Chapter 13

FACE VERIFICATION FOR ACCESS CONTROL

Wen Gao[1,2,3] and Shiguang Shan[1,3]
1. Institute of Computing Technology, Chinese Academy of Sciences
2. Department of Computer Science, Harbin Institute of Technology
3. Graduate School, Chinese Academy of Sciences
{wgao,sgshan}@ict.ac.cn

Abstract In this chapter, we discuss the research issues and state-of-the-art of face verification for access control. We start by analyzing a typical face verification system for access control, and then explore the dominant technologies in the field. We introduce most of the typical and popular algorithms, listing the open research issues or technical challenges. Some available commercial systems in this field are also listed, and three standard performance evaluations, FERET, XM2VTS and FRVT2000 are simply introduced.

Keywords: Face recognition, Access control, Face detection, Face verification, Template Matching, Deformable template, Eigenface, Fisherface, linear subspace, Active shape model, Active appearance model, Illumination cones, Photometric alignment, Elastic graph matching, Hidden Markov Models, Quotient image, linear object class, Principle Component Analysis,

13.1. Introduction

The human face is another attractive source of biometric information, from which discriminatory measurements can be acquired intuitively and naturally without much user interaction relative to other biometric information. The recognition of faces has been a well-established field of research, with a large number of algorithms proposed in the past thirty years. Some commercial systems have also emerged in recent years. Several survey papers are available [12, 14, 23, 26, 70].

Though it an easy task for human beings to recognize a person from just one or several photos, this is not the case for computers. Automatic face recognition has been recognized as one the most challenging tasks in pattern recognition and artificial intelligence. In this section, we briefly introduce some background and definitions of face verification, from a computational perspective.

13.1.1 Motivation

Verification of identity based on biometric information is essential for many security applications, since the conventional authentication approaches, e.g. user/password mechanisms, have proved unreliable and inconvenient. Examples include access control to physical facilities, security systems or information databases. Suspect tracking, surveillance and intrusion detection are also potential applications [26].

In addition to the wide range of commercial and law enforcement applications mentioned above, there are many emerging fields that can benefit from face verification technology, such as the new generation of intelligent human-computer interfaces and e-services, including tele-shopping and tele-banking.

There exist many optional biometric cues for identity verification, such as the iris, fingerprint, voice, handprint, signature, and retina. The human face plays an irreplaceable role in biometrics technology due to some of its unique characteristics. First, most cameras are non-invasive; therefore face verification systems are one of the most publicly acceptable verification technologies in use. Another advantage is that face recognition systems can work mostly without the cooperation of the user concerned, which is therefore very convenient for the general users. Furthermore, they can even work in the situation where the subject concerned is not aware of the procedure. This point greatly facilitates applications such as criminal hunting, surveillance, tracking shoplifters, suspect tracking and investigation, etc. This is why the U.S. has decided to install the FaceFINDER™ real-time face recognition system at an undisclosed major U. S. airport, after the 9.11 terrorism attacks.

13.1.2 Computational Solutions to Access Control Based on Face Verification

Access control is an important application of Biometric products. It is obvious that controlling the access to confidential physical buildings or information systems/databases, potentially dangerous vehicles (such as aeroplane, huge ship etc.), nuclear and biochemistry weapons etc., is essential for public security. However, conventional identity authentication mechanisms such as passwords or identification cards have proved unreliable. Biometrics is expected to solve these kinds of access control problems.

The general computational the machine recognition of faces can be formulated as follows: *given still or video images of a scene, identify one or more persons in the scene using a stored database of faces* [26]. Strictly speaking, there are two categories of applications in face recognition: face recognition/identification and face verification/authentication. In a recognition application, the input to the system is a face image, and the system reports the decided identity from a database of known individuals, whereas in verification application, the system confirms or rejects the claimed identity according to the input face image.

The basic architecture of face verification solutions for access control is illustrated in Figure 13.1. Here, one can see that an access control system based on face verification is commonly divided into several modules, including face detection, feature extraction, and face verification etc. The following sections describe these modules in detail.

13.1.3 Organization of this Chapter

The remainder of the chapter is organized as follows. First, a short review on face detection is given in Section 13.2. Then we describe three distinct aspects of face verification in detail in Section 13.3: models of identify, feature extraction, and classification. The solutions to illumination and pose problems are discussed in Section 13.4. An example access control system based on face verification is presented in Section 13.5. Several commercial systems are introduced and some famous evaluation protocols are addressed in Section 13.6. Some conclusions are drawn in last section.

13.2. Face Detection

Face detection aims to determine whether there are faces in an image, and, if any, determine how many there are and where each face is. The approaches to face detection may be divided into four categories [83]: knowledge based methods, template matching, invariant feature methods, and statistics/learning based methods. Although the boundaries between them are not always clear, we can still briefly review most of the approaches to face detection under this taxonomy.

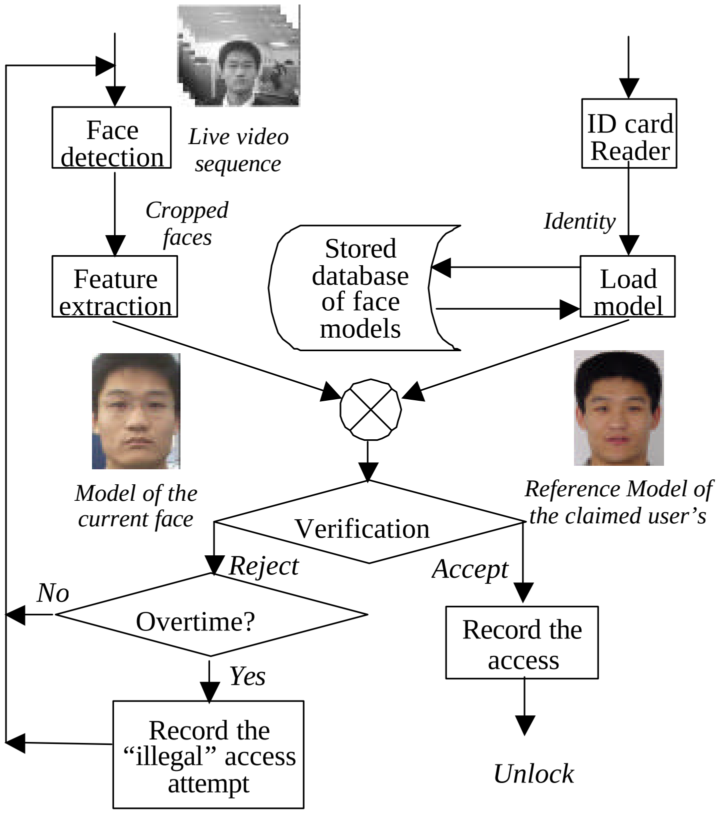


Figure 13.1. Access control system based on face verification.

1. Knowledge Based Methods

The classical work in this category is the multiple-rule based method proposed by Yang and Huang [22]. The main problem with knowledge-based methods is the difficulty of transforming human knowledge into rules described in computer languages, especially for 3-D rotated faces in different poses.

2. Template Matching Methods

Miao and Yin, et al. [67] proposed a mosaic Gravity-Center Template matching method. It can be observed that the main components of an upright human face, such as double eyebrows, double eyes, nose bottom and mouth, almost all orient in a horizontal direction and that the vertical scale of the features are approximately equal.

3. Invariant Feature Methods

There are many works using various invariant features including gray values, edges, textures, color or a combination of these features. Among them, color is most widely used for both face detection and lip-reading [84].

However, color information is not enough to correctly locate faces, although non-upright and non-frontal faces can be easily detected as candidates. It is therefore usually combined with other features such as edges or textures.

4. Statistics/Learning Based Methods

The methods in this category are the most widely used ones by a majority of researchers. They include the use of Eigenfaces [2, 5, 49], Fisherfaces [51], Neural Networks [59, 60], and Support Vector Machines (SVMs) [41, 61, 75, 76]. SVMs were first introduced by Osuna, Freud and Girosi [41] for face detection. Their system reports better performance than some neural network systems.

By using the segmentation techniques mentioned in above, given a still image or a live video sequence, faces can be cropped out and fed into the following face verification modules.

13.3. Face Verification

It is difficult to thoroughly review all the face verification technologies developed by thousands of researchers since the 1970s. Instead of discussing individual face recognition approaches in this chapter, we summarize the relevant technologies in terms of three aspects: the models of identity, feature extraction methods, and classification methods.

13.3.1 Models of Identity

Models of identity answer questions about what kind of representation should be exploited for faces, that is, what kind of "template" *T* should be extracted to represent a face. Typical models of identity are discussed below.

Geometric Features. Geometric features of a face were often exploited by early researches such as Poggio [8, 14] etc. The idea is to describe a face as the relative position of distinctive facial features such as eyes, mouth, nose and chin, as well as other parameters. Systems using geometric features for face recognition can be found in [8, 14, 9, 10, 12, 50]. Since pure geometry is not sufficient for recognition, it is generally combined with other features, such as gray-level templates etc.

Holistic or Analytic Templates. In the simplest version of template matching, the face is represented as a bi-dimensional array of intensity values sampled from the whole face region, that is, a template. A probe template is then compared using a suitable metric (typically the Euclidean distance) with stored ones [14, 28, 50, 71]. Several full templates per face may be used to account for the recognition from different viewpoints, as Beymer has done in

[28]. Another variation is to use multiple templates to cover different illumination variations, even for a single viewpoint.

Iso-Density Maps. Iso-density lines are the boundaries of constant gray level areas after quantifying an image. To understand the concept of iso-density lines better, one can consider the following analog from geology: If the brightness of a picture is viewed as the height of a mountain, then the equal altitude contour lines correspond to iso-density lines [4].

To extract isodensity lines, the gray level histogram is utilized. In [4], the histogram is divided into eight areas. Contour line tracing on iso-density levels is performed, and outlines based on 4-connectivity pixels are extracted as the iso-density lines, to represent the 3D structure of the face.

Statistical Principal Components. Eigenface, proposed by Turk and Pentland in 1991 [5], is a well-known model in the face recognition community, which is essentially PCA or KLT in a certain context. The basic idea is to regard face images as points in a high dimensional image space. These points approximately form a subspace, so called "face subspace", in the image space. A group of orthogonal bases of the "face subspace" can be estimated by eigenspace decomposition of the covariance matrix derived from a set of training face images. The bases are called "Eigenfaces" for their visual similarity to true faces. Any face image can then be represented approximately as a linear combination of Eigenfaces. The coefficients of the linear combination, computed by projecting the face image onto the subspace, are conventionally used as the features and can be fed into any classifier for recognition. Eigenfaces have been widely used in a variety of systems [4, 5, 19, 34, 37, 39, 49, 56, 62, 70, 72, 73], including some commercial systems.

