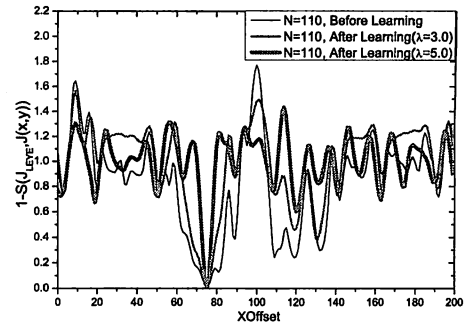


Fig. 10 Transition of similarity w.r.t. left eye for initial Gabor parameter, parameters learned by bunch depth N=110 and $\lambda = 3.0$,and parameters learned by N=110 and $\lambda = 5.0$.

4 Conclusions

We focus on the improvement of localization accuracy in the EBGM method by the optimization of Gabor parameters in the meaning of maximizing the degree of discriminant between Gabor jets from different fiducial points. We use the gradient descendent method for searching the local optimum of energy function, and derive the learning rule. According to our numerical experiments, the accuracy of localization is improved by optimized Gabor parameters when comparing with the conventional EBGM. This verifies the efficiency of the proposed approach. In the energy function we introduce a new hyper-parameter, which is set heuristically in the present work. Since the energy function could be extended to the framework of Markov random field, we will consider the possibility of automatically determining the hyper-parameter by probabilistic estimation, which is left as a topic for our future work. Further investigation regarding more efficient energy functions is another important future research topic.

References

[1] S. Arca, P. Campadelli, and R. Lanzarotti, "A face recognition system based on automatically determined facial fiducial points," *Pattern Recognition,* Vol.39, pp.432-443, 2006.

[2] L. Wiskott, J. M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. on PAMI,* Vol.19, pp.775-779, 1997.

[3] A. Tefas, C. Kotropoulos, and L. Pitas, "Using support vector machines to enhance the performance of elastic graph matching for frontal face authentication", *IEEE Trans. on PAMI,* Vol.23, pp.735-746, 2001.

[4] G. N. Stamou, N. Nikolaidis, I. Pitas, "Object tracking based on morphological elastic graph matching," *ICIP'05,* Vol.1, pp.709-712, 2005.

[5] P. Kalocsai, H. Neven, and J. Steffens, "Statistical analysis of Gabor-filter representation," *ICAFGR'98,* pp. 360-365, 1998.

[6] T. S. Lee, "Image representation using 2D Gabor wavelets," *IEEE Trans. on PAMI,* Vol.18, pp. 959-971, 1996.

[7] Z. Hou, and J. M. Parker, "Texture defect detection using support vector machines with adaptive Gabor wavelet features," *WACV'05,* Vol.19, pp.275-280, 2005.

[8] K. Kameyama, K. Mori, and Y. Kosugi, "A neural network incorporating adaptive Gabor filters for image texture classification," *ICNN'97,* Vol. 3, pp.1523-1528, 1997.

[9] P. Moreno, A. Bernardino, and J. S. Victor, "Gabor parameter selection for local feature detection," *IbPRIA'05,* Vol.1, pp.11-19, 2005.

[10] P. Viola, and M. Jones, "Robust real-time object detection," *Cambridge Research Laboratory, Technical Report Series,* Vol.2001-01, 2002.

[11] BioID Face Database, http://www.humanscan.de/support/downloads/facedb.php.