Computing and Informatics, Vol. 22, 2003, ??–??

# A SURVEY OF FACE DETECTION, EXTRACTION AND RECOGNITION

Yongzhong Lu, Jingli Zhou, Shengsheng Yu

*National Storage System Laboratory*
*School of Software Engineering*
*Huazhong University of Science and Technology*
*Wuhan, 430074, P. R. China*
*e-mail:* luyongz0@sohu.com

Manuscript received 23 June 2002; revised 27 January 2003
Communicated by Ladislav Hluchý

**Abstract.** The goal of this paper is to present a critical survey of existing literatures on human face recognition over the last 4–5 years. Interest and research activities in face recognition have increased significantly over the past few years, especially after the American airliner tragedy on September 11 in 2001. While this growth largely is driven by growing application demands, such as static matching of controlled photographs as in mug shots matching, credit card verification to surveillance video images, identification for law enforcement and authentication for banking and security system access, advances in signal analysis techniques, such as wavelets and neural networks, are also important catalysts. As the number of proposed techniques increases, survey and evaluation becomes important.

**Keywords:** Face recognition, eigenface, elastic matching, neural networks, pattern recognition

## 1 INTRODUCTION

Face recognition is becoming an active research area spanning several disciplines such as image processing, pattern recognition, computer vision, neural networks, cognitive science, neuroscience, psychology and physiology. It is a dedicated process, not merely an application of the general object recognition process. It is also the representation of the most splendid capacities of human vision.

As one of the most successful applications of pattern recognition, image analysis and understanding, face recognition has recently received significant attention, especially during the past several years. This is evidenced by the emergence of face recognition conferences such as AFGR [1], AVBPA [2], CV [3], PR [4], CVPR [5], WACV [6], CGIPV [7], WCVBVSMA [8], CCVHM [9], SSWNN [10], IP [11], and systematic empirical evaluations of face recognition technology (FRT), including the FERET [12–20] and XM2VTS [21, 22] protocols. There are at least two reasons for this trend: the first is the wide range of commercial and law enforcement applications, and the second is the availability of feasible technologies after 30 years of research.

The previous literatures on systematic empirical evaluations of face recognition are chiefly the earlier surveys by Samal & Iyengar 92 [23] on nonconnectionist approaches, by Valentin et al. in 94 [24] on connectionist schemes, the abundant and thorough survey by Chellappa et al. in 1995 [25] on 20 years of face recognition, the longer and more comprehensive surveys by Fromherz in 1997 [26], and by Zhou et al. in 2000 [27]. They focused on the development of face recognition before the mid 1997. During the past several years, face recognition has received increased attention and has advanced technically. Many commercial systems [28–31] using face recognition are now available. Significant research efforts have been focused on video-based face modeling, processing and recognition.

This review paper reports on research results mostly published in the mid 1997 and later, thus complementing the earlier surveys by Samal & Iyengar 92, Valentin et al. in 94, Chellappa et al., Fromherz in 1997, and Zhou et al. in 2000.

In this paper we provide a critical review of the most recent development in face recognition. This paper is organized as follows: In Section 2 we briefly review issues that are relevant to face recognition system. Section 3 provides a detailed review of the development in face detection and localization techniques using grayscale, rang and other images. In Section 4 and 5 face feature extraction and recognition techniques is detailed and discussed, respectively. Finally, summary and conclusions are in Section 6.

## 2 FACE RECOGNITION SYSTEMS

In general, an automatic face recognition systems are comprised of three steps. Their basic flowchart is given in Figure 1. Among them, detection may include face edge detection, segmentation and localization, namely obtaining a pre-processed intensity face image from an input scene, either simple or cluttered, locating its position and segmenting the image out of the background. Feature extraction may denote the acquirement of the image features from the image such as visual features, statistical pixel features, transform coefficient features, and algebraic features, with emphasis on the algebraic features, which represent the intrinsic attributes of an image. Face recognition may represent to perform the classification to the above image features in terms of a certain criterion. Segmentation among three steps is considered to be trivial, easy and simple for many applications such as mug shots, drivers licenses, personal ID card, and passport pictures. Thus this problem did not receive much attention. Scholars have given more interest on addressing other problems. However, recently more effort is devoted to the segmentation problem with the advancement of face recognition systems under complex background.


