

Chapter 2

Remote Identification of Faces

Vishal M. Patel, Jie Ni and Rama Chellappa

Abstract Face recognition in unconstrained acquisition conditions is one of the most challenging problems that has been actively researched in recent years. It is well known that many state-of-the-art still image-based face recognition algorithms perform well, when constrained (frontal, well illuminated, high-resolution, sharp, and full) face images are acquired. However, their performance degrades significantly when the test images contain variations that are not present in the training images. In this chapter, we highlight some of the key issues in remote face recognition. We define the remote face recognition as one where faces are several tens of meters (10–250 m) from the cameras. We then describe a remote face database which has been acquired in an unconstrained outdoor maritime environment. The face images in this database suffer from variations due to blur, poor illumination, pose, etc. Recognition performance of a subset of existing still image-based face recognition algorithms is discussed on the remote face data set. It is demonstrated that in addition to applying a good classification algorithm, finding features that are robust to variations mentioned above and developing statistical models which can account for these variations are very important for remote face recognition.

This work was partially supported by a MURI grant N00014-08-1-0638 from the Office of Naval Research.

V. M. Patel (✉) · J. Ni · R. Chellappa
Department of Electrical and Computer Engineering, Center for Automation Research,
University of Maryland, College Park, MD 20742, USA
e-mail: pvishalm@umiacs.umd.edu

J. Ni
e-mail: jni@umiacs.umd.edu

R. Chellappa
e-mail: rama@umiacs.umd.edu

2.1 Introduction

During the past two decades, face recognition (FR) has received great attention and tremendous progress has been made [66]. Numerous image-based algorithms [6, 8, 20, 39, 56, 58, 59, 66] and video-based algorithms [19, 33, 68] have been developed in the FR community. Currently, most of the existing FR algorithms have been evaluated using databases which are collected at close range (less than a few meters) and under different levels of controlled acquisition conditions. Some of the most extensively used face datasets such as CMU PIE [53], FRGC/FERET [46] and YaleB [23] were captured in constrained settings, with studio lights to control the illumination variations while pose variations are controlled by cooperative subjects.

While FR techniques on these datasets have reached a high level of recognition performance over the years, research in remote unconstrained FR field is still at a nascent stage. Recently, a new database called “Labeled Faces in the Wild” (LFW) [29] whose images are collected from the web, has been frequently used to address some of the issues in unconstrained FR problem. Yet concerns have been raised that these images are typically posed and framed by photographers and there is no guarantee that such a set accurately captures the range of variations found in the real world setting [47]. [61] describe a face video database, UTK-LRHM, acquired from long distances with high magnifications. And the magnification blur is described as a major source of degradation in their data. Figure 2.1 shows some sample images from the FERET, FRGC and LFW datasets, respectively.

In this chapter, we address some of the issues related to the problem of FR when face images are captured in unconstrained and remote setting. As one has a little control of the acquisition of the face images, the images one gets often suffer from low resolution, poor illumination, blur, pose variation and occlusion. These variations present serious challenges to existing FR algorithms. We provide a brief review of developments and progress in remote face recognition. Furthermore, we introduce a new dataset which was collected in a remote maritime environment [17]. We provide some preliminary experimental studies on this dataset and offer insights and suggestions for the remote FR problem.

The organization of this chapter is as follows: In Sect. 2.2, we describe some of the issues that arise in remote face recognition. Section 2.3 briefly discusses image quality measures for face images. In Sect. 2.4, we describe the remote face database collected by the authors’ group. Section 2.5 describes the algorithms that are evaluated and corresponding recognition results are discussed in Sect. 2.6. Finally, conclusions are given in Sect. 2.7.

2.2 Face Recognition at a Distance

Reliable extraction and matching of biometric signatures from face acquired at a distance is a challenging problem [55]. First, as the subjects may not be cooperative, the pose of the face and body relative to the sensor is likely to vary greatly. Second,

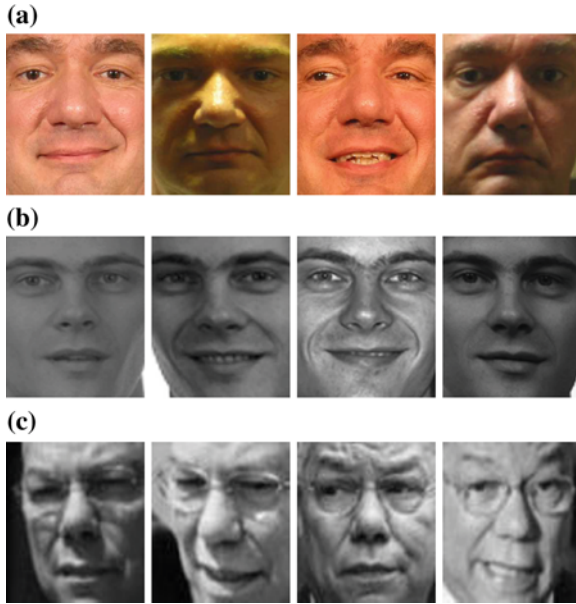


Fig. 2.1 **a** Sample images from the FRGC dataset. **b** Sample images from the FERET dataset. **c** Sample images from the LFW dataset

the lighting is uncontrolled and could be extreme in its variation. Third, when the subjects are at great distances, the effects of a scattering media (static: fog and mist, dynamic: rain, sleet, or sea spray) are greatly amplified. Fourth, the relative motion between the subjects and the sensors produce jitter and motion blur in the images. In this section, we investigate various artifacts introduced as a result of long range acquisition of face images. Some of the factors that can affect long range FR system performance can be summarized into four types [55]: (1) technology (dealing with face image quality, heterogeneous face images, etc.), (2) environment (lighting, etc.), (3) user (expression, facial hair, facial ware etc.), and (4) user-system (pose, height, etc.). In what follows, we discuss some of these factors in detail.

