Face Hallucination and Recognition

Xiaogang Wang and Xiaoou Tang

Department of Information Engineering, The Chinese University of Hong Kong
{xgwang1, xtang}@ie.cuhk.edu.hk
http://mmlab.ie.cuhk.edu.hk

Abstract. In video surveillance, the faces of interest are often of small size. Image resolution is an important factor affecting face recognition by human and computer. In this paper, we study the face recognition performance using different image resolutions. For automatic face recognition, a low resolution bound is found through experiments. We use an eigentransformation based hallucination method to improve the image resolution. The hallucinated face images are not only much helpful for recognition by human, but also make the automatic recognition procedure easier, since they emphasize the face difference by adding some high frequency details.

1 Introduction

In video surveillance, the faces of interest are often of small size because of the great distance between the camera and the objects. Image resolution is a potential factor affecting face recognition performance. In the low-resolution face images, many detailed facial features are lost and faces are indiscernible to human. We also notice that in many automatic face recognition systems, face images are down sampled to small size, and also achieve satisfied performance. But how will the image resolution affect recognition accuracy is still open to discussion.

Several algorithms have been proposed to render a high-resolution face image from the low-resolution one. This technique is called hallucination [4]. Since face images are well structured and have similar appearance, they span a small subset in the high dimensional image space [3]. This implies that the high frequency detail can be inferred from the low frequency components, utilizing the face structural similarity.

The simplest way to increase resolution is direct interpolation of input images with such algorithms as nearest neighbour, cubic spline. But its performance is poor if the image size is too small. Baker and Kanade [4] develop a hallucination method based on the property of face image. It infers the high frequency component from a parent structure by recognizing the local features from the training set. Liu et. al. [1] develop a two-step statistical modeling approach integrating global and local parameter models. Hallucination has effectively improved the resolution of face images thus makes it much easier for a human being to recognize a face. However, how much informa-

J. Kittler and M.S. Nixon (Eds.): AVBPA 2003, LNCS 2688, pp. 486-494, 2003.

tion has been extracted from the low-resolution image by the hallucination process and its contribution to automatic face recognition has not been studied in previous works.

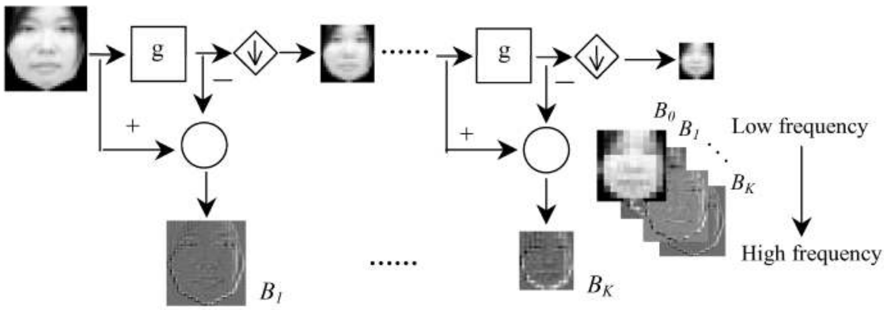


Fig. 1. Multi-resolution analysis in spatial domain. g is the smoothing function, and, are different frequency bands

In this paper, we study the face recognition performance using different image resolutions. We use a novel hallucination method based on eigentransformation [6]. It is closely related to the work in [5], in which an eigentransformation approach was developed for sketch recognition. In our method, PCA is applied to the low-resolution face image. In the PCA space, different frequency components are independent. By selecting the number of eigenfaces, we could extract the maximum amount of facial information from the low-resolution face image and remove the noise. The new hallucinated face image is rendered by mapping between the low- and high- resolution training pairs. We also study the effect of hallucination on automatic face recognition. Since hallucination emphasizes the face difference by adding some high frequency details, it may help the automatic recognition process. Experiments are conducted on a database containing images of 188 people and the XM2VTS face database [2].

2 Multiresolution Analysis

Viewing a 2D image as a vector, the process of getting a low-resolution face image from the high-resolution face image can be formulated as

$$\vec{I}_l = H\vec{I}_h + \vec{n} . \tag{1}$$

Here, $\vec{I}_h$ and $\vec{I}_l$ represent the high- and low-resolution face image vectors respectively. H is the transformation matrix involving blurring and downsampling process, and $\vec{n}$ is the noise perturbation to the low-resolution face image captured by camera.