Singular Values. Singular values of an image intensity array imply algebraic features of the face image, and can be used to represent the face. The Singular Value Decomposition (SVD) decomposes a matrix A into its left singular vectors L, right singular vectors R and the corresponding singular values V. It has been shown that the singular vectors represent "shape" information while the singular values are representative of the "gain" in the image.

In [3], Hong shows that the singular values (SVs) extracted by SVD perform well as shape descriptors. He also proves their stability and invariance to proportional variation of image intensity in the optimal discriminating vector space, and to transposition, rotation, translation and reflection. When an image is represented as an n-dimensional SV feature vector, the recognition problem can be solved in an n-dimensional feature space.

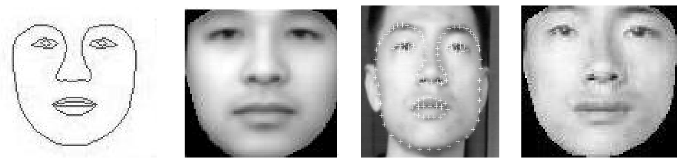
2D Shape Model. It is believed that the contours of the whole face and its salient features (eye, nose and mouth) are important for recognition in the biological vision system. This is why many choose to represent a face by its shape.

One model for shape is the Point Distribution Model (PDM) [22]. In PDM, shape is generally represented by a vector of length *2n*, consisting of the concatenation of the *x* and *y* coordinate values of *n* predefined landmarks in the object image. The PDM models assume the existence of a set of *M* annotated examples from which to derive a statistical description of shape variation. As a powerful shape description, a PDM can subsequently be used to locate new instances of such shapes in other images. It is most useful for describing shapes that have a well understood "general" configuration, but which cannot be easily described by a rigid model [22, 53, 55, 65, 77]. Obviously, a PDM sets up a sparse correspondence among shapes, while optical flow derives a pixel-wise, dense correspondence between shapes. However, both descriptions can be vectorized as a shape variable, denoted as *x*. It can be represented compactly by performing a Principal Component Analysis (PCA) as $b_s = P_s^T(x - \bar{x})$ *and* $x = \bar{x} + P_s b_s$, where P_s is the principal component matrix. The linear combination coefficients, b_s, are named statistical shape parameters [22, 53, 55, 65, 77].

Another available model to represent face shape utilizes correspondence field, e.g. *optical flow*, which describes the relative shape of an image measured with respect to an image with a standard reference shape. Optical flow is a pixel-wise representation for shape, defining a feature point at each pixel in a sub-image containing a face. The shape vector can be visualized as a vector field of correspondences between a face of standard shape and the given image being represented [28, 29, 45].

Statistical Texture Parameters. Texture is the geometrically normalized version of the facial image, or shape-free gray-level face patch. That is, geometrical differences among face images, i.e. spurious texture variation due to shape differences, are factored out by warping images to a predefined standard reference shape. This can be done by warping an example image so that its control points match the mean shape [22, 53, 55, 65, 77]. Figure 13.2 shows results from our experiments, where shape is modelled by a PDM.

Generally, texture is statistically modelled by PCA. Any texture can be represented by statistical texture parameters, b_g, as follow: $b_g = P_g^T(g - \bar{g})$ and $g = \bar{g} + P_g b_g$, where P_g is the principal component matrix. In this case, the linear combination coefficients, *bg*, are called statistical texture parameters.



(a) Reference shape (b) Reference texture (c) Annotated face (d) Getting texture

Figure 13.2. Image warping and texture.

Statistical Appearance Model. Shape and texture are two distinct aspects of facial appearance. They can be combined together appropriately to model a face for representation and/or recognition.

If both shape and texture are modeled with PCA, the shape and texture of any example can be summarized by its shape parameter vector and texture parameter vector. Since there may be correlations between shape and texture variations, a further PCA can be applied to the data. For each example, a vector can be generated by concatenating the shape and texture parameters with weights applied to each shape parameter, allowing for difference in units between shape and texture models. Then PCA can be applied on these vectors to build a combined model [55, 65, 77].

Formally, for each example, the shape parameter *bs* and the texture parameter b_g are combined as: $b = \begin{pmatrix} w_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} w_s p_s^T (x - \bar{x}) \\ p_g^T (g - \bar{g}) \end{pmatrix}$, where w_s is a diagonal matrix of weights for each shape parameter. PCA is further applied to generate $b = Qc$ *and* $Q = \begin{pmatrix} Q_s \\ Q_g \end{pmatrix}$, where Q is a matrix composed of eigenvectors and c is a vector of appearance parameters combined with the shape and texture information. Then one can express shape and texture as functions of *c* by: $x = \bar{x} + P_s W_s Q_s c$ and $g = \bar{g} + P_g Q_g c$.

Statistical appearance models can be used for both image analysis and image synthesis. Active appearance models (AAMs) [55, 65, 77], in which statistical appearance models have been applied to image analysis and synthesis by T.F.Cootes and Taylor, has attracted more and more attention in both the computer graphics and face recognition community in recent years.

Labelled Graph. Faces can be represented as labelled graphs [33, 46, 49], with nodes positioned at fiducial points and labelled with local texture information, and edges labelled with the 2D distance between fiducial points. Typical local texture features are Gabor wavelet features, in which each node

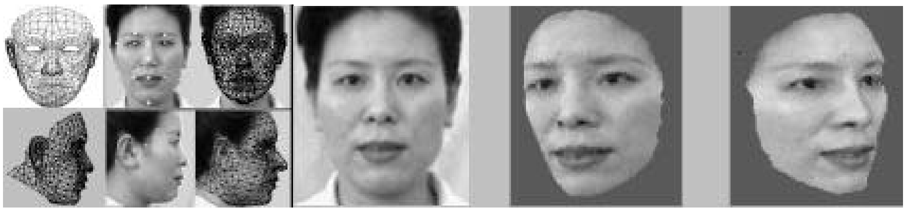
contains a set of complex Gabor wavelet coefficients, known as a *jet*. A jet can be expressed as $J_j = a_j \exp(i\phi_j)$, where magnitudes, $a_j(\vec{x})$, vary slowly with position, and phases, $\phi_j(\vec{x})$, rotate at a rate approximately determined by the frequency of the kernel.
Theoretically, the labelled graph is an attractive model to represent a face, since it not only models geometric (configurable) information derived from the landmarks, but local features corresponding to predefined fiducial points and their neighbours.

Surface Property (Normal & Albedo). Ideally, the face surface can be approximated using Lambertian model. The image brightness of a Lambertian surface element illuminated by a point light source is $e(x) = n(x) \cdot s$, where $n(x) = \rho(x) \cdot \hat{n}(x)$ is the normal vector, $\rho(x)$ is the local surface albedo, and $\hat{n}(x)$ is the local unit surface normal. Similarly, $s = b \cdot \hat{s}$ is the source vector, where b represents the intensity of the source and $\hat{s}$ is a unit vector in the direction of the source. In Lambertian model, the intensity of a pixel in a face image depends on three factors: the surface normal, the albedo and the direction of the light source. Among them, the first two are inherent to the face surface, while the last one has nothing to do with the face property. Therefore, it is acceptable to recover the normal and albedo at each surface point from one or more images of a face viewed from a fixed viewpoint.

By factoring out the useless effect of lighting conditions in the face images, it is a theoretically perfect model to represent the face using illumination-free face properties such as surface normal and albedo, but it is more challenging to recover these surface properties. Nevertheless, the recent developments in shape-from-X (shading, motion, texture, contour etc.), illumination cones [82], and photometric stereo [43] will greatly facilitate the application of the model.

3D Face Model. By providing a prior knowledge, 3D face models can facilitate the face recognition in both facial image analysis and virtual view generation. A technique to derive a specific person's 3D face model from a general 3D face model and two images (frontal and lateral) of the person is described in [80]. The model fitting procedure is illustrated in Figure 13.3 (a,b,c). Once a 3D model is generated, virtual views can be re-rendered for both analysis and synthesis. Figure 13.3 (e, f) shows two synthesized views from the 3D face model generated in Figure 13.3(c) after texture mapping [80].

Summary. This section surveys the representation problems of face images. Eleven distinct face models are introduced, from basic principles to main applications, from geometry to radiometry. A trend that can be seen clearly is that statistical appearance models are emerging as dominant models.



(a) (b) (c) (d) (e) (f)

Figure.13.3. Warping a Generic 3D-Face Model to an Individual One (a) Generic 3D-Face Model. (b) Extracted Feature Points in Frontal and Lateral Images. (c) Fitted Individual 3D-face model. (d) Input Given Person's Frontal Image. (e, f) Synthetic Face Rotated 20 Degree to the left and right respectively.

13.3.2 Feature Extraction

Models of identity deal with the problem of how to model a face. In this section, we concentrate on how to build and apply the models mentioned in above. Obviously, one-to-one correspondence between models of identity and model-building algorithms is difficult to setup. Specific models of identity may be built by using several different feature extraction algorithms, while one algorithms for feature extraction may provides results that can be used to build several different models of identity.

Template Matching. Template matching is one of the most typical techniques for feature extraction. Correlation is commonly exploited to measure the similarity between a stored template and the window image under consideration. Templates should be deliberately designed to cover variety of possible image variations. During the search in the whole image, scale and rotation should also be considered carefully to speed up the process [14].

Beymer [28] exploited template matching to locate salient features (eyes, nose and mouth) in a face image. Affine transforms are introduced to process matching with pose variations.

Hough Transform. The Hough Transform (HT) is a well-accepted feature detection method. It requires an explicitly chosen class of objects for detection (e.g., lines, circles, or ellipses) and the parameterization of this class that describes all possible "ideal" instances of the object. Since in the contour of a face, few "ideal" mathematical curves exist, its application in face recognition is limited. However, it can be used to detect the irises of the eyes in the detected face region.

Deformable Template. Deformable template (DT), proposed by Yuille [6, 11], is a feature extraction algorithm that extracts geometric features from an image. As an algorithm that makes use of the global information, DT is effective in accurately and robustly extracting the locations and the shapes of salient facial organs such as eyes, mouth and chin against noises. In DT, templates are specified by a set of parameters that enables the priori knowledge about the expected shape of the features to guide the detection process. The templates are flexible enough to change their sizes and other parameters to match themselves to the data. The final values of the parameters are used to describe the features. This method works despite of the variations in scale, tilt, rotation and lighting conditions. Variations of the parameters allow the template to fit any normal instance of the feature.