2.2.1 Illumination

Variation in illumination conditions is one of the major challenges in remote face recognition. In particular, when images are captured at long ranges, one does not have any control over lighting conditions. As a result, the captured images often suffer from extreme (due to sun) or low light conditions (due to shadow, bad weather, evening, etc.).

The performance of most existing FR algorithms is highly sensitive to illumination variation. Various methods have been introduced to deal with this illumination

problem in FR. They are based on illumination cone [9, 22], spherical harmonics [7, 49, 62], quotient images [51, 57], gradient faces [63], logarithmic total variation [18], albedo estimation [10], photometric stereo [67], and dictionaries [32, 45].

Changes induced by illumination can usually make face images of the same subject far apart than images of different subjects. One can use estimates of albedo to mitigate the illumination effect. Albedo is the fraction of light that a surface point reflects when it is illuminated. It is an intrinsic property that depends on the material properties of the surface and it is invariant to changes in illumination. Assuming the Lambertian reflectance model for the facial surface, which is a restrictive assumption, one can relate the surface normals, albedo and the intensity image by an image formation model. The diffused component of the surface reflection is given by

$$x_{i,j} = \rho_{i,j} \max(\mathbf{n}_{i,j}^T \mathbf{s}, 0), \quad (2.1)$$

where x is the pixel intensity, $\mathbf{s}$ is the light source direction, $\rho_{i,j}$ is the surface albedo at position i, j , $\mathbf{n}_{i,j}$ is the surface normal of the corresponding surface point and $1 \leq i, j \leq N$. The max function in (2.1) accounts for the formation of attached shadows. Neglecting the attached shadows, (2.1) can be approximated as

$$x_{i,j} = \rho_{i,j} \max(\mathbf{n}_{i,j}^T \mathbf{s}, 0) \approx \rho_{i,j} \mathbf{n}_{i,j}^T \mathbf{s}. \quad (2.2)$$

Let $\mathbf{n}_{i,j}^{(0)}$ and $\mathbf{s}^{(0)}$ be the initial values of the surface normal and illumination direction. These initial values can be domain dependent average values. The Lambertian assumption imposes the following constraints on the initial albedo

$$\rho_{i,j}^{(0)} = \frac{x_{i,j}}{\mathbf{n}_{i,j}^{(0)} \cdot \mathbf{s}^{(0)}}, \quad (2.3)$$

where $\cdot$ is the standard dot product operator. Using (2.2), Eq. (2.3) can be re-written as

$$\rho_{i,j}^{(0)} = \rho_{i,j} \frac{\mathbf{n}_{i,j} \cdot \mathbf{s}}{\mathbf{n}_{i,j}^{(0)} \cdot \mathbf{s}^{(0)}} = \rho_{i,j} + \frac{\mathbf{n}_{i,j} \cdot \mathbf{s} - \mathbf{n}_{i,j}^{(0)} \cdot \mathbf{s}^{(0)}}{\mathbf{n}_{i,j}^{(0)} \cdot \mathbf{s}^{(0)}} \rho_{i,j} \quad (2.4)$$

$$= \rho_{i,j} + \omega_{i,j}, \quad (2.5)$$

where $\omega_{i,j} = \frac{\mathbf{n}_{i,j} \cdot \mathbf{s} - \mathbf{n}_{i,j}^{(0)} \cdot \mathbf{s}^{(0)}}{\mathbf{n}_{i,j}^{(0)} \cdot \mathbf{s}^{(0)}} \rho_{i,j}$. This can be viewed as a signal estimation problem

where $\rho_{i,j}$ is the original signal, $\rho^{(0)}$ is the degraded signal and ω is the signal dependent noise. Using this model, the albedo map can be estimated using the method of minimum mean squared error criterion [10]. The illumination-free albedo image can then be used for recognition. Figure 2.2 shows the results of albedo estimation for two face images acquired at a distance using the method presented in [10].

Fig. 2.2 Results of albedo estimation for remotely acquired images. *Left* Original images; *Right* Estimated albedo images



2.2.2 Pose Variation

Pose variation can be considered as one of the most important and challenging problems in face recognition. Magnitudes of variations of innate characteristics, which distinguish one face from another, are often smaller than magnitudes of image variations caused by pose variations [64]. Popular frontal face recognition algorithms, such as Eigenfaces [56] or Fisherfaces [8, 20], usually have low recognition rates under pose changes as they do not take into account the 3D alignment issue when creating the feature vectors for matching.

Existing methods for face recognition across pose can be roughly divided into two broad categories: techniques that rely on 3D models and 2D techniques. Some of the methods that use 3D information include [12] and [13]. One of the advantages of using 2D methods is that they do not require the 3D prior information for performing pose-invariant face recognition [15, 24, 48]. Image patch-based approaches have also received significant attention in recent years [4, 5, 16, 31, 35]. Note that some of these methods are highly sensitive to variations in illumination, resolution, blur and occlusion.