As shown in Figure 1, a process of iterative smoothing and downsampling decomposes the face image into different bands, $B_0,\ldots,B_K$. In this decomposition, different frequency bands are not independent. Some components of the high-frequency bands, $B_1,\ldots,B_K$, can be inferred from the low frequency band B_0 . This is a starting point for hallucination. Many super-resolution algorithms assume the dependency as homogeneous Markov Random Fields (MRFs), i.e. the pixel only relies on the pixels in its neighborhood. This is an assumption for general images. It is not optimal for the face class without considering face stuctural similarity. A better way to address the dependency is using PCA, in which different frequency components are independent.

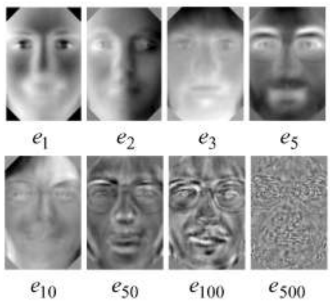


Fig. 2. Eigenfaces sorted by eigenvalues. e_i is the ith eigenface

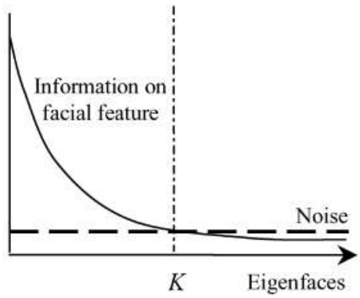


Fig. 3. Extract facial information in the PCA space of low-resolution face images

3 Hallucination and Recognition

Face image can be reconstructed from some eigenfaces in the PCA representation. PCA also decomposes face image into different frequency components, but encoding facial information in a more compact way, since it takes into account of the face distribution. Our algorithm first employs PCA to extract as much useful information from low-resolution face image as possible, and then renders a high-resolution face image by eigentransformation. A detailed description for eigentransformation can be found in [5].

3.1 Principle Component Analysis

We represent a set of face images by a N by M matrix, $\left[\vec{l}_1,\ldots\vec{l}_M\right]$, where $\vec{l}_i$ is the image vector, N is the number of image pixel, and M is the number of the training samples ($N >> M$). In PCA, a set of eigenvectors $E_l = \left[e_1,\ldots,e_K\right]$, also called eigenfaces, are computed from the ensemble covariance matrix,

$$C = \sum_{i=1}^{M}\left(\vec{l}_i - \vec{m}_l\right)\left(\vec{l}_i - \vec{m}_l\right)^T = LL^T \tag{2}$$

where $\vec{m}_l$ is the mean face computed from the sample set, and L is the sample matrix,

$$L = \left[\vec{l}_1 - \vec{m}_1,\ldots,\vec{l}_M - \vec{m}_M\right] = \left[\vec{l}\,'_1,\ldots\vec{l}\,'_M\right]. \tag{3}$$

For a face image $\vec{x}_l$, a weight vector is computed by projecting it onto eigenfaces,

$$\vec{w}_l = E_l^T(\vec{x}_l - \vec{m}_l). \tag{4}$$

This is a face representation based on eigenfaces. A face can be reconstructed from the *K* eigenfaces,

$$\vec{r}_l = E_l \vec{w}_l + \vec{m}_l. \tag{5}$$

Figure 2 shows some eigenfaces sorted by eigenvalues. Eigenfaces with large eigenvalues are “face-like”, and characterize low frequency components. Eigenfaces with small eigenvalues are “noise-like”, and characterize high frequency details.

