Extracting Micro-Structural Gabor Features for Face Recognition

Dian Gong [1,2] Qiong Yang [1], Xiaoou Tang [1] and Jianhua Lu [2]

[1]Microsoft Research Asia, Beijing Sigma Centre,
Beijing 100080, China
t-qiyang@microsoft.com

[2]Department of Electronic Engineering,
Tsinghua University,
Beijing 100084, China

***Abstract*—Robustness and discriminability are two key issues in face recognition. In this paper, we propose a new algorithm which extracts micro-structural Gabor feature to achieve good robustness and discriminability simultaneously. We first design a family of directional block partitions to compute the block-level directional projections of the classical Gabor feature. Then we use two statistical kernels, i.e, the mean kernel and the variance kernel, to extract the micro-structural statistics. Analysis of both robustness and discriminability is conducted to show that the new feature is not only more robust to misalignment, but also more discriminative than the classical down-sampling Gabor feature, which is further demonstrated by three groups of experiments on the BANCA dataset.**

Keywords-Micro-structural Gabor feature; Face recognition; Statistical kernel

I. INTRODUCTION

In recent years, Gabor feature has become one of the most effective features for face recognition[1~11]. However, the complete Gabor feature set is huge and redundant. This will not only increase much burden in computation, but also introduce high risk of "curse-of-dimensionality". So it is necessary to reduce the dimensionality of Gabor features.

At present, there are at least three kinds of methods to deal with the dimensionality problem of Gabor features. The first kind is based on graph. Lades[1] and Wiskott[2] constructed the grid graph and the labeled graph respectively to extract a family of Gabor features. The second is uniform down-sampling. For example, Liu et al. [5] reduced the Gabor feature dimension by only convolving Gabor filters on uniformly down-sampled pixels. In the third category of methods, an objective function is generally designed to optimize the selection of Gabor filters, including selection of their locations, scales, and orientations. Among them, Gökberk et al. [7] learned the importance of each grid point. The least important grids were discarded, and the remaining grids were weighted according to their recognition performance. They also used a number of feature selection algorithms and a genetic algorithm to find the optimal Gabor kernel location for face recognition [8]. In [9], Wu et al. selected optimal Gabor filters for high-speed face identification. AdaBoost was also used by Yang et al.[10] to choose the most discriminative Gabor feature set.

In all the above methods, only the Gabor feature on a set of individual sites is considered, so they have two drawbacks. Firstly, the Gabor feature on individual sites is sensitive to misalignment, such as shift and size variation. Secondly, it cannot reflect the spatial correlation between Gabor features at the neighboring sites, which is quite important for discriminating different persons.

In this paper, we propose a new algorithm which extracts micro-structural Gabor features (MSGF) to enhance the performance of face recognition. The key idea originates from our observation that micro-structural characteristics will play an important role in the face recognition problem. Although there are differences in the global face patterns among different persons, the spatial correlation between neighboring sites may contain much discriminative information when human recognizes persons. Therefore, we compute the micro-structural Gabor features to reflect the local statistics of Gabor features at neighboring sites. We first design a family of directional block partitions to compute the block-level directional projections of the classical Gabor features. Then we present two kernels, the mean kernel and the variance kernel, to extract the statistical characteristics of those block-level directional projections. The mean kernel improves the feature's robustness, and the variance kernel enhances its discrimination capability.

II. CLASSICAL GABOR FEATURE

Gabor wavelet representation was introduced to image analysis due to its biological relevance and its outstanding capacity on spatial locality, scale selectivity, and orientation selectivity. It can be denoted as follows:

$$O_{\gamma,\varphi}(x,y) = I(x,y) \otimes h_{\gamma\varphi}(x,y), \tag{1}$$

where $\boldsymbol{I} = [I(x,y)]_{H\times W}$ is an image with height H and width W,

$$\begin{aligned} h_{\gamma\varphi}(x,y) &= \exp\left\{-\frac{1}{2}\left[\frac{R_{1\gamma}^2 + R_{2\gamma}^2}{\sigma_\varphi^2}\right]\right\} \cdot \exp\left[i \cdot \frac{2\pi R_{1\gamma}}{\lambda_\varphi}\right], \\ R_{1\gamma} &= x\cdot\cos\phi_\gamma + y\cdot\sin\phi_\gamma, \\ R_{2\gamma} &= -x\cdot\sin\phi_\gamma + y\cdot\cos\phi_\gamma. \end{aligned} \tag{2}$$