Let $\bar{\mathbf{n}}_{i,j}$, $\bar{\mathbf{s}}$ and $\bar{\Theta}$ be some initial estimates of the surface normals, illumination direction and initial estimate of surface normals in pose Θ , respectively. Then, the initial albedo at pixel (i, j) can be obtained by

$$\bar{\rho}_{i,j} = \frac{x_{i,j}}{\bar{\mathbf{n}}_{i,j}^{\bar{\Theta}} \cdot \bar{\mathbf{s}}}, \quad (2.6)$$

where $\bar{\mathbf{n}}_{i,j}^{\bar{\Theta}}$ denotes the initial estimate of surface normals in pose $\bar{\Theta}$. Using this model, we can re-formulate the problem of recovering albedo as a signal estimation problem. Using arguments similar to Eq. (2.3), we get the following formulation for the albedo estimation problem in the presence of pose



Fig. 2.3 Pose normalization. *Left column* Original input images. *Middle column* Recovered albedoes corresponding to frontal face images. *Right column* Pose normalized relighted images

$$\bar{\rho}_{i,j} = \rho_{i,j} h_{i,j} + \omega_{i,j}, \quad (2.7)$$

where $w_{i,j} = \frac{\bar{\mathbf{n}}_{i,j}^\Theta \cdot \mathbf{s} - \bar{\mathbf{n}}_{i,j}^\Theta \cdot \bar{\mathbf{s}}}{\bar{\mathbf{n}}_{i,j}^\Theta \cdot \bar{\mathbf{s}}}$, $h_{i,j} = \frac{\bar{\mathbf{n}}_{i,j}^\Theta \cdot \bar{\mathbf{s}}}{\bar{\mathbf{n}}_{i,j}^\Theta \cdot \bar{\mathbf{s}}}$, $\rho_{i,j}$ is the true albedo and $\bar{\rho}_{i,j}$ is the degraded albedo. Using this model, a pose-robust albedo estimation method was proposed in [12]. Figure 2.3 shows some examples of pose normalized images using this method. Once pose and illumination have been normalized, one can use these images for illumination and pose-robust recognition [44, 54].

Face recognition methods require gallery and probe faces to be aligned. Methods for pose alignment need the extraction of feature points. It is our experience that existing methods for point feature extraction are not very effective on remote faces, mainly due to illumination and blur.

2.2.3 Occlusion

Another challenge in remote FR is that since long-range face is usually for non-cooperative subjects, acquired images are often contaminated by occlusion. The occlusion may be the result of subject wearing sunglasses, scarf, hat or a mask. Recognizing subjects in the presence of occlusion requires robust techniques for classification. One such technique was developed for the principle component analysis in [14]. The recently developed algorithm for FR using sparse representation is also shown to be robust to occlusion [59]. Figure 2.4 shows some images with occlusion from the remote face dataset.

Fig. 2.4 Some occluded face images in remote face dataset



2.2.4 Blur

In remote FR, the distance between the subject and the sensor is in a spacial extent. This in turn results in out of focus blur effects in the captured image. Often times abbreviations of the imaging optics causes non-ideal focusing. Motion blur is another phenomenon that occurs when the subject is moving rapidly or the camera is shaking. [2, 43] are some of the methods that attempt to address this issue in face recognition. In [2] local phase quantization is used to recognize blurred face images. Local phase quantization is based on quantizing the Fourier transform phase in local neighborhoods. It is shown that the quantized phase is blur invariant when certain conditions are met. In [43], a method of inferring a point spread function (PSF) representing the process of blur on faces is presented. The method uses learned prior information derived from a training set of blurred faces to make the ill-posed problem more tractable.

In the remote acquisition setting, often times blur is coupled with illumination variations. It might be desirable to develop an algorithm that can compensate for both blur and illumination simultaneously. Given the $N \times N$ arrays y and x , representing the observed image and the image to be estimated, respectively, the matrix deconvolution problem can be described as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \gamma, \quad (2.8)$$

where $\mathbf{y}$, $\mathbf{x}$, and γ are $N^2 \times 1$ column vectors representing the arrays y , x , and γ lexicographically ordered, $\mathbf{H}$ is the $N^2 \times N^2$ matrix that models the blur operator and γ denotes an $N \times N$ array of noise samples. Using the Lambertian model (2.2), Eq. (2.8) can be re-written as

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x} + \gamma = \mathbf{H}\Phi\rho + \gamma \\ &= \mathbf{G}\rho + \gamma, \end{aligned} \quad (2.9)$$

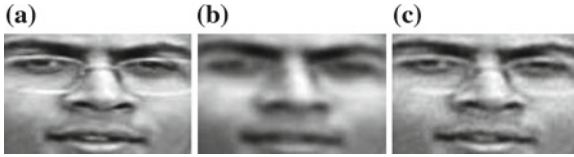


Fig. 2.5 Image deconvolution experiment with a face image from the remote face dataset. **a** Original image. **b** Noisy blurred image. **c** Parametric manifold-based estimate



Fig. 2.6 A typical low-resolution face image in remote face dataset

where $\Phi = \text{diag}(\mathbf{n}_{i,j}^T \mathbf{s})$ of size $N^2 \times N^2$, ρ is $N^2 \times 1$ vector representing ρ and $\mathbf{G} = \mathbf{H}\Phi$. Having observed $\mathbf{y}$, the general inverse problem is to estimate ρ with incomplete information of $\mathbf{G}$. It is well-known that regularization is often used to find a unique and stable solution to the ill-posed inverse problems. One such regularization method using patch manifold prior was developed in [42]. Figure 2.5 shows an example of a face image that is deblurred using the method in [42].

2.2.5 Low Resolution

Image resolution is an important parameter in remote face acquisition, where there is no control over the distance of human from the camera. Figure 2.6 illustrates a practical scenario where one is faced with a challenging problem of recognizing humans when the captured face images are of very low resolution (LR). Many methods have been proposed in the vision literature that can deal with this resolution problem in face recognition. Most of these methods are based on some application of super-resolution (SR) technique to increase the resolution of images so that the recovered higher-resolution (HR) images can be used for recognition. One of the major drawbacks of applying SR techniques is that there is a possibility that recovered HR images may contain some serious artifacts. This is often the case when the

resolution of the image is very low. As a result, these recovered images may not look like the images of the same person and the recognition performance may degrade significantly.

An Eigen-face domain SR method for FR was proposed in [25]. This method proposes to solve the FR at LR using super-resolution (SR) of multiple LR images using their PCA domain representation. Given a LR face image, [30] proposes to directly compute a maximum likelihood identity parameter vector in the HR tensor space that can be used for SR and recognition. A Tikhonov regularization method that can combine the different steps of SR and recognition in one step was proposed in [27].

