THE HONG KONG
POLYTECHNIC UNIVERSITY
香港理工大學
Pao Yue-kong Library
包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

Low-Resolution Face Recognition

QU, Tong

Master of Science in Electronic and Information Engineering

The Hong Kong Polytechnic University

August 2013

Statement of Authorship

Except where reference is made in the text of this dissertation, this dissertation contains no material published elsewhere or extracted in whole or in part from a dissertation presented by me for another degree or diploma.

No other person’s work has been used without due acknowledgement in the main text of the dissertation.

This dissertation has not been submitted for the award of any other degree or diploma in any other tertiary institution.

Qu Tong

Name: QU, Tong

Dated: 27-August-2013

Dedication

To my parents

Abstract

Abstract of dissertation entitled: Low-Resolution Face Recognition submitted by QU, Tong for the degree of MSc in Electronic and Information Engineering at The Hong Kong Polytechnic University in August 2013.

Among face-recognition (FR) problems, the identification of low-resolution (LR) face images is still a challenging task. Traditional FR algorithms cannot work satisfactorily in matching LR probe images to high-resolution (HR) gallery images. To perform this matching, there are three standard approaches: (1) down-sample the gallery images and then perform the matching of LR face images; (2) upscale the probe images using super-resolution (SR) methods and then perform the matching of HR face images; and (3) project the LR probe images and the HR gallery images into a common subspace and then perform matching in the subspace. In this project, traditional algorithms based on the first two approaches will first be introduced and evaluated under different resolutions. The four baseline FR algorithms are PCA, also known as eigenfaces, combined PCA and LDA (PCA+LDA, a variant of fisherfaces), the PCA+LDA-based FR algorithm based on Gabor features (G-PCA+LDA), and LGBPHS. The three baseline SR algorithms are the bicubic interpolation, eigentransformation and Coherent Local Linear Reconstruction Super-resolution (CLLR-SR). After that, a coupled-projection method based on Canonical Correlation Analysis (CCA) is proposed and evaluated. Experiments show that the coupled-projection method produces higher identification rates than other FR methods do.

Acknowledgements

Many people have offered me indispensable help to make this dissertation project possible. I would like to thank my project supervisor Prof. Kenneth Kin-Man Lam for his offering me the opportunity to work on this prestigious topic and subsequently providing me very useful guidance and instructions as well as encouragement throughout my study on this project. Without his precious help, I could not have been able to accomplish this dissertation. Also, I would like to express my gratitude to the Department of Electronic and Information and Engineering for the arrangements and support for carrying out this project.

Table of Contents

List of Figures

List of Tables

List of Acronyms

FR	Face Recognition, Face-Recognition
LR	Low-Resolution
HR	High-Resolution
SR	Super-Resolution
PCA	Principal Component Analysis
LDA	Linear Discriminant Analysis
GMP	Gabor Magnitude Picture
LBP	Local Binary Pattern
LGBP	Local Gabor Binary Pattern
LGBPHS	Local Gabor Binary Pattern Histogram Sequence
CCA	Canonical Correlation Analysis
LLE	Local Linear Embedding
CLLR	Coherent Local Linear Reconstruction
CLLR-SR	Coherent Local Linear Reconstruction Super-Resolution
CLDA	Coherent Linear Discriminant Analysis

Chapter 1 Introduction

1.1 Background

Face recognition (FR) has received great interest and achieved impressive success during the past three decades[1], especially in applications under controlled environment. However, the performance of face recognition in uncontrolled environment is still far from being satisfactory. Usually, these environments involve subjects that are uncooperative or unaware that an image is being captured. In some scenarios, it can be desirable to photograph subjects without their active participation or knowledge, such as in video surveillance. In cases where the image acquisition environment is not ideal, captured faces can have a much lower resolution than faces captured in a controlled setting because of the large distance between the camera and the objects.

Face recognition based on low-resolution (LR) images is a very challenging problem because compared with high-resolution (HR) images, LR ones lose much facial information which is important for high recognition accuracy. Hence, it is impossible for traditional face recognition algorithms to compare LR images captured at testing to HR images directly, as the LR and HR images do not share a common feature representation.

To deal with the problem, there are three standard approaches: (1) down-sample the entire gallery and then perform matching in LR feature space; (2) use super-resolution (SR) or interpolation methods to reconstruct the HR version of the LR probe and then perform matching in the usual way in HR feature space; and (3) project the HR gallery images and LR probe images into a common subspace and then perform matching in the common feature space. Figure 1-1 shows the relationship among the three standard approaches.

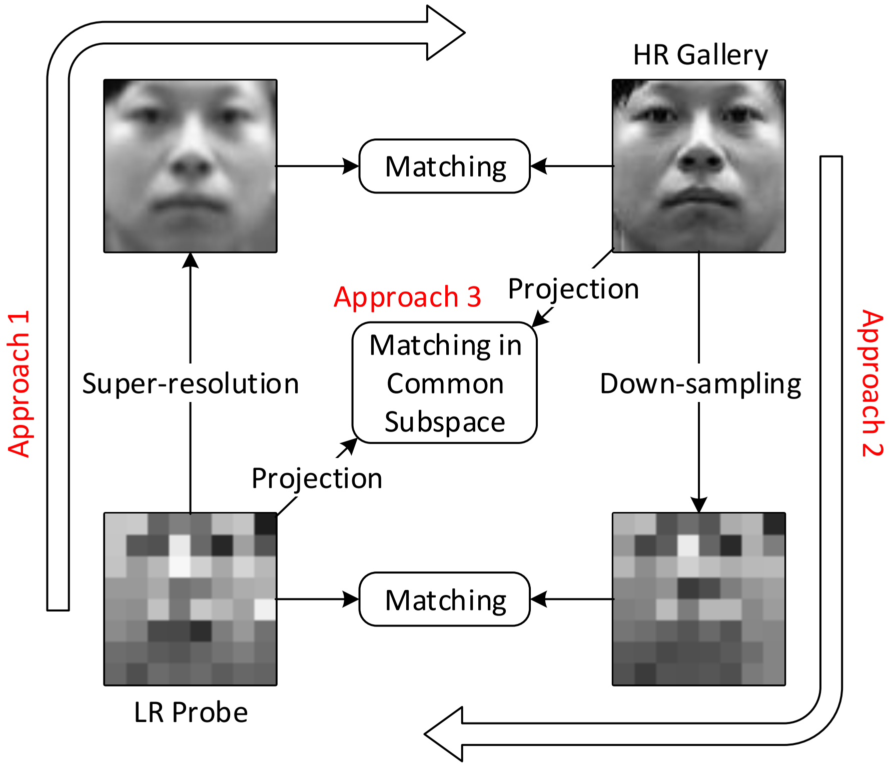


Figure 1-1 Standard approaches to matching a LR probe to a HR gallery.

(1) Down-sampling the gallery and then matching; (2) super-resolving the probe and then matching; (3) projecting LR probe and HR gallery into a common subspace and then marching.

Adapted from [2].

Reducing the size of the gallery images is the simplest approach to matching LR probe images. Once gallery images are down-sampled, traditional linear projection methods for face recognition, including Principal Component Analysis (PCA)[3] and Linear Discriminant Analysis (LDA)[4], can be used. While these approaches can be effective for comparing images of the same resolution, down-sampling the gallery images needlessly discards information in the data.

The opposite approach is to increase the resolution of the probe images. Performing some form of SR makes the probe images the same dimensionality as the gallery images, and standard single resolution FR methods can once again be used. Simple methods such as bilinear or bicubic interpolation do not require any training. Other SR methods can be trained to learn the relationship between HR and LR images. These approaches can be effective for reconstructing HR images, and while they can produce visually appealing results, they often lack the high frequency components of true HR images to be very effective for recognition tasks.

1.2 Overview

In this project, four baseline FR algorithms are firstly introduced in Chapter 2, and they are evaluated in subsequent chapters. The four baseline FR algorithms are Principal Component Analysis (PCA)[3], combined PCA and Linear Discriminant Analysis (PCA+LDA)[5], the PCA+LDA algorithm based on Gabor features (G-PCA+LDA)[6], and the Local Gabor Binary Pattern Histogram Sequence (LGBPHS)[7]. The PCA- and the PCA+LDA-based face recognition algorithms are both fundamental and have been well-studied. 2-D Gabor wavelets[8] and local binary patterns (LBP)[9] are extensively used for local feature representation and extraction, and demonstrate their success in face recognition. Therefore, the PCA+LDA algorithm based on Gabor features and the LGBPHS are also used as baseline algorithms. Furthermore, in contrast to the other three baseline algorithms, LGBPHS is not a statistical-learning-based method. As a result, it will not be tuned to a specific training set, and dose not suffer from the generalizability problem.

The three baseline SR methods evaluated are bicubic interpolation, eigentransformation[10], and Coherent Local Linear Reconstruction (CLLR)[11]. The bicubic interpolation method is widely used for image resampling, which does not require any training, while eigentransformation and CLLR are methods dedicated to hallucinating face images by learning the structural relationship between LR and HR training samples, which are detailed in Chapter 3. Their algorithms are closely related to PCA, and CLLR is also based on Canonical Correlation Analysis (CCA)[12] and Local Linear

Embedding (LLE)[13]. The experimental results in Chapter 5 show that the example-based super-resolution methods outperform the bicubic interpolation for face recognition purpose.

Experiments are conducted to test the above mentioned FR and SR algorithms on different face image resolutions. By comparing results of the first standard approaches and the second standard approaches described in Section 1.1, we can find out that for most FR algorithms, using SR methods to upscale the LR probe image does not actually increase the identification rate, since the these methods are aimed at improve the visual quality of the images rather than improving the classification performance. The super-resolved images often lack the high frequency components of true HR images to be very effective for recognition tasks.

Therefore, in Chapter 4, a coupled-projection method is proposed: Coherent Linear Discriminant Analysis (CLDA). This method is inspired by the CLLR super-resolution algorithm and CCA is used as a tool to project the LR and HR images into a common subspace.

1.3 Objectives and Organization

In this project, the existing FR algorithms on the LR face images are first evaluated. Then, the performances of SR methods on FR are evaluated. Finally, we consider one simple coupled-projection method for LR probe identification and evaluate the performances of them.

This dissertation is organized as follows: After a brief introduction to the background and objectives of the project in Chapter 1, Chapter 2-3 introduce the baseline FR algorithms and the baseline SR algorithms. Chapter 4 presents a coupled-projection method for the recognition of the LR facial images, which are inspired by the methods introduced in Chapter 2-3. Chapter 5 proposes an evaluation protocol and illustrates the evaluation results of the algorithms proposed in Chapter 2-4. Finally, Chapter 6 summarizes all the evaluation results and draws a conclusion for this project.

Chapter 2 Baseline Face Recognition Algorithms

2.1 Principal Component Analysis (PCA)

PCA, also known as the eigenfaces[3] is commonly used for dimensionality reduction in face recognition. PCA chooses projection directions $\boldsymbol{W}_{PCA}$ that maximize the total scatter across all images of all faces in the training set.

Let us consider a training set that contains N sample images $\{x_1, x_2, \dots, x_N\} \in \mathbb{R}^n$, and assume that each image belongs to one of the C classes $\{X_1, X_2, \dots, X_C\}$. Then, consider a linear transformation mapping the original n-dimensional image space into an m-dimensional feature space, where $m < n$. The new feature vectors $y_k \in \mathbb{R}^m$ are defined by the following linear transformation:

$$y_k = \boldsymbol{W}^T x_k,\ k = 1, 2, \dots, N, \tag{1}$$

where $\boldsymbol{W} \in \mathbb{R}^{n \times m}$ is a matrix with orthonormal columns.

The total scatter matrix $\boldsymbol{S}_T$ is defined as follows:

$$\boldsymbol{S}_T = \frac{1}{N} \sum_{k=1}^{N} (x_k - \mu)(x_k - \mu)^T, \tag{2}$$

where $\mu \in \mathbb{R}^n$ is the mean vector of all sample images in the training set.

After applying the linear transformation $\boldsymbol{W}^T$, the scatter of the transformed feature vectors $\{y_1, y_2, \dots, y_N\}$ is $\boldsymbol{W}^T \boldsymbol{S}_T \boldsymbol{W}$. Then, the projection matrix $\boldsymbol{W}_{PCA}$ can be chosen as follows:

$$\boldsymbol{W}_{PCA} = \arg\max_{\boldsymbol{W}} |\boldsymbol{W}^T \boldsymbol{S}_T \boldsymbol{W}| = [\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad \cdots \quad \boldsymbol{w}_m], \tag{3}$$

where $\{\boldsymbol{w}_i | i = 1, 2, \ldots, m\}$ is the set of n-dimensional eigenvectors of $\boldsymbol{S}_T$ corresponding to the m largest eigenvalues. In most circumstances, m can be chosen far less than n without significantly decreasing the recognition rates. Then, classification is performed using a nearest neighbour classifier in the reduced feature space.

2.2 Combined PCA and LDA (PCA+LDA)

2.2.1 Linear Discriminant Analysis (LDA)

LDA[4],[5] is a widely used method for feature extraction and dimensionality reduction in pattern recognition and has been proposed in face recognition. LDA tries to find the "best" project direction in which training samples belonging to different classes are best separated. Mathematically, it selects the projection matrix $\boldsymbol{W}_{LDA}$ in such a way that the ratio of the determinant of the between-class scatter matrix $\boldsymbol{S}_B$ of the projected samples and the within-class scatter matrix $\boldsymbol{S}_W$ of the projected samples is maximized.

Let the between-class scatter matrix be defined as

$$\boldsymbol{S}_B = \sum_{i=1}^{C} N_i (\mu_i - \mu)(\mu_i - \mu)^T, \tag{4}$$

and the within-class scatter matrix be defined as

$$\boldsymbol{S}_W = \sum_{i=1}^{C} \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T, \tag{5}$$

where μ_i is the mean image of class X_i, and N_i is the number of samples in class X_i. If $\boldsymbol{S}_W$ is non-singular, the projection matrix $\boldsymbol{W}_{LDA}$ is chosen as the matrix with orthonormal columns which maximizes the ratio of the projected samples to the determinant of the within-class scatter matrix of the projection samples, i.e.,

$$\boldsymbol{W}_{LDA} = \arg\max_{\boldsymbol{W}} \frac{|\boldsymbol{W}^T \boldsymbol{S}_B \boldsymbol{W}|}{|\boldsymbol{W}^T \boldsymbol{S}_W \boldsymbol{W}|} = [\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad \dots \quad \boldsymbol{w}_m], \tag{6}$$

where $\{\boldsymbol{w}_i | i = 1, 2, \dots, m\}$ is the set of generalized eigenvectors of $\boldsymbol{S}_B$ and $\boldsymbol{S}_W$ corresponding to the m largest generalized eigenvalues $\{\lambda_i | i = 1, 2, \dots, m\}$, i.e.,

$$\boldsymbol{S}_B \boldsymbol{w}_i = \lambda_i \boldsymbol{S}_W \boldsymbol{w}_i, \ i = 1, 2, \dots, m. \tag{7}$$

Note that there are at most $C - 1$ nonzero generalized eigenvalues, and so an upper bound on m is $C - 1$ where C is the number of classes.

2.2.2 PCA+LDA Method

In the FR problem, one difficulty is that the within-class scatter matrix $\boldsymbol{S}_W \in \mathbb{R}^{n \times n}$ is always singular. This is because of the fact that the rank of $\boldsymbol{S}_W$ is at most $N - C$, and, in general, the number of images in the training set N is much smaller than the number of pixels in each image n. This means that it is possible to choose the matrix $\boldsymbol{W}$ such that the within-class scatter of the projected samples can be made exactly zero.

In order to avoid the problem stated above, a combined PCA and LDA method, which is also called the Fisherfaces, is used by projecting the image set to a lower dimensional space so that the resulting within-class scatter matrix $\boldsymbol{S}_W$ is non-singular. In this method, PCA is used to reduce the dimension of the feature space to $N - C$, and then LDA defined by (6) is applied to reduce the dimension to $C - 1$.

More formally, $\boldsymbol{W}_{PCA+LDA}$ is given by

$$\boldsymbol{W}_{PCA+LDA}^T = \boldsymbol{W}_{PCA}^T \boldsymbol{W}_{LDA}^T, \tag{8}$$

where

$$\boldsymbol{W}_{PCA} = \arg\max_{\boldsymbol{W}} |\boldsymbol{W}^T \boldsymbol{S}_T \boldsymbol{W}|, \tag{9}$$

$$\boldsymbol{W}_{LDA} = \arg\max_{\boldsymbol{W}} \frac{|\boldsymbol{W}^T \boldsymbol{W}_{PCA}^T \boldsymbol{S}_B \boldsymbol{W}_{PCA} \boldsymbol{W}|}{|\boldsymbol{W}^T \boldsymbol{W}_{PCA}^T \boldsymbol{S}_W \boldsymbol{W}_{PCA} \boldsymbol{W}|}. \tag{10}$$

Note that the optimization for $\boldsymbol{W}_{PCA}$ is performed over $n \times (N - C)$ matrices with orthonormal columns, while the optimization for $\boldsymbol{W}_{LDA}$ is performed over $(N - C) \times m$ matrices with orthonormal columns. In computing $\boldsymbol{W}_{PCA}$, the smallest $C - 1$ principal components are discarded.

After the projecting the images into the reduced feature space, same as PCA, the classification is performed by using the nearest neighbour classifier in the feature space.

2.3 PCA+LDA Algorithm based on Gabor Features (G-PCA+LDA)

2.3.1 Gabor Wavelets

Instead of using the original grayscale image as the input in the previous two algorithms, the input in this algorithm is the Gabor wavelet transformed image from the original one. Gabor wavelets are biologically motivated convolution kernels which are plane waves restricted by a Gaussian envelope function, and those kernels demonstrate spatial locality and orientation selectivity. In face recognition, Gabor wavelets exhibit robustness to moderated lighting changes, small shifts and deformations.

The Gabor wavelets (kernels, filters) can be defined as follows[14],[15],[16]:

$$\psi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu,v}\|^2 \|z\|^2}{2\sigma^2}} \left(e^{i\vec{k}_{\mu,v} z} - e^{-\frac{\sigma^2}{2}}\right), \tag{11}$$

where μ and v define the orientation and the scale of the Gabor kernels, $z = (x, y)$, $\|\cdot\|$ denotes the norm operator, and the wave vector $k_{\mu,v}$ is defined as follows:

$$k_{\mu,v} = k_v e^{i\varphi_\mu}, \tag{12}$$

where $k_v = k_{max}/f^v$ and $\varphi_\mu = \pi\mu/8$. k_{max} is maximum frequency, and f is the spacing factor between kernels in the frequency domain[16].

The Gabor kernels in (11) are all self-similar since they can be generated from one filter, the mother wavelet, by scaling and rotation via the wave vector $k_{\mu,v}$. Each kernel is a product of a Gaussian envelope and a complex plane wave, while the first term in the brackets in (11) is determines the oscillatory part of the kernel and the second term compensates for the DC value. The effect of the DC term becomes negligible when the parameter σ, which determines the ratio of the Gaussian window width to wavelength, has sufficiently large values.

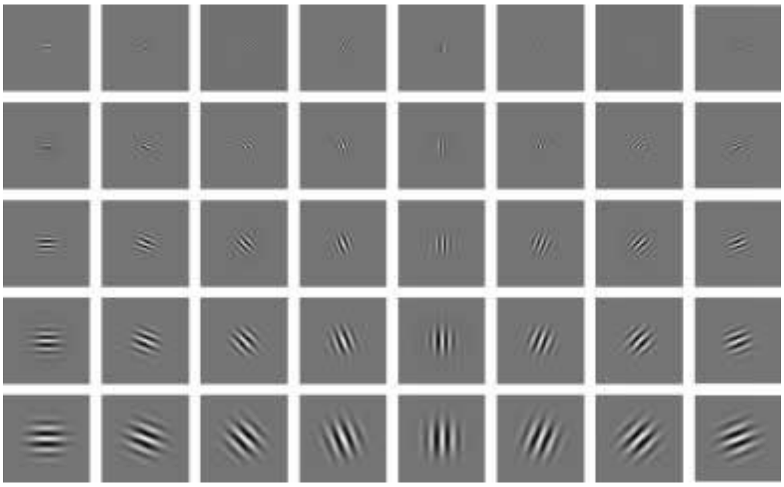


Figure 2-1 The real part of Gabor kernels at five scales and eight orientations.

Retrieved from [6].

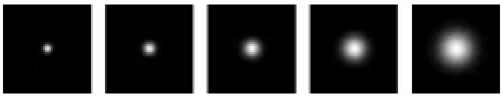


Figure 2-2 The magnitude of Gabor kernels at five different scales.

Retrieved from [6].

In most cases one would use Gabor wavelets of five different scales, $v \in \{0, 1, \ldots, 4\}$, and eight orientations, $\mu \in \{0, 1, \ldots, 7\}$, Figure 2-1 and Figure 2-2 shows the real part of the Gabor kernels at five scales and eight orientations and the magnitude of five different scales, with the following parameters[1]: $\sigma = 2\pi$, $k_{max} = \pi/2$, and $f = \sqrt{2}$. The kernels exhibit desirable characteristics of spatial frequency, spatial locality, and orientation selectivity.

2.3.2 Gabor Feature Representation

The Gabor wavelet representation of an image is the convolution of the image with a family of Gabor kernels as defined by (11). Let $I(x, y)$ be the grayscale image. The convolution of image I and a Gabor kernel $\psi_{\mu,\upsilon}$ is defined as follows[6]:

$$G_{\mu,\upsilon}(z) = I(z) * \psi_{\mu,\upsilon}(z), \tag{13}$$

where $z = (x, y)$, $*$ denotes the convolution operator, and $G_{\mu,\upsilon}(z)$ is the convolution result corresponding to the Gabor kernel at orientation μ and scale υ. Therefore, the set $S = \left\{G_{\mu,\upsilon}(z) \,\middle|\, \mu \in \{0, 1, \dots, 4\}, \nu \in \{0, 1, \dots, 4\}\right\}$ forms the Gabor wavelet representation of the image $I(z)$.