$h_{\gamma\varphi}(x,y)$ is the impulse response function of Gabor filter with parameters $(\lambda_\varphi, \phi_\gamma, \sigma_\varphi)$, and $O_{\gamma,\varphi}(x,y)$ is its output at site (x,y). Among them, $\lambda_\varphi = \lambda_{\max}/f^\varphi$ and $\phi_\gamma = \pi\gamma/n$ give the wavelength and orientation of the Gabor filter respectively, and $\sigma_\varphi = \beta\lambda_\varphi$ gives the deviance of Gaussian function. If $n=8$, $\gamma \in \{0,1,2,3,4\}$, $\varphi \in \{0,1,...,7\}$, $f=\sqrt{2}$, $\lambda_{\max}=8$, and $\beta=1$, a family of Gabor filters with 5 scales and 8 orientations are generated.

Usually, the amplitude or the absolute value of the real part of $O_{\gamma,\varphi}(x,y)$ is computed to extract the Gabor feature. Take the amplitude as an example. It can be computed by

$$G_{\gamma,\varphi}(x,y) = \left\| O_{\gamma,\varphi}(x,y) \right\|. \tag{3}$$

III. MICRO-STRUCTURAL GABOR FEATURE

The classical Gabor feature in Eq.3 is based on an individual site, and it is sensitive to misalignment [11], such as shift and size variation, which often occurs in fully automatic face recognition system. Moreover, the feature based on individual sites cannot describe the spatial correlation between the Gabor features at neighboring sites, which is significant for discriminating different persons. As a result, in the next part, a new algorithm will be proposed to extract micro-structural Gabor features, which will improve both the robustness and discrimination capability of the classical Gabor feature.

A. Block-Level Directional Projection

The first step of the new algorithm is the block-level directional projection of the classical Gabor feature. It first divides the image into sub-blocks as shown in Fig. 1. Then it designs a family of directional block partitions. Each of them is designed for a corresponding Gabor filter with specific orientation (Fig.2) due to Gabor's property on orientation selectivity. After that, the block-level directional projections of the classical Gabor feature are computed independently for each scale and each orientation. Suppose there are a family of Gabor filters with 4 orientations, e.g. $\phi_\gamma \in \{0, \pi/4, \pi/2, 3\pi/4\}$, and the block size is 3×3, then the directional block partitions are designed as shown in Fig. 2 (pixels bounded by the same color belong to the same partition). The directional projection computes the sum of Gabor features in each partition, as depicted in Fig. 2.

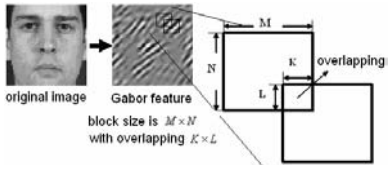


Fig. 1 Block-level representation

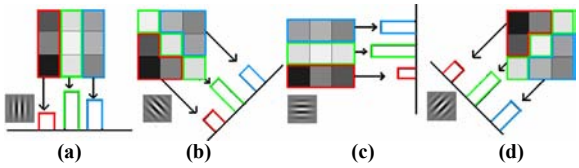


Fig. 2 The family of directional block partitions and their corresponding block-level directional projections. The orientation of the corresponding Gabor filter is displayed in the left-bottom corner of each sub-figure.

B. Statistical Kernel

In the second step, the new algorithm presents two statistical kernels to extract the statistics of the above block-level directional projections. One is the mean kernel, and the other is the variance kernel. Suppose the block size is $M\times M$ ($M=N$), and the block-level directional projections of the Gabor feature with $(\lambda_\varphi, \phi_\gamma, \sigma_\varphi)$ in the i -th block are $\boldsymbol{P}_i^{\gamma\varphi} = \left[p_{i1}^{\gamma\varphi}, p_{i2}^{\gamma\varphi}, ..., p_{iM}^{\gamma\varphi} \right]^T$. The directional mean $\mu_i^{\gamma\varphi}$ and the directional variance $(\sigma_i^{\gamma\varphi})^2$ can be computed as:

$$\mu_i^{\gamma\varphi} = \frac{1}{M}\sum_{j=1}^{M} p_{ij}^{\gamma\varphi} = \boldsymbol{K}_\mu^{\ T} \cdot \boldsymbol{P}_i^{\gamma\varphi}, \tag{4}$$