Fig. 1. The basic flowchart of a face recognition

In face recognition systems, it is clear that the evaluation and benchmarking of the algorithms is crucial. Previous work on the evaluations [12–22, 32–33] provides insights into how the evaluation of recognition algorithms and systems can be performed efficiently. The most important facts learned in previous evaluations are as follows: (1) large sets of test images are essential for adequate evaluation; (2) The sample should be statistically as similar as possible to the images that arise in the application being considered; (3) Scoring should be done in a way which reflects the costs or other system requirement changes that result from errors in recognition; (4) System reject-error behavior should be studied, not just forced recognition; (5) The most useful form of evaluation is that based as closely as possible on a specific application; (6) The accuracy, samples, speed and hardware, and human interface are extremely required for the face recognition.

Whether a face recognition system is better was measured using two basic methods. The first one measured identification performance, where the primary statistic is the percentage of probes that are correctly identified by the algorithm. The second one measured verification, where the performance measure is the equal error rate between probability of false alarm and of correct verification. (A more complete method of reporting identification performance is a cumulative match characteristic. For verification performance it is a receiver operating characteristic (ROC).)

A standard database of face imagery was essential to the success of the FERET program, both to supply standard imagery to the algorithm developers and to supply a sufficient number of images to allow testing of these algorithms. Until recently, there exist some common and large databases for FRT evaluation test, which embrace MIT Face Database (about 2592 images), UMD Face Database (about 7100 images), USC Face Database (about 100 images), ARPA/ARL Face Database (about 8500 images), ORL Face Database (about 400 images), Yale Face Database (about 165 images), Stirling University Face Database (about 1580 images), UMIST Face Database (about 564 images), Oulu University Physics-Based Face Database (about 2100 images), M2VTS Multi-modal Face Database (about 10 000 images), CMU Test Images for Face Detection (about 5000 images), and UMASS Face Database (about 150 images), etc. Among them ARPA/ARL Face Database is one of the most famous of them, since it possesses those traits such as multi-race, multi-aged phase, diverse expressions, illumination change, pose variation, large number, many test participants. It consists of about 1100 image sets of two frontals, and pairs of 1/4, of 1/2, of 3/4, and of full profiles. It totals about 8500 images. It has gradually become the most complete gallery for face recognition test. In addition, some algorithms performed very well under a certain database: Subspace LDA from UMD [34], Probabilistic Eigenface from MIT [35], and Elastic Graph Matching from USC [36].

## 3 DETECTION AND LOCALIZATION TECHNIQUES

Face detection and localization from images is a key problem and a necessary first step in face recognition systems, with the purpose of localizing and extracting the face region from the background. It also has several applications in areas such as content-based image retrieval, video coding, video conferencing, crowd surveillance, and intelligent human-computer interfaces. However, it was not until recently that the face detection problem received considerable attention among researchers. The human face is a dynamic object and has a high degree of variability in its appearance, which makes face detection a difficult problem in computer vision. It is an essential step in face recognition. During the past several years, a wide variety of face detection and localization techniques have been growing fast. Many progresses on it have been made and reported recently. The associated studies are elaborated below.