Though the LR images are directly not suitable for face recognition purpose, it is also not necessary to super-resolve them before recognition, as the problem of recognition is not the same as super-resolution. Based on this motivation, some different approaches to this problem have been suggested. Coupled Metric Learning [36] attempts to solve this problem by mapping the LR image to a new subspace, where higher recognition can be achieved. Extension of this method was recently proposed in [37]. A similar approach for improving the matching performance of the LR images using multidimensional scaling was recently proposed in [11]. A log-polar domain-based method was proposed in [28]. Additional methods for LR FR include correlation filter-based approach [1], a support vector data description method [34] and a dictionary-based method [52]. 3D face modeling has also been used to address the LR face recognition problem [38, 50]. There have been works to solve the problem of unconstrained FR using videos [3].

In practical scenarios, the resolution change is also coupled with other parameters such as pose change, illumination variations and expression. Algorithms specifically designed to deal with LR images quite often fail in dealing with these variations. Hence, it is essential to include these parameters while designing a robust method for LR face recognition.

2.2.6 Atmospheric and Weather Artifacts

Most of the current vision algorithms and applications are applied to the images that are captured under clear and nice weather conditions. However, often times in outdoor applications, one faces adverse weather conditions such as extreme illumination, fog, haze, rain and snow [40, 41, 55]. These extreme conditions can also present additional difficulties in developing robust algorithms for face recognition. Figure 2.7, shows an image that is collected in remote setting where the illumination condition caused by the sun is extreme.



Fig. 2.7 Extreme illumination conditions caused by the sun

2.3 Long Range Facial Image Quality

In FR systems, the ultimate recognition performance depends on how cooperative the subject is, as well as the resolution in addition to illumination variations that invariably are present in outdoors. For non-cooperative situations, one can increase the performance by combining tracking technology and recognition technology together. For instance, the system would first track the subject's face. Then, it would get a series of image of the same person. Using multiple images to recognize an individual can provide better recognition accuracy than just using a single image. However, in order to detect and track without false alarms, the system must acquire images of the subject with sufficient quality and resolution [55].

As discussed in the previous section, various factors could affect the quality of remotely acquired images. Hence it is essential to derive an image quality measure to study the relation between image quality and recognition performance. To this end, a blind signal-to-noise ratio estimator has been defined for facial image quality [55]. This measure is defined based on the concept that the statistics of image edge intensity are correlated with noise and SNR estimation [65].

Suppose the probability density function $f_{\|\nabla I\|}(r)$ of the edge intensity image $\|\nabla I\|$ can be calculated as a mixture of Rayleigh probability density functions. Consider the quantity

$$Q = \int_{2\mu}^{\infty} f_{\|\nabla I\|}(r) dr,$$

where μ is the mean of $\|\nabla I\|$. It has been shown that the value of Q for the noisy image is always smaller than the value of Q for the image without noise [65]. Then, the face image quality is defined as

$$Q' = \frac{\sum \text{edge above } 2\mu\text{'s pixels}}{\sum \text{edge pixels}} \simeq \int_{2\mu}^{\infty} f \|\nabla \mathbf{I}\|(r) dr.$$

It has been experimentally verified that the estimator Q' is well correlated with the recognition performance in FR [55]. Hence, setting up a comprehensive metric to evaluate the quality of face images is essential in remote face recognition. Also, these measures can be used to reject images that are of low quality. One can also derive these type of metric based on sparse representation [44, 59].

2.4 Remote Face Database

In this section, we introduce a remote face database in which a significant number of images are taken from long distances and under unconstrained outdoor maritime environment [17]. As discussed in Sect. 2.2, the quality of the images differs in the following aspects: the illumination is not controlled and is often severe; there are pose variations and occluded faces due to non-cooperative subjects; finally, the effects of scattering and high magnification resulting from long distance contribute to the blurriness of face images.

The distance from which the face images were taken varies from 5 to 250m under different scenarios. Since the extraction of faces in the data set using the existing state-of-the-art face detection algorithms was not reliable and the faces only occupied small regions in large background scenes, the faces were manually cropped and rescaled to a fixed size. The resulting database for still color face images contains 17 different individuals and 2,106 face images in total.

The faces were manually labeled according to their type (i.e. different illumination conditions, occlusion, blur etc.). In total, the database contains 688 clear images, 85 partially occluded images, 37 severely occluded images, 540 images with medium blur, 245 with sever blur, and 244 in poor illumination conditions. The remaining images have two or more coupled conditions, such as poor lighting and blur, occlusion and blur etc. Figure 2.8 shows two sample images acquired in a remote maritime setting. Some of the extracted images from the database are shown in Fig. 2.9. Note that some of these face images contain extreme variations which makes recognition very difficult even for humans.

2.5 Algorithms

We present the results of two state-of-the-art FR algorithms on the remote face database and compare their recognition performance [17]. More experimental results using a dictionary-based method can be found in [44].



Fig. 2.8 Typical images illustrating the different scenarios in the maritime domain

Fig. 2.9 Cropped face images with different variations from the remote face database



2.5.1 Baseline Algorithm

The baseline recognition algorithm used performs Principle Component Analysis (PCA) [60] followed by Linear Discriminate Analysis (LDA) [8, 20] and a Support Vector Machine (SVM) [26].

LDA is a well-known method for feature extraction and dimensionality reduction for pattern recognition and classification tasks. It uses class specific linear methods for dimensionality reduction. It selects projection matrix $\mathbf{A}$ in such a way that the ratio of the between-class scatter and the within-class scatter is maximized [20]. The criterion function is defined as

$$\mathbf{A}_{opt} = \arg \max_{\mathbf{A}} \frac{|\mathbf{A}^T \Sigma_B \mathbf{A}|}{|\mathbf{A}^T \Sigma_W \mathbf{A}|}$$

where $|\cdot|$ denotes determinant of a matrix, Σ_B and Σ_W are between-class and within-class scatter matrices, respectively.