3.2 Eigentransformation

Given the low-resolution sample set *L*, according to singular value decomposition theorem, E_l also can be computed from,

$$E_l = L V_l \Lambda_l^{-1/2}, \tag{6}$$

where V_l and Λ_l are the eigenvector and eigenvalue matrix for $L^T L$. From (5) and (6), the reconstructed face image can be represented by

$$\vec{r}_l = L V_l \Lambda_l^{-1/2} \vec{w}_l + \vec{m}_l = L\vec{c} + \vec{m}_l, \tag{7}$$

where $\vec{c} = V_l \Lambda_l^{-1/2} \vec{w}_l = [c_1, c_2, \ldots, c_M]^T$. Equation (7) can be rewritten as,

$$\vec{r}_l = L\vec{c} + \vec{m}_l = \sum_{i=1}^{M} c_i \vec{l}'_i + \vec{m}_l. \tag{8}$$

This shows that the input low-resolution face image can be reconstructed from the optimal linear combination of the *M* low-resolution training face images. Replacing each low-resolution image $\vec{l}'_i$ by its high-resolution sample $\vec{h}'_i$, and replacing $\vec{m}_l$ with the high-resolution mean face $\vec{m}_h$, we get $\vec{x}_h$, which is expected to be an approximation to the real high-resolution face image.

3.3 Recognition

In our algorithm, the hallucinated face image is synthesized by the linear combination of high-resolution training images and the coefficients come from the low-resolution face images using the PCA method. Because of the structural similarity among face images, in multiresolution analysis, there exists strong correlation between the high frequency band and low frequency band. For high-resolution face images, PCA can compact these correlated information onto a small number of principle components. Then, in the eigentransformation process, these principle components can be inferred from those of the low-resolution face image by mapping between the high- and low-resolution training pairs. Therefore, some information in the high frequency band bands are partially recovered.

In practice, the low-resolution image is often disturbed by noise which has a flat distribution on all the axes. For low-resolution face images, the energy on small eigenvectors is small, thus is overwhelmed by noise. By selecting an optimal eigenface number *K*, we can extract the facial information and remove the noise. The information on these noisy components (eigenfaces after *K* in Fig. 3) is lost, and cannot be recovered since the components on different eigenvectors are independent in PCA space. In this sense, our hallucination method has extracted the maximum amount of facial information exists in the low-resolution face images.

Given the significant improvement of the face appearance by the hallucination process, it is interesting to investigate whether the hallucination helps automatic recognition. Since more high frequency details are recovered, we expect the ballucination process to help the recognition performance.

4 Experiment

4.1 Hallucination Experiment

Our hallucination experiment is conducted on a data set containing 188 individuals with one face image for each individual. Using the "leave-one-out" methodology, at each time, one image is selected for testing and the remaining are used for training. In preprocessing, the face images are aligned by the two eyes. The distance between the eye centers is fixed at 50 pixels, and the image size is fixed at 117×125. Images are blurred by averaging neighbour pixels and down sampled to low-resolution images. Here, we use the eye center distance de to measure the face resolution.

Some hallucination results are shown in Fig. 4. The input face images are down sampled to 23×25, with de equal to 10. Compared with the Cubic B-Spline interpolation result, the hallucinated face images have much clearer detail features. They are good approximation to the original high-resolution images.

Figure 5 reports the hallucination performance for different input resolutions. The eye center distance is down sampled to 20, 10, 7, and 5. Figure 6 repots the average RMS error per pixel in intensity for the 188 face images. Under a very low resolution, the low-resolution and direct interpolated face images are almost indiscernible, and the RMS error of Cubic B-spline interpolation increases quickly. The performance of hallucination by eigentransformation is much better. When de is down sampled to 10, the result of eigentransformation is still satisfactory. For further lower resolutions, there are some distortions on the eyes and mouth.