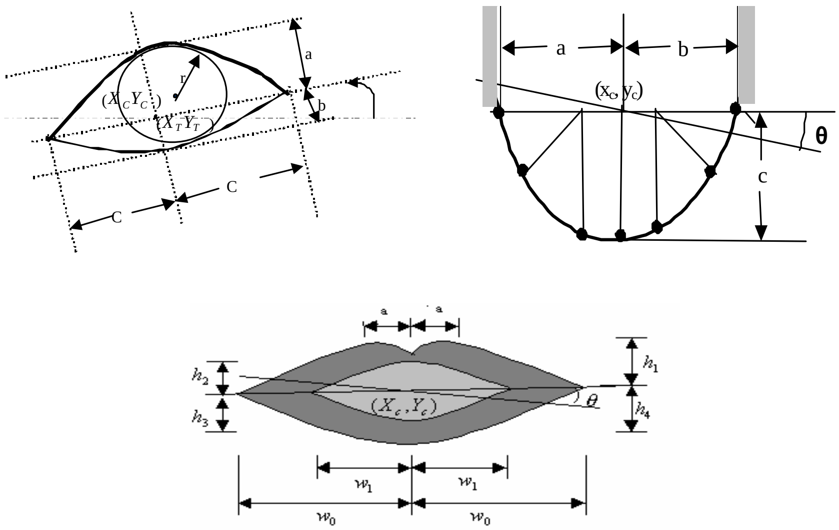


Figure 13.4. Template for the eyes, chin and mouth.

The deformable templates interact with the image in a dynamic manner. An energy function is defined, which contains terms attracting the template to salient features, such as edges, peaks, valleys and the intensity itself. The minimum of the energy function corresponds to the best fit of the image. The parameters of the template are then updated using steepest descent or other optimal algorithms. This corresponds to following a path in parameter space, and contrasts with traditional methods of template matching, which would involve sampling the parameter space to find the best match. Changing these parameters corresponds to altering the position, orientation, size and other properties of the template. The initial values of the template are determined by a preprocessing [6, 11]. Typical templates for the eyes, mouth and chin are illustrated in Figure 13.4.

Active Contour. Active Contour Model proposed by Kass et al. [1] is a sophisticated approach to contour extraction in image understanding. It is defined as an energy minimizing process of a contour. If the contour is described parametrically by $v(s) = (x(s), y(s))$, where $x(s), y(s)$ are *x, y* coordinates along the contour and $s \in [0, 1]$, the energy function is defined as follows:

$$E_{snake} = \int_0^1 E_{snake}(v(s))ds = \int_0^1 \left[E_{\text{int}}(v(s)) + E_{image}(v(s)) + E_{con}(v(s))\right]ds$$

where E_{con} represents external constraint forces and E_{int} represents the internal energy of the contour. Generally, E_{int} is defined as $E_{\text{int}} = \alpha(s)|\frac{dv}{ds}|^2 + \beta(s)|\frac{d^2v}{ds^2}|^2$, where $\alpha(s)$, $\beta(s)$ specify the elasticity and stiffness of the active contour. The second term E_{image} is the image energy that attracts the contour to the desired feature in the image. As an example, the following definition of E_{image} will attract the active contour to lines, edges, and terminations: $E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term}$.

In face recognition, it is important to extract the whole face from the image so that the changing background will not influence the recognition results. While the contour of a face is usually difficult to be parameterized, the active contour model is suitable to extract the contour of a face.

Principal Component Analysis (PCA). PCA estimates a group of orthogonal and linearly independent bases of the "face subspace" using eigenspace decomposition of the covariance matrix derived from a set of training face images. Formally, let $u_1, u_2, \ldots, u_m$ be the leading eigenvectors corresponding to the first m maximum eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ of the covariance matrix, thus $U_f = [u_1 u_2 \cdots u_d]$ expands the "face subspace". Then any face image x can be approximately represented as a linear combination of the Eigenfaces. The coefficients of the linear combination can be computed by projecting the face image to the subspace through $W = U_f^T x$.

Conventionally, $W = [w_1 w_2 \cdots w_d]$ is used as the features extracted from the input face image x and can be fed into any classifier (Nearest

Neighbours, Linear Discriminant Analysis (LDA) [39, 62, 70], Support Vector Machine (SVM) [75] and Artificial Neural Network (ANN) etc.) for recognition.

It is well known that Eigenface transform is the optimal transform in the sense of Minimum Square Error (MSE), but not Most Discriminating Features (MDF). Many efforts have been done to seek MDF from the grey-level information or the Eigenface transforms [51, 62, 70]. LDA is the most popular one.

In practice, face images should be processed carefully to align all the training examples as well as the new given images under recognition. To alleviate the influence of background, translation, rotation, lighting, scale variance, geometric and intensity normalization should be carefully adopted.

Active Shape Model. Active shape model [22, 53] is an optimal procedure to extract 2D shape representation from an input facial image based on the statistical shape model trained by PCA from a set of annotated facial images.

Given a rough starting approximation, an instance of a shape model can be fitted to a new given image by searching in the space of shape parameters. An iterative approach is employed to improve the fitting of the instance to an image. In practice, one can look along profiles normal to the model boundary at each model point. If we expect the model boundary to correspond to an edge, we can simply locate the strongest edge (including the orientation if known) along the profile. This position indicates a new location for the model point. PCA is then conducted to revise the new shape according to the statistical prior. To speed up the searching procedure, multi-resolution strategy should be adopted as in [22, 53]. An example of ASM is illustrated in Figure 13.5, where 55 landmarks are used.

(a) Initial shape (b) Result of ASM search

Figure.13.5. Active Shape Model.

Active Appearance Model. By training on a set of facial image annotated by predefined landmarks, statistical appearance model is finally modeled as:

$$x ? \bar{x} ? P_s W_s Q_s c\,,\ g ? \bar{g} ? P_g Q_g c\,,\ Q ? \begin{matrix} ?Q_s ? \\ ? \quad ? \\ ?Q_g ? \end{matrix},$$

where c is the final appearance model.

Active Appearance Model (AAM) [55, 65] then deals with the kernel problem: given an image to be interpreted, an appearance model described above and a reasonable starting approximation, AAM adjusts the model parameters to generate a synthetic example that matches the new image as much as possible. Basically, AAM is an optimization algorithm that searches in both the parameter space of the appearance model and the global transformation parameter space. The interpretation procedure can be considered as an optimization, where the difference between the new given image and the model image synthesized by the appearance model is minimized. A difference vector $?I$ can be defined as $?I ? I_n ? I_m$, where I_n is the vector of grey-level values in the new given image, and I_m is the vector of grey-level values for the current model parameters. Then, searching the best match between the model and the given image is equivalent to minimizing the magnitude of the difference vector, $? ? \|?I\|^2$, by varying c, the model parameters. Cootes etc. [55, 65, 77] further proposed to learn some priori knowledge about how to adjust the model parameters during an image search to speed up the search procedure. Their method is to model the relationship between $?I$ and $?c$ by using a linear regression model: $?c ? ? ?I$. Then $?c$ is used in the iterative algorithm for minimizing $?$. Figure 13.6 illustrates the procedure of AAM and the final experimental results.

(a) (b) (c) (d)

Figure 13.6. AAM procedure and results (a) Input new given image (b) Initial model image overlapped (c) Intermediate model image overlapped (d) Final model image overlapped.

Optical Flow-Based Shape and Texture Extraction. Beymer [29] proposed an algorithm, named *Vectorizer*, to build the correspondence field between a new given image and a pre-defined reference image with standard shape. In other words, it tries to extract the relative shape representation of the face and its shape-free grey-level intensity.

The algorithm is first initialized by optical flow to build the pixel-wise correspondence between the new given image and the reference image. Then an iterative process of interleaving shape and texture computation is carried out to tune the correspondence. The face *Vectorizer* alternates back and forth between the shape computation and texture computation. The key idea here is to couple the two computations so that each of them uses the other's output, that is, the texture computation uses shape for geometrical normalization, and the shape computation uses the texture analysis to synthesize a "reference" image for finding correspondence.

Since the *Vectorizer* finally builds a pixel-wise dense correspondence between the two images, it can be applied to the problems of facial feature detection and registration of two arbitrary faces.

Elastic Bunch Graph Matching. To extract a graph model from a face, a data structure named Face Bunch Graph (FBG) is developed to represent the faces in general [46]. FBG is constituted by multiple face graphs. Each graph has the same topologic structure, with the nodes referring to the identical fiducial points from different faces. A set of *jets*, referring to the same fiducial points from different faces, is called a bunch. To represent wide ranges of local features, different jets should be stored at each node. The graph edges are labelled with the averages of the distance between the fiducial points. Generally, FBG is constructed manually or self-automatically by deliberately designing a bootstrapping algorithm.

So far, a face is modelled as a labeled graph, and faces in general are represented by FBG as mentioned above. Then, given a new image, the goal of Elastic Bunch Graph Matching (EBGM) is to build its labeled graph by finding the fiducial points, and thus to extract a graph from the image which maximizes its similarity with the Face Bunch Graph based on certain similarity function.

In practice, a heuristic algorithm is generally needed to approach the optimum within a reasonable time. A coarse to fine approach can be adopted by introducing the degrees of freedom of the FBG progressively: translation, scale, aspect ratio, and local distortion.

Photometric Stereo. **A**ssuming a known reflectance function, photometric stereo recovers surface orientation unambiguously. Take a particular Lambertian surface with varying albedo $?(x, y)$ as an example. The key idea of photometric stereo is to look at the surface from one fixed viewing direction while changing the direction of incident illumination. Suppose we have three or more such Lambertian surface images, the surface normal can be uniquely determined based on the shading variations of the observed image.

In traditional photometric stereo, to recover shape and albedo under Lambertian surface assumption, we need at least three images illuminated by three different point light sources with known intensities and incident directions. Quite recently, Georghiades, Belhumeur and Kriegman [52, 54, 57, 82] proposed the illumination cone technology to relax these constraints. Their method is to reconstruct the shape and albedo for each face by using seven face images taken in a fixed pose but illuminated by point light sources at varying, unknown positions. The surface geometry and albedo map is estimated up to a generalized bas-relief (GBR) transformation. The symmetries and similarities in faces can be utilized to solve the three parameters specifying the GBR transformation.

Summary. In this section, techniques for building the face models described in Section 13.3.1 are further discussed. Typical features extraction algorithms include template matching, Hough transform, deformable template, active contour, active shape model and EGBM. They all can be used to extract geometric features and/or 2D shape models. Both AAM and *Vetorizer* can provide the information needed to build the statistical appearance model. SFS and Photometric aim at the recovery of surface properties.