The within-class scatter matrix becomes singular when the dimension of the data is larger than the number of training samples. In order to deal with this, we first use PCA as a dimensionality reduction method to project the raw data onto an intermediate feature space with much lower dimension. Then, LDA is applied on features from this intermediate PCA space. It is well known that LDA is not feasible when there is only one image per subject. To further mitigate this small sample size problem of LDA, we impose a regularizer in the LDA objective function

$$\mathbf{a}_{opt} = \arg \max_{\mathbf{a}} \frac{\mathbf{a}^T \Sigma_B \mathbf{a}}{\mathbf{a}^T \Sigma_W \mathbf{a} + \alpha J(\mathbf{a})}$$

where the optimal $\mathbf{a}$'s form the columns of the optimal projection matrix $\mathbf{A}_{opt}$. In our application, we use the Tikhonov regularizer $J(\mathbf{a}) = \|\mathbf{a}\|_2^2$. The resulting method is often known as the Regularized Discriminate Analysis (RDA) [21]. Then, the low-dimensional discriminant features from RDA are fed into a linear SVM for classification.

2.5.2 Sparse Representation-Based Algorithm

A sparse representation-based classification (SRC) algorithm for FR was proposed in [59] which was shown to be robust to noise and occlusion. The comparison with other existing FR methods in [59] suggests that the SRC algorithm is among the best. Hence, we treat it as state-of-the-art and use it for comparisons in this chapter.

The idea proposed is to create a dictionary matrix of the training samples as column vectors. The test sample is also represented as a column vector. Different dimensionality reduction methods are used to reduce the dimension of both the test

vector and the vectors in the dictionary. It is then simply a matter of solving an ℓ_1 minimization problem in order to obtain the sparse solution. Once the sparse solution is obtained, it can provide information as to which training sample the test vector most closely relates to.

Let each image be represented as a vector in $\mathbb{R}^n$, $\mathbf{D}$ be the dictionary (i.e. training set) and $\mathbf{y}$ be the test image. The SRC algorithm is as follows:

1. Create a matrix of training samples $\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_k]$ for k classes, where $\mathbf{D}_i$, $i = 1, \dots, k$ are the set of images of each class.
2. Reduce the dimension of the training images and a test image by any dimensionality reduction method. Denote the resulting dictionary and the test vector as $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{y}}$, respectively.
3. Normalize the columns of $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{y}}$.
4. Solve the following ℓ_1 minimization problem

$$\hat{\alpha} = \arg \min_{\alpha'} \|\alpha'\|_1 \quad \text{subject to} \quad \tilde{\mathbf{y}} = \tilde{\mathbf{D}}\alpha', \quad (2.10)$$

5. Calculate the residuals

$$r_i(\tilde{\mathbf{y}}) = \|\tilde{\mathbf{y}} - \tilde{\mathbf{D}}\delta_i(\hat{\alpha})\|_2,$$

for $i = 1, \dots, k$ where δ_i a characteristic function that selects the coefficients associated with the i th class.

6. Identity($\mathbf{y}$) = $\arg \min_i r_i(\tilde{\mathbf{y}})$.

The assumption made in this method is that given sufficient training samples of the k th class, $\tilde{\mathbf{D}}_k$, any new test image $\mathbf{y}$ that belongs to the same class will approximately lie in the linear span of the training samples from the class k . This implies that most of the coefficients not associated with class k in $\hat{\alpha}$ will be close to zero. Hence, α' is a sparse vector. This algorithm can also be extended to deal with occlusions and random noise. Furthermore, a method of rejecting invalid test samples can also be introduced within this framework. In particular, to decide whether a given test sample is a valid sample or not, the notion of Sparsity Concentration Index (SCI) has been proposed in [59]. See [59] for more details.

2.6 Experimental Results

In this section, we report experimental results using the algorithms described in Sect. 2.5 on the remote face dataset [17].

The first set of experiments was designed to test the effectiveness of albedo maps [10]. We select the gallery set from clear images, and gradually increase the number of gallery of faces from one to fifteen images per subject. Each time the gallery images are chosen randomly. We repeat the experiments five times and take the average to arrive at the final recognition result. All the remaining clear images are selected

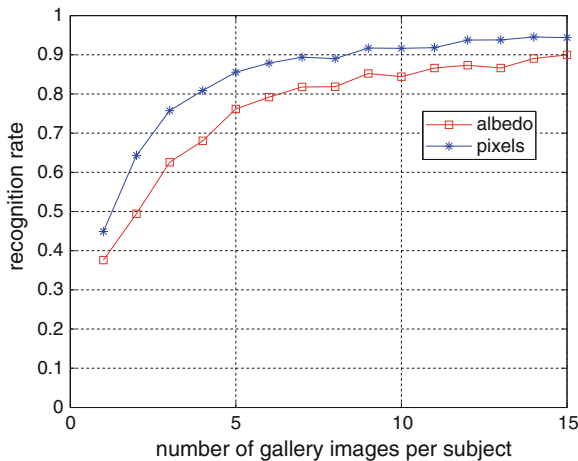


Fig. 2.10 Comparison of intensity images and albedo maps using baseline

for testing. We compare the albedo maps with the intensity images as input to the baseline. All the parameters of PCA, LDA and SVM are fine tuned. The results are shown in Fig. 2.10. We found that intensity images outperform albedo maps although the albedoes are intended to compensate for illumination variations. One possible reason is that, the face images in the database are sometimes a bit away from frontal. As albedo estimation needs a good alignment between the observed images and the ensemble mean, the estimated albedo map is erroneous. These artifacts are also seen in Fig. 2.2. On the other hand, intensity images contain texture information which can partly counteract variations induced by pose.