$$\begin{aligned}(\sigma_i^{\gamma\varphi})^2 &= \sum_{j=1}^{M}\left(p_{ij}^{\gamma\varphi}\right)^2 - \left(\frac{1}{M}\sum_{j=1}^{M} p_{ij}^{\gamma\varphi}\right)^2 \\ &= \sum_{j=k}(1-\frac{1}{M})\left(p_{ij}^{\gamma\varphi}\right)^2 + \sum_{j\neq k}(-\frac{1}{M}) p_{ij}^{\gamma\varphi} p_{ik}^{\gamma\varphi}, \\ &= \left(\boldsymbol{P}_i^{\gamma\varphi}\right)^T \cdot \boldsymbol{K}_\sigma \cdot \boldsymbol{P}_i^{\gamma\varphi}\end{aligned} \tag{5}$$

where $\boldsymbol{K}_\mu = \frac{1}{M}\boldsymbol{I}_M$, $\boldsymbol{K}_\sigma = \boldsymbol{I}_{M\times M} - \frac{1}{M}\boldsymbol{I}_M \boldsymbol{I}_M^{\ T}$.

From the above, we can see that if we define $\boldsymbol{K}_\mu$ as the mean kernel and $\boldsymbol{K}_\sigma$ as the variance kernel, $\mu_i^{\gamma\varphi}$ and $(\sigma_i^{\gamma\varphi})^2$ can be computed by a 1-order multiplication with the mean kernel (Eq.4) and a 2-order multiplication with the variance kernel (Eq.5) respectively. If the block size is 3×3, the mean kernel and the variance kernel can be depicted as shown in Fig. 3.

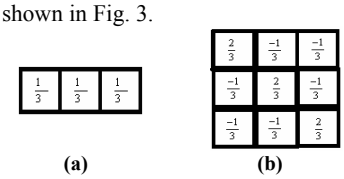


Fig. 3 The statistical kernel for micro-structural Gabor features. (a) The mean kernel. (b) The variance kernel.

In this way, the micro-structural Gabor feature $\boldsymbol{Z}$ can be extracted as $\boldsymbol{Z} = \left\{\boldsymbol{Z}_i^{\gamma\varphi}\right\}$, where $\boldsymbol{Z}_i^{\gamma\varphi} = \left[\mu_i^{\gamma\varphi}, \sigma_i^{\gamma\varphi}\right]^T$ is computed from the Gabor feature with parameters $(\lambda_\varphi, \phi_\gamma, \sigma_\varphi)$, and $\sigma_i^{\gamma\varphi} = \sqrt{(\sigma_i^{\gamma\varphi})^2}$. Since the mean kernel can be regarded as a low-pass filter in the frequency domain, $(\mu_i^{\gamma\varphi})^2$ represents the energy of the low-frequency part. From $(\sigma_i^{\gamma\varphi})^2 = \sum_{j=1}^{M}(p_{ij}^{\gamma\varphi})^2 - (\mu_i^{\gamma\varphi})^2$ where the first item on the right hand represents the total energy in the block, we can deduce that $(\sigma_i^{\gamma\varphi})^2$ indicates the sum of the energy in the high-frequency part. Therefore, the micro-structural Gabor feature basically decomposes the energy into two parts: the low-frequency part which is emphasized by $\mu_i^{\gamma\varphi}$ to improve the feature's robustness to misalignment, and the high-frequency part which is represented by $\sigma_i^{\gamma\varphi}$ to depict the statistical variation of the classical Gabor features in the block area and thereby enhance the discrimination capability of features. By extracting micro-structural Gabor feature, both robustness and discriminability are achieved simultaneously.

To better understand the newly proposed micro-structural Gabor feature, we investigate both robustness analysis and discriminability analysis on the new feature in the next part.

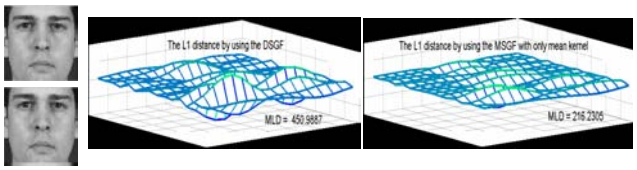
C. Robustness Analysis

To analyze the robustness, we compare our MSGF feature with the classical down-sampling Gabor feature (DSGF) which is down-sampled by $D_x \times D_y$. Only directional mean is used here, since the robustness is mainly achieved by the mean kernel implementation. Fig. 4 shows that our new feature is much more robust to shift than DSGF, and Fig. 5 demonstrates its better robustness on size variation.