Up to the close years, there exist a variety of approaches to face detection. Categorization of them may depend on different criteria. In terms of modeling process used, the approaches to face detection may fall into two main categories: (1) local feature-based ones; (2) global methods. Their face detection regions are required by the comparative matching between the detecting region and constructed template based on modeling. In the former ones, salient features such as the eyes, nose, and mouth are first located. Various measurements of these facial components are used to construct feature vectors. These approaches to face recognition basically rely on the detection and characterization of above individual facial features and their geometrical relationships. The latter ones, on the other hand, take a holistic view towards face recognition without explicitly finding facial features. They involve encoding the entire facial image and treating the resulting facial code as a point in a high-dimensional space and assume that all faces are constrained to particular positions, orientations, and scales.

### 3.1 Approaches Based on Features

#### 3.1.1 Geometrical Method

The method is based on face geometrical configuration. Generic knowledge about faces employed is facial organs' position, symmetry, and edge shape as follows: a face contains four main organs, i.e., eyebrows, eyes, nose and mouth; a face image is symmetric in the left and right directions; eyes are below two eyebrows; nose lies between and below two eyes; lips lie below nose; the contour of a human head can be approximated by an ellipse, and so on. By using the facial components as well as positional relationship between them we can locate the faces easily. When a face image is fed into the system, a preprocessing step will be applied to remove small light details and to enhance the contrast. Then, the processed image will be the threshold to produce a binary image. Finally, a labeling step and a grouping algorithm will be used to group detected features block by block to locate the faces. Many detailed studies may refer to the literatures [37] (Jeng et al. 1998), [38] (Berngger et al. 1998) [39] (Feng and Yuen, 1998), [40] (OHTA et al., 1998), [41] (Wang et al., 1999), [42] (Reisfeld and Teshurun, 1998), [43] (Jie Zhou et al., 1999), [44] (Lv et al., 2000), [45] (Kwon, and Lobo, 1999), [46] (Tao et al., 1999), [47] (Tian et al., 1999), [48] (Decarlo and Metaxas, 2000), [49] (Wang and Tan, 2000), [50] (Roach et al., 2000), [51] (Lin and Fan, 2000), [52] (Jing and Mariani, 2000), [53] (Wang et al., 2001), [54] (Li et al., 2001), [55] (Sclaroff and Liu, 2001), [56] (Ho and Huang, 2001), [57] (Wong et al., 2001). Face detection can often be achieved by detecting geometrical relationships among facial organs as mentioned, because they are simple, straightforward and efficient. Jeng et al. 1998 [37] proposed a useful geometrical face model and an efficient facial feature detection approach, which is based on the fact that human faces are constructed in the same geometrical configuration and could accurately detect facial features, especially the eyes, even when the images have complex backgrounds such as bad lighting condition, skew face orientation, and facial expression. The geometrical face model in [37] was constructed to evaluate which combination of feature blocks was a face as Figure 2, according to the average proportion between each facial organ obtained by estimating several real faces. Facial symmetry is another geometrical characteristics. The context free generalized symmetry transform is an interest operator which is motivated by the biological mechanisms of attention and fixation, and is inspired by the intuitive notion of symmetry. It assigns a symmetry magnitude and a symmetry orientation to every pixel. The input to the transform is an edge map — the gradients of intensity at each pixel and its output is a symmetry map, which is a new kind of an edge map, where the magnitude and orientation of an edge depend on the symmetry associated with the pixel. In [42] Reisfeld and Yeshurum 1998 proposed a method for automatic and robust detection of eyes and mouth using the aforementioned transform. In most cases, the overall shape of the face or the whole head is very similar to an ellipse. Some researchers have been using this shape information for facial detection. [49] extended previous shape-based face detection methods by applying a special elliptical template containing directional information of edges. An elliptical ring in [49] was used as the template as illustrated in Figure 3. Figure 3 (a) is a normal upright face; (b) is the binary image after edge linking; (c) is an elliptical ring representing the contour. Recently some scholars, including Ho and Huang 2001 [56], and Wong et al. 2001 [57] have studied the genetic algorithm-based detection. Ho and Huang 2001 [56] have presented a genetic algorithm-based optimization approach for facial modeling from an un-calibrated face image using a flexible generic parameterized geometrical facial model (FGPFM). The notations for the model ratios in [56] are illustrated in Figure 4. $L_e$ is the distance between two far corners of the eyes; $L_f$ is the distance between the mouth-line and eyes-line; $L_w$ is the distance between the nose base and the mouth-line; $L_m$ is the width of the mouth; $L_y$ is the width of one eye; Ln is the width of the nose. Figure 5 in [57] showed an example of a selected face region based on the location of an eye pair. A square block is used to represent the detected face region.