13.3.3 Classification

In Section 13.3.1 and 13.3.2, we discussed the problem of modeling a face. So far, we are able to model both training images and testing images. In this section, we will concentrate on how to classify a testing image by comparing it with the stored models.

Nearest Neighbour (NN). The simplest classification scheme is the nearest neighbor classifier in the model space. Under this scheme, an image in the test set is recognized by assigning it the label of the closest point in the learning set. Here, distances are measured in the model space.

This procedure, which is also referred as correlation, has several well-known disadvantages. First, if the images in the learning set and test set are collected under varying lighting conditions, then the corresponding points in the image space will not be tightly clustered. Second, correlation is computationally expensive. Third, it requires large amount of storage since the learning set must contain numerous images for each person who registered.

Bayesian Inference. Moghaddam et al. [34, 37, 56, 73] formulate a probabilistic similarity measure based on the probability of the image intensity differences, denoted by $\Delta = I_1 - I_2$, that is, the characteristic of typical variations in appearance of the same object. Two mutually exclusive classes are defined: within-class differences Ω_I and between-class differences Ω_E.

In terms of the within-class differences, given a posterior probability determined by Bayesian rules, the similarity measure between two facial images can be directly defined as $S(I_1, I_2) = P(\Delta \in \Omega_I) = P(\Omega_I|\Delta) = \frac{P(\Delta|\Omega_I)P(\Omega_I)}{P(\Delta|\Omega_I)P(\Omega_I) + P(\Delta|\Omega_E)P(\Omega_E)}$, where the priors $P(\Omega_I)$ and $P(\Omega_E)$ can be set as default setting of equal priors. Moghaddam et al. model each of the classes as Gaussian density [73], and the class-conditional densities are defined as:

$$P(\Delta|\Omega_I) = \frac{1}{(2\pi)^{d/2}|\Sigma_I|^{1/2}} \exp\left(-\frac{1}{2}\Delta^T \Sigma_I^{-1} \Delta\right), \quad P(\Delta|\Omega_E) = \frac{1}{(2\pi)^{d/2}|\Sigma_E|^{1/2}} \exp\left(-\frac{1}{2}\Delta^T \Sigma_E^{-1} \Delta\right).$$

An alternative probabilistic similarity measure can be defined in a simpler form by only exploiting the within-class likelihood using the ML rule instead of the MAP rule. It is represented as $S(I_1, I_2) = P(\Delta|\Omega_I)$.

For identification problem, there is a gallery $\{g_j\}$ of K known individuals and a to-be-identified probe p. The similarity score between p and each g_j is $S(p, g_j)$. The probe is identified as person k who has the maximum similarity score, that is, $k = \arg\max_j S(p, g_j)$.

The performance advantage of the probabilistic matching technique has been demonstrated in an independent double-blind test on a large (800+) database as part of ARPA's September 1996 "FERET" competition, where Bayesian similarity outperformed competing algorithms [71, 72].

Artificial Neural Network (ANN). The application of Artificial Neural Networks (ANN) in face recognition has addressed several problems: gender classification, face recognition and classification of facial expressions [9, 23, 48, 49, 59, and 69]. In [48], Lawrence applies a Convolutional Neural-Network approach to face recognition.

Theoretically, any models of identity can be fed into a Neural Network for classification. Geometric features, templates, statistical shape/texture/

appearance models and singular values etc. can all be trained by NN for face verification.

Nearest Feature Line Method. The basic assumption of the nearest feature line (NFL) method is that at least two distinct prototype feature points are available for each class. This is usually satisfied in most cases. Working in a feature space, NFL method uses a linear model to interpolate and extrapolate each pair of prototype feature points that belong to the same class. More specifically, two prototype feature points are generalized by a feature line (FL) that passes through the two feature points. The FL approximates variants of the two prototypes under variations in poses, illuminations and expressions, i.e. all face images that could possibly derived from the prototypes. It virtually provides an infinite number of prototype feature points of the class. The capacity of the prototype set is thus expanded. The classification is done using the minimum distance between the feature points of the query and the FLs. The classification result also provides a quantitative position number as a by-product, which can be used to indicate the relative changes (in terms of poses, illuminations and expressions) between the query face and the two associated faces [85].

Linear Subspace Method. In [51], a linear subspace method is proposed based on the observation that for a Lambertian surface without self-shadowing, the images of a particular face lie in a 3-D linear subspace. For each face, three or more images are taken under different lighting directions to construct a 3-D basis for the linear subspace. To perform recognition, simply compute the distances between the new image and each linear subspace and choose the face with the shortest distance. This method is actually a variant of the photometric alignment method in [43].

In [86 and 87], a variant of the linear subspace method is proposed. It estimates the class conditional probability of each class based on the computations of linear subspace for each face. A face-unlock screen saver is developed in the spirit of this method. In [88 and 89], by deriving multiple virtual images from one face image, the subspace method is further extended to accommodate the case when only one sample image is available for each face.

Linear Discriminant Analysis (LDA). To use the class membership information and find eigenfeatures that emphasize the variations of different faces images while de-emphasize the variations of the same face due to different illuminations and facial expression etc., Linear Discriminant Analysis (LDA) was proposed in [51]. This method is class specific. It chooses an optimal projection W_{opt} as follows:

$$W_{opt} = \arg\max_{W} \frac{\left|W^T S_B W\right|}{\left|W^T S_W W\right|}$$

where S_B is the between-class scatter matrix and assumed to be non-singular, S_W is the within-class scatter matrix. In the projected space, the points corresponding to the images of the same face are clustered while those corresponding to the images of different faces are separated.

When LDA is applied to the face recognition, one difficulty arises: the within-class scatter matrix S_W is always singular. More specifically, the rank of S_W is at most $N - c$ with N being the total number of learning images and c being the number of different faces, but in general, N is much smaller than the pixels number n in each image. To overcome this difficulty, PCA is first used to reduce the dimension of the feature space from n to $N - c$ or less before the standard LDA method is used. This combined method is known as Fisherfaces. It partly solves the generalization problem and has demonstrated excellent performance [39, 62, 70].

Support Vector Machines (SVMs). Support Vector Machines (SVMs) have been recently proposed by Vapnik and his co-workers as a very effective method for general-purpose pattern recognition. Intuitively, given a set of points belonging to two classes, a SVM finds the hyper-plane that separates the largest possible fraction of points of the same class to the same side while maximizing the distances from either class to the hyper-plane. This hyper-plane is called Optimal Separating Hyper-plane (OSH). It minimizes the risk of misclassifying not only the samples in the training set but also the unseen samples in the test set [75].

The application of SVMs to computer vision area has emerged recently. Osuna et al. [41] train a SVM for face detection, where the discrimination is between two classes: face and non-face, each with thousands of samples. Guo and Stan [75] show that the SVMs can be effectively trained for face recognition and is a better learning algorithm than the nearest center approach.

Graph Matching. After all images, including the gallery images and the probe images, are extracted using EBGM procedure, the faces are represented as labelled face graphs. The matching procedure then involves the distance computation of the jets between different graphs, which is represented as:

$$S_a(J, J') = \frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}}$$

On the FERET dataset, the algorithm performs impressively well for the frontal images with recognition accuracy of 98%. For half rotated and profile images, the performance degrades to 57% and 84%, respectively; however, since these are difficult cases in face detection and recognition systems, the results are still comparatively good [71].

Hidden Markov Models (HMMs). HMMs are generally used for the statistical modelling of non-stationary vector time series. By considering the facial configurable information as a time varying sequence, HMMs can be applied to face recognition [25].

The most significant facial features of a frontal face image, including the hair, forehead, eyes, nose and mouth, occur in a natural order from top to bottom, even if the image has small rotations in the image plane, and/or rotations in the plane perpendicular to the image plane. Based on this observation, the image of a face may be modelled using a one-dimensional HMM by assigning each of these regions a state as illustrated in Figure 13.7 [25].

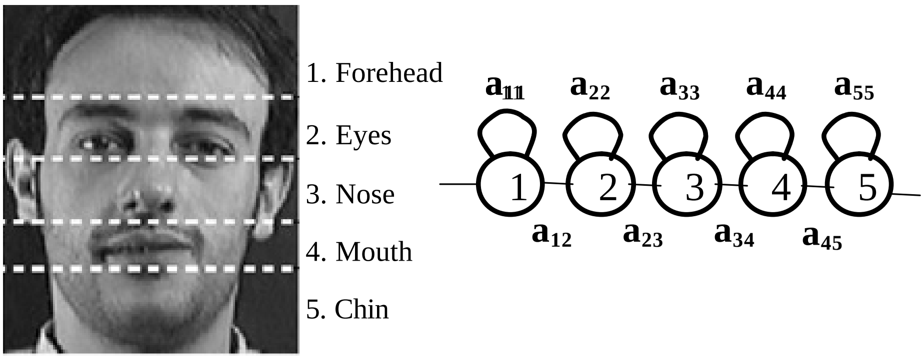


Figure 13.7. A Top-to-Bottom 5 states HMM in [25].

Given a face image for one subject in the training set, the goal of the training stage is to optimise the parameters to best describe the observation. Recognition is carried out by matching the test image against each of the trained models. To complete this procedure, the image is converted to an observation sequence and the likelihood is computed for each stored model. The model with the highest likelihood reveals the identity of the unknown face.

The HMM approach has shown the ability to yield satisfactory recognition rates. However, HMMs are processor intensive models, which implies that the algorithm may run slowly.

Summary. We have discussed several classifiers in this section. It should be noted that, except the graph matching and HMMs-based approaches, the input to other classifiers generally could be any of the models of identity discussed in Section 13.3.1. Therefore, many face recognition systems can be easily setup by combining different models with various classifiers. Such examples include PCA+NN, PCA+LDA, PCA+SVM, PCA+ANN, PCA+Bayesian, AAM+LDA, AAM+SVM, AAM+ANN etc. Among these algorithms, PCA-based Bayesian inference [73], PCA-based LDA [70] and PCA-based SVM [75] have demonstrated impressive performance in the experiments.

13.4. Major Challenges

Robust face recognition is still a challenging problem, even many techniques have been proposed, and some significant progress has been made. Up to now, at least two major challenges need to be emphasized [71], that is, **Illumination Variation Problem** and **Pose Variation Problem**. Either one of the problems can cause serious performance degradation in most of the existing systems. An even more difficult situation would be from the combined problem of pose and illumination variations. Unfortunately, this often happens when face images are acquired in an uncontrolled practical environment such as in the case of surveillance [70].