Applying the convolution theorem, we can derive each $G_{\mu,\upsilon}(z)$ for (13) via the Fast Fourier Transform (FFT), i.e.

$$\mathcal{F}\{G_{\mu,\upsilon}(z)\} = \mathcal{F}\{I(z)\}\mathcal{F}\{\psi_{\mu,\upsilon}(z)\}, \tag{14}$$

$$G_{\mu,\upsilon}(z) = \mathcal{F}^{-1}\{\mathcal{F}\{I(z)\}\mathcal{F}\{\psi_{\mu,\upsilon}(z)\}\}, \tag{15}$$

where $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ denote the Fourier and inverse Fourier transform, respectively.

---

[1] These parameters can be adjusted according to the size of the normalized face images.

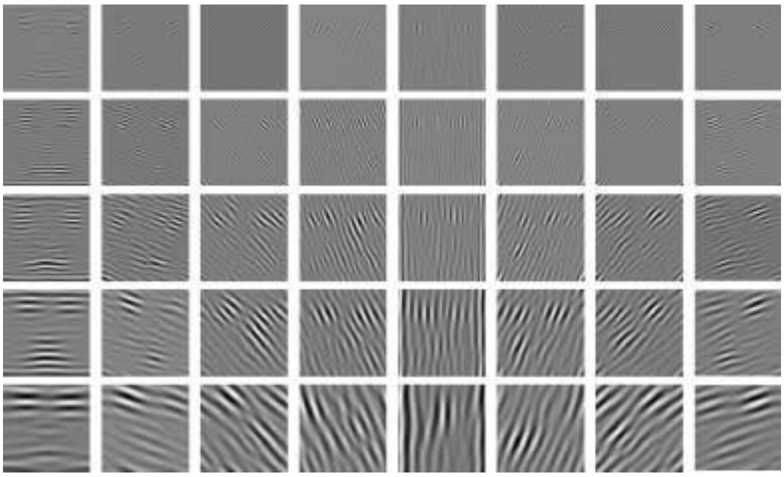


Figure 2-3 The real part of the Gabor wavelet representation.

Retrieved from [6].

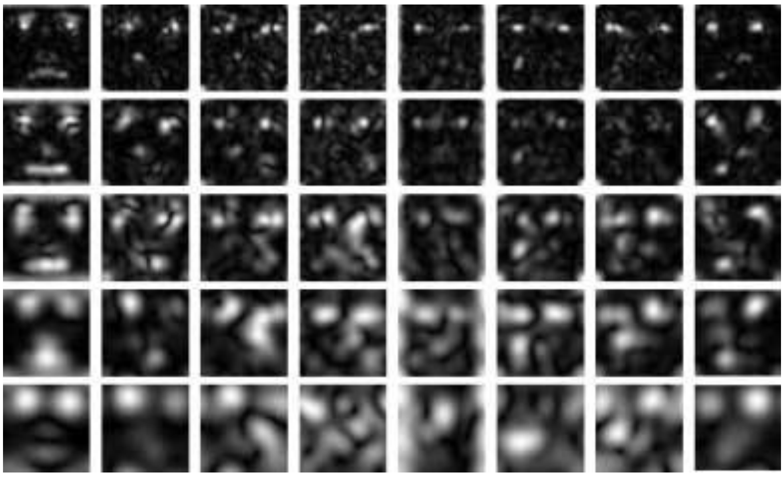


Figure 2-4 The magnitude of the Gabor wavelet representation.

Retrieved from [6].

Figure 2-3 and Figure 2-4 show the Gabor wavelet representation[1] of the real part and the magnitude, respectively, of a sample image. These representation results display scale, locality, and orientation properties corresponding to those displayed by the Gabor wavelets in Figure 2-1 and Figure 2-2. To encompass different spatial frequencies

---

[1] Note that, because the phase information of the transform is time-varying, only the magnitude of the Gabor wavelet representation is used as the feature extraction result instead of the complex representation, which is known as the Gabor Magnitude Pictures (GMPs).

(scales), spatial localities, and orientation selectivity, all these representation results are concatenated to derive an augmented feature vector F_G. Before the concatenation, each representation is normalized to zero mean and unit variance. Then, a vector out of the $G_{\mu,\upsilon}(z)$ is constructed by concatenating its rows (or columns). The augmented Gabor feature vector F_G is defined as follows[6]:

$$F_G = \left[G_{0,0}^T \quad G_{0,1}^T \quad \dots \quad G_{4,7}^T\right]^T, \tag{16}$$

where $(\cdot)^T$ is the transpose operator. The augmented Gabor feature vector thus encompasses all the elements of the Gabor wavelet representation set, $S = \left\{G_{\mu,\upsilon}(z) \,\middle|\, \mu \in \{0, 1, \dots, 4\}, \nu \in \{0, 1, \dots, 4\}\right\}$, as important discriminating information.

2.3.3 G-PCA+LDA Method

The G-PCA+LDA method[6] is similar to the PCA+LDA method. The only difference is that the G-PCA+LDA method uses Gabor features instead of the image-pixel features. After projecting the augmented Gabor feature vectors into the reduced feature space using the projection matrix $\boldsymbol{W}_{PCA+LDA}$, the nearest neighbour classifier is used to classify the face images.

2.4 Local Gabor Binary Pattern Histogram Sequence (LGBPHS)

The overall framework of the face-image representation approached based on the Local Gabor Binary Pattern Histogram Sequence is illustrated in Figure 2-5. In this approach, a face image is modelled as a “histogram sequence” by the following procedure[7]: (1) an input face image is normalized and transformed to obtain multiple Gabor Magnitude Pictures (GMPs) in the frequency domain by applying multi-scale and multi-orientation Gabor filters; (2) each GMP is converted to a Local Gabor Binary

Pattern (LGBP) map; (3) each LGBP map is further divided into non-overlapping rectangular regions with a specific size, and histogram is computed for each region; and (4) the LGBP histograms of all the LGBP maps are concatenated to form the final histogram sequence as the model of the face. The following sub-sections will describe the procedure in detail.

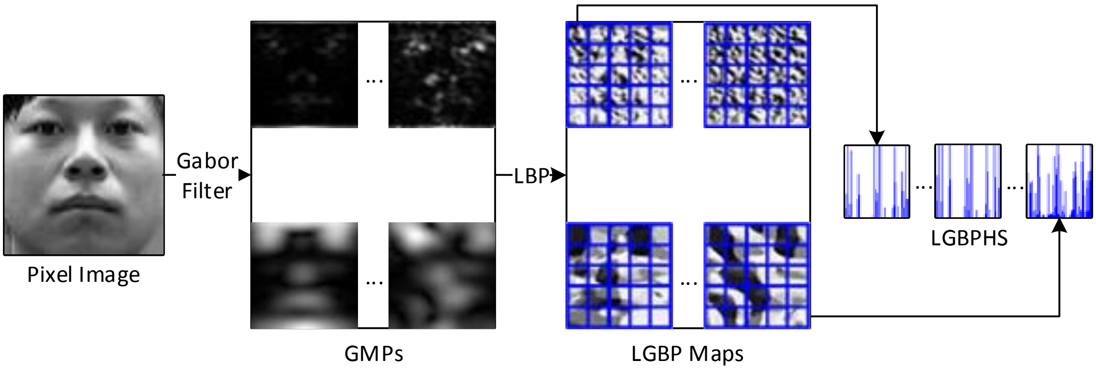


Figure 2-5 Framework of the LGBPHS approach.

Adapted from [7].

2.4.1 Gabor Magnitude Picture (GMP)

The Gabor Magnitude Picture is the magnitude of the Gabor wavelet representation. Gabor wavelets and Gabor feature representation are introduced in Section 2.3.1-2. In the LGBPHS approach, the same Gabor kernels are used and for each Gabor kernel, one magnitude value will be computed at each pixel position, which will totally result in 40 Gabor Magnitude Pictures (GMPs), which are denoted as follows[7]:

$$\left\{GMP_{\mu,v}\middle|\mu \in \{0,1,\dots,4\}, v \in \{0,1,\dots,4\}\right\}. \tag{17}$$

2.4.2 Local Gabor Binary Pattern (LGBP)

The magnitude values of the Gabor transform change very slowly with displacement, so they can be further encoded. In order to enhance the information in the GMPs, the magnitude values are encoded with Local Binary Pattern (LBP) operator. The original

LBP operator labels the pixels of an image by thresholding the 3×3-neighbourhood of each pixel $\{f_p | p = 0, 1, \dots, 7\}$ with the centre value f_c and considering the result as a binary number[17], i.e.

$$S(f_p - f_c) = f(x) = \begin{cases} 1, & f_p \geq f_c \\ 0, & f_p < f_c \end{cases}. \tag{18}$$

Then, by assigning a binomial factor 2^p for each $S(f_p - f_c)$, the LBP pattern at the pixel is achieved as

$$LBP = \sum_{p=0}^{7} S(f_p - f_c)\, 2^p, \tag{19}$$

which characterizes the spatial structure of the local image texture. The operator $LGBP$ denotes the LBP operates on GMP. The transform result at position (x, y) of $GMP_{\mu,v}$ is denoted as $LGBP_{\mu,v}(x, y)$, which composes the $LGBP_{\mu,v}$ map[7].

2.4.3 LGBP Histogram Sequence

FR under varying imaging conditions such as illumination and expression is a very difficult problem. Usually, the variations will appear more on some specific regions in face image. Therefore, local feature histogram is exploited to summarize the region property of the LGBPs by the following procedure: (1) each LGBP map is spatially divided into multiple non-overlapping regions; (2) histogram is extracted from each region; and (3) all the histograms estimated from the regions of all the LGBP maps are concatenated into a single histogram sequence to represent the given face image. The above process is formulated as follows[7]:

The histogram h_i of an image $I(x, y)$ with grey levels in the range $[0, L-1]$ could be defined as

$$h_i = \sum_{(x,y)} Comp\{I(x,y), i\}, \ i = 0, 1, \dots, L-1, \tag{20}$$

where i is the i-th grey level, h_i is the number of pixels in the image with grey level i and

$$Comp\{a,b\} = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases}. \tag{21}$$

Assume each LGBP map is divided into m regions $R_0, R_1, \dots, R_{m-1}$. The histogram of r-th region of the specific LGBP map is computed by

$$H_{\mu,v,r} = \left(h_{\mu,v,r,0}, h_{\mu,v,r,1}, \dots, h_{\mu,v,r,L-1}\right), \tag{22}$$

where

$$h_{\mu,v,r,i} = \sum_{(x,y)\in R_r} Comp\left\{LGBP_{\mu,v}(x,y), i\right\}, \tag{23}$$

Finally, the histogram pieces computed from the regions of all the 40 LGBP maps are concatenated to a histogram sequence, F_{LGBPHS}, as the final face representation, i.e.

$$F_{LGBPHS} = \left(H_{0,0,0}, \dots, H_{0,0,m-1}, H_{0,1,0}, \dots, H_{0,1,m-1}, \dots, H_{7,4,m-1}\right). \tag{24}$$

2.4.4 Direct LGBPHS Matching

For histogram matching[18], the histogram intersection $\Psi(H^1, H^2)$ is used as the similarity measurement of two histograms[19], i.e.

$$\Psi(H^1, H^2) = \sum_{i=1}^{M} \min(h_i^1, h_i^2), \tag{25}$$

where h^1 and h^2 are two histograms, and M is the number of bins in the histogram. The intuitive motivation for this measurement is the calculation of the common part of two histograms. Using this measurement, the similarity of two face images based on the LGBPHS face representation is computed by

$$\mathrm{Sim}(F_{LGBPHS}^{1}, F_{LGBPHS}^{2}) = \sum_{\mu=0}^{7} \sum_{\nu=0}^{4} \sum_{r=0}^{m-1} \Psi\left(H_{\mu,\nu,r}^{1}, H_{\mu,\nu,r}^{2}\right), \tag{26}$$

where

$$F_{LGBPHS}^{1} = \left(H_{0,0,0}^{1}, \dots, H_{0,0,m-1}^{1}, H_{0,1,0}^{1}, \dots, H_{0,1,m-1}^{1}, \dots, H_{7,4,m-1}^{1}\right), \tag{27}$$

$$F_{LGBPHS}^{2} = \left(H_{0,0,0}^{2}, \dots, H_{0,0,m-1}^{2}, H_{0,1,0}^{2}, \dots, H_{0,1,m-1}^{2}, \dots, H_{7,4,m-1}^{2}\right). \tag{28}$$

Given the histogram sequence, the computation of (26) is very efficient, since there is no float multiplication operator in the procedure. In addition, from the computation procedure of both the LGBPHS and the similarity measurement, it is clear that no statistical or learning stage is involved, which has naturally obviated the inherited generalizability problem based on the statistical learning approaches.

Chapter 3 Baseline Super-Resolution Methods

3.1 Bicubic Interpolation

In image processing, bicubic interpolation is often chosen over bilinear interpolation or nearest neighbour in image resampling, when speed is not an issue. In contrast to bilinear interpolation, which only takes 4 pixels (2-by-2) into account, bicubic interpolation considers 16 pixels (4-by-4). Images resampled with bicubic interpolation are smoother and have fewer interpolation artefacts[20].

Suppose the function values f and the derivatives f_x, f_y and f_{xy} are known at the four corners $(0,0)$, $(1,0)$, $(0,1)$, and $(1,1)$ of the unit square. The interpolated surface can then be written as:

$$p(x,y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^j. \tag{29}$$

The interpolation problem consists of determining the 16 coefficients a_{ij}. Matching $p(x,y)$ with the function values yields four equations:

$$f(0,0) = p(0,0) = a_{00},$$
$$f(1,0) = p(1,0) = a_{00} + a_{10} + a_{20} + a_{30},$$
$$f(0,1) = p(0,1) = a_{00} + a_{01} + a_{02} + a_{03},$$
$$f(1,1) = p(1,1) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij},$$

Likewise, eight equations for the derivatives in the x-direction and the y-direction:

$$f_x(0,0) = p_x(0,0) = a_{10},$$
$$f_x(1,0) = p_x(1,0) = a_{10} + 2a_{20} + 3a_{30},$$
$$f_x(0,1) = p_x(0,1) = a_{10} + a_{11} + a_{12} + a_{13},$$

$$f_x(1,1) = p_x(1,1) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} i,$$

$$f_y(0,0) = p_y(0,0) = a_{01},$$

$$f_y(1,0) = p_y(1,0) = a_{01} + a_{11} + a_{21} + a_{31},$$

$$f_y(0,1) = p_y(0,1) = a_{01} + 2a_{02} + 3a_{03},$$

$$f_y(1,1) = p_y(1,1) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} j,$$

and four equations for the cross derivative f_{xy}:

$$f_{xy}(0,0) = p_{xy}(0,0) = a_{11},$$

$$f_{xy}(1,0) = p_{xy}(1,0) = a_{11} + 2a_{21} + 3a_{31},$$

$$f_{xy}(0,1) = p_{xy}(0,1) = a_{11} + 2a_{12} + 3a_{13},$$

$$f_{xy}(1,1) = p_{xy}(1,1) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} ij,$$

where the expressions above have used the following identities:

$$p_x(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} i x^{i-1} y^{j}, \tag{30}$$

$$p_y(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} x^{i} j y^{j-1}, \tag{31}$$

$$p_{xy}(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} i x^{i-1} j y^{j-1}. \tag{32}$$

This procedure yields a surface $p(x, y)$ on the unit square $[0,1] \times [0,1]$ which is continuous and with continuous derivatives. Bicubic interpolation on an arbitrarily sized regular grid can then be accomplished by patching together such bicubic surfaces, ensuring that the derivatives match on the boundaries.

If the derivatives are unknown, they are typically approximated from the function values at points neighbouring the corners of the unit square, e.g. using finite differences.

Grouping the unknown parameters a_{ij} in a vector:

$$\alpha = [a_{00}\ a_{10}\ a_{20}\ a_{30}\ a_{01}\ a_{11}\ a_{21}\ a_{31}\ a_{02}\ a_{12}\ a_{22}\ a_{32}\ a_{03}\ a_{13}\ a_{23}\ a_{33}]^T,$$

and letting

$$x = \begin{bmatrix} F^T & F_x^T & F_y^T & F_{xy}^T \end{bmatrix}^T,$$

where

$$\begin{aligned} F &= [f(0{,}0) \quad f(1{,}0) \quad f(0{,}1) \quad f(1{,}1)]^T, \\ F_x &= [f_x(0{,}0) \quad f_x(1{,}0) \quad f_x(0{,}1) \quad f_x(1{,}1)]^T, \\ F_y &= [f_y(0{,}0) \quad f_y(1{,}0) \quad f_y(0{,}1) \quad f_y(1{,}1)]^T, \\ F_{xy} &= [f_{xy}(0{,}0) \quad f_{xy}(1{,}0) \quad f_{xy}(0{,}1) \quad f_{xy}(1{,}1)]^T, \end{aligned}$$

the problem can be reformulated into a linear equation:

$$A\alpha = x, \tag{33}$$

where the inverse of A is:

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 & -2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & -2 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 & 0 & 0 & -2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 & 1 & 1 & 0 & 0 \\ -3 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & -1 & 0 \\ 9 & -9 & -9 & 9 & 6 & 3 & -6 & -3 & 6 & -6 & 3 & -3 & 4 & 2 & 2 & 1 \\ -6 & 6 & 6 & -6 & -3 & -3 & 3 & 3 & -4 & 4 & -2 & 2 & -2 & -2 & -1 & -1 \\ 2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ -6 & 6 & 6 & -6 & -4 & -2 & 4 & 2 & -3 & 3 & -3 & 3 & -2 & -1 & -2 & -1 \\ 4 & -4 & -4 & 4 & 2 & 2 & -2 & -2 & 2 & -2 & 2 & -2 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

3.2 Eigentransformation

3.2.1 A Review of PCA

PCA represents face images using a weighted combination of eigenfaces. Let a training set of HR images and the corresponding LR images denoted by[10]:

$$\begin{aligned} \{\vec{h}_i\}_{i=1}^{M} &= [\vec{h}_1, \dots, \vec{h}_M] \in \mathbb{R}^{N^H \times M}, \\ \{\vec{l}_i\}_{i=1}^{M} &= [\vec{l}_1, \dots, \vec{l}_M] \in \mathbb{R}^{N^L \times M}, \end{aligned} \tag{34}$$

where $\vec{h}_i$ and $\vec{l}_i$ are the corresponding HR and LR image vector, N^H and N^L are the numbers of HR and LR image pixels, respectively, and M is the number of training samples.

The mean faces are computed as:

$$\vec{m}_h = \frac{1}{M}\sum_{i=1}^{M}\vec{h}_i\,,\ \vec{m}_l = \frac{1}{M}\sum_{i=1}^{M}\vec{l}_i. \tag{35}$$

For LR images, removing the mean faces from each image, we have:

$$L = [\vec{l}_1 - \vec{m}_l, \dots, \vec{l}_M - \vec{m}_l] = [\vec{l}'_1, \dots, \vec{l}'_M]. \tag{36}$$

A set of eigenvectors, also called eigenfaces, are computed from the eigenvectors of the ensemble covariance matrix:

$$C = \sum_{i=1}^{M}(\vec{l}_i - \vec{m}_l)(\vec{l}_i - \vec{m}_l)^T = LL^T. \tag{37}$$

Directly computing the eigenvectors of C is not practical because of the large size of the matrix. Alternatively, the eigenvectors of a smaller matrix $R = L^T L$ can be first computed[21]:

$$(L^T L)V_l = V_l \Lambda_l, \tag{38}$$

where V_l is the eigenvector matrix and Λ_l is the eigenvalue matrix.

