

FACE IMAGE RESOLUTION VERSUS FACE RECOGNITION PERFORMANCE BASED ON TWO GLOBAL METHODS

Jingdong Wang* Changshui Zhang* Heung-Yeung Shum⁺

*State Key Laboratory of Intelligent Technology and Systems
Department of Automation, Tsinghua University, Beijing
⁺Microsoft Research Asia

ABSTRACT

Face recognition is an interesting topic in computer vision and object recognition. Researchers have proposed numeric recognition methods under the various conditions such as different pose, illumination and expression. However no research has been done to depict the relation between face image resolution and face recognition performance so far.

In this paper, we divide the face image information into the discriminative and structure information, and conjecture that the face recognition rate will level off when face image resolution arrives at one certain resolution threshold. Then, we propose several PCA and LDA based face recognition experiments and present statistical results to depict how face recognition rate changes with face image resolution. The experimental results validate the conjecture. Moreover, we propose several experiments to demonstrate that only enhancing the structure information could not help to improve the recognition performance.

1. INTRODUCTION

Face recognition is a popular topic in computer vision and object recognition in recent twenty years. It has a number of real-world applications such as surveillance, secure access and human/computer interface. Also, many methods successfully applied in face recognition can be eventually transferred to other object recognition problems. A detail survey can be found in [1].

In the following, we will first give a brief review on recognition methods and then present some research works about the recognition robustness against the various conditions. Finally, we introduce the related work.

Face Recognition Methods

For review clarity, we divide face recognition techniques roughly into two categories: a) the global approach, and b) the component-based approach.

a) In the former category, the input to a classifier is a single feature vector that represents the whole face image. Several classifiers have been proposed in the literature e.g. min-

imum distance classification in the eigenspace [2] [3] based on PCA, the independent face space based on ICA [4] and the discriminative subspace based on LDA [2] [3], neural networks [5], and probabilistic matching based on intrapersonal/extrapersonal image difference [6] [7]. The global techniques work well for classifying frontal views of face. For all the global techniques an alignment stage needs to be added before classifying the face. In [7] the correspondences between two face images are first built by labeling some key points and then an affine transform is computed to warp an input image into a reference face image. Active morphable model is adopted in [8] to match the input face with the reference face. Active shape models are used in [9] to align the input face with the model face. A semiautomatic alignment step in combination with SVM classification was proposed in [10].

b) For the component-based algorithms, the main idea is to compensate for pose changes by allowing a flexible geometrical relation between the components in the classification stage. In [11] face recognition was performed by independently matching templates of three facial regions(eyes, nose and mouth). The configuration of the components during classification was unstrained since the system did not include a geometrical model of the face. A similar approach with an additional alignment stage was proposed in [12]. In [13] a geometrical model of a face was implemented by a 2-D elastic graph. The recognition was based on wavelet coefficients that were computed on the nodes of the elastic graph. In [14] a window was shifted over the face image and the DCT coefficients computed within the window were fed into a 2-D Hidden Markov Model.

Face Recognition Robustness

Although some of the previous techniques have achieved successes in constrained scenarios, the general task of face recognition still poses a number of challenges with respect to the changes in illumination, facial expression and pose. Therefore currently researchers pay more attention to the study of the robustness against the changes in pose, illumination and expression.

In [15] Bernd Heisele et al. presented a component-based and two global recognition methods with multi-class support vector classifier and evaluated them with respect to the robustness against the pose changes. Takeshi Shakunaga and Kaz-uma Shigenari proposed a decomposed eigenface based face recognition method for realizing robust face recognition under various lighting conditions when too few images are available for registration in [18]. In [16] Geof Givens and his collaborators considered 11 factors that influence PCA-based face recognition performance including race, gender, age, glasses, facial hair, bangs, mouth, eyes, complexion, makeup and expression. They built an ANOVA system to analyze the relation between these subject covariates and the distance of the image pairs of the same subject. Some researchers are also concerned about how to optimize some specific face recognition methods, e.g. in [17] Wendy et al. analyzed PCA-based face recognition algorithms and studied the relation between recognition performance and the selection of eigenvector and distance measures.

When studying face recognition, researchers all run across one problem: what resolution face image is proper for face recognition. However, it is a pity that no researchers have worked on this problem. Fortunately, some researchers have done some related work. Simon Baker et al. thought that the enhanced information of the high resolution over the low resolution face could be decided by the low resolution face and built the corresponding face hallucination algorithm [19]. Ce Liu et al. first constructed a "recognized" global model to hallucinate the individual global face, then built a local model to enhance the local face feature [20].