D. Discriminability Analysis

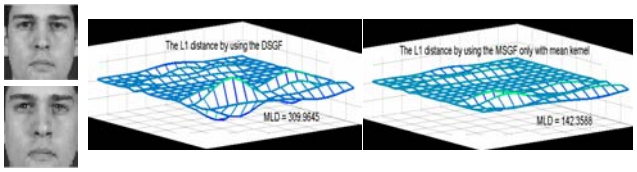
Discussions on discriminability analysis are necessary. Although $\mu_i^{\gamma\varphi}$ represents the global characteristics of the Gabor features in the whole block and it improves the robustness, it cannot reflect the spatial correlation between Gabor features at neighboring sites. Adding $\sigma_i^{\gamma\varphi}$ will emphasize the statistical variation in the block, and thereby enhances the discrimination capability.

Take Fig. 6 as an example. Two images manifest different patterns, but they take the same value when only applying the mean kernel. However, the variance kernel will discriminate these two patterns. Fig. 7 is another demonstration of the stronger discriminability of the new feature. In the figure, the new feature with both the mean kernel and the variance kernel shows more discrimination



(a) (b) (c)

Fig. 4 Comparison of the robustness on shift between DSGF and MSGF with only mean kernel. (a) The original image and the shifted face with 3-pixel vertical shift. (b) The L1-distance between DSGF of the two images. The mean L1 distance (MLD) is 450.99. (c) The L1-distance between MSGF of the two images. MLD=216.23. ($\lambda = 5\sqrt{2}$, $\phi = \pi/2$, $\beta = 0.9$, $D_x = D_y = 4$, $M = N = 4$, $K = L = 0$)



(a) (b) (c)

Fig. 5 Comparison of the robustness on size variation between DSGF and MSGF with only mean kernel. (a) The original image and the zoomed face with a 4%-zooming rate. (b) The L1-distance between DSGF of the two images. The mean L1 distance (MLD) is 309.96. (c) The L1-distance between MSGF of the two images. MLD=142.36. ($\lambda = 5\sqrt{2}$, $\phi = \pi/2$, $\beta = 0.9$, $D_x = D_y = M = N = 4$, $K = L = 0$)

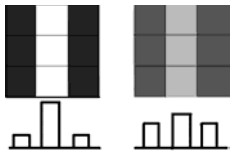
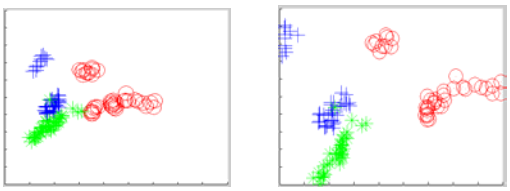


Fig. 6 Two patterns with the same directional mean but the different directional variance.



(a) (b)

Fig. 7 Comparison of discriminability between DSGF and MSGF with both mean and variance kernels. (a) Distribution of samples using DSGF. (b) Distribution of samples using MSGF with both kernels. Data are taken from three persons in the BANCA database[12]. The first two PCA components of the two kinds of features are displayed. Parameters of Gabor feature are the same as in Sec. IV.

capability than the classical down-sampling Gabor feature.

IV. Experiments and Analysis

Three groups of experiments on the BANCA database[12] are conducted to evaluate the performance of our new feature. The gray-scale feature is used as a baseline, and the classical down-sampling Gabor feature (DSGF) in Liu's work [5] is used as a benchmark for comparison. The database contains 52 subjects. In the experiments, the MC test configuration is employed, thereby 5 images/subject in Session 1 are used for training, and the other 35 images/subject in Session 1~4 are used for testing. So there are altogether 260 images for training and 1820 images for testing. Because the size of the training set is small, Gabor filters with only 2 scales and 4 orientations have been used on the original grey-level images in order to balance the feature dimension against the sample number.

A. Experiments on Precisely Aligned Data

The first group of experiments is tested on precisely aligned face images, each of which is cropped and normalized to the size of 55×51 based on the manual registration of eyes. Four kinds of features, the grayscale feature, the classical down-sampling Gabor feature in Liu's work [5] (DSGF), the micro-structural Gabor feature (MSGF) with only mean kernel (MSGF-1), and MSGF with both kernels (MSGF-2), are extracted. For each kind of feature, PCA is applied to reduce the feature dimension. Lastly, a nearest-neighbor (NN) decision is used for face recognition. In the experiments, we have $\lambda_\varphi \in \{6.5, 6.5\sqrt{2}\}$, $\phi_\gamma \in \{0, \pi/4, \pi/2, 3\pi/4\}$, $\beta = 1$, $D_x = D_y = 2$, $M = N = 3$, and $K = L = 0$. The experimental results are shown in Fig. 8.