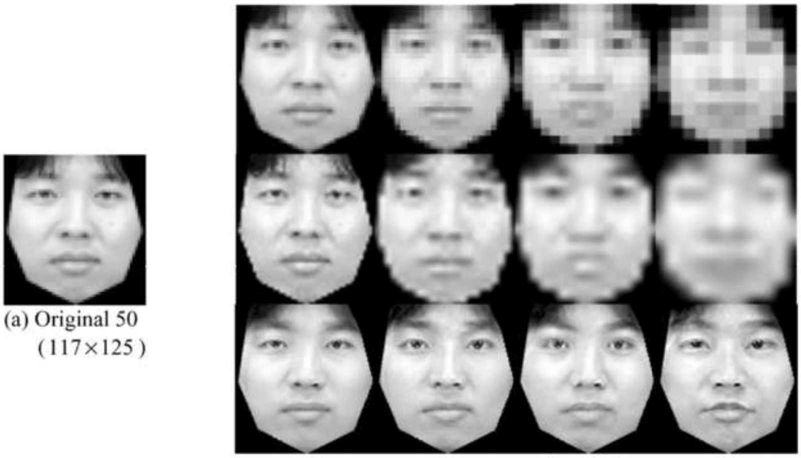
As discussed in Section 3, some high frequency detail is lost in the process of blur and downsampling, or is overwhelmed by noise. Selecting the eigenface number in eigentransformation, we could control the detail level by keeping maximum facial information while removing the noise. This point can be illustrated in the experiment reported by Figure 7. We add zero mean, white Gaussian noise with five different standard deviations (σ) to the low-resolution face image, and then use different eigenface number (*K*) for hallucination. The optimal eigenface number decreases as the increase of noise. Using 180 eigenfaces, the hallucinated face images are noisy and distorted for all the five levels of noise. When *K* is reduced to 100, face images under small noise ($\sigma = 0.03, 0.05$) are well hallucinated. but results under more noise ($\sigma = 0.07, 0.1, 0.12$) have a larger distortion. Using 50 eigenfaces, all of the images show little noise effect. So eigenface number can control the detail level to make the hallucinated face images robust to noise.

Fig. 4. Hallucinated face images by eigentransformation

4.2 Recognition Experiment

We study the recognition performance using low-resolution face images and hallucinated face images. Two hundred and ninety five individuals from the XM2VTS face database are selected, with two face images in different sessions for each individual. One image is used as reference, and the other is used for testing. We use direct correlation for recognition, which is perhaps the simplest face recognition algorithm. The recognition accuracies over different resolutions are plotted in Figure 8. When de is reduced from 50 to 10, there is only slight fluctuation on recognition accuracy using low-resolution face images. When de is further reduced to 7 and 5, the recognition accuracy for low-resolution face images drops greatly. Resolution with de equal to 10 is perhaps a lower bound for recognition. Below this level there may not be enough information for recognition. This is also consistent with the hallucination experiment in 4.1. Satisfactory hallucination results can be obtained when de is larger than 10.

We also try to explore whether hallucination can contribute to automatic face recognition. We expect hallucination make the recognition procedure easier, since it emphasizes the face difference by adding some high frequency details. In this experiment, the low-resolution testing image is hallucinated by reference face images, but the face image of the testing individual is excluded from the training set. As shown in Figure 8, the hallucination improved the recognition accuracy when the input face images have very low resolutions.



(a) Original 50 (117×125)

20 (47×50) 10 (23×25) 7 (16×17) 5 (11×12)

(b) The first row is the input face images, for which *de* is 20, 10, 7, 5 respectively; the second row is the hallucinated face images.

Fig. 5. Hallucinated face images using input images of different resolutions

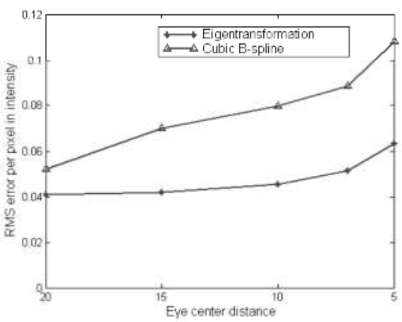


Fig. 6. RMS error per pixel in intensity using Cubic-spline interpolation and hallucination by eigentransformation The intensity is between 0 and 1

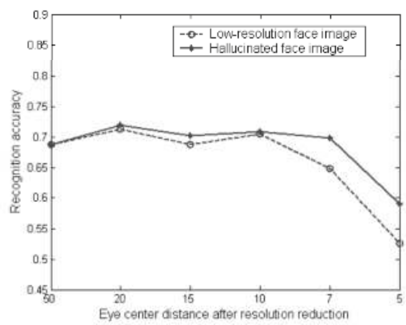


Fig. 8. Recognition accuracy using low-resolution face images and hallucinated face images based on XM2VTS database