Motivated by their work, we divide face image information into two kinds of information: the discriminative information and the structure information. The former represents the individual information compared with other face images. The latter is the common information of all face images under the same resolution. Then one conjecture is given that face recognition rate will level off when the face image resolution arrives at one certain value.

Finally, we perform PCA and LDA-based recognition algorithms on the AR-face database. The relation curves between face recognition rate and face resolution validate that the recognition rate will not increase until the face resolution arrives at some certain value. Moreover we propose experiments to indicate that it does not help to improve face recognition performance only to enhance the structure information with the bilinear interpolation.

The rest of this paper is organized as follows. Section 2 presents some analysis of the relation between face image resolution and face recognition performance, then Section 3 describes two global recognition methods: PCA and LDA. The experimental results are presented in Section 4. Conclusions and discussion are given in the final section.

2. FACE IMAGE RESOLUTION ANALYSIS

In this section, we first introduce the generalized Gaussian Pyramid, which is used to generate the multi-resolution face images. Then, we analyze the face image information constitution and conjecture the relation between face recognition performance and face image resolution. Finally, we give an example on OCR to illuminate the face information constitution.

Generalized Gaussian Pyramid

The Gaussian Pyramid [21] starting at level $l = k$ of an image I is the set of images $G_k(I), G_{k+1}(I), \dots, G_N(I)$

$$G_l(I) = \begin{cases} I & \text{if } l = k \\ REDUCE(G_{l-1}(I)) & \text{if } k < l \leq N \end{cases}$$

and N is chosen so that $G_N(I)$ is some fixed size. Traditionally the size of $G_l(I)$ is one-fourth of $G_{l-1}(I)$. In this paper, we want to reduce the high resolution face image (e.g. 256×192) to any lower resolution one. Firstly, we choose

$$REDUCE(I)(m, n) = \frac{1}{4} \sum_{i=0}^1 \sum_{j=0}^1 I(2m+i, 2n+j)$$

as the pixel averaging function to generate Gaussian interpolation to get the face images with the resolution 128×96 , 64×48 , 32×24 , 16×12 . Secondly, we apply the bilinear down-sampling upon the higher resolution face image $G_l(I)$ near to $h \times w$ and generate the generalized Gaussian face image with the resolution $h \times w$ (The higher resolution face image near to 72×54 is $G_1(I)$ with the resolution 128×96). The generalized Gaussian images are shown in Figure 1 and Figure 2.



Fig. 1. The Gaussian Pyramid

Face Image Information Analysis

When observing an object, people will usually see it more clearly when moving nearer and see it most clearly when the distance is (smaller than) one fixed value d_0 . If the object is known to be a face, we will recognize it when the distance is another fixed value d_1 , which is usually larger than d_0 . The former phenomenon can be interpreted by Shannon Sampling Theory. According to Shannon Sampling Theory,



Fig. 2. The lower resolution face images from 120×96 to 72×54 using bilinear downsampling on the image of 128×96

the image with one certain resolution can retain all the original image information, i.e., the corresponding reconstruction image will not lose any information. In the following, we will analyze the latter phenomenon.

The face image information can be partitioned into two kinds of information: the discriminative information and the structure information. The discriminative information represents the information distinguishing from other face images. The structure information represents the common information of all face images under the same resolution, which can be roughly estimated from the lower resolution face image. Accordingly, the difference of the discriminative information between the higher resolution face image and the lower one is called the enhanced discriminative information, and the difference of the structure information is called the enhanced structure information.

Conjecture The discriminative information will increase with the face image resolution increasing and remain stable when reaching one specific resolution.

According to this conjecture, the face recognition rate will increase with the face image resolution and will level off when arriving at one certain resolution. In Section 4 we will propose experiments to validate this conjecture.

One Illuminative Example

We give a simple example to illuminate the structure and discriminative information. When performing optical character recognition (OCR) with known font definition, we also observe the similar phenomenon. Before reaching the resolution threshold, the enhanced information is mainly character individual information which helps to recognize it more correctly. When reaching the threshold, most of the enhanced information corresponds to the font information. In this problem, the individual character information corresponds to the discriminative information and the font information corresponds to the structure information.

3. PCA AND LDA BASED FACE RECOGNITION METHODS

In this section, we introduce two global face recognition methods: PCA and LDA. Then we present some distance measures adopted in the two methods.