In the second set of tests, we repeat the gallery selection as the first set of experiments, and select the test images to be clear, poorly illuminated, medium blurred, severely blurred, partially occluded and severely occluded respectively. The intensity images are used as input. The rank-1 recognition results using baseline are given in Fig. 2.11. We observe that the degradations in the test images decrease the performance, especially when the faces are occluded and severely blurred.

In the third set of experiments, we compare the SRC method and baseline algorithm. We selected 14 subjects with 10 clear images per subject to form the gallery set. The test images are selected to contain clear, blurred, poorly illuminated and occluded images respectively. For the SRC method, we compute the SCI value of each image which can be used as a measure to reject images of low quality.

From the comparison results reported in Fig. 2.12, we observe that when no rejection of test images is allowed, the recognition accuracy of baseline is superior to the SRC method. This may be the case because when gallery images do not contain variations that occur in the test images, the SRC method can not approximate the test images correctly through linear span of gallery images. However, when rejection of images is allowed, we remove images whose SCI values are below certain threshold, and the performance of SRC method increases accordingly. The rejection rates in

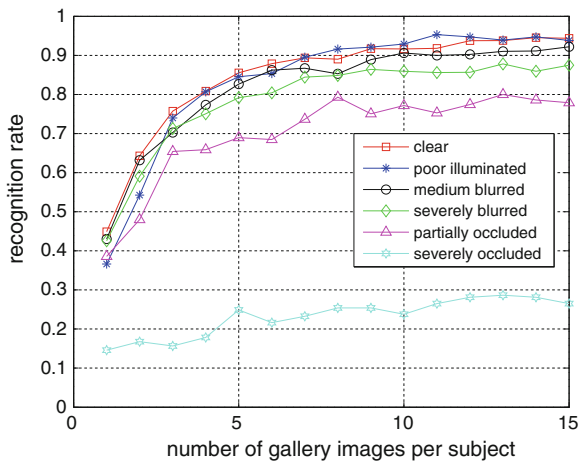


Fig. 2.11 Performance of the baseline algorithm as the condition of probe varies

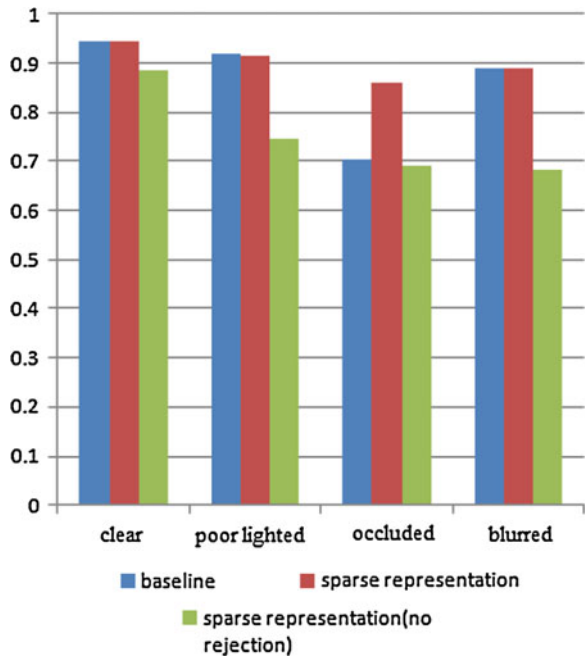


Fig. 2.12 Comparison between SRC and baseline algorithms

Fig. 2.12 are 6, 25.11, 38.46 and 17.33 % when the test images are clear, poorly lighted, occluded and blurred, respectively. Besides, we do observe the advantage of SRC method for handling occluded images.

2.7 Discussion and Conclusion

In this chapter, we briefly discussed some of the key issues in remote face recognition. We then described a remote face database collected by the authors' group and reported the performance of state-of-the-art FR algorithms on it. The results demonstrate that recognition rate decreases as the remotely acquired face images are affected by illumination variation, blur, occlusion, pose variation.

The evaluations reported here can provide guidance for further research in remote face recognition. In particular, the following are some areas where one can address the challenges in remote face recognition:

- Develop face recognition algorithms which can account for multiple sources of variation. For instance, low resolution is usually coupled with illumination and blur. The coupling among different variation factors makes the remote face recognition problem extremely difficult. Robust methods are needed to deal with these coupling effects in remote face recognition.
- Incorporate features which are robust to these variations. For instance, features which are robust to blur might not be able to deal with illumination variations. Hence, the design of features that can normalize the coupling effect due to various variations is needed.
- In order to develop robust recognition algorithm under these conditions, a more comprehensive quality metric to reject low quality images is needed. As a result this metric can make the recognition system more effective in practical acquisition conditions.

References

1. Abiantun R, Savvides M, Vijaya Kumar BVK (2006) How low can you go? low resolution face recognition study using kernel correlation feature analysis on the frgc2 dataset. In: Biometric consortium conference, 2006 biometrics symposium: special session on research at the, pp 1–6
2. Ahonen T, Rahtu E, Ojansivu V, Heikkilä J (2008) Recognition of blurred faces using local phase quantization. In: International conference on pattern recognition, pp 1–4
3. Arandjelovic OD, Cipolla R (2006) Face recognition from video using the generic shape-illumination manifold. In: European conference on computer vision, vol IV, pp 27–40
4. Arashloo SR, Kittler J (2011) Energy normalization for pose-invariant face recognition based on mrf model image matching. *IEEE Trans Pattern Anal Mach Intell* 33(6):1274–1280
5. Ashraf AB, Lucey S, Chen T (2008) Learning patch correspondences for improved viewpoint invariant face recognition. In: IEEE conference on computer vision and pattern recognition, pp 1–8
6. Bartlett MS, Movellan JR, Sejnowski TJ (2002) Face recognition by independent component analysis. *IEEE Trans Neural Networks* 13(6):1450–1464
7. Basri R, Jacobs DW (2003) Lambertian reflectance and linear subspaces. *IEEE Trans Pattern Anal Mach Intell* 25(2):218–233
8. Belhumeur P, Hespanha J, Kriegman D (1997) Eigenfaces versus fisherfaces: recognition using class specific linear projection. *IEEE Trans Pattern Anal Mach Intell* 19(7):711–720
9. Belhumeur PN, Kriegman DJ (1996) What is the set of images of an object under all possible lighting conditions? In: IEEE conference on computer vision and pattern recognition