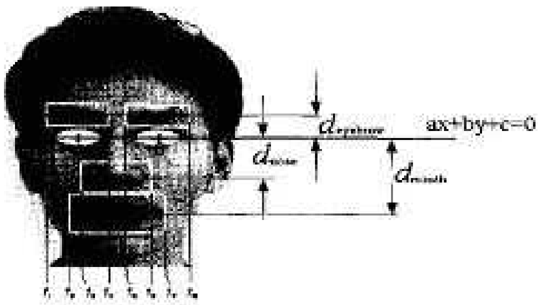


Fig. 2. The geometrical face model

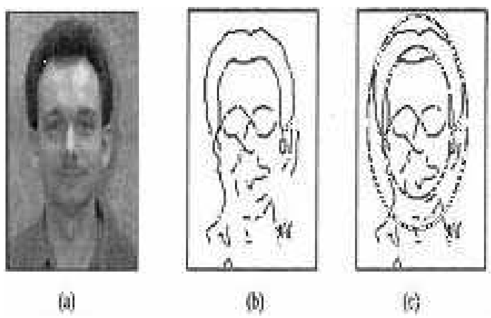


Fig. 3. Deformable template

The Sequential Testing Method also belongs to this category. The approach is coarse-to-fine in both the exploration of poses and the representation of objects. Features are spatial arrangements of edge fragments, induced from training faces at a reference pose, and computation is minimized via a generalized Hough transform; there is no on-line optimization and no segmentation apart from visual selection itself. All tests are binary and indicate the presence or absence of loose spatial arrangements of oriented edge fragments. Detection means finding a sufficient number of arrangements of each size along a decreasing sequence of pose cells. At the beginning, the tests are simple and universal, accommodating many poses simultaneously, but the false alarm rate is relatively high. Eventually, the tests are more discriminating, but also more complex and dedicated to specific poses. As a result,

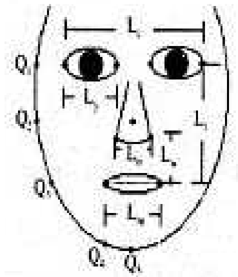

Fig. 4. The notions of the model ratios

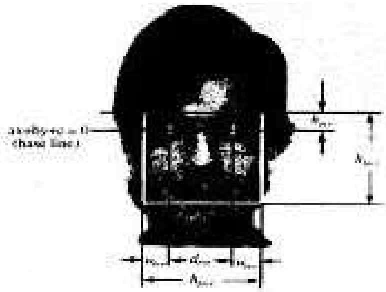

Fig. 5. The head of geometry of our head model

the spatial distribution of processing is highly skewed and detection is rapid, but at the expense of (isolated) false alarms which, presumably, could be eliminated with localized, more intensive, processing. The pertinent documents are in [58–59]. In [59], Figure 6 (left: the grey level was proportional to this count; right: the scan line corresponding to the arrow; it covered three faces) showed an illustration of the spatial distribution of processing corresponding to the scene shown in Figure 7.