Principal Component Analysis

Given a t -dimensional vector representation of each face, the Principal Component Analysis (PCA) can be used to find a subspace whose basis vectors correspond to the maximum-variance directions in the original space. Let W represent the linear transformation that maps the original t -dimensional space onto a f -dimensional feature subspace where normally $f \ll t$. The new feature vectors $y_i \in \mathbb{R}^f$ are defined by $y_i = W^T x_i, i = 1, \dots, N$. The columns of W are the eigenvectors e_i obtained by solving the eigenvalue decomposition $\lambda_i e_i = Q e_i$, where $Q = XX^T/N$ is the covariance matrix, and λ_i the eigenvalue associated with the eigenvector e_i . Usually, before obtaining the eigenvectors of Q , the average of all image vectors is subtracted from all image vectors to ensure that eigenvector with the highest eigenvalue represents the dimension in the eigenspace in which variance of vectors is maximum in a correlation sense.

Pentland et al. [22] have empirically pointed out that superior face recognition results are achieved when the first two eigenvectors are discarded. This is because the first two eigenvectors seem to represent changes in illumination. Figure 3 shows the average face together with the first four eigenfaces (from left to right) on the AR-face database, which demonstrate the first two eigenfaces represent changes in illumination. In this paper, we will perform face recognition with eliminating the first two eigenvectors.



Fig. 3. Average face + first 4 eigenfaces

Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) searches for those vectors in the underlying space that best discriminate among classes (rather than those that best represent the data). More formally, given a number of independent features relative to which the data is described, LDA creates a linear combination of these which yields the largest mean differences between the desired classes. Mathematically speaking, for all the samples of all classes we define two measures: (i)

one called within-class scatter matrix, as given by

$$S_w = \sum_{j=1}^c \sum_{i=1}^{N_j} (x_i^j - \mu_j)(x_i^j - \mu_j)^T$$

where x_i^j is the i th sample of class j , μ_j is the mean of class j , c is the number of classes, and N_j the number of samples in class j ; (ii) the other is called between-class scatter matrix

$$S_b = \sum_{j=1}^c (\mu_j - \mu)(\mu_j - \mu)^T$$

where μ represents the mean of all classes. The goal is to maximize the between-class measure while minimizing the within-class measure. One way to do this is to maximize the ratio $\frac{\det|S_b|}{\det|S_w|}$. Like in [3], we also compute the projection matrix W with performing SVD on $S_w^{-1}S_b$. In this paper, we first choose the PCA subspace as the intermediate space and then obtain the final discriminating subspace by using LDA. The first five fisherfaces are shown in Figure4 (from left to right).



Fig. 4. The first 5 fisherfaces

Distance Measures

There are mainly four traditional distance measures used in the subspace obtained by PCA or LDA: City-block(L1), Euclidean(L2), Angle and Mahalanobis, which are defined as follows:

L1 City Block Distance

$$d(x, y) = |x - y| = \sum_{i=1}^k |x_i - y_i|$$

L2 Euclidean Distance(Squared)

$$d(x, y) = \|x - y\|^2 = \sum_{i=1}^k (x_i - y_i)^2$$

Angle Negative Angle Between Image Vectors

$$d(x, y) = -\frac{x \cdot y}{\|x\| \|y\|} = -\frac{\sum_{i=1}^k x_i y_i}{\sqrt{\sum_{i=1}^k (x_i)^2 \sum_{i=1}^k (y_i)^2}}$$

Mahalanobis Mahalanobis Distance

$$d(x, y) = -\sum_{i=1}^k \frac{1}{\sqrt{\lambda_i}} x_i y_i$$

where λ_i is the i th Eigenvalue corresponding to the i th Eigenvector.

For simplicity, in this paper we choose the block distance measure and use Nearest Neighbor classifier.

4. EXPERIMENTS

The experimental results to be presented in this section are obtained using the AR-face database [23]. This database consists of over 3200 color images of the frontal faces of 126 subjects. There are 26 different images for each subject. For each subject, these images were recorded in two different sessions separated by two weeks, each session consisting of 13 images. Each image in the database consists of a 768×576 array of pixels.

In this experiment, we choose the first 7 face images of the first session by eliminating occluded face images for each subject, as shown in Figure5. Then we have 126×7 face images. We observe that the size of the face part is almost above 256×192 . We manually locate the centers of eye and then perform geometric normalization with the eye locations fixed to get geometric normalized face image with the size 256×192 . One example of normalized faces is in Figure6. Then we use the generalized Gaussian Pyramid to generate the multi-resolution face images as shown Figure1 and Figure2.