Multiplying both sides by L, we have:

$$(LL^T)LV_l = LV_l \Lambda_l. \tag{39}$$

Therefore, the orthonormal eigenvector matrix of $C = LL^T$ can be computed from:

$$E_l = LV_l \Lambda_l^{-\frac{1}{2}}. \tag{40}$$

For a novel face image $\vec{x}_l$, a weight vector can be computed by projecting it onto the eigenfaces:

$$\vec{w}_l = E_l^T(\vec{x}_l - \vec{m}_l). \tag{41}$$

This is a face representation based on eigenfaces. A face can be reconstructed from K eigenfaces, $E_l = [e_1, \dots, e_K]$:

$$\vec{r}_l = E_l \vec{w}_l + \vec{m}_l. \tag{42}$$

3.2.2 Eigentransformation

For eigentransformation[10], a training set containing LR face images, and the corresponding HR face images are used. Following the previous discussion, applying PCA to the input LR face image $\vec{x}_l$ using (49) and (51), the reconstructed face image can be represented by:

$$\vec{r}_l = LV_l \Lambda_l^{-\frac{1}{2}} \vec{w}_l + \vec{m}_l = L\vec{c} + \vec{m}_l, \tag{43}$$

where $\vec{c} = V_l \Lambda_l^{-\frac{1}{2}} \vec{w}_l = [c_1, \ldots, c_M]^T$. Equation (52) can be rewritten as:

$$\vec{r}_l = L\vec{c} + \vec{m}_l = \sum_{i=1}^{M} c_i \vec{l}_i' + \vec{m}_l. \tag{44}$$

This shows that the input LR face image can be reconstructed from the optimal linear combination of M LR training face images. Here, $\vec{c}$ describes the weight that each training face contributes in reconstructing the input face. The sample face that is more similar to the input face has a greater weight contribution. Replacing each LR resolution image $\vec{l}_i'$ by its corresponding HR version $\vec{h}_i'$, and replacing $\vec{m}_l$ with the HR mean face $\vec{m}_h$, we have:

$$\vec{x}_h = \sum_{i=1}^{M} c_i \vec{h}_i' + \vec{m}_h. \tag{45}$$

$\vec{x}_h$ is expected to be an approximation to the real HR face image. However, due to the combination coefficient c_i is not computed from the HR training data, some non-face-like distortions may be involved. These non-face-like distortions can be reduced by reconstructing $\vec{x}_h$ from the HR eigenfaces. Let E_h and $\Lambda_h = diag(\lambda_1, \ldots, \lambda_K)$ be the eigenface and eigenvalue matrices computed from the HR training images. The principal components of $\vec{x}_h$ projecting on the HR eigenfaces are:

$$\vec{w}_h = E_h^T(\vec{x}_h - \vec{m}_h). \tag{46}$$

The eigenvalue λ_i is variance of HR face images on the ith eigenface. If the principal component $w_h(i)$ is much larger than λ_i, non-face-like distortion may be involved for the ith eigenface dimension. To reduce the distortion, we apply constraints on the principal components according to the eigenvalues:

$$\vec{w}_h'(i) = \begin{cases} w_h(i), & |w_h(i)| \leq a\sqrt{\lambda_i} \\ \operatorname{sign}(w_h(i))\, a\sqrt{\lambda_i}, & |w_h(i)| > a\sqrt{\lambda_i} \end{cases}, \quad a > 0, \tag{47}$$

where $a\sqrt{\lambda_i}$ is used to bound the principal components and a is a positive scale parameter. The final super-resolved face image is reconstructed by:

$$\vec{x}_h' = E_h \vec{w}_h' + \vec{m}_h. \tag{48}$$

The diagram of the algorithm based on eigentransformation is shown in Figure 3-1. When a LR image $\vec{x}_l$ is input, it is approximated by a linear combination of the LR images using the PCA method, and a set of coefficients $[c_1, \ldots, c_M]^T$ on the training set is obtained. Keeping the coefficients and replacing the LR training images with the corresponding HR ones, a new HR face image can be synthesized. The synthesized face image is projected onto the HR eigenfaces and reconstructed with constraints on the principal components. This transformation procedure is called eigentransformation, since it uses the eigenfaces to transform the input image to output result.

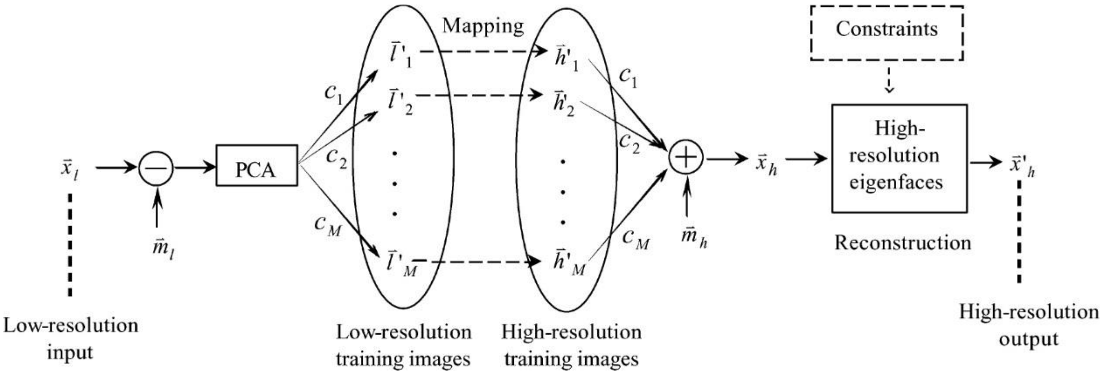


Figure 3-1 System diagram using eigentransformation for super-resolution.

Retrieved from [10].

3.3 Coherent Local Linear Reconstruction (CLLR)

3.3.1 Canonical Correlation Analysis (CCA)

CCA is a way of measuring the linear correlation between two sets of variables[1]. CCA finds two bases, one for each variable, in which the corresponding correlation coefficients of the two sets of variables are maximized[12],[22]. The LR and HR variables are transformed into correlated spaces, called coherent subspaces, when projected into these new bases.

Suppose we are given two sets of m vectors P and Q[2] with zero means:

$$\begin{aligned} P &= \{P_j\}_{j=1}^{m} = [P_1, \dots, P_m], \\ Q &= \{Q_j\}_{j=1}^{m} = [Q_1, \dots, Q_m]. \end{aligned} \tag{49}$$

The transform vectors u and v are defined as:

$$\begin{aligned} u &= W_p^T P, \\ v &= W_q^T Q. \end{aligned} \tag{50}$$

where W_p and W_q denote the corresponding basis vectors in the coherent subspaces. The objective of CCA is to obtain basis vectors W_p and W_q which maximize the correlation coefficient of u and v:

---

[1] In this SR context, the variables are vectors each of which represents a single LR or HR image, such as its PCA coefficients, or residual image patches.

[2] In this SR problem, P_j and Q_j, respectively, represent the PCA coefficients of LR and HR versions of a face image during global face recognition, or LR and HR residual patches when performing residual compensation.

$$\rho(u,v) = \frac{W_p^T C_{PQ} W_q}{\sqrt{W_p^T C_{PP} W_p W_q^T C_{QQ} W_q}}, \tag{51}$$

where $C_{PP} = E[PP^T]$ and $C_{QQ} = E[QQ^T]$ represent the within-set covariance matrices of P and Q respectively, while $C_{PQ} = E[PQ^T]$ and $C_{QP} = E[QP^T]$ denote their covariance matrices. $E[\cdot]$ is mathematical expectation.

To find W_p and W_q, we first compute:

$$\begin{aligned} R_1 &= C_{PP}^{-1} C_{PQ} C_{QQ}^{-1} C_{QP}, \\ R_2 &= C_{QQ}^{-1} C_{QP} C_{PP}^{-1} C_{PQ}. \end{aligned} \tag{52}$$

It can then be shown that the eigenvectors of R_1 and R_2 give the two desired sets of basis vectors W_p and W_q[12].

3.3.2 CLLR-SR Algorithm

The coherent local linear reconstruction super-resolution (CLLR-SR) method performs neighbourhood-based reconstruction in the coherent subspaces using a two-step SR framework, as illustrated in Figure 3-2[11]. LR and HR image features, either PCA coefficients or residual patch vectors, can be projected into a coherent subspace using the base vectors obtained via CCA. CCA is used in both steps of the face SR algorithm to improve the similarity between the HR and LR feature neighbourhoods.

In the first step, both LR and HR images are projected into the PCA subspaces and the PCA coefficients are considered to be the set of face features. To promote similarity of neighbourhood structures, the LR and HR face features are projected into coherent subspaces using the basis vectors calculated from the PCA coefficients from training samples. Given a novel LR image, neighbourhood-based reconstruction can be used with these coefficients to obtain the HR face features from LR and HR face features of the training faces in the coherent subspace.

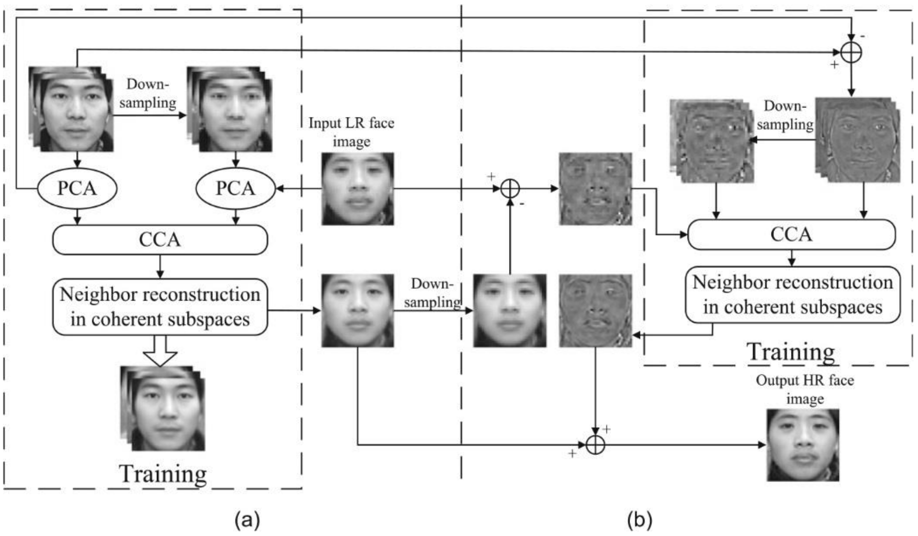


Figure 3-2 Architecture of the CLLR-SR algorithm.

(a) Global face reconstruction; (b) residual compensation.

Retrieved from [11].

In order to further recover high-frequency information, residual face images are divided into overlapping patches which are used as the face features in the second step. Again, coherent subspaces are learnt for LR and HR versions of each training patch. Then, for a novel LR image, LR patches are used in the coherent subspace to provide a neighbourhood-based reconstruction of each HR residual image patch from the corresponding input LR residual image patch. After computing all HR residual patches, pixels in the overlapping areas are averaged to give the overall HR residual image.[1] The final HR face image is obtained by adding the HR global face image and the residual face image.

---

[1] As this is a patch-based approach, blocky artefacts may appear in the reconstructed HR images. In order to reduce the artefacts, the pixels in the overlapped regions are merged based on a central distance dependent smoothing filter, which is further explained in the subsequent section.

3.3.2.1 Global Reconstruction

Let a training set of HR images and the corresponding LR images denoted by:

$$\begin{aligned} \boldsymbol{I}^H &= \{I_i^H\}_{i=1}^m = [I_1^H, \dots, I_m^H] \in \mathbb{R}^{n^H \times m}, \\ \boldsymbol{I}^L &= \{I_i^L\}_{i=1}^m = [I_1^L, \dots, I_m^L] \in \mathbb{R}^{n^L \times m}, \end{aligned} \tag{53}$$

and their corresponding PCA coefficients:

$$\begin{aligned} \boldsymbol{X}^H &= \{x_i^H\}_{i=1}^m = [x_1^H, \dots, x_m^H] \in \mathbb{R}^{p \times m}, \\ \boldsymbol{X}^L &= \{x_i^L\}_{i=1}^m = [x_1^L, \dots, x_m^L] \in \mathbb{R}^{q \times m}. \end{aligned} \tag{54}$$

During training for global reconstruction, the LR and HR face images are first projected into the PCA subspaces:

$$\begin{aligned} x_i^H &= (\boldsymbol{W}_{PCA}^H)^T (I_i^H - \mu^H), \\ x_i^L &= (\boldsymbol{W}_{PCA}^L)^L (I_i^L - \mu^L), \end{aligned} \tag{55}$$

where μ^H and μ^L are the HR and LR mean faces, and $\boldsymbol{W}_{PCA}^H$ and $\boldsymbol{W}_{PCA}^L$ are the PCA projection matrices for HR and LR training images respectively.

Then, CCA is applied to capture the coherent subspaces relating the two data sets $\boldsymbol{X}^H$ and $\boldsymbol{X}^L$. First, the two data sets are subtracted by their mean values $\bar{x}^H$ and $\bar{x}^L$, respectively, to give the centralized data sets $\widehat{\boldsymbol{X}}^H = [\hat{x}_1^H, \dots, \hat{x}_m^H]$ and $\widehat{\boldsymbol{X}}^L = [\hat{x}_1^L, \dots, \hat{x}_m^L]$. CCA finds two basis vectors V^H and V^L for the data sets $\widehat{\boldsymbol{X}}^H$ and $\widehat{\boldsymbol{X}}^L$ in order to maximize the correlation coefficient between vectors $\boldsymbol{C}^H = (V^H)^T \widehat{\boldsymbol{X}}^H$ and $\boldsymbol{C}^L = (V^L)^T \widehat{\boldsymbol{X}}^L$. In other words, we maximize:

$$\rho = \frac{E\left[\boldsymbol{C}^H (\boldsymbol{C}^L)^T\right]}{\sqrt{E[(\boldsymbol{C}^H)^2]\, E[(\boldsymbol{C}^L)^2]}} = \frac{E\left[(V^H)^T \widehat{\boldsymbol{X}}^H (\widehat{\boldsymbol{X}}^L)^T V^L\right]}{\sqrt{E\left[(V^H)^T \widehat{\boldsymbol{X}}^H (\widehat{\boldsymbol{X}}^H)^T V^H\right] E\left[(V^L)^T \widehat{\boldsymbol{X}}^L (\widehat{\boldsymbol{X}}^L)^T V^L\right]}}. \tag{56}$$

Based on (61), where

$$\mathrm{C}_{PP} = E\left[\widehat{\boldsymbol{X}}^H (\widehat{\boldsymbol{X}}^L)^T\right],$$
$$\mathrm{C}_{PQ} = E\left[\widehat{\boldsymbol{X}}^H (\widehat{\boldsymbol{X}}^L)^T\right],$$
$$\mathrm{C}_{QP} = E\left[\widehat{\boldsymbol{X}}^L (\widehat{\boldsymbol{X}}^H)^T\right],$$
$$\mathrm{C}_{QQ} = E\left[\widehat{\boldsymbol{X}}^L (\widehat{\boldsymbol{X}}^L)^T\right],$$

the basis vectors V^H and V^L in (65) are calculated: these are the eigenvectors of R_1 and R_2.

Transforming the PCA coefficients $\boldsymbol{X}^H$, $\boldsymbol{X}^L$ into the coherent subspaces using the basis vectors gives the corresponding projected CCA coefficient sets $\boldsymbol{C}^H = \{c_i^H\}_{i=1}^m$ and $\boldsymbol{C}^L = \{c_i^L\}_{i=1}^m$:

$$\begin{aligned} c_i^H &= (V^H)^T \hat{x}_i^H, \\ c_i^L &= (V^L)^T \hat{x}_i^L. \end{aligned} \tag{57}$$

There is a similarity between the internal structures of the two data sets $\boldsymbol{X}^H$ and $\boldsymbol{X}^L$, and after transforming them into coherent subspaces, the correlation between the two sets is maximized. Neighbourhoods are more consistent between $\boldsymbol{C}^H$ and $\boldsymbol{C}^L$: CCA provides an optimal linear relationship between the neighbourhood structures for the two sets.

To perform reconstruction, for a novel LR face image I_l, its PCA coefficient vector x_l are computed:

$$x_l = (\boldsymbol{W}_{PCA}^L)^T (I_l - \mu^L). \tag{58}$$

Then x_l is projected into coherent subspace using:

$$c_l = (V^L)^T (x_l - \bar{x}^L). \tag{59}$$

For c_l, we seek K^G-nearest neighbours $\{c_i^L\}_{i=1}^{K^G}$ in C^L and corresponding optimal weights $W^G = \{w_i^G\}_{i=1}^{K^G}$ which minimize the reconstruction error:

$$\varepsilon = \left\| c_l - \sum_{i=1}^{K^G} w_i^G c_i^L \right\|^2, \tag{60}$$

subject to the constraint:

$$\sum_{i=1}^{K^G} w_i^G = 1. \tag{61}$$

Minimizing the above objective function is a constrained least squares problem[13]. To solve this problem, define the local Gram matrix as:

$$G_{ij} = (c_l - c_i^L)(c_l - c_j^L). \tag{62}$$

By reconstruction, this Gram matrix is symmetric and semipositive definite. The reconstruction error can be minimized analytically using a Lagrange multiplier to enforce the constraint in (70). In terms of the inverse Gram matrix, the optimal weights are given by:

$$w_i^G = \frac{\sum_j G_{ij}^{-1}}{\sum_{lm} G_{lm}^{-1}}. \tag{63}$$

The solution as written in (72), appears to require an explicit inversion of the Gram matrix. In practice, a more efficient and numerically stable way to minimize the error (which yields the same result as above) is simply to solve the linear system of equations:

$$\sum_j G_{ij} w_j^G = 1, \tag{64}$$

and then to rescale the weights so that they sum to one.

In unusual cases, it can arise that the Gram matrix in (71) is singular or nearly singular, for example, when there are more neighbours than input dimensions, or when the data points are not in general position. In this case, the least square problem for finding the weights does not have a unique solution, and the Gram matrix must be conditioned (before solving the linear system) by adding a small multiple of the identity matrix[13]:

$$G_{ij} \leftarrow G_{ij} + \delta_{ij}\left(\frac{\varDelta^2}{K^G}\right)Tr(G),$$

where

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & otherwise \end{cases}.$$

$Tr(G)$ denotes the trace of G, and $\varDelta^2 \ll 1$. This amounts to adding a regularization term to the reconstruction cost that measures the summed squared magnitude of the weights.[1]

The regularization term acts to penalize large weights that exploit correlations beyond some level of precision in the data sampling process. It may also introduce some robustness to noise and outliers. This form of regularization (with $\Delta = 0.1$) is used in the experiments.

After computing the optimal weights, by applying these weights to $\{c_i^H\}_{i=1}^{K^G}$ in C^H, the corresponding HR image feature in the coherent subspace can be reconstructed:

$$c_h = \sum_{i=1}^{K^G} w_i^G c_i^H. \tag{65}$$

---

[1] One can also consider the effect of this term in the limit $\Delta \rightarrow 0$.

Next, using an inverse transformation, the PCA coefficients of the corresponding HR face image are obtained:

$$x_h = ((V^H)^T)^{\dagger} c_h + \bar{x}^H, \tag{66}$$

where $(\cdot)^{\dagger}$ denotes pseudoinverse, so $((V^H)^T)^{\dagger} = (V^H (V^H)^T)^{-1} V^H$.

Finally, the HR global face image is obtained by computing:

$$G_h = \boldsymbol{W}_{PCA}^{H} x_h + \mu^H. \tag{67}$$

3.3.2.2 Residual Compensation

As the global face image is reconstructed from an eigenmodel, it mainly contains low and middle frequency components. In order to provide additional details in the output face image, it is necessary to carry out residual compensation on the global reconstructed face.

In the residual compensation step of CLLR-SR algorithm, residual face images, the difference between original face images and their globally reconstructed face images, are divided into square patches to facilitate recovery of further detail[23],[24]. Overlapping patches are used in order to enforce smoothness constraints. In order to apply neighbourhood reconstruction to residual compensation, LR and HR image features should have similar locally geometries. Again, CCA is used as before, this time to maximize the correlation between LR and HR residual patches for residual compensation. This time, for nearest neighbour search, the current position and its eight connected neighbour are considered for speed, and to enforce smoothness constraints.

During training, for LR face images in the training set $\boldsymbol{I}^L = [I_1^L, \dots, I_m^L]$, the globally reconstructed face images $\boldsymbol{G}^H = [G_1^H, \dots, G_m^H]$ are calculated using the method in Section 3.3.2.1. The corresponding HR residual face images are defined as:

$$R_i^H = I_i^H - G_i^H. \tag{68}$$

The training set of HR residual face images is thus $\boldsymbol{R}^H = \{R_i^H\}_{i=1}^m$. Similarly, define the LR residual face images as:

$$R_i^L = I_i^L - G_i^L, \tag{69}$$

where $G_i^L = D[G_i^H]$, and $D[\cdot]$ represents down-sampling[1]. The training set of LR residual face images is in turn $\boldsymbol{R}^L = \{R_i^L\}_{i=1}^m$.

Each image in the residual image training sets $\boldsymbol{R}^L$ and $\boldsymbol{R}^H$ is divided into overlapping square patches. For novel LR face image I_l, the corresponding global SR face image is found using the method in Section 3.3.2.1. Then, based on (78), the corresponding residual face image R_l is obtained, which is also divided into overlapping patches.

Subsequently, CCA is applied and basis vectors for training LR and HR residual patches are calculated for each novel LR residual patch. Then, the training LR and HR residual patches are projected into coherent subspaces. Next, each novel LR residual patch is transformed into coherent subspace by these basis vectors. Finally, neighbourhood-based reconstruction is used to determine the corresponding HR residual patch in coherent subspace.

In detail, the residual patch vector of novel face image $(R_l)_j$, where j represents the position information of each patch, and residual training patch vectors $(\boldsymbol{R}^L)_j$ and $(\boldsymbol{R}^H)_j$ are transformed into coherent subspaces as $(\hat{R}_l)_j$, $(\hat{\boldsymbol{R}}^L)_j$ and $(\hat{\boldsymbol{R}}^H)_j$. The training patch vectors $(\boldsymbol{R}^L)_j$ and $(\boldsymbol{R}^H)_j$ for each LR residual patch vector $(R_l)_j$

---

[1] In the experiments, bicubic interpolation is used as the down-sampling method.

are the ones at the current position and its eight connected neighbouring patches, determined by the LR residual patch location in the LR face image, in all patches comprising $\boldsymbol{R}^L$ and $\boldsymbol{R}^H$.