In addition, occlusion and make-up are other sources of appearance variation. Glasses, especially black-frame glasses or sunglasses, will greatly change the appearance of a face image, not to say the modern make-up techniques such as pasting black beards or other accessories. Since quite little work has been done in this area, this section will mainly concentrate on the solutions to these problems.

13.4.1 Illumination Problem

Solutions to illumination problem include invariant features-based methods, parameterized illumination manifold, photometric alignment, linear illumination subspace, quotient images and illumination cones etc.

Invariant Features-Based Methods. Some image representations are considered to be illumination invariant to some extent. These include edge maps, derivatives of the grey level, images filtered with 2D Gabor-like functions, and a representation that combines a log function of the intensity with these representations [47]. But none of these representations is sufficient to overcome the image variations.

Illumination and Pose Manifolds. Murase and Nayar [17] proposed a continuous and compact representation of object appearance, the ***parametric eigenspace***, which is parameterized by the variables, namely, object pose and illumination.

In this method, an image set of the object is first obtained by varying pose and illumination in small increments. The image set is then normalized in brightness and scaled to achieve invariance to sensor magnification and illumination intensity. The eigenspace for the image set is constructed and all object images (learning samples) are projected onto this space to obtain a set of points. These points lie on a ***manifold*** that is parameterized by pose and illumination and can be constructed from the discrete points using spline interpolation [27].

Recognition, pose and illumination direction estimation can then be achieved as follows: given an image consisting of an interested object, the segmented object region is normalized in scale and brightness such that it has the same size and brightness range as the images used in the learning stage. This normalized image is projected onto the eigenspace. The closest manifold reveals the identity of the object, and exact position of the closest point on the manifold determines pose and illumination direction.

Photometric Alignment. Proposed by Shashua et al. in [43], the basic idea of photometric alignment is to find an algebraic connection between all images of an object taken under varying illumination conditions. In [43], they prove that "an image of an object with an order *k* linear reflection model, $I(p) ? x(p) \cdot ?$, can be represented as a linear combination of a fixed set of *k* images of the object". The Lambertian model of reflection is a typical case of order *3* linear reflectance models, and face surface can be approximated as a Lambertian model. Therefore, assume we take three pictures of a face I_1, I_2, I_3 from light source directions s_1, s_2, and s_3, respectively, then, any new given image *I* of the face, taken from a new given setting of lighting sources, can be simply represented as a linear combination of the three pictures, that is, $I(p) ? ?_1 I_1(p) ? ?_2 I_2(p) ? ?_3 I_3(p)$, for some coefficients $?_1, ?_3, ?_3$. The coefficients can be solved by observing the grey-values of three points. Using more than three points will provide a least squares solution. The solution is unique, offering that *s1, s2,* and *s3* are linearly independent, and the normal directions of the three sampled points span all other surface normalization.

Alignment-based recognition under changing illuminations can then proceed in the following way. Let the images I_1, I_2, I_3 be the model images of the face. For any new given image *I*, rather than matching it directly to previously seen images (the model images), a number of points (at least 3) is first selected to solve the coefficients $?_1, ?_3, ?_3$, and then synthesize an image by $I' ? ?_1 I_1(p) ? ?_2 I_2 ? ?_3 I_3$. If the image I is of the same face, and the only change is in illumination, then I and I' should perfectly match (the matching is not necessarily done at the image intensity level, one can match the edges of I against the edges of I', for example). This procedure has factored out the effects of changing illuminations from the recognition process without recovering scene information, i.e. surface albedo or surface normal, and without assuming knowledge of directions of light sources. Another property of this method is that one can easily find a least squares solution for the reconstruction of the synthesized image, thereby being less sensitive to errors in the model or noise in the input.

Shashua et al. [43] further address the problems that arise when some of the objects points are occluded from some of the light sources, when the surface reflects light specularly, and when spectral composition of light sources is changing.

3D Linear Illumination Subspaces. This method, as a variant of photometric alignment methods, also exploits the observation that, for a Lambertian surface without self-shadowing, the images of a particular face lie in a 3-D linear subspace [51]. For classification, this observation suggests a simple classification algorithm to recognize Lambertian surfaces invariant to different lighting conditions. For each face, use three or more images taken under different lighting directions to construct a 3-D basis for the linear subspace. To perform recognition, we can simply compute the distances between the new image and each linear subspace and choose the face corresponding to the shortest distance. This recognition scheme is called Linear Subspace method.

If there were no noise or self-shadowing, Linear Subspace algorithm would achieve error-free classification under any lighting conditions, provided that the surfaces obey the Lambertian reflectance model. Nevertheless, there are several compelling reasons to look elsewhere. First, due to self-shadowing, specularities and facial expressions, some regions of the face may have variability that does not satisfy the linear subspace model. Second, to recognize a test image, we must measure the distance in the linear subspace for each person's data. While this is an improvement over a correlation scheme that needs a large number of images for each class, it is still computationally expensive. Finally, from the storage point of view, the linear subspace algorithm must keep three images in memory for every person, which is space intensive.

Quotient Images Based Method. More recently, Shashua et al. [81] propose a Quotient images based method to address the problem of "class-based" image-based recognition and rendering with varying illumination.

Their key result is based on the definition of an illumination invariant signature image, which enables an analytic generation of the image space with varying illuminations.

They show that the set of all images, generated by varying lighting conditions on a collection of Lambertian objects that have the same shape but different surface albedoes, can be characterized analytically using images of a prototype object and an illumination invariant "signature" image per object of the class. The Cartesian product between the signature image of an object y and the linear subspace determined by the images of the prototype object generates the image space of y. They also show how to obtain the signature image from a database of example images of several objects, and prove that the signature image obtained is invariant to illumination conditions.

The method works remarkably well on real face images using a very small set of example objects, as few as two example objects. In many cases, the re-rendering results are indistinguishable from the "real" objects, and the recognition results outperform conventional methods by far.

Illumination Cones. In the last few years, Belhumeur and Kriegman et al. [52, 54, 57, 82] have proposed a generative appearance-based method, named illumination cones, for recognizing human faces under variations of lighting and viewpoint. Their work is well summarized in [82].

Belhumeur et al. [52] first prove that the set of images of an object in a fixed pose, seen under all possible illumination conditions, is a convex cone in the space of images. Particularly, the illumination cone of a convex object with Lambertian reflectance can be completely determined by several properly chosen images. Although faces are neither Lambertian surfaces nor convex, experimental results show that the illumination cone of a face can be also established from a few images acquired under different lighting conditions. To construct the illumination cone of a face, its shape and albedo should be recovered first. Georghiades et al. [47] use seven images of a face in a fixed pose, but under different and unknown lighting conditions, to reconstruct its surface geometry and albedo map. In turn, this reconstruction serves as a generative model to render or synthesize images of the face under new given poses and illumination conditions. The pose space is then sampled and, for each pose, the corresponding illumination cone is approximated by a low-dimensional linear subspace whose basis vectors are estimated using the generative model.

Once the illumination cone of a specific face is constructed, recognition can be achieved by assigning to a test image the identity of the closest approximated illumination cone, based on Euclidean distance within the image space. Because illumination cones represent the whole image set of an object under all possible configurations of point light sources at infinity, nearly perfect recognition rates can be still achieved even under extreme illumination conditions.

13.4.2 Pose Problem

Various methods have been proposed to handle the pose problem. Basically, these methods can be divided into three categories: 1) multi-view based methods when multiple images per person are available, 2) hybrid methods when multiple training images are available for training but only one test image per person is available for recognition, and 3) single image based methods when no training is carried out. Up to now, the second type of approach is most popular [70].

Multi-View Based Methods. One of the earliest efforts in multi-view based methods was from Beymer in [28], where a multi-view component template-based correlation-matching scheme was proposed. In his work, pose estimation and face recognition are coupled in an iterative loop. For each hypothesized pose, the input image is aligned to the database images with a selected pose. The alignment is first carried out through 2D affine transformation based on three key feature points (eyes and nose), and then optical flow is used to refine the alignment of each template. The correlation scores of all pairs of matching templates are used to perform recognition.

View-based Eigenface [19] is another approach of the multi-view strategy. It explicitly encodes the pose information by constructing an individual eigenspace for each pose, and uses these pose eigenspaces for a given image to estimate the pose of the face and thus recognizes in a subspace specific to the estimated pose.

Linear Object Classes. The work described in [44, 45] by Vetter and Poggio is based on the idea of linear object classes. Linear object classes are 3D objects whose 3D shape can be represented as a linear combination of a sufficiently small number of prototypical objects. They have the property that new orthographic views of any object of the class under uniform affine 3D transformations, and in particular rigid transformations in 3D, can be generated exactly if the corresponding transformed views are known for the set of prototypes. Based on this property, it is possible to "learn" a direct mapping from standard pose to a particular virtual pose as follows. Using multiple prototype objects, first de-compose the new given face as a linear combination of prototypes at the standard pose, yielding a set of linear prototype coefficients. Then, by taking the linear combination of prototype objects at the virtual pose using the same set of coefficients, the new given face at the virtual pose can be synthesized.

This approach accomplishes more than just object recognition tasks. It can provide additional artificial example images of an object when only a single image is given. On the other hand, the coefficients, which result from a decomposition of shape and texture into example shapes and textures, provide a representation of the object that is invariant to any affine transformations.

In [58], Vetter et al. further extend the Linear Object Class to combine the prior knowledge of 3D shape by introducing a generic 3D model of human head. Example images are used to "learn" a pose-invariant shape and texture description of a new face. And 3D model is used to solve the correspondence problem between images with faces in different poses.

Parallel Deformation. Parallel deformation is another example-based technique to represent prior knowledge by using 2D example views of prototype faces under different poses and apply the rotation seen in the prototypes to "rotate" the given single real view [29, 30].

It works as follows: using only one prototype object, measure the 2D deformation of object features from the standard to virtual view. Then map this 2D deformation onto the new given object and use the deformation to distort, or warp the new given image from the standard pose to the virtual one.

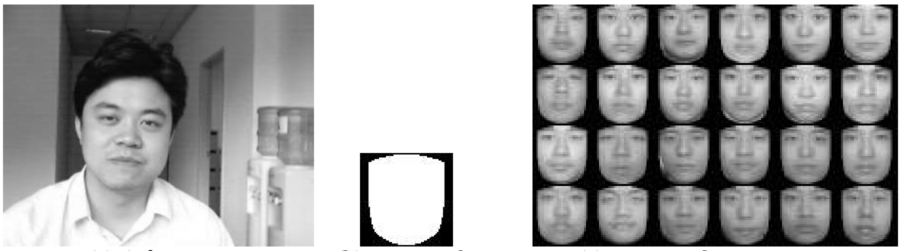
3D Model. Generic 3D models of the human face can be used to predict the appearance of a face under different poses, expressions and lighting conditions. 3D face shape is represented either by a polygonal model or by a more complicated multi-layer mesh that simulates tissue. Once a 2D face image is texture mapped onto the 3D model, the face can be considered as a traditional 3D object in computer graphics subjecting 3D rotations or changes in light source positions. Faces are texture mapped onto 3D model either by specifying corresponding facial features in both the image and 3D model or by recording both 3D depth and color image data simultaneously by using specialized equipment like the Cyberware scanner.