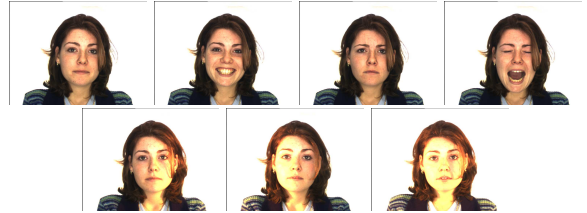


Fig. 5. The first 7 faces of the first session in the AR face database for one subject

The Relation Curve Between Face Recognition Performance and Face Image Resolution

In this experiment, we use a leave-one-out methodology to obtain statistical results. We randomly choose one of seven images for each subject as the testing part and the remainders as the training part. Because this process is quite time-consuming, we performed the randomly choosing method seven times with assuring that iff each of seven face images is chosen once as the testing face image. We test PCA and LDA face recognition methods with Nearest Neighbor



Fig. 6. The normalized faces

classifier. The result is depicted in Figure7. From the curve changing trend, we can observe that the threshold resolution is about 64×48 , i.e., before 64×48 the recognition rate increases roughly with the increasing resolution and when reaching 64×48 , the rate keeps stable when the resolution increasing.

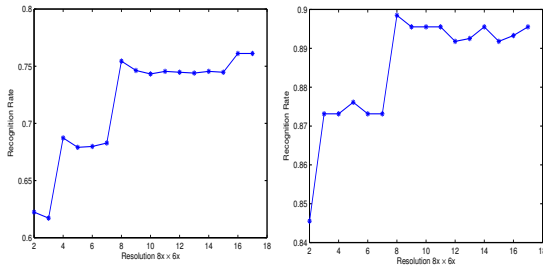


Fig. 7. Performance curve with the resolution changing from 16×12 to 136×102 based on PCA (the left figure) and LDA (the right figure)

Structure Information versus Face Recognition Performance

We are also concerned about how the enhanced structure information influences recognition performance. So we expand the face images with the resolution 16×12 , 32×24 and 64×48 to the corresponding higher resolution face images with the resolution 32×24 , 64×48 and 128×96 by the bilinear interpolation¹. Then we perform PCA and LDA based classifiers while keeping the energy ratio 90% and keeping the eigenvector number fixed.

From Figure8, we have the following two observations: (1) the recognition rate of expanded faces will decrease in comparison with unexpanded faces when keeping the energy ratio fixed and keeping the eigenvector number fixed

¹Strictly speaking, the bilinear interpolation only adds a little structure information. A reasonable method to enhance the structure information should incorporate the prior information of the higher resolution faces.

using PCA, and (2) for LDA when keeping the energy ratio fixed the recognition rate will decrease, however when keeping the eigenvector number fixed the recognition will increase a little and still be less than the near higher resolution face recognition rate (64×48 is an exception because recognition rate in unexpanded 64×48 is higher than unexpanded 128×96 , which is seen in Figure7). From the two observations, we conclude that only enhancing the structure information does not help to improve recognition performance.

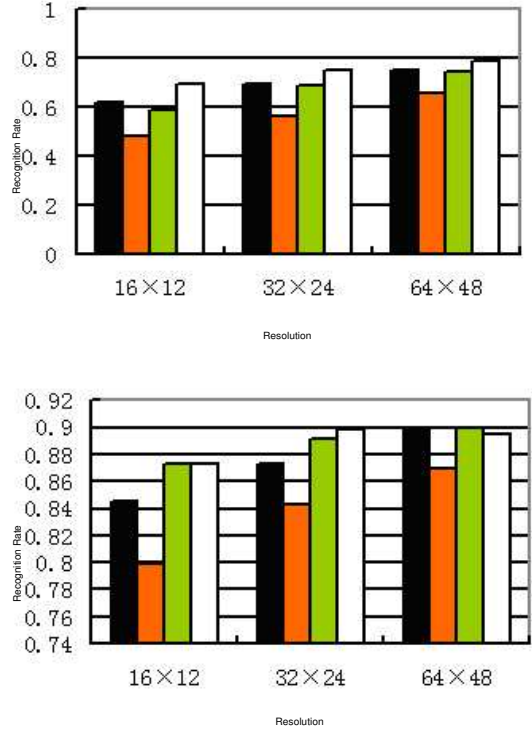


Fig. 8. Performance comparison. The first column of each resolution represents real $h \times w$, the second represents expanded $2h \times 2w$ while keeping the energy ratio fixed, the third represents expanded $2h \times 2w$ while keeping the eigenvector number fixed and the fourth represents real $2h \times 2w$. The top figure is for PCA, and the bottom for LDA.

5. CONCLUSION

In this paper, we first analyze the relation between face recognition rate and face image resolution by dividing face image information into the discriminative and structure information. Then, we test two global face recognition algorithms on the AR-face database to obtain the relation curve, which validates our conjecture and also gives the threshold resolution of 64×48 .