Relying on the relatively fixed geometric structure of human face images, for each $(\hat{R}_l)_j$ in $\hat{R}_l$, we find its K^R-nearest neighbours $\{(\hat{R}_i^L)_j\}_{i=1}^{K^R}$ in $(\hat{\boldsymbol{R}}^L)_j$. Then, the corresponding optimal weights $W^R = \{w_i^R\}_{i=1}^{K^R}$ are determined by minimizing:

$$\varepsilon = \left\|(\hat{R}_l)_j - \textstyle\sum_{i=1}^{K^R} w_i^R (\hat{R}_i^L)_j\right\|^2, \tag{70}$$

subject to the constraint:

$$\textstyle\sum_{i=1}^{K^R} w_i^R = 1. \tag{71}$$

Using the method in Section 3.3.2.1, the reconstruction weights $W^R = \{w_i^R\}_{i=1}^{K^R}$ can be found. Then these weights are used for the corresponding HR feature $(\hat{R}_i^H)_j$ corresponding to $(\hat{R}_i^L)_j$:

$$(\hat{R}_h)_j = \textstyle\sum_{i=1}^{K^R} w_i^R (\hat{R}_i^H)_j. \tag{72}$$

Finally, the HR residual patches $(R_h)_j$ are obtained by transforming the HR features into the pixel domain. Using all the residual patches $(R_h)_j$, with pixels in the overlapping areas averaged, the HR residual image R_h is obtained. The final HR face image is then

$$I_h = G_h + R_h. \tag{73}$$

3.3.3 Improvements

Often times, the resultant image of the CLLR-SR algorithms may involve non-face-like distortions, such as multiple facial contours, because the images in both training set and probe set are not well-aligned. Even though the face images are aligned by the position of the eyes, due to the individual differences of the layout of the face priors, non-face-distortions may also appear. Of course, this problem may be solved by advanced face alignment techniques, but here, inspired by the method introduced in Section 3.3.2.2, a patch-based method is used to improve the performances. Experiments shows that the quality of the super-resolved images is much improved.

The patch-based method can also be utilized in the first step of the CLLR-SR algorithm (global reconstruction). For a novel LR face image, it is divided into overlapped patches, and we super-resolve this image patch by patch using the method in Section 3.3.2.2. After all the LR patches are super-resolved, we combine the resultant HR patches by merging the overlapped regions.

Usually, directly averaging the overlapped area of the patches may result in blocky artefacts in the merged image, because the edge of the patches may not be smooth after averaging. To deal with this problem, a smoothing mask is designed and applied to each patch before merging.

In detail, the smoothing mask $w_i(x, y)$ is a weighting map of the same size of an image patch $P_i(x, y)$, where (x, y) denotes the position of the pixel and i represents the position information. For each image patch $P_i(x, y)$, the value of weights in the weighting map are designed in the following criterions:

(1) If $(x, y)_i$ is not in the overlapped region, $w_i(x, y) = 1$.

(2) If $(x,y)_i$ is in the overlapped region, $w_i(x,y)=\frac{\alpha}{d+1}$, where d is Euclidean distance between $(x,y)_i$ and the centre of the patch $(x_c,y_c)_i$, α is the scaling factor to make sure the summation of all the masking weights of a certain pixel (x,y) is 1.

(3) For each pixel (x,y) of the image, $\sum_i w_i(x,y)=1$.

After all the weights are calculated, we have the smoothing masks $\{w_i(x,y)\}_{i=1}^{N^P}$ for all the corresponding patches $\{P_i(x,y)\}_{i=1}^{N^P}$, where N^P denotes the number of overlapped patches for a face image. Then, the patches are multiplied with the corresponding smoothing masks pixel by pixel, respectively:

$$P_i'(x,y)=P_i(x,y)\,w_i(x,y). \tag{74}$$

Finally, the merged image is obtained by combining all the weighted patches by adding the values of each weighted pixel:

$$I(x,y)=\sum_i P_i'(x,y)=\sum_i P_i(x,y)\,w_i(x,y). \tag{75}$$

By multiplying the pixel values in overlapped region with central distance dependent weights, the merged image appears to be more smoothing.

The above improvements may also be applied to SR method based on eigentransformation and experiments shows the patch-based SR methods are generally robust to partially misalignment of face images.

As shown in Figure 3-3, the patch-based CLLR-SR method can effectively eliminate the false-contour distortion caused by the misalignment of the training dataset, while obtaining a much sharper estimation of the HR face image than the bicubic interpolation method.

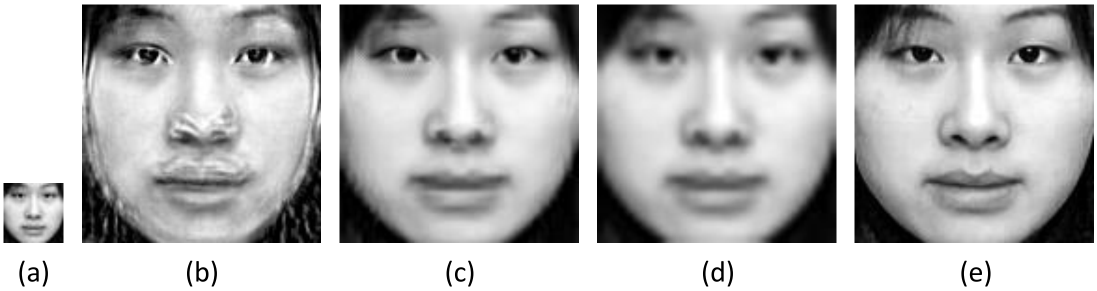


Figure 3-3 Comparison of super-resolved face image results.

(a) Low resolution input test image of size 32×32; (b) super-resolved face image of size 128×128 produced by CLLR-SR method; (c) super-resolved face image of size 128×128 produced by patch-based CLLR-SR method; (d) super-resolved face image of size 128×128 produced by bicubic interpolation; (e) grand-true high-resolution face image of size 128×128.

Chapter 4 Coupled-Projection Method

Inspired by the CLLR-SR method and the CCA algorithm, a coupled-projection method based on CCA is proposed in this chapter. As stated in Section 1.1, LR probe images cannot be matched with HR gallery images directly, because the dimensions of classic LR and HR features usually do not match. Hence, ideas are brought up to find a coupled-projection method which projects both the LR and HR images into a common subspace. The dimensions of the LR and HR features in the common subspace can be matched with each other. Consequently, traditional FR algorithms, such as LDA, can be employed in this common subspace.

4.1 Framework

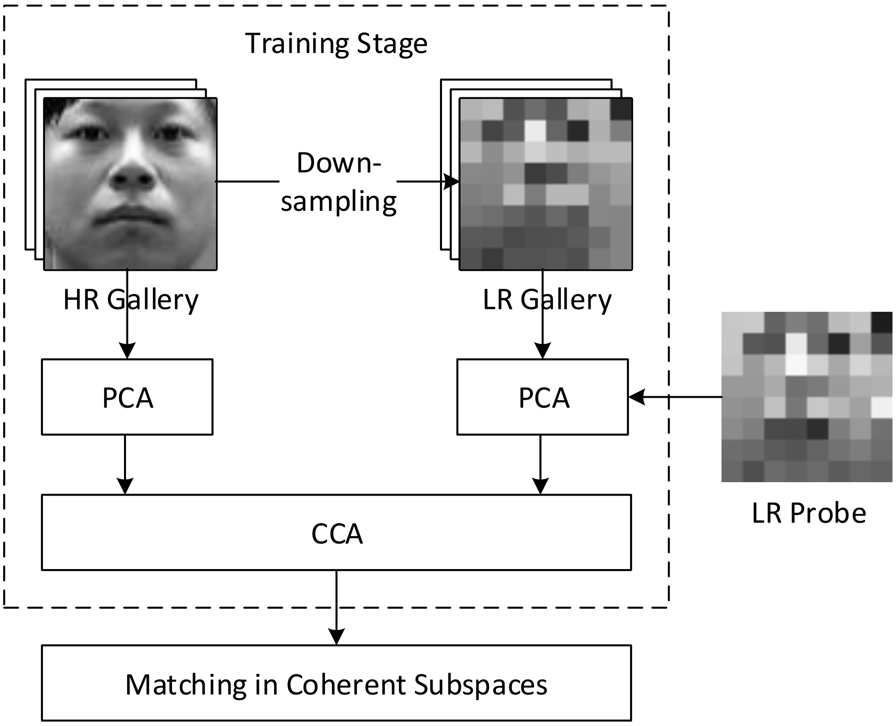


Figure 4-1 Framework of the coupled-projection method.

The framework of the coupled-projection method is illustrated in Figure 4-1. In the training stage, PCA is first used to project the image set to a lower dimensional subspace for both the LR and HR image sets.[1] Then, CCA is applied to find two sets of basis vectors to project the LR and HR PCA features into coherent subspaces, respectively. For a novel LR image, we first calculate its PCA feature, and then project it into the coherent subspace using the corresponding basis vectors. Then, the matching is undertaken in the coherent subspaces. Since the dimensions of the LR and HR features are the same in the coherent subspaces, traditional FR algorithms can be used.

4.2 Algorithm

In detail, denote a training set of HR images and the corresponding LR images as:

$$\begin{aligned} \boldsymbol{I}^H &= \{I_i^H\}_{i=1}^m = [I_1^H, \dots, I_m^H], \\ \boldsymbol{I}^L &= \{I_i^L\}_{i=1}^m = [I_1^L, \dots, I_m^L], \end{aligned} \tag{76}$$

and their corresponding PCA coefficients:

$$\begin{aligned} \boldsymbol{X}^H &= \{x_i^H\}_{i=1}^m = [x_1^H, \dots, x_m^H], \\ \boldsymbol{X}^L &= \{x_i^L\}_{i=1}^m = [x_1^L, \dots, x_m^L]. \end{aligned} \tag{77}$$

In the training stage, the LR and HR face images are first projected into the PCA subspaces as follows[3]:

$$\begin{aligned} x_i^H &= (\boldsymbol{W}_{PCA}^H)^T (I_i^H - \mu^H) \\ x_i^L &= (\boldsymbol{W}_{PCA}^L)^L (I_i^L - \mu^L) \end{aligned} \tag{78}$$

where μ^H and μ^L are the HR and LR mean faces, respectively, and $\boldsymbol{W}_{PCA}^H$ and

---

[1] Except for pixel images, Gabor features may also be considered.

$\boldsymbol{W}_{PCA}^{L}$ are the PCA projection matrices for the HR and LR training images, respectively.

Then, using the method in described in Section 3.3.2.1, CCA is applied to capture the coherent subspaces relating the two data sets $\boldsymbol{X}^H$ and $\boldsymbol{X}^L$, and the corresponding projected CCA coefficient sets $\boldsymbol{C}^H = \{c_i^H\}_{i=1}^m$ and $\boldsymbol{C}^L = \{c_i^L\}_{i=1}^m$ are obtained as follows[11]:

$$\begin{aligned} c_i^H &= (V^H)^T(x_i^H - \bar{x}^H), \\ c_i^L &= (V^L)^T(x_i^L - \bar{x}^L), \end{aligned} \tag{79}$$

where V^H and V^L are the two basis vectors for the HR and LR data sets, respectively, and $\bar{x}^H$ and $\bar{x}^L$ are the means of the data sets, respectively.

Next, we input a novel LR image I_l, its PCA coefficient vector x_l are computed as follows:

$$x_l = (\boldsymbol{W}_{PCA}^{L})^T(I_l - \mu^L). \tag{80}$$

Then, x_l is projected into the coherent subspace using:

$$c_l = (V^L)^T(x_l - \bar{x}^L). \tag{81}$$

Finally, c_l is matched against either $\boldsymbol{C}^H = \{c_i^H\}_{i=1}^m$ or $\boldsymbol{C}^L = \{c_i^L\}_{i=1}^m$ using traditional FR algorithms, such as LDA. Usually, the results between the matching against $\boldsymbol{C}^H$ or $\boldsymbol{C}^L$ have little difference, since the correlation between $\boldsymbol{C}^H$ or $\boldsymbol{C}^L$ are maximized by CCA. In the experiments, a combined PCA and Coherent Linear Discriminant Analysis algorithm based on Gabor features will be evaluated.

To be precise, there is no literally a common subspace in the G-PCA+CLDA algorithm. However, CCA is used to capture two coherent subspaces, in which the correlation between the HR and the corresponding LR features are maximized, and the dimension of the two subspaces are the same. Hence, the LR feature projection can be compared with the HR feature projection directly.

4.3 Advantages

The advantages of the proposed coupled-projection method over the first two standard approaches mentioned in Section 1.1 are stated below.

The coupled-projection method learns from the HR galleries directly, and also captures the structural relationship between the HR and LR images. Hence, useful information will not be discarded as it happens in the first standard approach when down-sampling the HR gallery images.

Despite the fact that SR algorithms can be used to obtain the visually appealing HR version of the LR probe images, the super-resolved images often lack the high frequency components of true HR images, which is highly required in the recognition step of face images. Although some complex SR algorithms, such as CLLR-SR, consider a second step to compensate for the high frequency components, the resultant HR images are still not satisfactory. Moreover, the performance of learning-based SR algorithms always depends on the training stage. If a statistical-learning-based SR algorithm is not trained sufficiently, such as ill-posed parameter settings, the performance may drop radically. As what the experimental results is going to illustrate, the performance of the statistical-learning-based SR algorithm can be unstable when evaluated on different databases. To summarize, the SR-FR approach (the second standard approach stated in Section 1.1) is not a general and efficient solution to the low-resolution face-recognition problem.

Comparing with the SR-FR approach, the coupled-projection approach does not involve complex training procedures and employs CCA to learn two projection matrices for the LR and HR images. Subsequently, the LR and HR images are projected into coherent subspaces where matching are performed. The approach considers the similarity between the LR and HR images as the statistical-learning-based SR algorithms do and does not suffer the insufficient training problem.

In addition, in the image-feature representation step, traditional features, such as Gabor features, and dimension reduction methods, such as PCA, can be employed; after coupled-projection, the traditional classifiers, such as LDA, can be used for classification. Hence, the coupled-projection method is also compatible with other traditional algorithms.

In the experiments, the PCA+CLDA and G-PCA+CLDA approaches perform well on subsets of all databases.

Chapter 5 Experiments and Analysis

5.1 Evaluation Protocol

5.1.1 Data Sets

In the experiments, subsets of the CAS-PEAL-R1 database[25], AR face database[26], Yale B face database[27], and Caltech face database[28] are used to evaluate the FR and SR algorithms. Table 5-1 is an overview of the training and probe subsets of the above mentioned databases used in the experiments.

Table 5-1 Overview of the training and probe subsets

Database Subsets	#Subjects	Images	Training Samples	Probe Samples
CAS-PEAL-R1	300	2606	1200	1406
AR	99	1265	693	572
Yale B	10	650	200	450
Caltech	26	443	212	231

CAS-PEAL-R1 Subset: The subset of the CAS-PEAL-R1 database used in the experiments includes variations of expression, accessory, lighting, time, background and distance. The training set contains 1200 images of 300 subjects, which are randomly selected from the 1040 subjects in the CAS-PEAL-R1 database, with each subject contributing four images randomly selected from the frontal subset of the CAS-PEAL-R1 database. Sample images are shown in Figure 5-1.

AR Subset: The subset of the AR face database contains 99 subjects with variations including expression, lighting and accessory. For each subject, 13 frontal images or so are selected, in which 7 images or so are randomly selected for training and the remaining images are used for testing. Sample images are shown in Figure 5-2.

Figure 5-1 Samples from the CAS-PEAL-R1 subset.

Figure 5-2 Samples from the AR subset.

Figure 5-3 Samples from the Yale B subset.

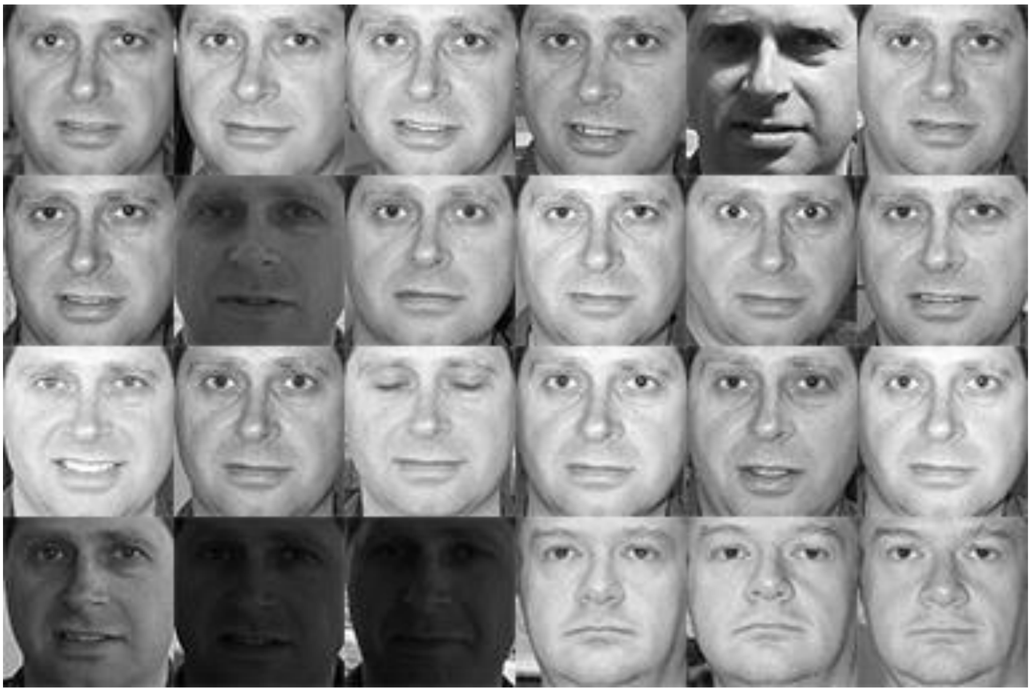


Figure 5-4 Samples from the Caltech subset.

Yale B Subset: The subset of the Yale B face database contains 10 subjects. Each subject have 65 frontal images of different lighting conditions. For each subject, 20 images are randomly selected for training and the remaining 45 images are used for testing. Sample images are shown in Figure 5-3.

Caltech Subset: The subset of the Caltech face database used in the experiments contains 443 frontal images of 26 subjects. For each subject 10 images or so are randomly selected for training, and the remaining images are used for testing. Sample images are shown in Figure 5-4.

By comparing the four datasets, the CAS-PEAL-R1 subset and the AR subset contain relatively more subjects than the Yale B subset and the Caltech subset. Also, the CAS-PEAL-R1 and the AR subset include face images with occlusions, while the Yale B subset and the Caltech subset don't. Further we compare the CAS-PEAL-R1 subset and the AR subset and find out that the occlusions in the AR subset cover more area of the face image than those in the CAS-PEAL-R1 subset do. And by comparing the Yale B subset and the Caltech subset, the Yale B subset contains more variations of lighting conditions than the Caltech subset does. The major comparisons of the four evaluated face datasets are summarized in Table 5-2.

Table 5-2 Major comparisons of database subsets

Database Subset	#Subjects	Occlusions	Degree of Illumination Changes
CAS-PEAL-R1	Large (300)	Glasses, Hats	Medium
AR	Large (99)	Glasses, Scarf	Low
Yale B	Small (10)	No	High
Caltech	Small (26)	No	Low/Medium

5.1.2 Image Preprocessing

In the experiments, the preprocessing process of the face images is divided into two steps: geometric normalization and illumination normalization[25]. Geometric normalization is to provide features that are invariant to geometric transformations of the face images, such as the location, the rotation, and the scale of the face in the image, and to remove irrelevant information about the face images, such as the background and the hair of a subject. Illumination normalization is to decrease the variations of a face image induced by lighting changes while keeping distinguishing features. The details of the two steps are described below.

In the geometric normalization step, each face image is scaled and rotated so that the eyes are positioned in a horizontal line, and the distance between them equals a predefined length. Then, the face image is cropped to include only the face region with little hair and background. The size of a cropped face image is 64×64. Typically, the hairstyle of a specific subject and the background are constant in a face database; thus, better performance can be obtained with larger face regions.

In the illumination normalization step, the histogram equalization (HE) method is used. The diagram of the face image processing procedures is shown in Figure 5-5.

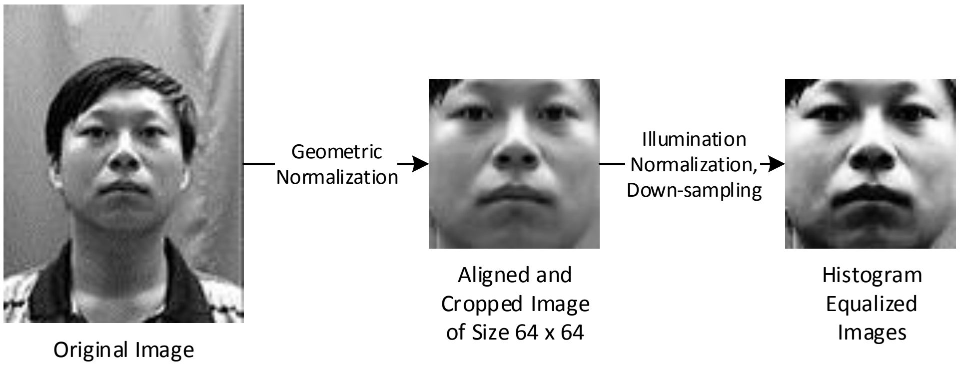


Figure 5-5 Diagram of face image preprocessing procedures.

5.2 Experimental Results

5.2.1 Evaluation of FR Algorithms on LR Face Images

5.2.1.1 Parameter Settings

In this section, experiments are undertaken to test the performances of tradition face-recognition algorithms on low-resolution face images directly. As is described in Section 1.1, for the first standard approach, we use the LR probes to match the down-sampled gallery images. In the experiments, images with resolutions of 4×4, 5×5, …, 64×64 are obtained, and the evaluated baseline FR algorithms are PCA, PCA+LDA, G-PCA+LDA and LGBPHS.

For the PCA projection step in the PCA methods, the eigenvectors corresponding to the largest eigenvalues are first selected such that 99% of the total energy are retained. However, in choosing the projection which maximizes the total scatter, PCA retains unwanted variations due to lighting, occlusions and other facial variations[5]. The variation between the images of the same face due to illumination and occlusions are almost always larger than image variations due to change in face identity[29]. Consequently, the points in the projected space will not be well clustered, and worse, the classes may be smeared together. It has been suggested that by discarding a certain number of the most significant principal components, the variation due to lighting and occlusions is reduced. This number of the most significant principal components varies with different database subsets and different resolution of face images. Hence, in the experiments, we test the PCA method by further discarding different number of the most significant principal components for each resolution of each database subset, and the best results are presented. Both the number of principal components used for classification and the number of discarded principal components are given in the appendix.