10. Biswas S, Aggarwal G, Chellappa R (2009) Robust estimation of albedo for illumination-invariant matching and shape recovery. *IEEE Trans Pattern Anal Mach Intell* 29(2):884–899
11. Biswas S, Bowyer KW, Flynn PJ (2010) Multidimensional scaling for matching low-resolution facial images. In: *IEEE international conference on biometrics: theory applications and systems*, pp 1–6
12. Biswas S, Chellappa R (2010) Pose-robust albedo estimation from a single image. In: *IEEE conference on computer vision and pattern recognition*, pp 2683–2690
13. Blanz Volker, Vetter Thomas (2003) Face recognition based on fitting a 3d morphable model. *IEEE Trans Pattern Anal Mach Intell* 25:1063–1074
14. Candès EJ, Li X, Ma Y, Wright J (2011) Robust principal component analysis? *J ACM* 58(1): 1–37
15. Castillo CD Jacobs DW (2009) Using stereo matching with general epipolar geometry for 2d face recognition across pose. *IEEE Trans Pattern Anal Mach Intell* 31(12):2298–2304
16. Chai Xiujuan, Shan Shiguang, Chen Xilin, Gao Wen (2007) Locally linear regression for pose-invariant face recognition. *IEEE Trans Image Process* 16(7):1716–1725
17. Chellappa R, Ni J, Patel VM (2012) Remote identification of faces: problems, prospects, and progress. *Pattern Recogn Lett* 33(15):1849–1859
18. Chen T, Yin W, Zhou XS, Comaniciu D, Huang TS (2006) Total variation models for variable lighting face recognition. *IEEE Trans Pattern Anal Mach Intell* 28(9):1519–1524
19. Chen Y-C, Patel VM, Phillips PJ, Chellappa R (2012) Dictionary-based face recognition from video. In: *European conference on computer vision*
20. Etemad K, Chellappa R (1997) Discriminant analysis for recognition of human face images. *J Opt Soc Am* 14:1724–1733
21. Friedman J (1989) Regularized discriminant analysis. *J Am Stat Assoc* 84:165–175
22. Georgiades AS, Belhumeur PN, Kriegman DJ (2001) From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Trans Pattern Anal Mach Intell* 23(6):643–660
23. Georgiades AS, Belhumeur PN, Kriegman DJ (2001) From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Trans Pattern Anal Mach Intell* 23(6):643–660
24. Gross R, Matthews I, Baker S (2004) Appearance-based face recognition and light-fields. *IEEE Trans Pattern Anal Mach Intell* 26(4):449–465
25. Gunturk BK, Batur AU, Altunbasak Y, Hayes III MH, Mersereau RM (2003) Eigenface-domain super-resolution for face recognition. *IEEE Trans Image Process* 12(5):597–606
26. Guo G, Li SZ, Chan K (2000) Face recognition by support vector machines. In: *IEEE international conference on automatic face and gesture recognition*, Grenoble, France, pp 196–201
27. Hennings-Yeomans PH, Baker S, Kumar BVKV (2008) Simultaneous super-resolution and feature extraction for recognition of low-resolution faces. In: *IEEE international conference on computer vision and pattern recognition*, pp 1–8
28. Hotta K, Kurita T, Mishima T (1998) Scale invariant face detection method using higher-order local autocorrelation features extracted from log-polar image. In: *IEEE international conference on automatic face and gesture recognition*, pp 70–75
29. Huang G, Ramesh M, Berg T, Learned-Miller E (2007) Labeled faces in the wild: a database for studying face recognition in unconstrained environments. University of Massachusetts, Amherst, Technical Report, pp 07–49
30. Jia K, Gong S (2005) Multi-modal tensor face for simultaneous super-resolution and recognition. In: *IEEE international conference on computer vision*, vol 2, pp 1683–1690
31. Kanade T, Yamada A (2003) Multi-subregion based probabilistic approach toward pose-invariant face recognition. In: *IEEE international symposium on computational intelligence in robotics and automation*, vol 2, pp 954–959
32. Lee K, Ho J, Kriegman DJ (2005) Acquiring linear subspaces for face recognition under variable lighting. *IEEE Trans Pattern Anal Mach Intell* 27(5):684–698
33. Lee K-C, Ho J, Yang M-H, Kriegman D (2005) Visual tracking and recognition using probabilistic appearance manifolds. *Comput Vis Image Underst* 99:303–331