### 3.1.2 Color-Based or Texture-Based Method

Color and texture are two important modalities in many images processing tasks, ranging from remote sensing to medical imaging, robot vision, face recognition, etc. By now their analysis methods have been widely utilizing to detect faces for different races, sexes, and ages. Some research results show that human skin colors cluster in a small region only in the GRB color space instead of the HIS color space; human skin colors differ more in brightness than in colors; and every texture is distinctive and distinguishable from one another. Therefore, the normalized GRB or texture models are considered to be capable of characterizing human face with less variance in color or texture. Recently, many intensive studies have been reported, i.e., [60, 61] (Terillon et al., 1998), [62] (Yang and Ahuja, 1998), [63] (nakia and Stockmann, 1998), [64] (Albol et al., 199), [65] (Garcia and Tziritas, 1999), [66] (Lu et al., 1999), [67] (Xie et al., 1999), [68] (Li et al., 2000), [69] (Wu and Huang, 2000), [70] (Wei et al., 2000), [71] (Störring et al., 20000, [72] (Schwerdt and Crowley, 2000), [73] (Liang et al., 2000), [74] (Dass and Jain, 2001), [75] (Zhang et al., 2001), [76] (Duta and Jain, 1998), [77] (Cascia and Sclaoff, 1999), [78] (Decsombes et al., 1999), [79] (Fan and Sung, 2000), [80] (Li and Peng, 2001). Terrillon et al. [60, 61] used a skin color model based on the Mahalanobis metric and a shape analysis based on invariant Fourier-Mellin moments to automatically detect and locate human faces in two-dimensional complex scene images. During the detection, color segmentation of an input image is performed by threshold in a normalized hue-saturation color space where the effects of the variability of human skin color and the dependency of chrominance on changes in illumination are reduced. Literature [73] employed a multi-modal face tracker which integrated eye blink detection, cross-correlation, and robust tracking of skin colored regions. A new technique was developed which replaced threshold and connected components with the moments of color pixels weighted by a Gaussian density function. The paper [77] explored new ways of learning and retrieving the appearance of human faces in black, white and still images by gray-tone texture model. An example of the inverse texture mapping of a cylinder in arbitrary 3D position in [77] is shown in Figure 8. Fan and Sung [79] proposed a feature based similarity measure (FBSM) to take into account the spatial differences between feature points of two poses. The feature-texture similarity measure (FTSM) was sensitive to pose differences between two face images and could be used to directly determine the best hill-climb directions in pose parameter space without computing gradients of error functions.

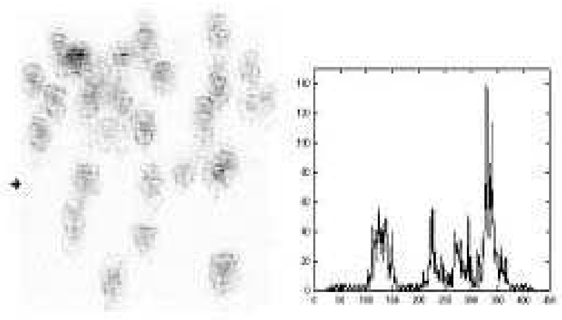

Fig. 6. The coarse-to-fine nature of the algorithm was illustrated by counting, for each pixel, the number of times the detector checks for the presence of an edge in its vicinity

### 3.1.3 Motion-Based Method

Human motion analysis is receiving increasing attention from computer vision researchers. This interest is motivated by applications over a wide spectrum of topics.

Fig. 7. Example of a scene

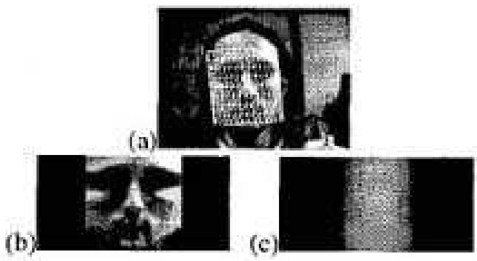


Fig. 8. (a) Input frame with the cylindrical model superimposed, (b) corresponding to texture map, (c) and confidence map