For the PCA projection step in the PCA+LDA and G-PCA+LDA methods, the eigenvectors corresponding to the largest eigenvalues are selected such that 99% of the total energy are retained. The number of principal components used for classification are given in the appendix.

To avoid unnecessary computation in the G-PCA+LDA method, for images with resolution greater than 16×16, the each GMP are down-sampled to 16×16, since higher resolution of GMP makes no contribution to improve the identification performance according to earlier experiments.

For the LGBPHS method, each LGBP map are divided into non-overlap regions with the size of 4×4 in the experiments.

5.2.1.2 Results and Discussion

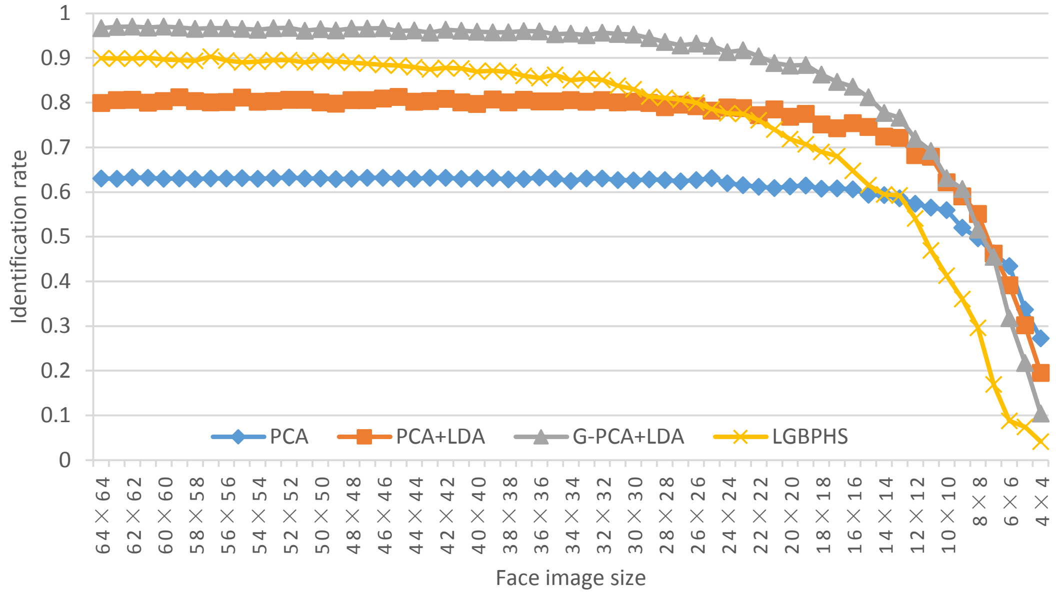


Figure 5-6 Identification performance of the four baseline FR algorithms on LR face images tested on the CAS-PEAL-R1 subset.

As shown in Figure 5-6, the identification performance of the four baseline FR algorithms drops when the resolution of face images drops, and LGBPHS, as a non-statistical-learning-based method, is more sensitive to resolution reduction. We can also observe the trend that PCA+LDA outperforms PCA, G-PCA+LDA outperforms PCA+LDA, while the performance of G-PCA+LDA is first to drop with the decreasing of the image size, then comes with PCA+LDA and PCA. Gabor feature may be helpful to improve the identification performance at high-resolution level, but it becomes more sensitive to the decreasing of the image resolution. The performance of LGBPHS is between those of PCA+LDA and G-PCA+LDA when image resolution is higher than

30 × 30. When face image size is smaller than 16 × 16, the identification performances of all the baseline algorithms begin to drop radically.

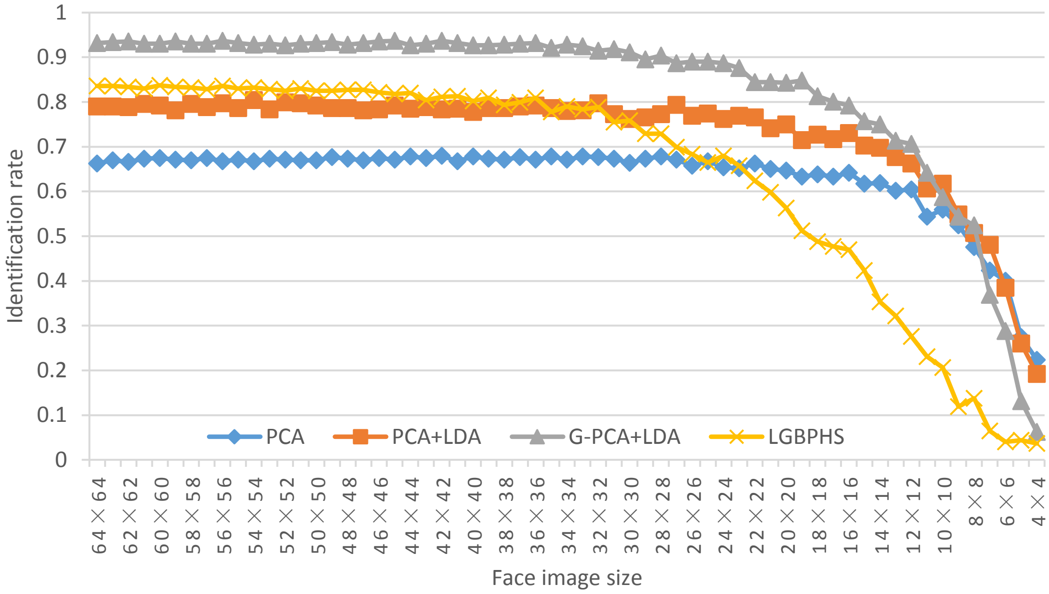


Figure 5-7 Identification performance of the four baseline FR algorithms on LR face images tested on the AR subset.

Figure 5-7 illustrates the identification performances of the baseline FR algorithms on LR face images tested on the AR subset and it shows the same trends as Figure 5-6 does. By comparing the results in Figure 5-6 and Figure 5-7 closely, we can find out that the PCA+LDA, G-PCA+LDA and LGBPHS methods produces lower identification rates when tested on the AR subset, while PCA produces higher identification rates. As we have discussed in Section 5.1.1, the AR subset contains occlusions (sunglasses, scarf) that might covers nearly half the region of the face image, which would seriously affect the performances of the baseline FR algorithms. We have mentioned in Section 5.2.1.1 that for PCA, by discarding a certain number of the most significant principal components, the variation due to lighting and occlusions can be reduced. Applying this trick, the performance of PCA can be improved. However, since LDA already considers maximizing the between-class scatter while minimizing the within-class scatter, the above mentioned trick is no longer necessarily applicable to the

PCA+LDA and G-PCA+LDA methods. Our experiments also show that the identification performances of the PCA+LDA and G-PCA+LDA methods will drop when the trick that discards a certain number of principal components is applied in the PCA step. For LGBPHS, it neither includes PCA processing nor considers the between- and within-class scatter as the LDA-based methods do, so LGBPHS produces lower identification rates when tested on the AR subset.

Figure 5-8 and figure 5-9 illustrate the identification performances of the four baseline FR algorithms on LR face images tested on the Yale B subset and the Caltech subset. According to the results, we can observe the trends that the identification rates drop when the image size decreases. However, the results are much better than those in Figure 5-6 and Figure 5-7. This is mainly because the number of subjects in these two database subsets are small, so it is easier for the evaluated FR algorithms to produce better classification results.

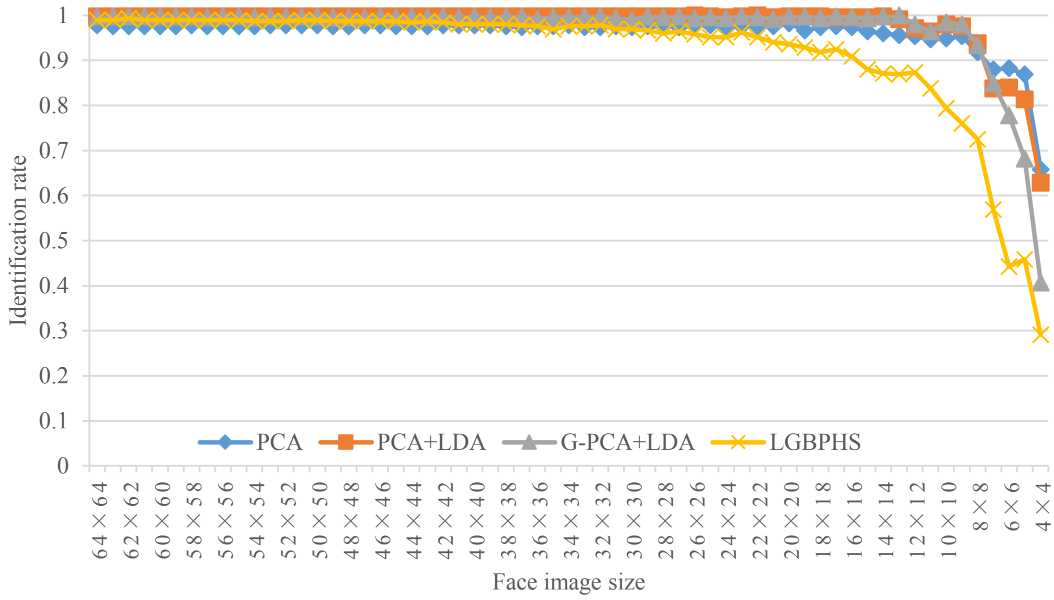


Figure 5-8 Identification performance of the four baseline FR algorithms on LR face images tested on the Yale B subset.

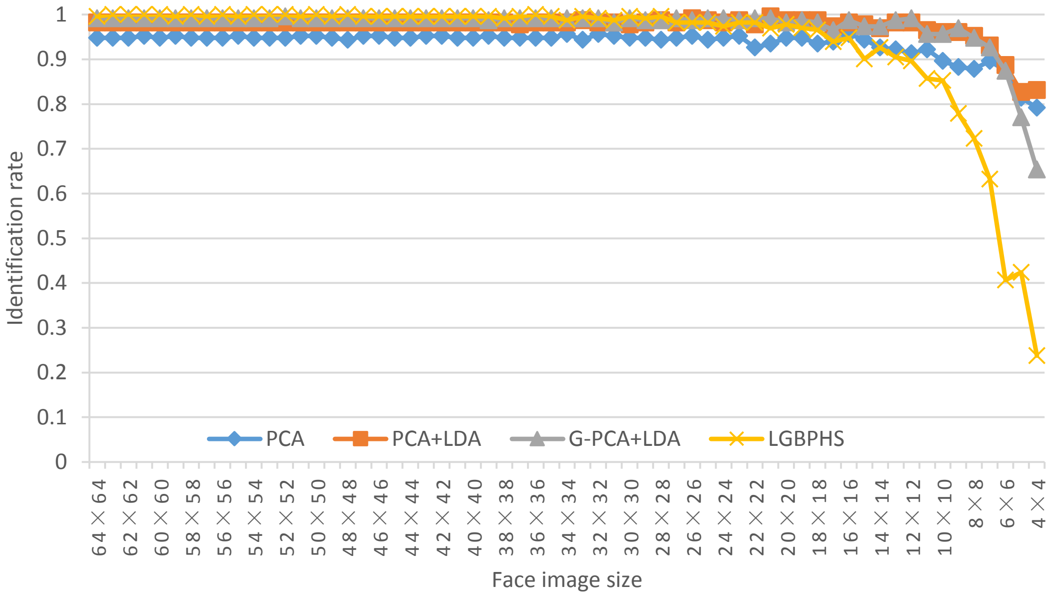


Figure 5-9 Identification performance of the four baseline FR algorithms on LR face images tested on the Caltech subset.

To better illustrate the algorithm performance on low-resolution face images, we consider continue to use the CAS-PEAL-R1 subset and the AR subset for evaluation in subsequent experiments.

5.2.2 Evaluation of SR Algorithms

5.2.2.1 Parameter Settings

In this section, we evaluate the performance of the baseline super-resolution algorithms on the CAS-PEAL-R1 subset and the AR subset. The three baseline SR algorithms are bicubic interpolation, eigentransformation and CLLR-SR. According to the experimental results in Section 5.2.1.2, we obtain face images with resolution of 16×16, 12×12, and 8×8, which are considered as low-resolutions. All the LR face images are first super-resolved to 48×48, which is considered as high-resolution, by the three baseline SR algorithms, respectively. Then, we evaluate the performances the SR algorithms by two measurements: Peak signal-to-noise ratio (PSNR)[30] and structural similarity (SSIM)[31].

For images super-resolved by the eigentransformation and the CLLR-SR methods, the patch-based can improve the performance and smoothing masks are used to merge the overlapped patches. In detail, LR images are divided into patches of size 4×4 with 50% overlap.

For the PCA used in the eigentransformation and the CLLR-SR methods, the number of the eigenvectors is selected such that 99% of the total energy is retained.

For a fair comparison, in eigentransformation, the positive scale parameter, a, is set to be 1, i.e.

$$a = 1.$$

In the CLLR-SR method, the neighbourhood sizes in the global reconstruction step and the residual compensation step are 80 and 100, respectively, i.e.

$$K^G = 80,\ K^R = 100.$$

5.2.2.2 Results and Discussion

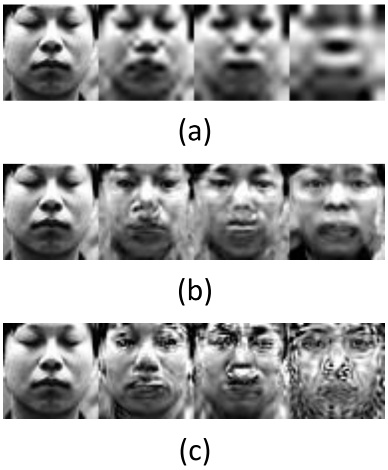
Table 5-3 Mean PSNR and SSIM of super-resolved images (CAS-PEAL-R1)

Probe Image Size	16×16		12×12		8×8	
SR Algorithms	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic	19.6786	0.7941	17.1842	0.6626	14.5991	0.4647
Eigentransformation	16.8320	0.6839	15.6176	0.6011	14.3213	0.5048
CLLR-SR	16.1310	0.6809	15.3689	0.5534	11.0385	0.2832

Table 5-4 Mean PSNR and SSIM of super-resolved images (AR)

Probe Image Size	16×16		12×12		8×8	
SR Algorithms	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic	19.6604	0.7381	17.5626	0.6203	15.3925	0.4656
Eigentransformation	17.2987	0.6212	16.0508	0.5347	14.8050	0.4435
CLLR-SR	17.4831	0.6263	15.4254	0.4789	10.9210	0.2316

Table 5-4, Table 5-5 show the results of the PSNR and SSIM measurements between the original HR images and the images super-resolved from different sizes by different SR algorithms. Note that only mean values of the PSNR and SSIM measurement results on each subset are presented in the tables above.

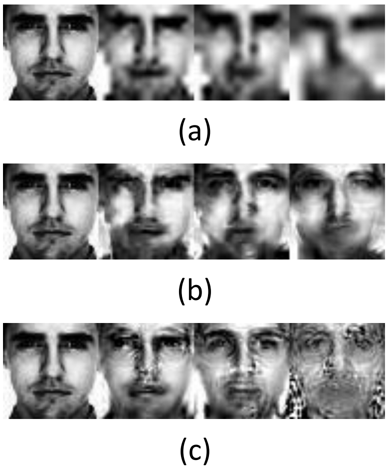


(a)

(b)

(c)

Figure 5-10 Super-resolution results based on the CAS-PEAL-R1 subset using different methods.
(a) Bicubic interpolation; (b) eigentransformation; (c) CLLR-SR.
The 1st column shows the original HR images, while the 2nd, 3rd, and 4th columns show the SR results from LR images of sizes 16×16, 12×12, and 8×8, respectively.



(a)

(b)

(c)

Figure 5-11 Super-resolution results based on the AR subset using different methods.
(a) Bicubic interpolation; (b) eigentransformation; (c) CLLR-SR.
The 1st column shows the original HR images, while the 2nd, 3rd, and 4th columns show the SR results from LR images of sizes 16×16, 12×12, and 8×8, respectively.

Figure 5-10 and Figure 5-11 show the sample super-resolution results based on the CAS-PEAL-R1 subset and the AR subset, respectively.

Referring to Table 5-3 and Table 5-4, bicubic interpolation produces the results with the best PSNR and SSIM performances. Referring to Figure 5-10 and Figure 5-11, we can observe that face images produced by eigentransformation and CLLR are sharper than those produced by bicubic interpolation, but with more non-face-like distortions. However, when the face image is small, neither bicubic interpolation nor the statistical-learning-based SR methods can produce a satisfactory estimation of the HR face image.

The reasons can be summarized in two aspects. (1) Bicubic interpolation is a non-statistical-learning-based SR method, which takes information only from the LR probe image itself to produce an HR estimation. Therefore, when the image size is too small to offer enough information for estimating the HR face image, bicubic interpolation failed to give a clear and sharp HR estimation. (2) Eigentransformation and CLLR-SR are two statistical-learning-based SR methods, both of which can learn information from HR training images to render an HR estimation of the LR probe image. Usually, in order to obtain satisfactory results, sufficient training of the SR algorithms is indispensable to acquire proper parameter settings. There is no way find general parameter settings for different probe images and different training sets, and finding proper parameter settings always need a lot of attempts, especially for complex SR algorithms, such as CLLR-SR.

5.2.3 Evaluation of FR Algorithms on SR Face Images

5.2.3.1 Parameter Settings

In this section, we evaluate the performances face super-resolution methods on face recognition. We test the four baseline FR algorithms using the super-resolved face images from Section 5.2.2.2, and compare the results with those obtained in Section 5.2.1.2.

For the PCA projection step in the PCA methods, the eigenvectors corresponding to the largest eigenvalues are first selected such that 99% of the total energy are retained. In the experiments, we test the PCA method by further discarding different number of the most significant principal components for each resolution of each database subset, and the best results are presented. Both the number of principal components used for classification and the number of discarded principal components are given in the appendix.

For the PCA projection step in the PCA+LDA and G-PCA+LDA methods, the eigenvectors corresponding to the largest eigenvalues are selected such that 99% of the total energy are retained. The number of principal components used for classification are given in the appendix.

To avoid unnecessary computation in the G-PCA+LDA method, for images with resolution greater than 16×16, the each GMP are down-sampled to 16×16, since higher resolution of GMP makes no contribution to improve the identification performance according to earlier experiments.

For the LGBPHS method, each LGBP map are divided into non-overlap regions with the size of 4×4 in the experiments.

5.2.3.2 Results and Discussion

Figure 5-12, Figure 5-13, Figure 5-14 and Figure 5-15 illustrate the identification performances of the four baseline algorithms with the super-resolved face images tested on the CAS-PEAL-R1 subset. We have already evaluated the SR algorithms and observed the SR results in Section 5.2.2.2, which are not satisfactory. Hence, as shown in the figures, the identification performances are not improved by applying super-resolution to the LR probe images. What’s more, for G-PCA+LDA, the identification rates drop to nearly zero.

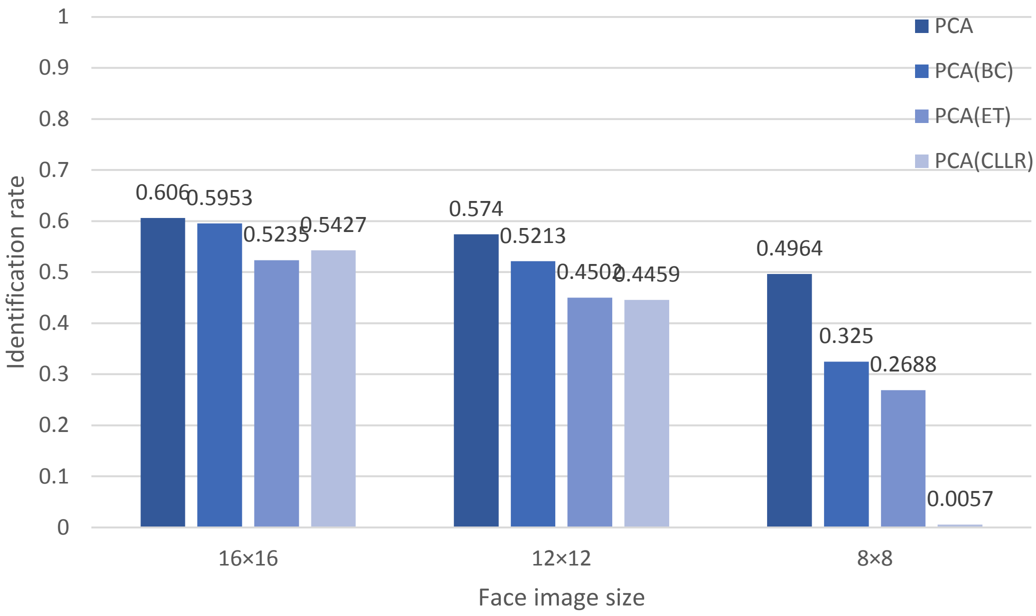


Figure 5-12 Identification performance of PCA on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset.

BC: bicubic interpolation; ET: eigentransformation; CLLR: CLLR-SR.