A generic 3D model could then be applied to the scenario of pose-invariant face recognition from one example view. The single view of each person could be texture mapped onto a 3D model, and then the 3D model could be rotated to novel poses to generate multiple virtual views for recognition purpose.

13.5. An Example Access Control System Based on Eigenface

This section will describe an example access control system based on Eigenface method, which, as mentioned above, is one of the most popular face recognition methods. In Eigenface method, a training set containing enough number of face examples are needed. The training set is also expected to cover all kinds of variations due to different lighting conditions, slight facial expressions and head poses.

In addition, to alleviate the influence of translation, ration, lighting and scale variance, geometric and grey normalization should be adopted. As to geometric normalization, generally, the two irises are fixed at specific locations using affine transformation. And a mask, as shown in Figure 13.8 (b), is covered over the face region to eliminate the alterable background and hairstyle. Finally, all faces are warped to a fixed size as shown in Figure 13.8 (c). Histogram equalization is also conducted to normalize illuminations, and all face data are vectorized to the unit length before they are fed into training or testing procedure.



(a) A face image (b) Face mask (c) Faces in the training set

Figure 13.8. Normalization and training set.

Given such a training set containing m training samples denoted as $\{f_1, f_2, \cdots, f_m\}$, Eigenfaces are learned as follows:

First, compute the covariance matrix of the training set as:

$$C \; ? \; \underset{i}{?} \; (f_i \; ? \; \bar{f})(f_i \; ? \; \bar{f})^T \, .$$

C is then decomposed by SVD as:

$C \; ? \; UDU^T$, or $C?_i \; ? \; ?_i ?_i$ for $i=1,2\ldots, m$ and $?_1 \; ? \; ?_2 \; ? \cdots \; ? \; ?_m$,

where $U \; ? \; [?_1, ?_2, \cdots, ?_m]$ and $D \; ? \begin{bmatrix} ?_1 & 0 & 0 \\ 0 & ?_i & 0 \\ 0 & 0 & ?_m \end{bmatrix}$.

In Eigenface method, $?_1, ?_2, \square\ , ?_m$ are called “Eigenfaces” because of their visual similarities to the face pattern. (In fact, it can be shown mathematically that each of them is the linear combination of the training images, which explains the similarity.) Some Eigenfaces learned from a training set are illustrated in Figure 13.9.

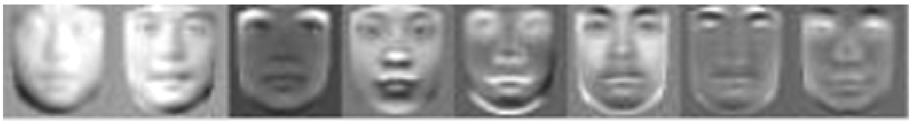


Figure 13.9. Leading Eigenfaces learnt from a training set of 350 faces.

Then, any input face f can be represented as the linear combination of these Eigenfacesby:

$$W \; ? \; U^T (f \; ? \; \bar{f}) .$$

In turn, f can be reconstructed by the inverse procedure as:

$$f \, ?? \; \bar{f} \; ? \, UW .$$

This procedure is visually illustrated in Figure 13.10.

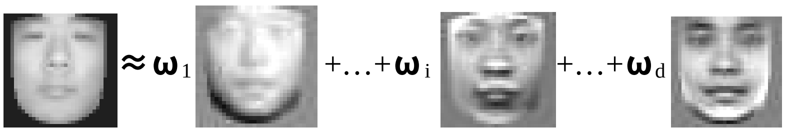


Figure.13.10. One face image is represented as the linear combination of *m* leading Eigenfaces.

The vector W computed in the equation described above is generally used as the feature of the input face f. Any classifier can be conducted on these features to achieve face recognition. Commonly, the similarity between two Eigenface features W_1 and W_2 can be defined as:

$$S(W_1, W_2) \; ? \; \frac{? \; W_1, W_2 \; ?}{\| W_1 \| ? \| W_2 \|} ,$$

where “$<W_1 \, , \, W_2>$” denotes the dot product of the two vectors, and “$\|.\|$” is the L_2 of the vector.

Finally, to determine whether the two Eigen-features W1and W2 are from the same face or not, a threshold T_k for each face should be learnt. This can be done by a simple Bayesian minimum probability of error rule, as illustrated in Figure 13.11.

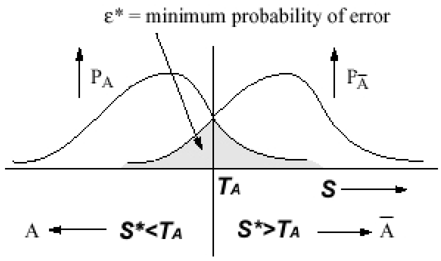


Figure 13.11. Determining threshold to Bayesian minimum probability of error rule.

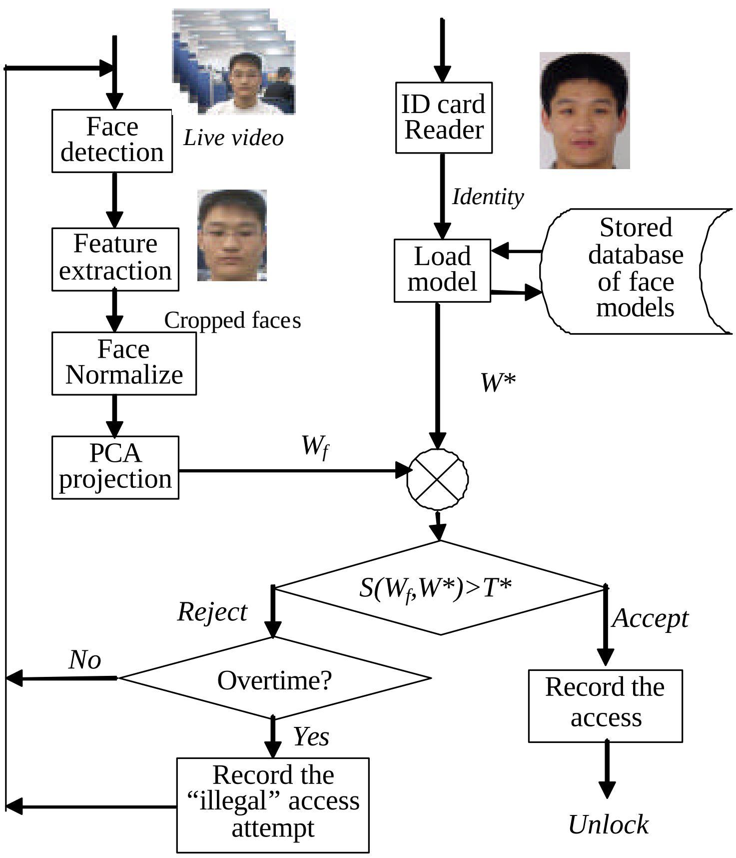


Figure. 13.12. Face verification for access control based on Eigenface method.

The architecture of the face verification system for access control based on Eigenface method is shown in Figure 13.12.

Note that, in the figure, the stored database of known face models is commonly trained offline. In addition, the records should be saved in the case that someone has passed the verification or cannot be verified in the given time. This would be quite useful for the later reference.

13.6. Commercial Systems and Performance Evaluation

Since the middle of the 1990s, some commercial face recognition products have emerged based on the techniques mentioned above. Along with the emergence of these commercial systems, how to evaluate these systems is attracting more and more attentions from researchers. In this section, we will present several commercial face recognition systems and a few performance evaluation benchmarks.

13.6.1 Existing Commercial Systems

FaceIt® is developed by Visionics Corporation. It is based on the Local Feature Analysis (LFA) method proposed by Atick [35] in Rockfeller University. It achieves excellent performance in the first FERET testing and is also highly ranked in the recent FRVT2000 testing.

FaceFINDER™ is from Viisage Technology Inc. Viisage developed a series of face recognition systems including FaceFINDER, FaceEXPLORER, FacePIN, FacePASS, FaceTOOLS etc. These systems are based on the well-known Eigenface method. *FaceFINDER*™ has recently been selected to deploy the first face-recognition technology system for security in a U.S. airport after the US 9.11 terrorism attacks.

Hunter™ is a facial recognition surveillance system from LAU Technologies. It is also based on the MIT Eigenface techniques. It participated the FRVT2000 evaluation.

FaceSnap® RECORDER is a turnkey solution for video surveillance, monitoring and law enforcement developed by C-VIS Computer Vision and Automation Gmbh. The kernel techniques are Elastic Graph Matching [46], which has been recognized as one of the most promising technologies in FERET testing [71].

TrueFace is developed by the eTrue Inc. and uses neural network technology. It has found several partners in e-commerce industry including Microsoft. The eTrue service has been chosen as the first biometric authentication service in Microsoft .NET enterprise servers.

In addition to the systems mentioned above, there are many other face recognition commercial systems available in the Biometrics market, including **SpotIt!** from ITC-irst (initialized by Roberto Brunelli), **Banque-**Tec International, **FaceVACS** from Cognitec AG, **BioID** (a system adopting sensor-fusion approach using face, speech and lip movement analysis by DCS AG, Germany) and ZN-Face etc.

13.6.2 Performance Evaluation

Given the numerous theories and techniques, as well as commercial systems that are applicable to face recognition, it is clear that evaluation of these algorithms is crucial. Obviously, large sets of test images are essential for adequate evaluation, while it is also extremely important that the sample be statistically as similar as possible to the images that arise in the application being considered. Scoring should be done in a way that reflects the costs or other system requirements.

During the past decade, some publicly available large face databases have been collected and corresponding evaluation protocols have been designed. Among them, in US, the series of FERET evaluations [71] have attracted many institutions and companies to participate. However, in Europe, (X)M2VTS evaluation is more attractive.

FERET. To facilitate the evaluation of various algorithms in a set-up very close to a real-world setting, and to identify potential problems that have not revealed in the researchers' small-scale tests, the FERET program was initiated, emphasizing the evaluation of FRT algorithms. Under the FERET program, a large number of face images were collected, and testing procedures were established. Up to date, 14,126 images from 1199 individuals are included in the FERET database [71].