From Fig. 8, we can observe that both MSGF features are better than the classical down-sampling Gabor feature. Among them, MSGF-2 is better than MSGF-1. When we use 20 features to test the probe set, the MSGF-2 achieves 85.88% accuracy rate, and DSGF gets only 81.76%. When 90 features are used, MSGF-2 achieves 91.98% accuracy rate, reducing the error rate by 12.64%. This demonstrates the effectiveness of the micro-structural Gabor features.

By further comparing the performance curves of the three kinds of Gabor feature, we see that, when the feature dimension is low, the performance improvement is mainly achieved by the mean kernel. Since the mean kernel acts as a low-pass filter, the block-level directional mean of the Gabor feature occupies a lower dimensionality than the classical Gabor feature. Thereby it achieves similar performance with much less features. When the feature dimension is high, the performance improvement is mainly achieved by the variance kernel, since it captures the statistical variance in the local block and thus enhances the discrimination capability.

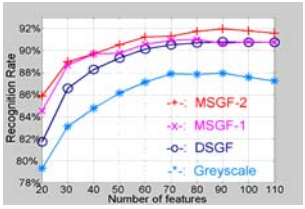
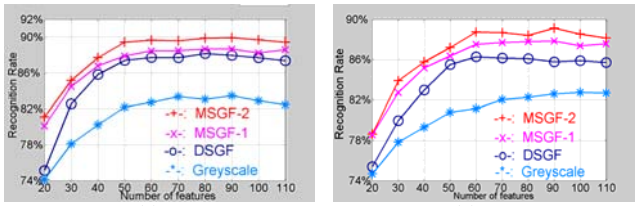


Fig.8 Experiments on precisely aligned data.

B. Experiments on Misaligned Data

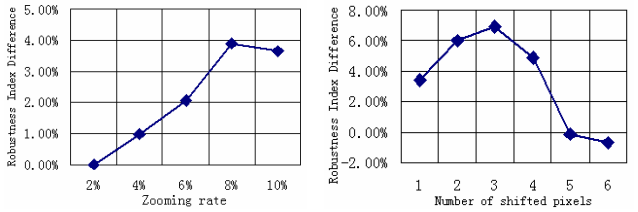
To test the performance on misaligned data, another group of two experiments are designed. The first is the test on zoomed faces, in which all the test images in Sec. IV (A) are zoomed by 4%. The second is a test on shifted faces, in which all the test images are vertically shifted by 1 pixel. The results are shown in Fig. 9.

From the figure, it is easy to find out that both MSGF-1 and MSGF-2 achieve better performance than DSGF, and again MSGF-2 is better than MSGF-1. In the 4%-zooming case, MSGF-2 achieves 10.05% error rate, 11.12% lower than that of MSGF-1, and 16.46% lower than that of DSGF. In the 1-pixel shift case, MSGF-2 achieves 89.12% recognition rate with 90 features, and its error rate is 23.54% lower than that of DSGF. The error reduction rate could also achieve 14.67% when compared with MSGF-1. The outperforming of MSGF-2 over MSGF-1 on all the experiments shows that the variance kernel enhances the discrimination capability with or without misalignment.



(a) (b)

Fig.9 Comparing MSGF with DSGF on misaligned data. (a) Comparison on zoomed images with 4% zooming rate (b). Comparison on shifted data with 1-pixel vertical shift.



(a) (b)

Fig.10 Robustness Analysis of MSGF and DSGF on misalignment. (a) Comparison on size variation. (b) Comparison on shift variation.

It is interesting to note that the mean kernel helps improve the robustness of features to misalignment. In the zoomed case, the performance of the three kinds of features drops by 31.05%, 21.20%, and 23.62% respectively. In the shift case, the error rate using DSGF increases by 55.01%, while using MSGF-1 only increases by 29.98%, and using MSGF-2 increases by 33.83%. In the zoomed case, the performance of the three kinds of features drops by 31.05%, 21.20%, and 23.62% respectively. From the above, we can see that both two kinds of MSGF features are more robust than the classical down-sampling Gabor feature, and the better robustness is mainly achieved by the mean kernel.