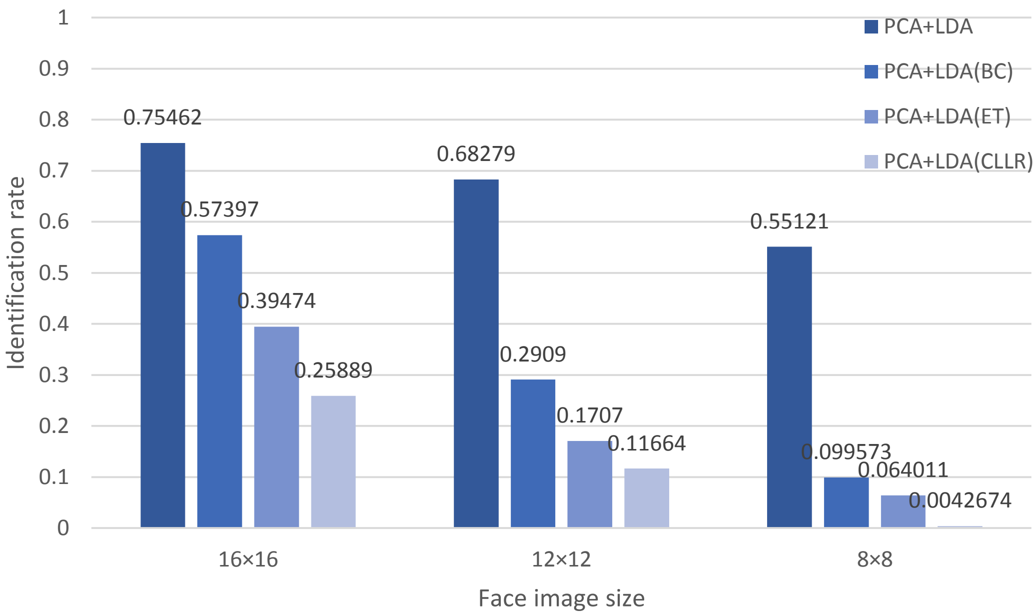


Figure 5-13 Identification performance of PCA+LDA on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset.

BC: bicubic interpolation; ET: eigentransformation; CLLR: CLLR-SR.

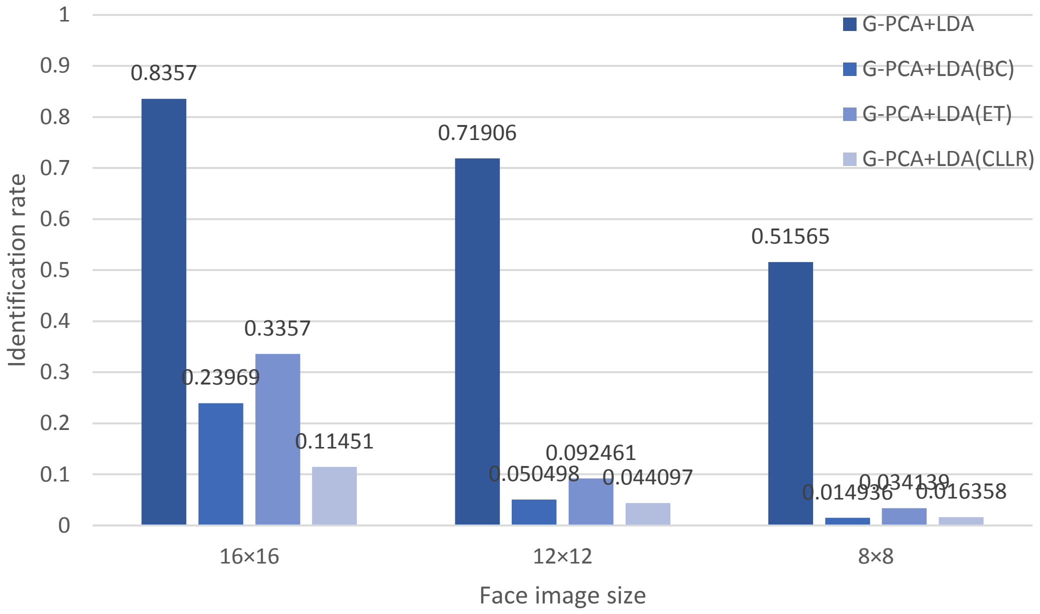


Figure 5-14 Identification performance of G-PCA+LDA on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset.

BC: bicubic interpolation; ET: eigentransformation; CLLR: CLLR-SR.

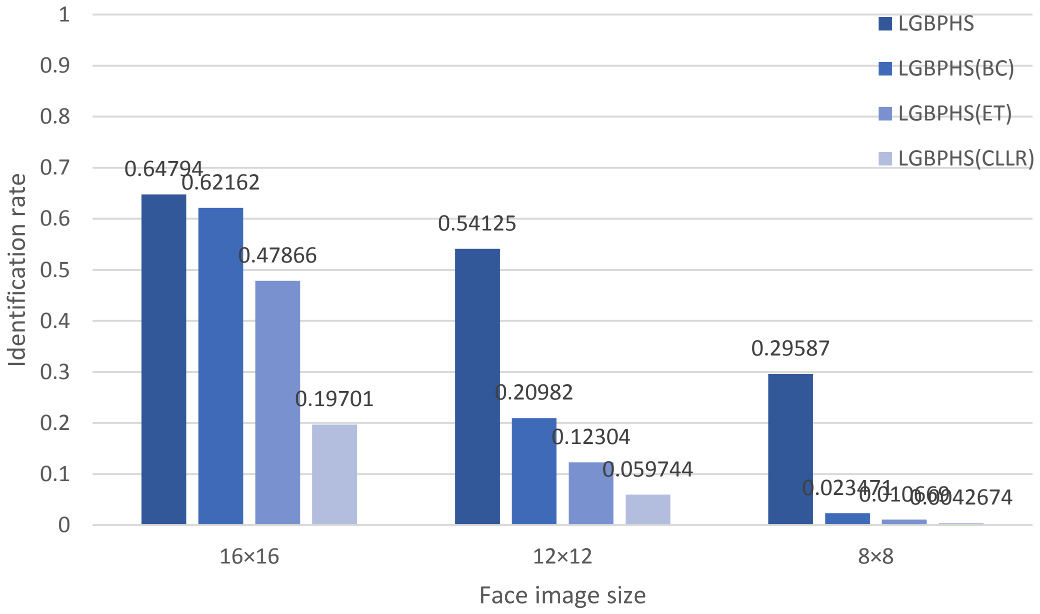


Figure 5-15 Identification performance of LGBPHS on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset.

BC: bicubic interpolation; ET: eigentransformation; CLLR: CLLR-SR.

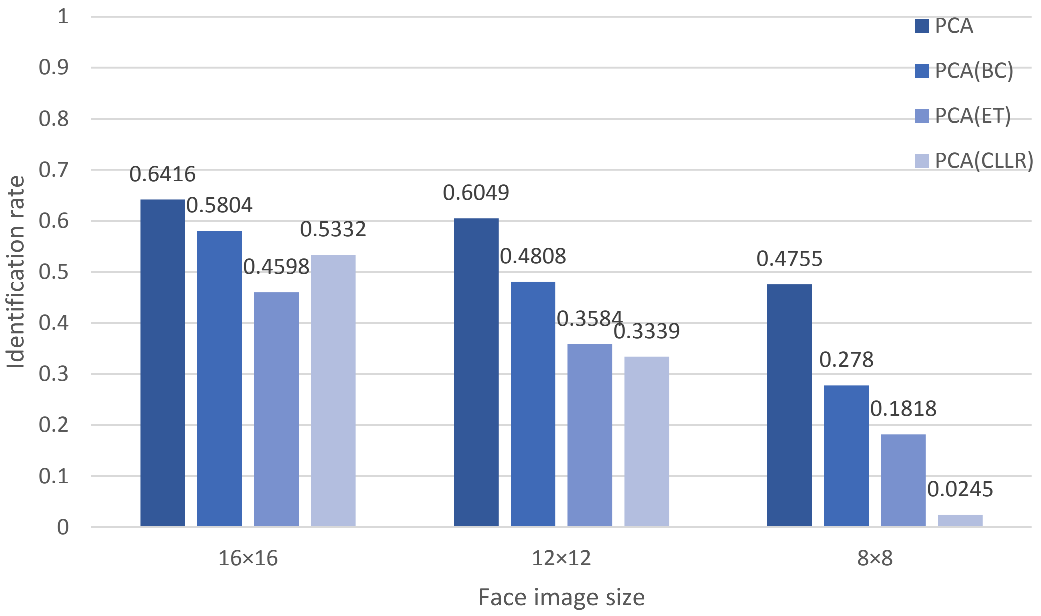


Figure 5-16 Identification performance of PCA on the three baseline SR algorithms tested on the AR subset.

BC: bicubic interpolation; ET: eigentransformation; CLLR: CLLR-SR.

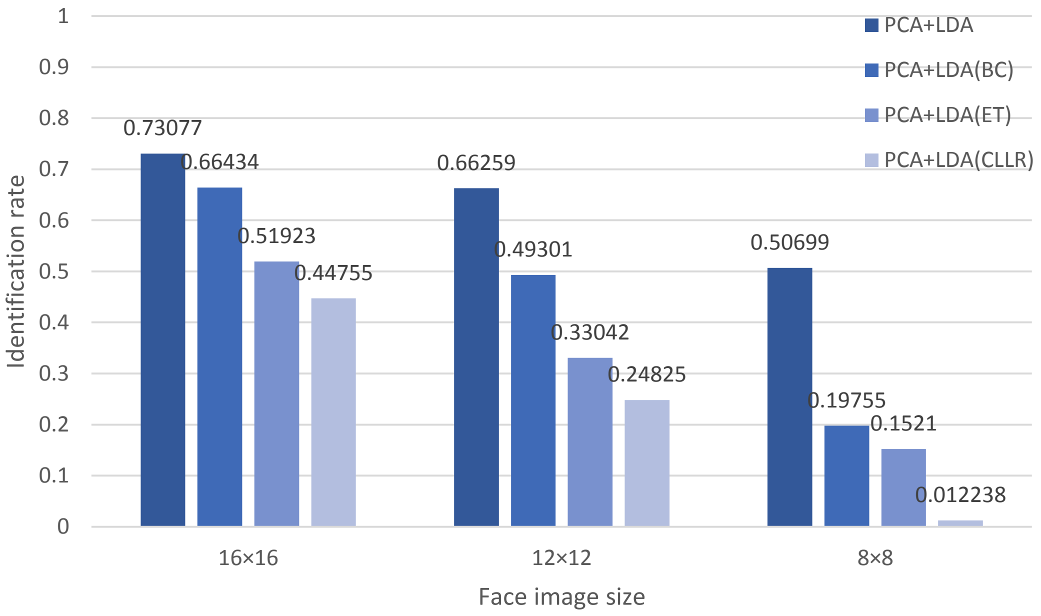


Figure 5-17 Identification performance of PCA+LDA on the three baseline SR algorithms tested on the AR subset.

BC: bicubic interpolation; ET: eigentransformation; CLLR: CLLR-SR.

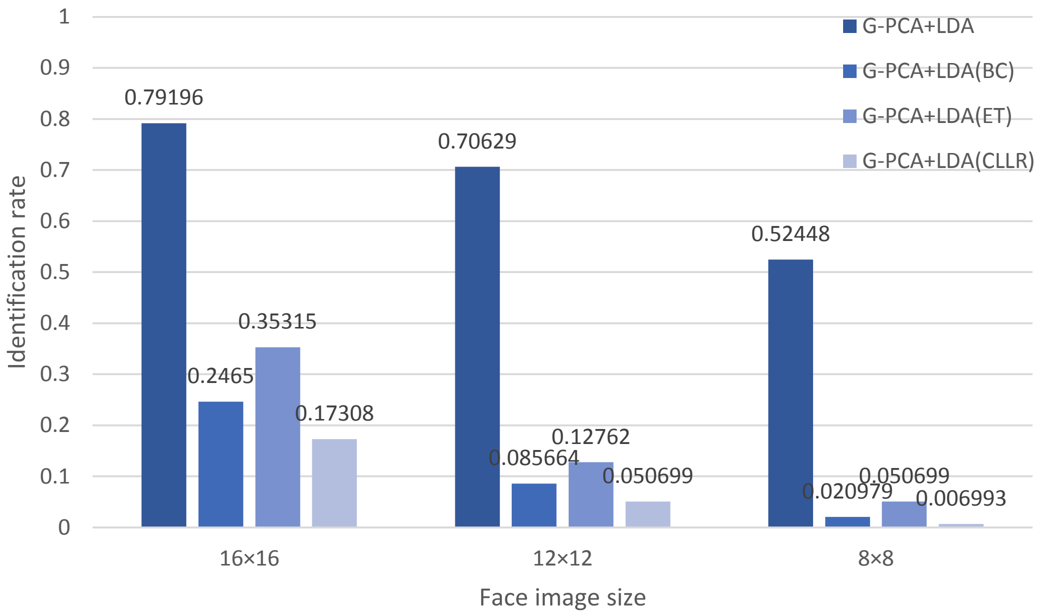


Figure 5-18 Identification performance of G-PCA+LDA on the three baseline SR algorithms tested on the AR subset.

BC: bicubic interpolation; ET: eigentransformation; CLLR: CLLR-SR.

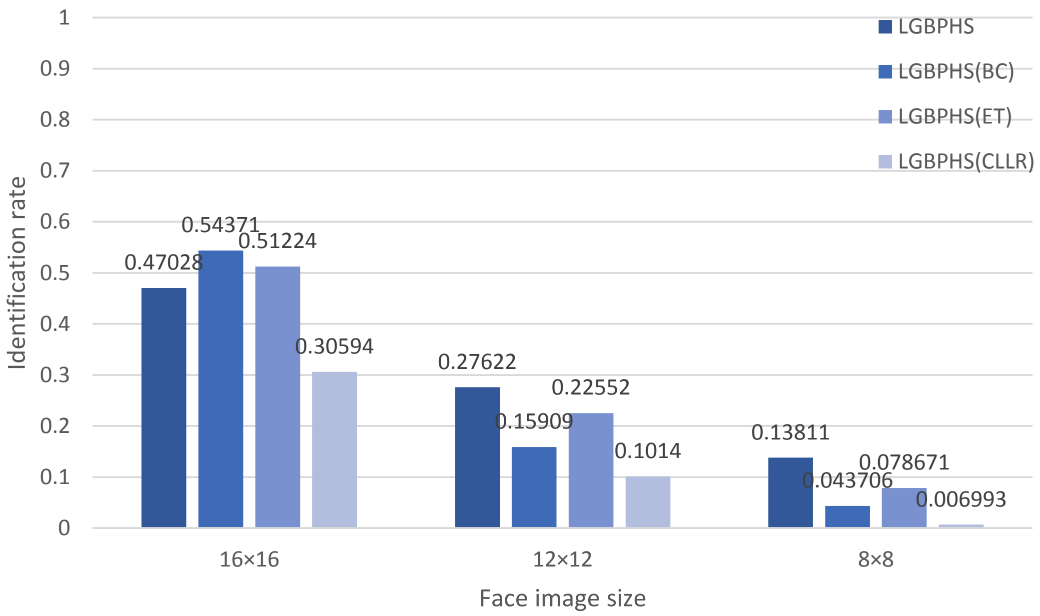


Figure 5-19 Identification performance of LGBPHS on the three baseline SR algorithms tested on the AR subset.

BC: bicubic interpolation; ET: eigentransformation; CLLR: CLLR-SR.

Figure 5-16, Figure 5-17, Figure 5-18 and Figure 5-19 illustrate the identification performances of the four baseline algorithms with the super-resolved face images tested on the AR subset. Again, the results show that in this experiment, super-resolving the LR probe images is not helpful to improve the performance of the face recognition algorithms on LR face images. Referring to Figure 5-19, there is an exception that LGBPHS produces better results with the probe face images super-resolved by bicubic interpolation and eigentransformation, when the resolution is 16×16. It is because that the super-resolved face image size is much larger than the size of original LR face images, thus resulting more regions and more LGBP map histograms for the matching step, which might be helpful to improve the performance of the LGBPHS algorithm.

In general, the bad identification performances illustrated in this section are due to unsatisfactory super-resolution results. We believe that if the SR algorithms can give a satisfactory estimation of the HR images, the identification performances can be improved. However, the SR algorithms are not suitable for face recognition, because the non-statistical-learning-based methods will fail to give a sharp HR estimation and the statistical-learning-based methods always involve complicated training procedures to obtain a low-distortion HR estimation, which is not recommended in practical applications.

5.2.4 Evaluation of Coupled-Projection Methods

5.2.4.1 Parameter Settings

In this section, we are going to evaluate the identification performance of two coupled-projection methods, PCA+CLDA and G-PCA+CLDA. The two methods are tested on the CAS-PEAL-R1 subset and the AR subset, and the results will be compared with those of PCA+LDA and G-PCA+LDA, respectively. In the experiments, we consider 48×48 as the high-resolution and 4×4, 5×5, $\ldots$, 32×32 as the low-resolutions.

For the PCA projection step in the PCA+CLDA and G-PCA+CLDA methods, the eigenvectors corresponding to the largest eigenvalues are selected such that 99% of the total energy are retained. The number of principal components used for classification are given in the appendix.

To avoid unnecessary computation in the G-PCA+CLDA method, for images with resolution greater than 16×16, the each GMP are down-sampled to 16×16, since higher resolution of GMP makes no contribution to improve the identification performance according to earlier experiments.

5.2.4.2 Results and Discussion

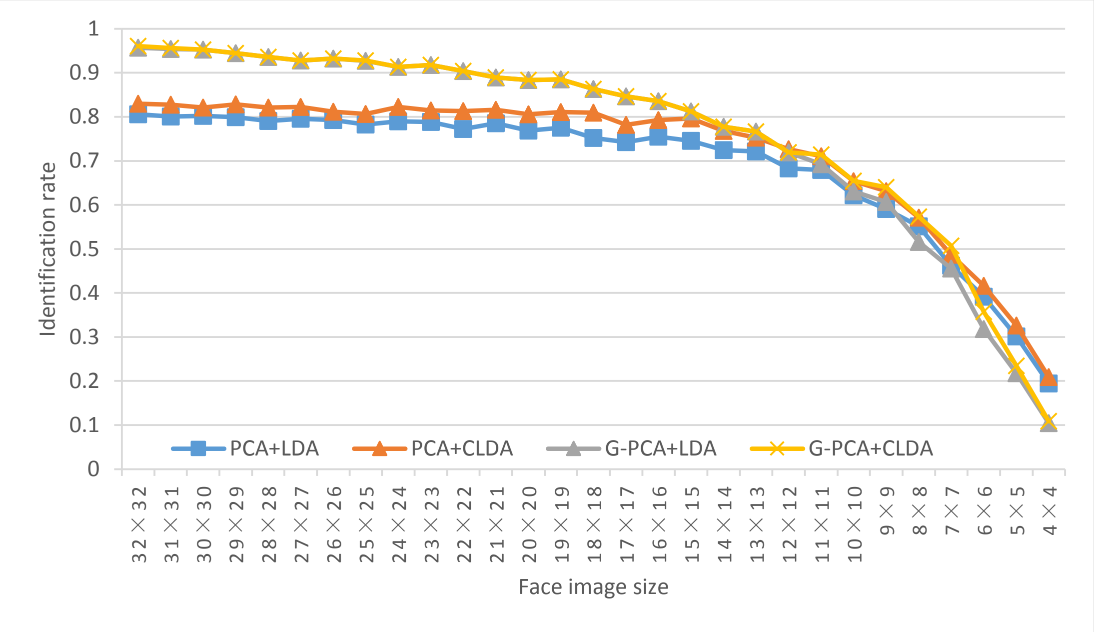


Figure 5-20 Identification performance of PCA+CLDA vs. PCA+LDA and G-PCA+CLDA vs. G-PCA+LDA tested on the CAS-PEAL-R1 subset.

Figure 5-20 and Figure 5-21 illustrate the identification performances of PCA+CLDA vs. PCA+LDA and G-PCA+CLDA vs. G-PCA+LDA experimented on the CAS-PEAL-R1 subset and the AR subset, respectively. The results show that the coupled-projection methods produce higher identification rates than the traditional FR methods do. CCA can be used to enhance the performance of the traditional FR algorithms, such as PCA+LDA and G-PCA+LDA, on LR face images. By applying CCA, the coupled-projection methods maximize the correlation between the HR and LR face images, thus improving the classification performance of the traditional FR algorithms. In a addition, in comparison with the CLLR-SR based FR approaches, the coupled –projection methods are less complex and include fewer computations. For example, the CLLR-SR algorithm has to back-project to image feature in the coherent subspace to the PCA space and then reconstruct the final face image, while the coupled-projection methods do not involve those back-projections.

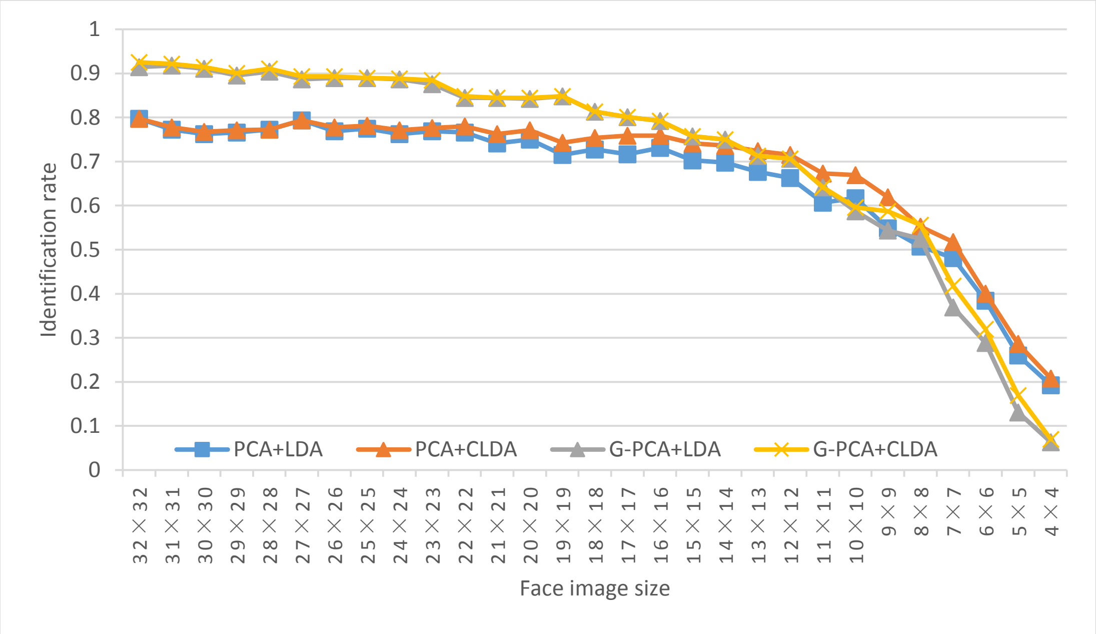


Figure 5-21 Identification performance of PCA+CLDA vs. PCA+LDA and G-PCA+CLDA vs. G-PCA+LDA tested on the AR subset.