Overall, three algorithms perform very well: Elastic Graph Matching from USC [46], Probabilistic Eigenface from MIT [56] and Subspace LDA from UMD [62]. On frontal images taken the same day, typical first-choice recognition accuracy is over 95%. For images taken with a different camera under various lighting conditions, typical accuracy drops to 80%~90%. For images taken one year later, the typical accuracy is approximately 50% [71, 72].

(X)M2VTS. The M2VTS project (Multi Modal Verification for Televices and Security Applications) deals with access control using multi-modal identification of human faces among different European ACTS (Advanced Communications Technologies & Services) projects. The M2VTS database contains 37 subjects and 5 shots for each person. The XM2VTS [69] multi-modal database is an expansion of the original M2VTS multi-modal database.

The XM2VTS database contains four recordings of 295 subjects taken over a period of four months. Each recording contains a speaking headshot and a rotating headshot. Sets of data taken from this database are available, including high-quality color images, 32 KHz 16-bit sound files, video sequences and a 3D model at production cost.

FRVT2000. The DoD Counter-drug Technology Development Program Office initialized the FRVT2000 in May and June 2000 to evaluate the performance of the commercial systems available in U.S. biometrics market. Besides the testing inherited from the FERET, FRVT2000 has further tested the performance of different system on the eight variations: compression, distance, expression, illumination, media, pose, resolution and temporal. Furthermore, FRVT2000 conducts a product usability test: access control with live subjects.

The results of the FRVT 2000 show that progress has been made in temporal changes, but developing algorithms that can handle temporal variations is still a necessary research area. In addition, developing algorithms that can compensate for pose variations, illuminations and distance changes are noted as other areas for future researches.

13.7. Summary and Conclusions

In this chapter, we survey the up-to-date technologies of automatic face verification for access control. We have focused on models of identify, features extraction and classification of the face verification problem. Major challenges and their corresponding solutions are discussed. An example access control system based on Eigenface is presented. Some commercial systems available in the industry market are introduced briefly along with the introductions of several famous evaluation projects on face recognition. As a summary based on the previous discussion, we've come to the following conclusions:

- Geometric feature based methods and template matching methods used to be popular technologies, in 1990s, ***appearance*** based technologies have become the dominant methods. While much recently, ***photometry*** based approaches have attracted much attention for their ability to deal with illumination and pose problems;
- Among these diverse techniques, those based on ***Eigenface*** have been recognized as the most popular and successful techniques. Eigenface based LDA, SVM and Bayesian inference are all among the techniques being and to be further studied. Recently, SVM has again provided a good development opportunity for Eigenface;

- The most promising techniques should perfectly combine the 2D shape, local features and holistic appearance features. Both ***Active Appearance Models*** and ***Elastic Graph Matching*** have achieved great success and attracted more and more attention.
- The state-of-the-art of Face Recognition Technology is still far from the public expectations, mainly due to the ***pose and illumination varying problems***. The most possible solutions to those open issues include illumination cone, linear subspace, illumination/pose manifolds, linear object class, photometric stereo, shape-from-shading, 3D models and Quotient image based methods.
- Industry calls for the face verification systems. Several face recognition ***commercial systems*** have been available in the Biometrics industry. Latest improvement in face recognition is expected to provide the industry more robust and reliable commercial systems.

Acknowledgment

This work is supported by National Science Foundation of China, National Hi-Tech Development Programm of China, 100 Outstanding Scientists foundation of Chinese Academy of Sciences, and YCNC Co. Ltd. The authors want to express thanks to these organizations.

Also, the authors would like to express their gratitude to Dr. Xilin Chen, Dr. Debin Zhao, Mr. Yan Liu, Ms. Hongxun Yao, Mr. Jun Miao, Ms. Ying Liu, Ms. Wenjie Sun, Dr. Jie Yan, Mr. Bo Cao, Mr. Pin Liao, Ms. Xiujuan Chai, Mr. Wei Wang, Mr. Guoqin Cui, Mr. Gang Deng, Mr. Feng Jiao, Ms. Wenjuan Lian and Mr. Hongming Zhang for their contributions on technical discussions, system implementation, materials preparation and revision for this chapter.

References

[1] M. Kass, A. Witkin and D. Terzopoulos. Snakes: Active contour models. Int. Journal of Computer Vision, pp.321-331, 1988.

[2] M. Kirby and L. Sirovich. Application of the Karhunen-Loeve procedure for the characterization of human faces, IEEE Trans. on PAMI, 12(1), 103-108, 1990.

[3] Zi-Quan Hong. Algebraic feature extraction of image for recognition. Pattern Recognition, 24(3), pp211-219, 1991.

[4] Osamu Nakamura, Shailendra Mathur, and Toshi Minami. Identification of human faces based on isodensity maps. Pattern Recognition, 24(3), pp263-272, 1991.

[5] M. Turk and A. Pentland. Eigen-faces for Recognition Journal of cognitive neuroscience, 3(1), pp71-86, 1991.

[6] A.L. Yuille. Deformable templates for face detection, J. Cog. Neu. 3, 59-70, 1991.

[7] A. Shashua. Geometry and Photometry in 3D Visual Recognition, PhD thesis, Massachusetts Inst. of Technology, Cambridge, 1992.

[8] R. Brunelli and T. Poggio, Face Recognition through Geometrical Features, Proceedings of ECCV '92, pp792-800, 1992.

[9] R.Brunelli and T. Poggio. HyperBF networks for gender classification, in Pro. DARPA Image Understanding Workshop, pp.311-314, 1992.

[10] B.S. Manjunath, R. Chellappa and C.V.D. Malsburg. A Feature based approach to face recognition, in Pro. IEEE Computer Soc. Con. On Computer Vision and Pattern Recognition, pp373-378, 1992.

[11] A.L. Yuille, P.W. Hallinan and D.S. Cohen. Feature extraction from faces using deformable templates. International Journal of Computer Vision, 8, 99-111. 1992.

[12] Ashok Samal and A. Prasana. IyenGar Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey, Pattern Recognition, 25(1), pp65-77, 1992.

[13] Martin Lades, Jan C. Vorbruggen, J. Buhmann, J. Lange, C.V.D. Malsburg, R.P. Wurtz and W. Konen. Distortion Invariant Object Recognition in the Dynamic Link Architecture, IEEE Trans. On Computers, 42(3), pp 300-311, 1993.

[14] R. Brunelli and T. Poggio. Face recognition: features vs. templates. IEEE Trans. on PAMI, 15(10): 1042-1052, 1993.

[15] G. Chow and X.B Li. Towards a system for automatic facial feature detection. Pattern Recognition, 26(12), pp1739-1755. 1993.

[16] Ke Liu, Yongqing Cheng and Jingyu Yang. Algebraic feature extraction for image recognition based on an optimal discriminant criterion. Pattern Recognition, 26(6): 903~911, 1993.

[17] S.K. Nayar and H. Murase. "Dimensionality Of Illumination Manifold In Eigenspace", TR-CUCS-021-94, Columbia University, 1994.

[18] Peter W. Hallinan. A low-dimensional representation of human faces for arbitrary lighting conditions. Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, Seattle, WA, 995~998, 1994.

[19] Alex Pentland, Baback Moghaddam and Thad starner. View-Based and Modular Eigenspaces for Face Recognition, IEEE Conference on Computer Vision & Pattern Recognition, 1994.

[20] G.Z. Yang and T.S. Huang, Human face detection in a complex background, Pattern Recognition, 27(1), 43-63, 1994.

[21] X. Xie, R. Sudhakar and H. Zhuang. Improving eye feature-extraction using deformable templates. Pattern Recognition, 27(6), 791-799, 1994.

[22] A. Lanitis, C.J. Taylor, and T.F. Cootes. An automatic face identification system using flexible appearance models. In British Machine Vision Conference, BMVA Press, 1: 65~74. 1994.

[23] D. Valentin, H. Abdi, A.J.O. Toole and Garrison W.Cottrell. Connectionist Models of Face Processing: A survey, Pattern Recognition. Vol.27, pp1209-1230, 1994.

[24] S.R. Gunn and M.S. Nixon. A Dual Active Contour. BMVC 94, September, York, U.K. 305-314, 1994.

[25] F.S. Samaria. Face Recognition Using Hidden Markov Models. PhD thesis, Trinity College, University of Cambridge, Cambridge, 1994.

[26] Rama Chellappa, Charles L. Wilson and Saad Sirohey. Human and Machine Recognition of faces: A survey, Proceedings of the IEEE, vol.83, no.5, pp. 705-740, 1995.5.

[27] H. Murase and S. Nayar. Visual Learning and recognition of 3D object from appearance, Int. J. Computer Vision, 14:5-24, 1995.

[28] D. Beymer and T. Poggio. Face Recognition from One Example View, Proc. Int'l Conf. Computer Vision, pp. 500-507, 1995.

[29] David Beymer. Vectorizing Face Images by Interleaving Shape and Texture Computation, A.I.Memo No. 1537, 1995.9.

[30] David Beymer and Tomaso Poggio. Image Representations for Visual Learning, Science, Vol.272, pp1905-1909, 1995.

[31] R. Epstein, P. Hallinan, and A. Yuille. 5+/-2 Eigenimages Suffice: An Empirical Investigation of Low-Dimensional Lighting Models, Proc. Physics-Based Modeling Workshop in Computer Vision, session 4, 1995.

[32] P. Hallinan, A Deformable Model for Face Recognition under Arbitrary Lighting Conditions, PhD thesis, Harvard Univ., 1995.

[33] T. Leung, M. Burl and P. Perona. Finding Faces in Cluttered Scenes Using Labeled Random Graph Matching, Proc. Int'l Conf.Computer Vision, pp. 637-644, 1995.

[34] B. Moghaddam and A. Pentland. Probabilistic Visual Learning for Object Detection, Proc. Int'l Conf. Computer Vision, pp. 786-793, 1995.

[35] P.Penev and J.Atick. "Local Feature Analysis: A General Statistical Theory for Object Representation," Network: Computation in Neural Systems, vol.7, pp.477-500, 1996.

[36] Liu Mingbao and Gao Wen. A Hierarchical Approach To Human Face Detection In a Complex Background. the First International Conference on Multimodal Interface, Beijing, 1996.

[37] Babak Moghaddam, Chahab Nastar and Alex Pentland. A Bayesian Similarity Measure for Direct Image Matching, International Conference on Pattern Recognition, Vienna, Austria, August 1996.