Motion analysis could extract the low-lever features such as body part segmentation, joint detection and identification and recover 3D structure from 2D projections in an image sequence. This motion information, which was comprised of position and velocity of moving eyes, speaking tone and expressions, etc., incorporated with intensity value, could be employed to easily locate the face. [81] proposed a combined expression recognition system based on the analysis of the dynamic expression image sequences. Here the authors took the face as being composed of several primary expression regions, in which the motion features could be extracted and constituted to eigen-sequences. The analysis of the arbitrary length of image sequences of facial expressions and combined expression recognition are proposed and implemented by analyzing the respective expression meaning and the expression contents of different primary regions and using the multi-feature fusion. Part of the dynamic expression image sequences in [81] shown in Figure 9. Bobick and Davis 2001 [82] presented a new view-based approach to the representation and recognition of human movement. The basis of the presentation is a temporal template — a static vector-image where the vector value at each point is a function of the motion properties at the

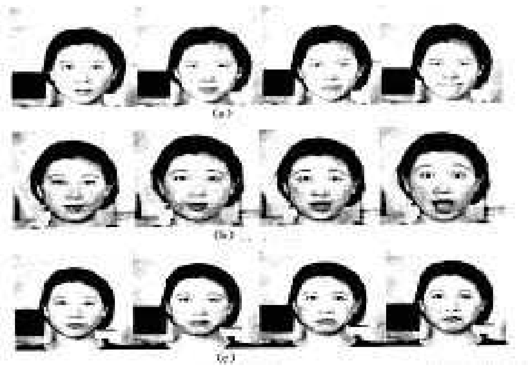


Fig. 9. Dynamic expression image sequences: (a) anger; (b) surprise; (c) sadness

corresponding spatial location in an image sequence. Other detailed studies may be referred to literature [83–86].

### 3.1.4 Other Methods

The above-mentioned methods are robust and effective for detecting faces. However, on the one hand, the geometrical, color-based or texture-based and motion-based methods show generally different sensitivity to illumination, pose, scale, resolution to some extent; the motion-based method is also closely related to human motion properties at every corresponding spatial location in an image sequence. On the other hand, the major difficulties are inevitably encountered in face recognition due to variation in luminance, facial expressions, visual angles and other potential features such as glasses, beard, etc. These lead to a need for employing multiple methods and various techniques, i.e. Bayesian, neural networks, fuzzy logic, and others. Consequently, some researchers have integrated multiple ways so as to achieve better performance and make best advantages of them [87–92]. [87] presented a novel analytically-based face recognition system where the eyes are detected using graph templates, the mouth is detected using deformable templates, and the location of the nose is found by using integral projections based on the mouth and eye locations. Using a 3D model of a head, the facial rotations are estimated in order for the system to compensate for rotation. T. Yokoyama et al. 1998 [88] proposed a facial contour extraction model in which three characteristics: global shape constraint of axis-symmetry, dual-scale filtering, and iterative initialization are considered. In [89] Jebara and Pentland described a real-time system initialized by using skin classification, symmetry operations, 3D warping and eigenfaces to find a face. Sun et al. 1998 presented a robust approach to face or facial features detection in which utilized color, local symmetry, geometry information of human face based various models. Literature [92] introduced a pose-invariant appearance model that utilized a generic-view shape template for alignment, texture nonlinearities across views of large pose variations, and a neural network for model fitting to new images.

It is worth noting that there sill exists another method belonging to the category. For instance, [93] and [94] are the related paradigms. In [93], Huang et al. 1998 proposed a combined approach to face detection. They first adopted the structure-based approach to obtaining the face location and facial components positions roughly, which is insensitive to the illumination, skin tone, and scale. Compared to template-based methods, structure-based approach could then be faster and more flexible to be extended to different scene variation. The symmetry of front face along the mid-axle is then another information which is used for validation of face, where the texture and image feature are used. Literature [94] presented a robust and precise scheme for face detection and precise facial feature location. The structural model is used to characterize the geometric pattern of facial components. The texture and feature models are used to verify the face candidates detected before. The center and the radius of the eyeballs of a person's eyes was detected using the face detected, the structural information extracted and the contour and region information. Figure 10 in [94] showed the faces detected with eyes marked.