34. Lee Sang-Woong, Park Jooyoung, Lee Seong-Whan (2006) Low resolution face recognition based on support vector data description. *Pattern Recognit* 39(9):1809–1812
35. Li A, Shan S, Chen X, Gao W (2009) Maximizing intra-individual correlations for face recognition across pose differences. In: *IEEE conference on computer vision and pattern recognition*, pp 605–611
36. Li B, Chang H, Shan S, Chen X (2009) Coupled metric learning for face recognition with degraded images. In: *Asian conference on machine learning: advances in machine learning*, Berlin, Heidelberg, Springer, pp 220–233
37. Li B, Chang H, Shan S, Chen X (2010) Low-resolution face recognition via coupled locality preserving mappings. *IEEE Signal Process Lett* 17(1):20–23
38. Medioni G, Choi J, Kuo C-H, Choudhury A, Zhang L, Fidaleo D (2007) Non-cooperative persons identification at a distance with 3d face modeling. In: *IEEE international conference on biometrics: theory, applications, and systems*, pp 1–6
39. Moghaddam B (2000) Bayesian face recognition. *Pattern Recognit* 33:1771–1782
40. Narasimhan SG, Nayar SK (2003) Shedding light on the weather. In: *IEEE conference on computer vision and pattern recognition*, vol 1, pp I-665–I-672
41. Nayar SK, Narasimhan SG (1999) Vision in bad weather. In: *IEEE international conference on computer vision*, vol 2, pp 820–827
42. Ni J, Turaga P, Patel VM, Chellappa R (2011) Example-driven manifold priors for image deconvolution. *IEEE Trans Image Process* 20(11):3086–3096
43. Nishiyama M, Hadid A, Takeshima H, Shotton J, Kozakaya T, Yamaguchi O (2011) Facial deblur inference using subspace analysis for recognition of blurred faces. *IEEE Trans Pattern Anal Mach Intell* 33(4):838–845
44. Patel VM, Wu T, Biswas S, Philips PJ, Chellappa R (2012) Dictionary-based face recognition under variable lighting and pose. *IEEE Trans Inf Forensics Secur* 7(3):954–965
45. Patel VM, Wu T, Biswas S, Phillips PJ, Chellappa R (2011) Illumination robust dictionary-based face recognition. In: *International conference on image processing*
46. Phillips PJ, Wechsler H, Huang J, Rauss PJ (1998) The feret database and evaluation procedure for face-recognition algorithms. *Image Vis Comput* 16:295–306
47. Pinto N, DiCarlo J, Cox D (2009) How far can you get with a modern face recognition test set using only simple features? In: *IEEE conference on computer vision and pattern recognition*, Miami, FL, pp 2591–2568
48. Prince SJD, Warrell J, Elder JH, Felisberti FM (2008) Tied factor analysis for face recognition across large pose differences. *IEEE Trans Pattern Anal Mach Intell* 30(6):970–984
49. Ramamoorthi R, Hanrahan P (2001) On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object. *J Opt Soc Am* 18(10):2448–2459
50. Rara HM, Elhabian SY, Ali AM, Miller M, Starr TL, Farag AA (2009) Distant face recognition based on sparse-stereo reconstruction. In: *IEEE international conference on image processing*, pp 4141–4144
51. Shashua A, Riklin-Raviv T (2001) The quotient image: class-based re-rendering and recognition with varying illuminations. *IEEE Trans Pattern Anal Mach Intell* 23(2):129–139
52. Shekhar S, Patel VM, Chellappa R (2011) Synthesis-based recognition of low resolution faces, *International Joint Conference on Biometrics (IJCB)* 1–6. Washington, DC
53. Sim T, Baker S, Bsat M (2003) The cmu pose, illumination, and expression database. *IEEE Trans Pattern Anal Mach Intell* 25:1615–1618
54. Taheri S, Sankaranarayanan AC, Chellappa R (2012) Joint albedo estimation and pose tracking from video. *IEEE Trans Pattern Anal Mach Intell* 35(7):1674–1689
55. Tistarelli M, Li SZ, Chellappa R (2009) *Handbook of remote biometrics: for surveillance and security*, 1st edn. Springer Publishing Company Incorporated, New York
56. Turk M, Pentland A (1991) Eigenfaces for recognition. *J Cogn Neurosci* 3:71–86
57. Wang H, Li SZ, Wang Y (2004) Generalized quotient image. In: *International conference on computer vision and pattern recognition*

58. Wiskott Laurenz, Fellous Jean-Marc, Krger Norbert, Von Der Malsburg Christoph (1997) Face recognition by elastic bunch graph matching. *IEEE Trans Pattern Anal Mach Intell* 19:775–779
59. Wright J, Ganesh A, Yang A, Ma Y (2009) Robust face recognition via sparse representation. *IEEE Trans Pattern Anal Mach Intell* 31:210–227
60. Yang M-H (2002) Kernel eigenfaces vs. kernel fisherfaces: face recognition using kernel methods. In: *IEEE international conference on automatic face and gesture recognition*, Washington, DC, pp 215–220
61. Yao Y, Abidi B, Kalka N, Schmid N, Abidi M (2008) Improving long range and high magnification face recognition: database acquisition, evaluation, and enhancement. *Comput Vis Image Underst* 111:111–125
62. Zhang L, Samaras D (2003) Face recognition under variable lighting using harmonic image exemplars. In: *International conference on computer vision and pattern recognition*
63. Zhang T, Tang YY, Fang B, Shang Z, Liu X (2009) Face recognition under varying illumination using gradientfaces. *IEEE Trans Image Process* 18(11):2599–2606
64. Zhang Xiaozheng, Gao Yongsheng (2009) Face recognition across pose: a review. *Pattern Recogn* 42:2876–2896
65. Zhang Z, Blum RS (1998) On estimating the quality of noisy images. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5:2897–2900. Seattle, Washington
66. Zhao W, Chellappa R, Phillips J, Rosenfeld A (2003) Face recognition: a literature survey. *ACM Comput Surv* 35(4):399–458
67. Zhou SK, Aggarwal G, Chellappa R, Jacobs DW (2007) Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination-invariant face recognition. *IEEE Trans Pattern Anal Mach Intell* 29(2):230–245
68. Zhou S, Krueger V, Chellappa R (2003) Probabilistic recognition of human faces from video. *Comput Vis Image Underst* 91:214–245

Signal and Image Processing for Biometrics

Scharcanski, J.; Proença, H.; Du, E. (Eds.)

2014, X, 331 p. 150 illus., 91 illus. in color., Hardcover

ISBN: 978-3-642-54079-0