[38] C.H. Lee, J.S. Kim and K.H. Park. Automatic human face location in complex background using motion and color information, Pattern Recognition, 29(11), 1877-1889, 1996.

[40] Daniel L. Swets and John Weng. Using discriminant Eigenfeatures for image retrieval. IEEE Trans. on PAMI, 18(8): 831~836, 1996.

[41] S. Nayar and H. Murase. Dimensionality of Illumination Manifolds in Appearance Matching, Proc. Int'l Workshop Object Representations for Computer Vision, p. 165, 1996.

[42] E. Osuna, R. Freund and F. Girosi. Training support vector machines: an application to face detection, Proc. CVPR, 130-136, 1997.

[43] A. Yuille and D. Snow. Shape and Albedo from Multiple Images Using Integrability, Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 158-164, 1997.

[44] A.Shashua. On Photometric Issues in 3D visual recognition from a single 2D Image, International Journal of Computer Vision, 21(1/2), 99-122, 1997.

[45] Thomas Vetter and Tomaso Poggio. Linear Object Classes and Image Synthesis from a Single Example Image, IEEE Trans. On PAMI, Vol.19, No. 7, pp733-742, 1997.

[46] Thomas Vetter, Michael J. Jones and Tomaso Poggio. A Bootstrapping Algorithm for Learning Linear Models of Object Classes, IEEE CVPR, pp40-46, Puerto Risco, USA, 1997.

[47] Laurenz Wiskott, Jean Marc Fellous, Norbert Kruger and Christoph von der Malsburg. Face Recogniton by Elastic Bunch Graph Matching, IEEE Trans. on PAMI, Vol.19, No. 7, pp775-779, 1997.

[48] Yael Adini, Yael Moses and Shimon Ullman. Face Recognition: The Problem of Compensting for changes in illumination Direction, IEEE Trans. on PAMI, Vol.19, No.7, pp721-732, 1997.

[49] Steve Lawrence, Lee Giles, Ah Chung Tsoi and Andrew D. Back. Face Recognition: A Convolutional Neural-Network Approach, IEEE Trans. on Neural Network, Vol.8, No.1, Jun. 1997.

[50] Jun Zhang, Yong Yan and Martin Lades. “Face Recognition: Eigenface, Elastic Matching and Neural Nets” , Proceedings of the IEEE, vol.85, no. 9, pp1422~1435, Sep. 1997.

[51] K.M.Lam and H.Yan. An Analytic-to-Holistic Approach for Face Recognition Based on a single Frontal view” , IEEE Trans. on PAMI, vol.20, no.7, July, 1997.

[52] P.N.Belhumeur, J.P.Hespanha and etc. Eigenfaces vs Fisherfaces: recognition using class specific linear projection. IEEE Trans. on PAMI, vol.20, no.7, July, 1997.

[53] P. Belhumeur, D. Kriegman, and A. Yuille. The Bas-Relief Ambiguity, Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1040-1046, 1997.

[54] A. Lanitis, C.J. Taylor and T.F. Cootes. Automatic interpretation and coding of face images using flexible models. IEEE Transactions on Pattern Analysis & Machine Intelligence, vol.19, no.7, pp.743-56, July 1997.

[55] P. Belhumeur and D. Kriegman. What Is the Set of Images of an Object under All Possible Illumination Conditions, Int'l J. Computer Vision, vol. 28, no. 3, pp. 245-260, July 1998.

[56] T.F. Cootes, G.J. Edwards and C.J. Taylor. Active Appearance Models, Proc. European Conf. Computer Vision, vol. 2, pp. 484-498, 1998.

[57] Baback Moghaddam, Wasiuddin Wahid and Alex Pentland. Beyond Eigenfaces: Probabilistic Matching for Face Recognition, the 3rd IEE Int. Con. On Auto. Face- and Gesture- Recognition, Nara, Japan, 1998.4.

[58] A.S. Georghiades, D.J. Kriegman and P.N. Belhumeur. “Illumination Cones For Recognition Under Variable Lighting: Faces”. Proc. of IEEE CVPR, pp52-58, 1998.

[59] Thomas Vetter, Synthesis of Novel Views from a Single Face Image, IJCV, 28(2), pp103-116, 1998.

[60] H.A. Rowley, S. Baluja, and T. Kanade. Neural network-based face detection, IEEE Trans. on PAMI, 20(1), 23-38, 1998.

[61] K.K. Sung and T. Poggio. Example-based learning for view-based human face detection, IEEE Trans. on PAMI, 20(1), 39-50, 1998.

[62] M. Pontil and A. Verri. Support Vector Machines for 3D Object Recognition, IEEE Trans. On PAMI, 20: 637-646, 1998.

[63] W. Zhao, R. Chellappa and A.Krishnaswamy. "Discriminant Analysis of Principal Components for Face Recognition", Proc. of Inter. Conf. On Auto. Face and Gesture Recognition, pp.336-341, 1998..

[64] D.W.Jacobs, P.N. Belhumeur and R. Basri. Comparing Images under Variable Illumination,in Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, pp. 610-617, 1998.

[65] Ruo Zhang, Ping-Sing Tai, James Edwin Cryer and Mubarak Sha. Shape from Shading: A Survey, IEEE Trans. On PAMI, 21(8). pp690-706, 1999.

[66] G. Edwards, T. Cootes, and C. Taylor. Advances in Active Appearance Models, Proc. Int'l Conf. Computer Vision, pp. 137-142, 1999.

[67] Volker Blanz and Thomas, Vetter. A Morphable Model For the Synthesis of 3D Faces, SIG'GRAPH'99, 1999.

[68] J. Miao, B.C. Yin and K.Q. Wang, et al. A hierarchical multiscale and multiangle system for human face detection in a complex background using gravity-center template, Pattern Recognitio n, 32(7), 1999.

[69] K. Jonsson, J. Matas, and J. Kittler. Learning salient features for real-time face verification. In S. Akunuri and C. Kull-man,editors, AVBPA'99, pages 60–65, 1999.

[70] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. XM2VTSDB: The Extended M2VTS Database, in Proceedings, International Conference on Audio- and Video-Based Person Authentication, pp. 72-77, 1999.

[71] W. Zhao and R. Chellappa. "Robust Image-Based 3D Face Recognition", CAR-TR-932, N00014-95-1-0521, CS-TR-4091, Center for Auto Research, UMD, 2000.1.

[72] P.J. Phillips, H. Moon and etc. "The FERET Evaluation Methodology for Face-Recognition Algorithms", IEEE Trans. on PAMI, Vol.22, No.10, pp1090-1104, 2000.

[73] Alex Pentland. Looking at People: Sensing for Ubiquitous and Wearable Computing, IEEE Trans. On PAMI, Vol.22, No.1, pp107-119, Jan. 2000.

[74] Baback Moghaddam, Tony Jebara and Alex Pentland. Bayesian Face Recognition, Pattern Recognition Vol.33, pp1771-1782, 2000.

[75] Constantine L. Kotropoulos and Anastasios Tefas. Ioannis Pitas Frontal Face Authentication Using Discriminating Grids with Morphological Feature Vectors, IEEE trans. On Multimedia, Vol.2, No.1, pp14-26 March, 2000.

[76] G. Guo, S. Z.Li and K. Chan. "Face Recognition by Support Vector Machines", Proc. of the 4th Int. Conf. on Auto. Face and Gesture Recognition. pp.196-201, Grenoble, 2000.3.

[77] K. Jonsson, J. Matas, J. Kittler and Y.P. Li. Learning Support Vectors for Face Verification and Recognition, Proceeding of the 4th International Conference on Face and Gesture Recognition, pp208-213, Grenoble, France, 2000.3.

[78] T.F. Cootes, K. Walker and C.J. Taylor. View-based Active Appearance Models, Proceeding of the 4th International Conference on Face and Gesture Recognition, pp227-232, Grenoble, France, 2000.3.

[79] Wenyi Zhao and Rama Chellappa. SFS Based View Synthesis for Robust Face Recognition, Proceeding of the 4th International Conference on Face and Gesture Recognition, pp285-292, Grenoble, France, 2000.3.

[80] Baback Moghaddam and Ming-Hsuan Yang. Gender Classification with Support Vector Machine, Proceeding of the 4th International Conference on Face and Gesture Recognition, pp306-311, Grenoble, France, 2000.3.

[81] Shiguang Shan, Wen Gao, Jie Yan and etc. Individual 3D face synthesis based on orthogonal photos and speech-driven facial animation. Proceedings of the International Conference on Image Processing (ICIP'2000), Vol. III, pp238-242, Vancouver, BC, Canada, 2000.9.

[81] A. Shashua and T. Riklin-Raviv. "The Quotient Image: Class-Based Re-Rendering And Recognition With Varying Illuminations" , IEEE Trans. on PAMI, 23(2): pp129-139, 2001.2.

[82] Athinodoros S. Georghiades, Peter N. Belhumeur and David J. Kriegman, From Few to Many: Illumination Cone Models for Face Recognition Under Variable Lighting and Pose, IEEE Trans. on PAMI, 23(6) pp.643660-139, 2001.6.

[83] M.H. Yang, N. Ahuja, D. Kriegman. Detecting faces in images: a survey, IEEE Trans. on PAMI, Vol.24, No.1, 2002.1

[84] Hongxun Yao, Wen Gao and etc. Face Detection and Location Based on Skin Chrominance and Lop Chrominance Transformation from Color Images Pattern Recognition, 2000.10.

[85] Stan Z. Li and Juwei Lu. Face Recognition Using the Nearest Feature Line Method, IEEE Transactions On Neural Networks, 10(2), PP.439-443, MARCH, 1999.

[86] Shiguang Shan, Wen Gao, Xilin Chen and Jiyong Ma. Novel Face Recognition Based on Individual Eigen-subspaces, Proceedings of International Conference on Signal Processing, ICSP2000, Vol. III, pp1522-1525, Beijing, China, 2000.

[87] Shiguang Shan, Wen Gao and etc. A Face-Unlock Screen Saver by Using Face Verification Based on Identity-Specific Subspaces, Proceeding of the Pacific-rim Conference on Multimedia (PCM'01), Beijing, China, pp1096-1101, 2001.

[88] Shiguang Shan and Wen Gao. Unified Framework For Classifying Facial Images Based On Facial Attribute-Specific Subspaces And Minimum Reconstruction Error, Proc. of ACCV'2002, Vol. II, pp858-863, Melbourne, Australia, Jan. 2002.

[89] Shiguang Shan, Wen Gao and Debin Zhao. Face Identification From A Single Example Image Based On Face-Specific Subspace (FSS), to appear in proc. of ICASSP2002, Orlando, Florida, USA, May 2002.