Fig. 10. Faces detected with eyes

### 3.2 Holistic Approaches

#### 3.2.1 Eigenface-Based Method

This method approximates the multi-template $T$ by a low-dimensional linear subspace $F$, usually called the face space. Images are initially classified as potential members of $T$, if their distance from $F$ is smaller than a certain threshold. The images which pass this test are projected on $F$ and these projections are compared to those in the training set. The method has been rather successful for various detection problems such as detecting frontal human faces [95–97]. However, it runs into problems if one tries to detect objects under arbitrary rotation and possible other distortions. In [97] Zhu et al. proposed a subspace approach to capture local

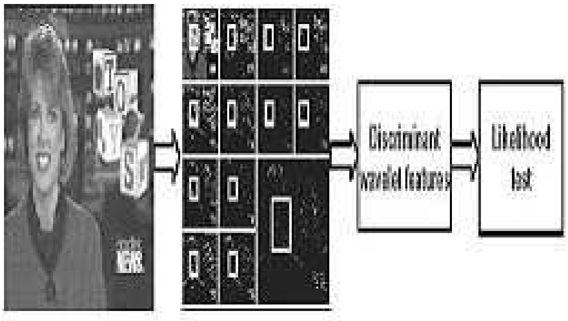


Fig. 11. Detection process

discriminative features in the space-frequency domain for fast face detection. Based on orthonormal wavelet packet analysis, the discriminant subspace algorithm was developed to search for the minimum cost subspace of the high-dimensional signal space, which led to a set of wavelet features with maximum class discrimination and dimensionality reduction. Figure 11 in [97] illustrated the detection process. The system decomposes an entire input image into subband images which contain the discriminant features. Multiple sliding windows within different subbands are aligned to the same spatial location. Features are selected from multiple subbands to calculate the likelihood ratios. Face locations are reported where the likelihood ratios exceed a fixed threshold.

### 3.2.2 Spatial Matching Detector Method

This approach embraces the Support Vector Machines, various template matching methods, other discriminable Kernel Cost Function methods, and so on. [101] offerred a novel detection method, which worked well even in the case of a complicated image collection of detected images which was called a multi-template. Only images which passed the threshold test imposed by the first detector were examined by the second detector, etc. The algorithm's performance compared favorably to the well-known eigenface and support vector machine based algorithms, but was substantially faster. A schematic description of the geometry behind anti-faces in [101] was presented in Figure 12. The algorithm's "positive set" (the images it classifies as members of the multi-template), is orthogonal to the direction around which random images cluster, hence, there are relatively few false alarms. Many associative references point to [98–103].

### 3.2.3 Neural Networks Method

Neural networks have been applied, with considerable success, to the problem of frontal face detection [104–110]. A neural networks based upright frontal face detection system was presented in [105]. The retinally connected neural network examined