Chapter 6 Conclusion

In this project, we focus on face recognition on low-resolution image. First, the problem of low-resolution face recognition is discussed and three standard approaches to recognizing a low-resolution face image are illustrated. Then, four baseline face recognition algorithms and three baseline super-resolution algorithms are introduced. After that, a coupled-projection method inspired by and based on the CCA algorithm is proposed.

The performances of the FR and SR algorithms are evaluated for face images at different resolution levels. The results shows that identification rate degrades as the face image resolution drops. However, the degree of the performance drop on different approaches changes sharply. In details, face recognition based on facial features, such as Gabor feature and LBP, appear more robust to the decreasing of the image size. Super-resolution of the LR probe images is considered as an approach to improve the performance of the evaluated FR algorithms, but the results are not ideal, and even worse. Since super-resolution may improve the quality of the image visually, it may not be useful enough to increase the identification rate. More efficiently, the coupled-projection method, which neither down-samples the gallery images nor super-resolves the LR probe images to a higher scale, project the LR and HR images into a common subspace, and then perform matching LR and HR representations in the subspace of the same dimension. Hence, traditional classifier, like LDA, can be used. Results demonstrate that the coupled-projection method combined with FR algorithms based on facial features, such as G-PCA+CLDA, outperforms other approaches on recognition of LR images.

Appendix I

Identification rates of the four baseline FR algorithms on LR face images tested on the CAS-PEAL-R1 subset (corresponding to Figure 5-6)

Image Size	PCA	PCA+LDA	G-PCA+LDA	LGBPHS
64 × 64	0.63016	0.79943	0.96728	0.89972
63 × 63	0.62945	0.80583	0.97013	0.899
62 × 62	0.63229	0.80654	0.97084	0.899
61 × 61	0.63158	0.80014	0.96871	0.90043
60 × 60	0.62945	0.8037	0.97084	0.89687
59 × 59	0.63016	0.81223	0.96871	0.89545
58 × 58	0.62873	0.8037	0.96515	0.89403
57 × 57	0.63016	0.80085	0.96728	0.90327
56 × 56	0.63016	0.80156	0.96657	0.89474
55 × 55	0.63087	0.81152	0.96515	0.89047
54 × 54	0.62945	0.80228	0.96302	0.89189
53 × 53	0.63087	0.8037	0.96728	0.89545
52 × 52	0.63229	0.80654	0.96799	0.89545
51 × 51	0.63016	0.80654	0.96088	0.89047
50 × 50	0.63016	0.80085	0.96444	0.89474
49 × 49	0.62873	0.79801	0.96159	0.8926
48 × 48	0.62945	0.80583	0.96586	0.88976
47 × 47	0.63158	0.80583	0.96586	0.88762
46 × 46	0.63158	0.80939	0.96728	0.88478
45 × 45	0.63016	0.81294	0.96017	0.88336
44 × 44	0.62945	0.80228	0.96159	0.8798
43 × 43	0.63158	0.8037	0.95661	0.87482
42 × 42	0.63158	0.80868	0.96302	0.87838
41 × 41	0.62945	0.80085	0.96088	0.87696
40 × 40	0.63016	0.7973	0.95875	0.86984
39 × 39	0.63087	0.80725	0.95733	0.87198
38 × 38	0.62802	0.80085	0.95733	0.86913
37 × 37	0.62873	0.80654	0.96017	0.8606
36 × 36	0.63229	0.80299	0.95946	0.85491
35 × 35	0.62945	0.80299	0.95306	0.86273
34 × 34	0.62447	0.80583	0.95448	0.84993
33 × 33	0.63016	0.80228	0.95092	0.85349
32 × 32	0.63016	0.80583	0.95661	0.85064
31 × 31	0.6266	0.80085	0.95377	0.83784

30 × 30	0.62589	0.80228	0.95235	0.83001
29 × 29	0.62802	0.79943	0.94452	0.81366
28 × 28	0.6266	0.79018	0.93599	0.81081
27 × 27	0.62376	0.79587	0.92817	0.80654
26 × 26	0.6266	0.79232	0.93243	0.80014
25 × 25	0.63087	0.78236	0.92745	0.78592
24 × 24	0.61949	0.78947	0.91323	0.77667
23 × 23	0.61522	0.78805	0.9175	0.77525
22 × 22	0.61166	0.7724	0.90398	0.76102
21 × 21	0.60882	0.78521	0.88905	0.73969
20 × 20	0.61238	0.76885	0.88336	0.71835
19 × 19	0.61451	0.77525	0.88478	0.70697
18 × 18	0.6074	0.75178	0.86344	0.6899
17 × 17	0.60811	0.74253	0.84637	0.68065
16 × 16	0.60597	0.75462	0.8357	0.64794
15 × 15	0.59388	0.74538	0.81223	0.61522
14 × 14	0.59317	0.72404	0.77738	0.59459
13 × 13	0.58606	0.72119	0.76671	0.59246
12 × 12	0.57397	0.68279	0.71906	0.54125
11 × 11	0.56543	0.67923	0.69203	0.46942
10 × 10	0.55974	0.62162	0.63087	0.41252
9 × 9	0.51991	0.59033	0.60669	0.3606
8 × 8	0.49644	0.55121	0.51565	0.29587
7 × 7	0.46302	0.4623	0.45519	0.16927
6 × 6	0.43457	0.39118	0.31792	0.087482
5 × 5	0.33713	0.30156	0.21764	0.073969
4 × 4	0.2724	0.19488	0.10455	0.041252

Identification rates of the four baseline FR algorithms on LR face images tested on the AR subset (corresponding to Figure 5-7)

Image Size	PCA	PCA+LDA	G-PCA+LDA	LGBPHS
64 × 64	0.662587	0.79021	0.93182	0.83566
63 × 63	0.66958	0.79021	0.93357	0.83566
62 × 62	0.666084	0.78846	0.93531	0.83392
61 × 61	0.673077	0.79545	0.93007	0.83042
60 × 60	0.674825	0.79196	0.93007	0.83741
59 × 59	0.671329	0.78147	0.93531	0.83392
58 × 58	0.66958	0.79545	0.93007	0.83217
57 × 57	0.674825	0.78846	0.93007	0.82867
56 × 56	0.667832	0.7972	0.93706	0.83566

55×55	0.671329	0.78671	0.93182	0.83042
54×54	0.667832	0.8042	0.92832	0.83217
53×53	0.673077	0.78322	0.93007	0.82867
52×52	0.671329	0.79895	0.92657	0.82517
51×51	0.66958	0.7972	0.93007	0.83042
50×50	0.66958	0.79196	0.93182	0.82517
49×49	0.676573	0.78671	0.93357	0.82517
48×48	0.673077	0.78671	0.92832	0.82692
47×47	0.66958	0.78147	0.93182	0.82692
46×46	0.674825	0.78322	0.93531	0.82168
45×45	0.671329	0.79196	0.93706	0.81818
44×44	0.678322	0.78497	0.92657	0.81993
43×43	0.674825	0.78846	0.93007	0.8042
42×42	0.68007	0.78322	0.93706	0.81119
41×41	0.667832	0.78497	0.93182	0.81294
40×40	0.678322	0.77797	0.92657	0.80245
39×39	0.673077	0.78671	0.92657	0.80944
38×38	0.671329	0.78671	0.92832	0.79371
37×37	0.676573	0.79021	0.93007	0.79895
36×36	0.671329	0.79196	0.93182	0.80944
35×35	0.678322	0.78671	0.92133	0.77797
34×34	0.671329	0.77972	0.92832	0.79021
33×33	0.678322	0.78147	0.92483	0.78322
32×32	0.676573	0.7972	0.91434	0.78846
31×31	0.673077	0.77273	0.91783	0.75524
30×30	0.664336	0.76224	0.91084	0.75874
29×29	0.673077	0.76573	0.8951	0.72902
28×28	0.678322	0.77273	0.90385	0.72902
27×27	0.671329	0.79371	0.88636	0.6993
26×26	0.657343	0.76923	0.88986	0.68357
25×25	0.667832	0.77448	0.88986	0.66434
24×24	0.653846	0.76224	0.88636	0.68007
23×23	0.652098	0.76923	0.87587	0.65734
22×22	0.662587	0.76573	0.84441	0.62413
21×21	0.65035	0.74126	0.84441	0.5979
20×20	0.646853	0.75	0.84266	0.56294
19×19	0.632867	0.71503	0.8479	0.51224
18×18	0.638112	0.72727	0.81294	0.48776
17×17	0.632867	0.71678	0.8007	0.47727
16×16	0.641608	0.73077	0.79196	0.47028

15 × 15	0.617133	0.7028	0.75699	0.42308
14 × 14	0.618881	0.69755	0.75	0.35315
13 × 13	0.601399	0.67657	0.71329	0.32168
12 × 12	0.604895	0.66259	0.70629	0.27622
11 × 11	0.543706	0.60664	0.64161	0.23077
10 × 10	0.559441	0.61713	0.58741	0.20629
9 × 9	0.524476	0.54895	0.54371	0.11888
8 × 8	0.475524	0.50699	0.52448	0.13811
7 × 7	0.423077	0.48077	0.36888	0.064685
6 × 6	0.40035	0.38462	0.28846	0.04021
5 × 5	0.274476	0.26049	0.13112	0.043706
4 × 4	0.223776	0.19231	0.062937	0.036713

Identification rates of the four baseline FR algorithms on LR face images tested on the Yale B subset (corresponding to Figure 5-8)

Image Size	PCA	PCA+LDA	G-PCA+LDA	LGBPHS
64 × 64	0.977778	0.997778	0.997778	0.98889
63 × 63	0.975556	0.997778	0.997778	0.98889
62 × 62	0.975556	0.997778	0.997778	0.99333
61 × 61	0.975556	0.997778	0.997778	0.98889
60 × 60	0.975556	0.997778	0.997778	0.98889
59 × 59	0.975556	0.997778	0.997778	0.98889
58 × 58	0.977778	0.997778	0.997778	0.98889
57 × 57	0.975556	0.997778	0.995556	0.98889
56 × 56	0.975556	0.997778	0.997778	0.98889
55 × 55	0.975556	0.997778	0.997778	0.98889
54 × 54	0.975556	0.997778	0.995556	0.98667
53 × 53	0.977778	0.997778	0.997778	0.98667
52 × 52	0.977778	0.997778	0.997778	0.98667
51 × 51	0.977778	0.997778	0.997778	0.98889
50 × 50	0.977778	0.997778	0.997778	0.98889
49 × 49	0.975556	0.997778	0.995556	0.98667
48 × 48	0.975556	0.997778	0.995556	0.98667
47 × 47	0.977778	0.997778	0.995556	0.98667
46 × 46	0.977778	0.997778	0.997778	0.98667
45 × 45	0.975556	0.997778	0.995556	0.98667
44 × 44	0.975556	0.997778	0.997778	0.98444
43 × 43	0.975556	0.997778	0.995556	0.98667
42 × 42	0.977778	0.997778	0.997778	0.98444
41 × 41	0.977778	0.997778	0.995556	0.98

40×40	0.977778	0.997778	0.997778	0.98
39×39	0.977778	0.997778	0.995556	0.98
38×38	0.975556	0.997778	0.997778	0.98
37×37	0.973333	0.997778	0.997778	0.97778
36×36	0.975556	0.997778	0.991111	0.97556
35×35	0.975556	0.997778	0.995556	0.96889
34×34	0.977778	0.997778	0.995556	0.97556
33×33	0.973333	0.997778	0.997778	0.97556
32×32	0.973333	0.997778	0.997778	0.97778
31×31	0.975556	0.997778	0.993333	0.96889
30×30	0.973333	0.997778	0.988889	0.96889
29×29	0.975556	0.997778	0.997778	0.96667
28×28	0.975556	0.997778	0.995556	0.96
27×27	0.973333	0.997778	0.997778	0.96444
26×26	0.975556	1	0.993333	0.95778
25×25	0.977778	0.997778	0.995556	0.95111
24×24	0.973333	0.995556	0.993333	0.95111
23×23	0.975556	0.997778	0.997778	0.96222
22×22	0.975556	1	0.995556	0.95111
21×21	0.975556	0.995556	0.991111	0.94
20×20	0.982222	0.997778	0.997778	0.93555
19×19	0.966667	0.997778	0.997778	0.92889
18×18	0.973333	0.997778	0.993333	0.91778
17×17	0.975556	0.995556	0.997778	0.92444
16×16	0.973333	0.995556	0.995556	0.90889
15×15	0.964444	0.995556	0.993333	0.88
14×14	0.96	0.997778	0.997778	0.87111
13×13	0.955556	0.991111	1	0.86889
12×12	0.953333	0.971111	0.98	0.87333
11×11	0.946667	0.964444	0.964444	0.83778
10×10	0.948889	0.98	0.984444	0.79333
9×9	0.953333	0.975556	0.98	0.76
8×8	0.917778	0.937778	0.933333	0.72444
7×7	0.88	0.837778	0.848889	0.56889
6×6	0.882222	0.84	0.777778	0.44222
5×5	0.868889	0.813333	0.682222	0.45778
4×4	0.657778	0.628889	0.406667	0.29111

Identification rates of the four baseline FR algorithms on LR face images tested on the Caltech subset (corresponding to Figure 5-9)

Image Size	PCA	PCA+LDA	G-PCA+LDA	LGBPHS
64 × 64	0.948052	0.982684	0.991342	0.99567
63 × 63	0.948052	0.982684	0.991342	1
62 × 62	0.948052	0.982684	0.991342	1
61 × 61	0.952381	0.982684	0.991342	1
60 × 60	0.948052	0.982684	0.991342	1
59 × 59	0.952381	0.982684	0.991342	0.99567
58 × 58	0.948052	0.982684	0.991342	1
57 × 57	0.948052	0.982684	0.991342	0.99567
56 × 56	0.948052	0.982684	0.991342	1
55 × 55	0.952381	0.982684	0.991342	0.99567
54 × 54	0.948052	0.982684	0.991342	1
53 × 53	0.948052	0.982684	0.991342	1
52 × 52	0.948052	0.982684	0.995671	1
51 × 51	0.952381	0.982684	0.991342	0.99567
50 × 50	0.952381	0.982684	0.991342	1
49 × 49	0.948052	0.982684	0.991342	0.99567
48 × 48	0.943723	0.982684	0.991342	1
47 × 47	0.952381	0.982684	0.991342	0.99567
46 × 46	0.952381	0.982684	0.991342	0.99567
45 × 45	0.948052	0.982684	0.991342	0.99567
44 × 44	0.948052	0.982684	0.991342	0.99567
43 × 43	0.952381	0.982684	0.991342	0.99567
42 × 42	0.952381	0.982684	0.991342	0.99567
41 × 41	0.948052	0.982684	0.991342	0.99567
40 × 40	0.948052	0.982684	0.991342	0.99567
39 × 39	0.952381	0.982684	0.987013	0.99567
38 × 38	0.948052	0.982684	0.991342	0.99134
37 × 37	0.948052	0.978355	0.991342	0.99567
36 × 36	0.948052	0.982684	0.991342	1
35 × 35	0.948052	0.982684	0.991342	0.99567
34 × 34	0.95671	0.982684	0.991342	0.98701
33 × 33	0.943723	0.987013	0.991342	0.99567
32 × 32	0.95671	0.982684	0.991342	0.99134
31 × 31	0.952381	0.982684	0.982684	0.98701
30 × 30	0.948052	0.978355	0.991342	0.99567
29 × 29	0.948052	0.982684	0.987013	0.99134
28 × 28	0.943723	0.987013	0.987013	0.99567

27 × 27	0.948052	0.982684	0.991342	0.98268
26 × 26	0.952381	0.991342	0.987013	0.98268
25 × 25	0.943723	0.987013	0.991342	0.98268
24 × 24	0.948052	0.982684	0.991342	0.97403
23 × 23	0.952381	0.987013	0.987013	0.98268
22 × 22	0.926407	0.978355	0.991342	0.98268
21 × 21	0.935065	0.995671	0.991342	0.9697
20 × 20	0.948052	0.987013	0.982684	0.97835
19 × 19	0.948052	0.987013	0.987013	0.97043
18 × 18	0.935065	0.987013	0.982684	0.96571
17 × 17	0.939394	0.974026	0.965368	0.93939
16 × 16	0.95671	0.982684	0.987013	0.94805
15 × 15	0.943723	0.978355	0.974026	0.90043
14 × 14	0.926407	0.969697	0.974026	0.92641
13 × 13	0.922078	0.982684	0.987013	0.90476
12 × 12	0.91342	0.982684	0.991342	0.8961
11 × 11	0.922078	0.965368	0.95671	0.85714
10 × 10	0.896104	0.961039	0.95671	0.85281
9 × 9	0.883117	0.961039	0.969697	0.77922
8 × 8	0.878788	0.952381	0.948052	0.72294
7 × 7	0.896104	0.930736	0.926407	0.63203
6 × 6	0.874459	0.887446	0.874459	0.40693
5 × 5	0.813853	0.82684	0.770563	0.42424
4 × 4	0.792208	0.831169	0.65368	0.2381

Identification rates of PCA on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset (corresponding to Figure 5-12)

Image Size	Direct	Bicubic	Eigentransformation	CLLR-SR
16 × 16	0.606	0.5953	0.5235	0.5427
12 × 12	0.574	0.5213	0.4502	0.4459
8 × 8	0.4964	0.325	0.2688	0.0057

Identification rates of PCA+LDA on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset (corresponding to Figure 5-13)

Image Size	Direct	Bicubic	Eigentransformation	CLLR-SR
16 × 16	0.75462	0.57397	0.39474	0.25889
12 × 12	0.68279	0.2909	0.1707	0.11664
8 × 8	0.55121	0.099573	0.064011	0.004267

Identification rates of G-PCA+LDA on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset (corresponding to Figure 5-14)

Image Size	Direct	Bicubic	Eigentransformation	CLLR-SR
16 × 16	0.8357	0.23969	0.3357	0.11451
12 × 12	0.71906	0.050498	0.092461	0.044097
8 × 8	0.51565	0.014936	0.034139	0.016358

Identification rates of LGBPHS on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset (corresponding to Figure 5-15)

Image Size	Direct	Bicubic	Eigentransformation	CLLR-SR
16 × 16	0.64794	0.62162	0.47866	0.19701
12 × 12	0.54125	0.20982	0.12304	0.059744
8 × 8	0.29587	0.023471	0.010669	0.004267

Identification rates of PCA on the three baseline SR algorithms tested on the AR subset (corresponding to Figure 5-16)

Image Size	Direct	Bicubic	Eigentransformation	CLLR-SR
16 × 16	0.6416	0.5804	0.4598	0.5332
12 × 12	0.6049	0.4808	0.3584	0.3339
8 × 8	0.4755	0.278	0.1818	0.0245

Identification rates PCA+LDA on the three baseline SR algorithms tested on the AR subset (corresponding to Figure 5-17)

Image Size	Direct	Bicubic	Eigentransformation	CLLR-SR
16 × 16	0.73077	0.66434	0.51923	0.44755
12 × 12	0.66259	0.49301	0.33042	0.24825
8 × 8	0.50699	0.19755	0.1521	0.012238

Identification rates G-PCA+LDA on the three baseline SR algorithms tested on the AR subset (corresponding to Figure 5-18)

Image Size	Direct	Bicubic	Eigentransformation	CLLR-SR
16 × 16	0.79196	0.2465	0.35315	0.17308
12 × 12	0.70629	0.085664	0.12762	0.050699
8 × 8	0.52448	0.020979	0.050699	0.006993

Identification rates LGBPHS on the three baseline SR algorithms tested on the AR subset (corresponding to Figure 5-19)

Image Size	Direct	Bicubic	Eigentransformation	CLLR-SR
16 × 16	0.47028	0.54371	0.51224	0.30594
12 × 12	0.27622	0.15909	0.22552	0.1014
8 × 8	0.13811	0.043706	0.078671	0.006993

Identification rates of PCA+CLDA vs. PCA+LDA and G-PCA+CLDA vs. G-PCA+LDA tested on the CAS-PEAL-R1 subset (corresponding to Figure 5-20)