The statistical results help the researchers on face recognition to choose the appropriate resolution face image. On the other hand, they also help to understand face hallucination algorithm proposed in [20] and provide a good advice that the global model embedding a middle-resolution face image and local model built from the middle-resolution to the high-resolution should be more reasonable.

6. REFERENCES

- [1] W. Zhao, R. Chellappa, A. Rosenfeld and P. Phillips, "Face Recognition: a literature survey," <http://citeseer.nj.nec.com/374297.html>, 2000, 67 pages.
- [2] P. N. Belhumeur, J.P. Hespanha and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE Trans. on PAMI, 1997.
- [3] A. M. Martinez and A. C. Kak, "PCA versus LDA," IEEE Trans. on PAMI, 2001.
- [4] B. A. Draper, K. Baek, M. S. Barlett and J.R. Beveridge, "Recognizing faces with PCA and ICA," Computer Vision and Image Understanding. To appear.
- [5] M. Fleming and G. Cottrell, "Categorization of faces using unsupervised feature extraction," In Proc. IEEE IJCNN International Joint Conference on Neural Networks, 1990.
- [6] M. L. Teixeira and J. R. Beveridge, "An Implementation and Study of the Moghaddam and Pentland Intrapersonal/Extrapersonal Image Difference Face Recognition Algorithm," CSU Computer Science Department Technical Report, May 2003.
- [7] B. Moghaddam, W. Wahid and A. Pentland, "Beyond Eigenfaces: Probabilistic Matching for Face Recognition," The 3rd IEEE Int'l Conference on Automatic Face and Gesture Recognition 1998.
- [8] X. Xu, "Active Morphable Model for the Analysis and Synthesis of Human Faces," Master Thesis, 2003
- [9] A. Lanitis, C. Taylor and T. Cootes, "Automatic interpretation and coding of face images using flexing models," IEEE Trans. on PAMI, 1997.
- [10] K. Jonsson, J. Matas, J. Kittler and Y. Li, "Learning support vector vectors for face verification and recognition," In Proc. IEEE International Conference on Automatic Face and Gesture Recognition 2000.
- [11] R. Brunelli and T. Poggio, "Face Recognition: Features versus templates," IEEE Trans. on PAMI, 1993.
- [12] D.J. Beymer, "Face recognition under varying pose," A.I. Memo 1461, Center for Biological and Computational Learning, MIT, Cambridge, MA, 1993.
- [13] L. Wiskott, J. Fellous, N. Kruger and C. von der malsburg, "Face recognition by elastic bunch graph matching," IEEE Trans. on PAMI, 1997.
- [14] A. Nefian and M. Hayes, "An embedded HMM-based approach for face detection and recognition," In Proc. IEEE international Conference on Acoustics, Speech, and Signal Processing, 1999.
- [15] B. Heisele, P. Ho and T. Poggio, "Face recognition with support vector machines: Global versus component-based approach," In Proc. IEEE International Conference on Computer Vision, 2001.
- [16] G. Givens, J. R. Beveridge, B. A. Draper and D. Bolme, "A Statistical Assessment of Subject Factors in the PCA Recognition of Human Faces," CVPR 2003 Workshop on Statistical Analysis in Computer Vision Workshop, June 2003.
- [17] W. S. Yambor, B. A. Draper and J.R. Beveridge, "Analyzing PCA-based face recognition algorithms: eigenvector selection and distance measures," in Empirical Evaluation Methods in Computer Vision, H. Christensen and J. Phillips (eds.), World Scientific Press, Singapore, 2002.
- [18] T. Shkunaga and K. Shigenari, "Decomposed eigenface for face recognition under various lighting conditions," IEEE Trans. on PAMI, 1993.
- [19] S. Baker and T. Kanade, "Hallucinating Faces," The fourth International Conference on Automatic face and Gesture Recognition, 2000.
- [20] C. Liu, H.Y. Shum and C.S. Zhang, "A Two-Step Approach to Hallucinating Faces: Global Parametric Model and Local nonparametric Model," In Proc. IEEE International Conference on Computer Vision and Pattern Recognition, 2001.
- [21] P. J. Burt, E. H. Adelson, "The Laplacian Pyramid as a Compact image Code," IEEE Trans. on Communications, 1983.
- [22] A. Pentland, T. Starner, N. Etcoff, N. Masouli, O. Oliyide and M. Turk, "Experiments with eigenfaces," Looking at people, workshop of IJCAI, 1993.
- [23] A.M. Martinez and R. Benavente, "The AR-face database," CVC Technical Report 24, June, 1998.