To further analyze the robustness on misalignment, a third group of two experiments are designed. The first is robustness analysis on size variation, and the second is on shift variation. Two algorithms are compared, DSGF and MSGF-2. The robustness index difference ΔR is computed by $\Delta R = R^{DSGF} - R^{MSGF2}$, where R^{DSGF} is the accuracy decreasing rate of DSGF, and R^{MSGF2} is that of MSGF-2. The results are shown in Fig. 10.

The experiments further demonstrate that MSGF-2 is more robust to misalignment than DSGF. Firstly, it manifests more robustness on size variation, since $\Delta R > 0$ when the zooming rate increases from 4% to 10%. Secondly, it is more robust to shift variation than DSGF when the shift is smaller than 4 pixels. In case that the shift is larger than 4 pixels, which exceeds both the half wavelength and the block size, the performance of Gabor feature drops drastically. It is difficult for the mean kernel to improve the performance of Gabor feature, since the shifted image will move out of the block. Generally, the 4-pixel shift seldom occurs in automatic eye location for an image with size 55×51. We tested the automatic eye locator[13] on the BANCA dataset, and the average location errors of 99.1% images are less than 4 pixels.

V. CONCLUSIONS AND FUTURE WORK

In this paper, a new algorithm is proposed to extract micro-structural Gabor features for face recognition. The main advantage of the new feature is that it achieves good robustness and discriminability simultaneously by decomposition into two parts: the low-frequency part improves the robustness to misalignment by applying mean kernel on the block-level directional projections; and the high-frequency part enhances the discrimination capability by using the variance kernel to capture the statistical variation in the local block area.

The new feature is of broad applicability. Although PCA is used in this paper to compress its dimension, the new feature can be combined with other methods such as Subspace LDA, Direct LDA, or Bayesian Subspace for higher recognition rate. It can also be applied on other datasets to test its robustness on various data. Moreover, it can be integrated into the current feature selection methods [7,8,10]. Finally, the idea of extracting micro-structural feature is valuable for pattern recognition. The block-level directional projection and the statistical kernel can be applied to other features, such as edge feature. The kernel is also not limited to the mean kernel and variance kernel in this paper. More effective and adaptive kernels could be designed for better performance, which will be an interesting direction of future work.

REFERENCES

[1] M. Lades, J.C. Vorbürggen, J. Buhmann, J. Lange, C. Malsburg, R.P. Würtz, and W. Konen. "Distortion invariant object recognition in the dynamic link architecture," IEEE Trans. on Computers, 42(3), pp.300-311, 1993.

[2] L. Wiskott, J. M. Fellous, N. Krüger, and C. Malsberg. "Face recognition by elastic bunch graph matching,"PAMI, 19(7),pp.775-779,1997.

[3] V. Krüger and G. Sommer, "Gabor wavelet networks for object representation," Journal of the Optical Society of America, 19(6), pp.1112-1119, 2002.

[4] L. Shen and L. Bai, "Gabor feature based face recognition using kernel methods," AFGR, pp.170-176, 2004.

[5] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition,"IEEE Trans. Image Processing,11(4):467-476,2002.

[6] D. Liu, K. Lam, and L. Shen, "Sampling Gabor features for face recognition," ICNNSP, pp.924-927, 2003.

[7] B. Gökberk, L. Akarun, and E. Alpaydin, "Feature selection for pose invariant face recognition," ICPR, pp.306-309, 2002.

[8] B. Gökberk, M. Irfanoglu, L. Akarun, and E. Alpaydin. "Optimal Gabor kernel location selection for face recognition," ICIP, pp.677-680, 2003.

[9] H. Wu, Y. Yoshida, and T. Shioyama, "Optimal Gabor filters for high speed face identification," ICPR, pp.107-110, 2002.

[10] P. Yang, S. Shan, W. Gao, S.Z. Li, and D. Zhang. "Face recognition using Ada-Boosted Gabor features," AFGR, pp.356-361, 2004.

[11] Q. Yang, "Research on appearance-based statistical face recognition," Ph. D Thesis at Tsinghua Univ. 2004.

[12] K. Messer, J. Kittler, M.Sadeghi, et al. "Face authentication test on the BANCA database," ICPR, pp.523-532, 2004.

[13] Y. Ma, X.-Q. Ding, Z. Wang, and N. Wang. "Robust Precise Eye Location Under Probabilistic Framework," AFGR, pp. 339-344, 2004.