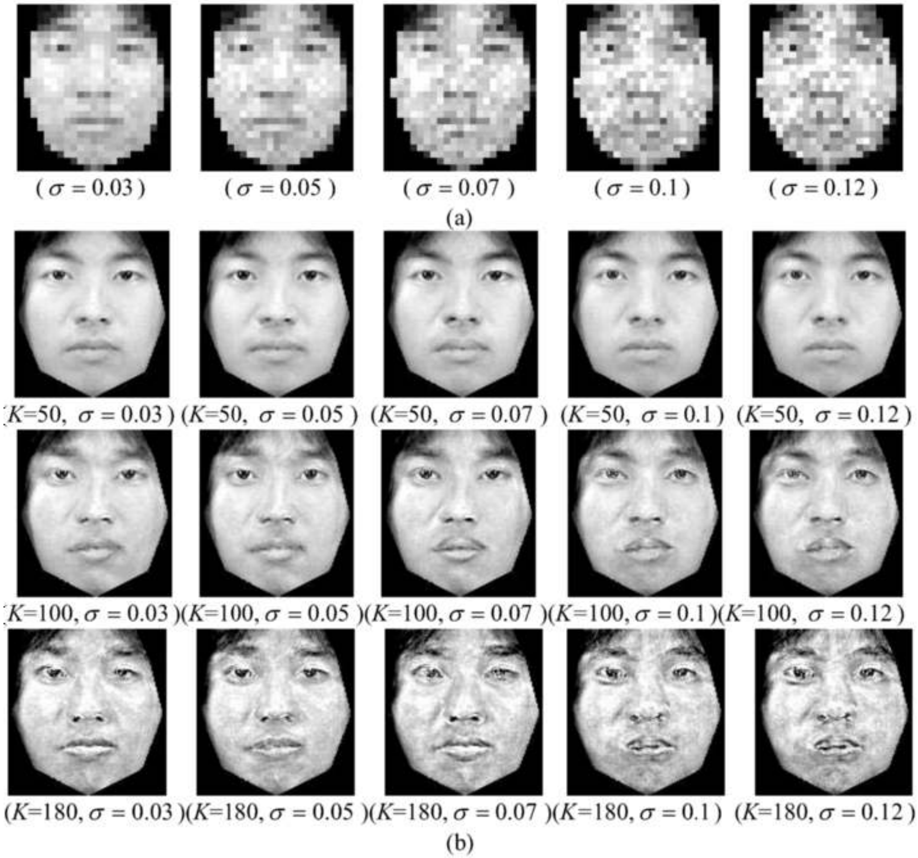


Fig. 9. RMS error per pixel in intensity using Cubic-spline interpolation and halluci-nation by eigentransformation The intensity is between 0 and 1

5 Conclusion

Our hallucination method based on eigentransformation could extract the maximum facial information from the low-resolution face images and render some high frequency facial feature to make the face image more discernible. It also makes the automatic face recognition more easier. We also study the face recognition performance over different resolutions. A low resolution bound for recognition is found in the experiment. This is only a preliminary study. The results need to be further confirmed using more face recognition algorithms and data sets.

Acknowledgement

This work was supported by the Research Grants Council of the Hong Kong SAR under Projects CUHK 4190/01E and AOE/E-01/99.

References

[1] C. Liu, H. Shum, and C. Zhang, "A Two-Step Approach to Hallucinating Faces: Global Parametric Model and Local Nonparametric Model, "*Proc. of IEEE International Conference on Computer Vision and Pattern Recognition",* pp. 192-198, 2001.

[2] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "XM2VTSDB: The Extended M2VTSDB," In the Second International Conference on Audio and Video-Based Biometric Person Authentication, pp. 72-77, March 1999.

[3] P. S. Penev, and L. Sirovich, "The Global Dimensionality of Face Space,"*Proc. of IEEE International Conference on Automatic Face and Gesture Recognition",* pp. 264-270, 2000.

[4] S. Baker, and T. Kanade, "Hallucinating Faces," Proceedings IEEE International Conference on Automatic Face and Gesture Recognition, pp. 83-88, 2000.

[5] X. Tang, and X. Wang, "Face Photo Recognition Using Sketch," *Proc. of ICIP,* 2002.

[6] X. Wang and X. Tang, "Hallucinating Face by Eigentransformation," *ICIP 2003*.