Image Size	PCA+LDA	PCA+CLDA	G-PCA+LDA	G-PCA+CLDA
32 × 32	0.80583	0.830014	0.95661	0.960882
31 × 31	0.80085	0.827881	0.95377	0.955903
30 × 30	0.80228	0.820768	0.95235	0.952347
29 × 29	0.79943	0.828592	0.94452	0.944523
28 × 28	0.79018	0.820768	0.93599	0.935989
27 × 27	0.79587	0.822191	0.92817	0.928165
26 × 26	0.79232	0.811522	0.93243	0.932432
25 × 25	0.78236	0.806543	0.92745	0.928165
24 × 24	0.78947	0.822191	0.91323	0.913229
23 × 23	0.78805	0.814367	0.9175	0.917496
22 × 22	0.7724	0.812945	0.90398	0.903983
21 × 21	0.78521	0.815789	0.88905	0.889047
20 × 20	0.76885	0.805832	0.88336	0.883357
19 × 19	0.77525	0.810811	0.88478	0.88478
18 × 18	0.75178	0.809388	0.86344	0.863442
17 × 17	0.74253	0.78165	0.84637	0.846373
16 × 16	0.75462	0.792319	0.8357	0.835704
15 × 15	0.74538	0.796586	0.81223	0.812233
14 × 14	0.72404	0.768137	0.77738	0.777383
13 × 13	0.72119	0.752489	0.76671	0.766714
12 × 12	0.68279	0.726174	0.71906	0.719061
11 × 11	0.67923	0.710526	0.69203	0.714083
10 × 10	0.62162	0.653627	0.63087	0.654339
9 × 9	0.59033	0.631579	0.60669	0.640114
8 × 8	0.55121	0.571124	0.51565	0.573257
7 × 7	0.4623	0.486486	0.45519	0.507112
6 × 6	0.39118	0.416074	0.31792	0.357041
5 × 5	0.30156	0.325747	0.21764	0.234708
4 × 4	0.19488	0.209104	0.10455	0.108819

Identification rates of PCA+CLDA vs. PCA+LDA and G-PCA+CLDA vs. G-PCA+LDA tested on the AR subset (corresponding to Figure 5-21)

Image Size	PCA+LDA	PCA+CLDA	G-PCA+LDA	G-PCA+CLDA
32 × 32	0.7972	0.797203	0.91434	0.924825
31 × 31	0.77273	0.777972	0.91783	0.921329
30 × 30	0.76224	0.767483	0.91084	0.914336
29 × 29	0.76573	0.770979	0.8951	0.90035
28 × 28	0.77273	0.772727	0.90385	0.910839
27 × 27	0.79371	0.793706	0.88636	0.893357
26 × 26	0.76923	0.777972	0.88986	0.893357
25 × 25	0.77448	0.781469	0.88986	0.88986
24 × 24	0.76224	0.770979	0.88636	0.888112
23 × 23	0.76923	0.776224	0.87587	0.884615
22 × 22	0.76573	0.77972	0.84441	0.847902
21 × 21	0.74126	0.762238	0.84441	0.844406
20 × 20	0.75	0.770979	0.84266	0.844406
19 × 19	0.71503	0.743007	0.8479	0.847902
18 × 18	0.72727	0.753497	0.81294	0.812937
17 × 17	0.71678	0.758741	0.8007	0.800699
16 × 16	0.73077	0.758741	0.79196	0.791958
15 × 15	0.7028	0.741259	0.75699	0.756993
14 × 14	0.69755	0.736014	0.75	0.75
13 × 13	0.67657	0.723776	0.71329	0.713287
12 × 12	0.66259	0.715035	0.70629	0.706294
11 × 11	0.60664	0.673077	0.64161	0.641608
10 × 10	0.61713	0.66958	0.58741	0.596154
9 × 9	0.54895	0.618881	0.54371	0.587413
8 × 8	0.50699	0.552448	0.52448	0.555944
7 × 7	0.48077	0.517483	0.36888	0.417832
6 × 6	0.38462	0.40035	0.28846	0.31993
5 × 5	0.26049	0.286713	0.13112	0.16958
4 × 4	0.19231	0.208042	0.062937	0.06993

Appendix II

Number of principal components used for classification (corresponding to Figure 5-6 Identification performance of the four baseline FR algorithms on LR face images tested on the CAS-PEAL-R1 subset)

Image Size	PCA		PCA+LDA	G-PCA+LDA
	K	D	K	K
64×64	528	0	528	1013
63×63	505	0	505	1001
62×62	499	0	499	1004
61×61	497	0	497	998
60×60	496	0	496	996
59×59	496	0	496	995
58×58	495	0	495	996
57×57	492	0	492	1003
56×56	493	0	493	1000
55×55	492	0	492	998
54×54	490	0	490	1002
53×53	490	0	490	1001
52×52	489	0	489	999
51×51	488	0	488	997
50×50	486	0	486	999
49×49	484	0	484	1001
48×48	483	0	483	999
47×47	481	0	481	1001
46×46	479	0	479	1004
45×45	477	0	477	1002
44×44	473	1	474	998
43×43	471	0	471	992
42×42	468	0	468	986
41×41	462	1	463	985
40×40	460	0	460	984
39×39	456	0	456	983
38×38	451	0	451	983
37×37	444	0	444	982
36×36	439	0	439	978
35×35	433	0	433	974
34×34	426	0	426	972
33×33	418	0	418	968

32 × 32	402	0	402	956
31 × 31	399	0	399	957
30 × 30	389	1	390	939
29 × 29	381	0	381	925
28 × 28	369	0	369	928
27 × 27	358	0	358	928
26 × 26	344	0	344	919
25 × 25	331	0	331	900
24 × 24	315	1	316	888
23 × 23	302	0	302	890
22 × 22	286	0	286	874
21 × 21	270	1	271	825
20 × 20	253	0	253	797
19 × 19	236	0	236	789
18 × 18	219	0	219	767
17 × 17	202	0	202	733
16 × 16	183	0	183	740
15 × 15	166	0	166	658
14 × 14	148	0	148	633
13 × 13	132	0	132	585
12 × 12	115	0	115	493
11 × 11	99	0	99	437
10 × 10	82	0	82	340
9 × 9	69	0	69	263
8 × 8	56	0	56	193
7 × 7	43	0	43	95
6 × 6	32	0	32	55
5 × 5	23	0	23	28
4 × 4	15	0	15	9
K: Number of principal components used for classification D: Number of discarded principal components				

Number of principal components used for classification (corresponding to Figure 5-7 Identification performance of the four baseline FR algorithms on LR face images tested on the AR subset)

Image Size	PCA		PCA+LDA	G-PCA+LDA
	K	D	K	K
64 × 64	427	6	433	631
63 × 63	406	6	412	627
62 × 62	404	6	410	627
61 × 61	402	6	408	627

60×60	400	6	406	626
59×59	402	6	408	625
58×58	398	6	404	627
57×57	400	6	406	627
56×56	399	6	405	627
55×55	400	6	406	627
54×54	399	6	405	627
53×53	397	6	403	628
52×52	397	6	403	627
51×51	396	6	402	627
50×50	395	6	401	627
49×49	394	6	400	627
48×48	394	6	400	627
47×47	391	7	398	629
46×46	391	6	397	629
45×45	390	6	396	628
44×44	388	6	394	628
43×43	386	6	392	626
42×42	384	6	390	625
41×41	383	6	389	624
40×40	380	6	386	623
39×39	378	6	384	623
38×38	374	6	380	623
37×37	372	6	378	623
36×36	369	6	375	621
35×35	365	6	371	620
34×34	361	6	367	620
33×33	354	6	360	620
32×32	341	6	347	615
31×31	343	7	350	615
30×30	334	11	345	611
29×29	333	6	339	606
28×28	324	6	330	606
27×27	316	6	322	605
26×26	307	6	313	600
25×25	298	5	303	596
24×24	289	5	294	594
23×23	277	5	282	595
22×22	265	6	271	588
21×21	253	5	258	572

20 × 20	237	6	243	559
19 × 19	224	5	229	554
18 × 18	208	6	214	544
17 × 17	190	8	198	530
16 × 16	175	5	180	535
15 × 15	164	2	166	502
14 × 14	147	2	149	479
13 × 13	128	5	133	452
12 × 12	114	2	116	409
11 × 11	98	2	100	370
10 × 10	82	2	84	303
9 × 9	68	2	70	235
8 × 8	51	5	56	180
7 × 7	39	5	44	95
6 × 6	31	2	33	55
5 × 5	20	3	23	27
4 × 4	13	2	15	9
K: Number of principal components used for classification D: Number of discarded principal components				

Number of principal components used for classification (corresponding to Figure 5-8 Identification performance of the four baseline FR algorithms on LR face images tested on the Yale B subset)

Image Size	PCA		PCA+LDA	G-PCA+LDA
	K	D	K	K
64 × 64	107	3	110	177
63 × 63	101	3	104	177
62 × 62	101	3	104	177
61 × 61	101	3	104	176
60 × 60	102	3	105	176
59 × 59	102	3	105	176
58 × 58	101	3	104	176
57 × 57	101	3	104	177
56 × 56	101	3	104	176
55 × 55	102	3	105	176
54 × 54	101	3	104	176
53 × 53	101	3	104	176
52 × 52	101	3	104	176
51 × 51	101	3	104	176
50 × 50	101	3	104	176
49 × 49	101	3	104	176

48×48	101	3	104	176
47×47	100	3	103	176
46×46	100	3	103	177
45×45	101	3	104	176
44×44	100	3	103	176
43×43	100	3	103	176
42×42	100	3	103	176
41×41	100	3	103	175
40×40	100	3	103	175
39×39	100	3	103	174
38×38	99	3	102	174
37×37	100	2	102	174
36×36	99	3	102	174
35×35	99	3	102	173
34×34	99	3	102	174
33×33	99	2	101	172
32×32	95	2	97	172
31×31	99	2	101	173
30×30	98	2	100	171
29×29	96	3	99	170
28×28	96	3	99	170
27×27	96	3	99	170
26×26	95	2	97	169
25×25	94	3	97	169
24×24	94	2	96	168
23×23	92	3	95	169
22×22	90	3	93	167
21×21	90	3	93	165
20×20	89	3	92	162
19×19	88	2	90	161
18×18	84	3	87	161
17×17	82	3	85	157
16×16	78	2	80	157
15×15	77	2	79	154
14×14	74	2	76	151
13×13	70	3	73	147
12×12	66	2	68	141
11×11	57	4	61	133
10×10	54	2	56	123
9×9	48	2	50	108

8 × 8	40	2	42	85
7 × 7	33	2	35	60
6 × 6	26	2	28	39
5 × 5	19	2	21	24
4 × 4	11	3	14	8
K: Number of principal components used for classification D: Number of discarded principal components				

Number of principal components used for classification (corresponding to Figure 5-9 Identification performance of the four baseline FR algorithms on LR face images tested on the Caltech subset)

Image Size	PCA		PCA+LDA	G-PCA+LDA
	K	D	K	K
64 × 64	175	0	175	199
63 × 63	169	0	169	198
62 × 62	168	0	168	199
61 × 61	169	0	169	198
60 × 60	169	0	169	198
59 × 59	168	0	168	198
58 × 58	168	0	168	198
57 × 57	168	0	168	199
56 × 56	168	0	168	198
55 × 55	168	0	168	198
54 × 54	168	0	168	198
53 × 53	168	0	168	198
52 × 52	167	0	167	198
51 × 51	167	0	167	198
50 × 50	167	0	167	198
49 × 49	167	0	167	198
48 × 48	167	0	167	198
47 × 47	167	0	167	199
46 × 46	166	0	166	199
45 × 45	166	0	166	199
44 × 44	166	0	166	199
43 × 43	166	0	166	199
42 × 42	166	0	166	198
41 × 41	165	0	165	198
40 × 40	165	0	165	198
39 × 39	165	0	165	198
38 × 38	164	0	164	198
37 × 37	164	0	164	198

36×36	163	0	163	197
35×35	163	0	163	197
34×34	162	0	162	198
33×33	162	0	162	198
32×32	155	0	155	195
31×31	161	0	161	197
30×30	160	0	160	197
29×29	158	0	158	195
28×28	157	0	157	196
27×27	156	0	156	195
26×26	155	0	155	194
25×25	153	0	153	194
24×24	152	0	152	194
23×23	150	0	150	193
22×22	146	2	148	193
21×21	145	0	145	191
20×20	142	0	142	188
19×19	138	0	138	188
18×18	134	0	134	188
17×17	130	0	130	186
16×16	121	0	121	184
15×15	118	0	118	182
14×14	112	0	112	180
13×13	103	0	103	175
12×12	93	0	93	167
11×11	83	0	83	162
10×10	71	3	74	150
9×9	61	2	63	130
8×8	49	2	51	106
7×7	42	0	42	74
6×6	32	0	32	46
5×5	22	0	22	27
4×4	15	0	15	9
K: Number of principal components used for classification D: Number of discarded principal components				

Number of principal components used for classification (corresponding to Figure 5-12 Identification performance of PCA on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset)

Image Size	Bicubic		Eigentransformation		CLLR-SR	
	K	D	K	D	K	D
16 × 16	483	0	483	0	483	0
12 × 12	483	0	483	0	483	0
8 × 8	483	0	483	0	483	0
K: Number of principal components used for classification D: Number of discarded principal components						

Number of principal components used for classification (corresponding to Figure 5-13 Identification performance of PCA+LDA on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset)

Image Size	Bicubic (K)	Eigentransformation (K)	CLLR-SR (K)
16 × 16	483	483	483
12 × 12	483	483	483
8 × 8	483	483	483
K: Number of principal components used for classification			

Number of principal components used for classification (corresponding to Figure 5-14 Identification performance of G-PCA+LDA on the three baseline SR algorithms tested on the CAS-PEAL-R1 subset)

Image Size	Bicubic (K)	Eigentransformation (K)	CLLR-SR (K)
16 × 16	999	999	999
12 × 12	999	999	999
8 × 8	999	999	999
K: Number of principal components used for classification			

Number of principal components used for classification (corresponding to Figure 5-16 Identification performance of PCA on the three baseline SR algorithms tested on the AR subset)

Image Size	Bicubic		Eigentransformation		CLLR-SR	
	K	D	K	D	K	D
16 × 16	394	6	394	6	394	6
12 × 12	394	6	394	6	394	6
8 × 8	394	6	394	6	394	6
K: Number of principal components used for classification D: Number of discarded principal components						

Number of principal components used for classification (corresponding to Figure 5-17 Identification performance of PCA+LDA on the three baseline SR algorithms tested on the AR subset)

Image Size	Bicubic (K)	Eigentransformation (K)	CLLR-SR (K)
16 × 16	400	400	400
12 × 12	400	400	400
8 × 8	400	400	400
K: Number of principal components used for classification			

Number of principal components used for classification (corresponding to Figure 5-18 Identification performance of G-PCA+LDA on the three baseline SR algorithms tested on the AR subset)

Image Size	Bicubic (K)	Eigentransformation (K)	CLLR-SR (K)
16 × 16	627	627	627
12 × 12	627	627	627
8 × 8	627	627	627
K: Number of principal components used for classification			

Number of principal components used for classification (corresponding to Figure 5-20 Identification performance of PCA+CLDA vs. PCA+LDA and G-PCA+CLDA vs. G-PCA+LDA tested on the CAS-PEAL-R1 subset)

Image Size	PCA+CLDA		G-PCA+CLDA	
	KL	KH	KL	KH
32 × 32	402	483	956	999
31 × 31	399	483	957	999
30 × 30	390	483	939	999
29 × 29	381	483	925	999
28 × 28	369	483	928	999
27 × 27	358	483	928	999
26 × 26	344	483	919	999
25 × 25	331	483	900	999
24 × 24	316	483	888	999
23 × 23	302	483	890	999
22 × 22	286	483	874	999
21 × 21	271	483	825	999
20 × 20	253	483	797	999
19 × 19	236	483	789	999
18 × 18	219	483	767	999
17 × 17	202	483	733	999
16 × 16	183	483	740	999

15 × 15	166	483	658	999
14 × 14	148	483	633	999
13 × 13	132	483	585	999
12 × 12	115	483	493	999
11 × 11	99	483	437	999
10 × 10	82	483	340	999
9 × 9	69	483	263	999
8 × 8	56	483	193	999
7 × 7	43	483	95	999
6 × 6	32	483	55	999
5 × 5	23	483	28	999
4 × 4	15	483	9	999
KL: Number of principal components used for CCA (low-resolution) KH: Number of principal components used for CCA (high-resolution)				

Number of principal components used for classification (corresponding to Figure 5-21 Identification performance of PCA+CLDA vs. PCA+LDA and G-PCA+CLDA vs. G-PCA+LDA tested on the AR subset)

Image Size	PCA+CLDA		G-PCA+CLDA	
	KL	KH	KL	KH
32 × 32	347	400	615	627
31 × 31	350	400	615	627
30 × 30	345	400	611	627
29 × 29	339	400	606	627
28 × 28	330	400	606	627
27 × 27	322	400	605	627
26 × 26	313	400	600	627
25 × 25	303	400	596	627
24 × 24	294	400	594	627
23 × 23	282	400	595	627
22 × 22	271	400	588	627
21 × 21	258	400	572	627
20 × 20	243	400	559	627
19 × 19	229	400	554	627
18 × 18	214	400	544	627
17 × 17	198	400	530	627
16 × 16	180	400	535	627
15 × 15	166	400	502	627
14 × 14	149	400	479	627
13 × 13	133	400	452	627

12×12	116	400	409	627
11×11	100	400	370	627
10×10	84	400	303	627
9×9	70	400	235	627
8×8	56	400	180	627
7×7	44	400	95	627
6×6	33	400	55	627
5×5	23	400	27	627
4×4	15	400	9	627
KL: Number of principal components used for CCA (low-resolution) KH: Number of principal components used for CCA (high-resolution)				

References

[1] Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *Acm Computing Surveys (CSUR)*, *35*(4), 399-458.

[2] Hennings-Yeomans, P. H., Baker, S., & Kumar, B. V. (2008, June). Simultaneous super-resolution and feature extraction for recognition of low-resolution faces. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on* (pp. 1-8). IEEE.

[3] Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of cognitive neuroscience*, *3*(1), 71-86.

[4] Etemad, K., & Chellappa, R. (1997). Discriminant analysis for recognition of human face images. *JOSA A*, *14*(8), 1724-1733.

[5] Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *19*(7), 711-720.

[6] Liu, C., & Wechsler, H. (2002). Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *Image processing, IEEE Transactions on*, *11*(4), 467-476.

[7] Zhang, W., Shan, S., Gao, W., Chen, X., & Zhang, H. (2005, October). Local Gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on* (Vol. 1, pp. 786-791). IEEE.

[8] Movellan, J. R. (2002). Tutorial on Gabor filters. *Open Source Document*.

[9] Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *28*(12), 2037-2041.

[10] Wang, X., & Tang, X. (2005). Hallucinating face by eigentransformation. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, *35*(3), 425-434.

[11] Huang, H., He, H., Fan, X., & Zhang, J. (2010). Super-resolution of human face image using canonical correlation analysis. *Pattern Recognition*, *43*(7), 2532-2543.

[12] Borga, M. (2001). Canonical correlation: a tutorial. *On line tutorial http://people.imt.liu.se/magnus/cca, 4*.

[13] Saul, L. K., & Roweis, S. T. (2003). Think globally, fit locally: unsupervised learning of low dimensional manifolds. *The Journal of Machine Learning Research*, *4*, 119-155.

[14] Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision research*, *20*(10), 847-856.

[15] Marčelja, S. (1980). Mathematical description of the responses of simple cortical cells. *JOSA*, *70*(11), 1297-1300.

[16] Lades, M., Vorbruggen, J. C., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, R. P., & Konen, W. (1993). Distortion invariant object recognition in the dynamic link architecture. *Computers, IEEE Transactions on*, *42*(3), 300-311.

[17] Ahonen, T., Hadid, A., & Pietikäinen, M. (2004). Face recognition with local binary patterns. In *Computer Vision-ECCV 2004* (pp. 469-481). Springer Berlin Heidelberg.

[18] Rubner, Y., Puzicha, J., Tomasi, C., & Buhmann, J. M. (2001). Empirical evaluation of dissimilarity measures for color and texture. *Computer Vision and Image Understanding*, *84*(1), 25-43.

[19] Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International journal of computer vision*, *7*(1), 11-32.

[20] Bicubic interpolation. (2013, June 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 11:28, August 3, 2013, from *http://en.wikipedia.org/w/index.php?title=Bicubic_interpolation&oldid=561554939*.

[21] Fukunaga, K. (1990). *Introduction to statistical pattern recognition*. Access Online via Elsevier.

[22] Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, *28*(3/4), 321-377.

[23] Liu, C., Shum, H. Y., & Zhang, C. S. (2001). A two-step approach to hallucinating faces: global parametric model and local nonparametric model. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (Vol. 1, pp. I-192). IEEE.

[24] Zhuang, Y., Zhang, J., & Wu, F. (2007). Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation. *Pattern Recognition*, *40*(11), 3178-3194.

[25] Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., & Zhao, D. (2008). The CAS-PEAL large-scale Chinese face database and baseline evaluations. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, *38*(1), 149-161.

[26] Martinez, A. M. (1998). The AR face database. *CVC Technical Report*, *24*.

[27] Georghiades, A. S., Belhumeur, P. N., & Kriegman, D. J. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *23*(6), 643-660.

[28] M. Weber. Frontal face dataset. *http://www.vision.caltech.edu/html-files/ archive, California Institute of Technology*, 1999.

[29] Moses, Y., Adini, Y., & Ullman, S. (1994). Face recognition: The problem of compensating for changes in illumination direction. In *Computer Vision—ECCV'94* (pp. 286-296). Springer Berlin Heidelberg.

[30] Peak signal-to-noise ratio. (2013, June 1). In *Wikipedia, The Free Encyclopedia*. Retrieved 12:03, August 3, 2013, from *http://en.wikipedia.org/w/index.php?title=Peak_signal-to-noise_ratio&oldid=557823718*.

[31] Wang, Z., & Bovik, A. C. (2009). Mean squared error: love it or leave it? A new look at signal fidelity measures. *Signal Processing Magazine, IEEE*, *26*(1), 98-117.