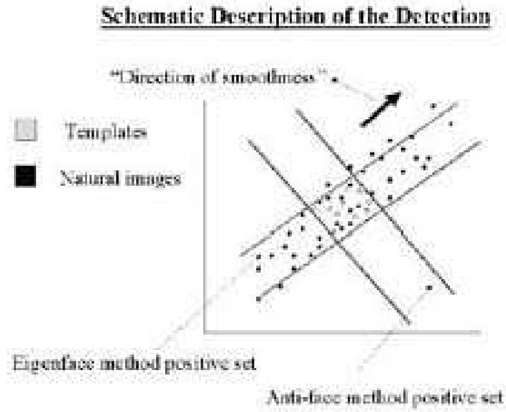


Fig. 12. Schematic description of the anti-face algorithm

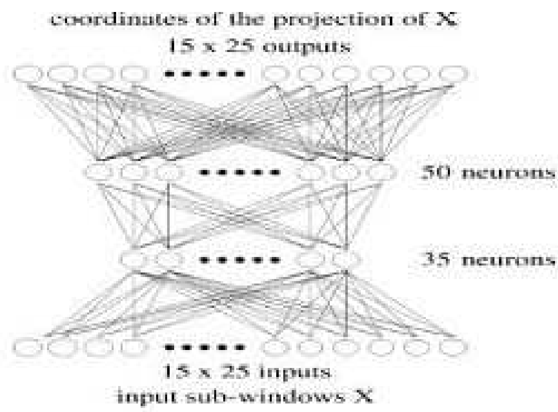


Fig. 13. Neural network layers

small windows of an image and decided whether each window contained a face. The system arbitrated between multiple networks to improve performance over a single network. In [106] hybrid neural method is proposed to locate human eyes. In [110] the new neural network model proposed, the Constrained Generative Model, performed an accurate estimation of the face set, using a small set of counter-examples. The neural network layers in [110] were shown in Figure 13. The use of three layers of weights allows to evaluate the distance between an input image and the set of face image.

### 3.2.4 Fuzzy Theory Based Method

This approach detects faces in color images based on the fuzzy theory. [111] is a typical example of fuzzy detection category. In this paper Wu et al. made two fuzzy models to describe the skin and hair color, in which used a perceptually uniform color space to describe the color information to increase the accuracy and stableness. Furthermore, the models were used to extract the skin and hair color regions, then comparing them with the pre-built head-shape models by using a fuzzy theory based pattern-matching method to detect face candidates. In [111], Figure 14a showed an input image. Figures 14b and 14c were grayscale images that indicated the skin color similarity map (SCSM) and hair color similarity map (HCSM) estimated from Figure 14a. Figure 14d showed the map of matching degree (MMD). Some experimental results were shown in 15.

### 3.2.5 Other Methods

This method integrates the above multiple methods and various techniques, i.e. Bayesian, neural networks, fuzzy logic, and others. In [112] three methods were developed to extract facial expression information for automatic recognition. The first one is facial feature point tracking using a coarse-to-fine pyramid method. The second one is dense flow tracking together with principal component analysis (PCA). The third one is high gradient component (i.e. Furrow) analysis in spatio-temporal domain.

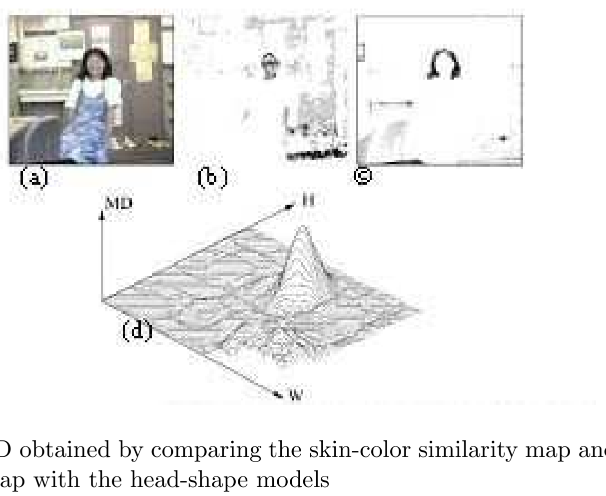


Fig. 14. An MMD obtained by comparing the skin-color similarity map and the hair-color similarity map with the head-shape models

Fig. 15. Experimental results of face-candidate detection

## 4 FEATURE EXTRACTION TECHNIQUES

The extraction of discriminant features is the most fundamental and important problem in face recognition. The image features generally may be divided into four groups: visual features, statistical pixel features, transform coefficient features, and algebraic features, with emphasis on the algebraic features, which represent the intrinsic attributes of an image. The visual features include edges, contours, textures and regions of an image. They are all visual features of a pixel; the statistical features of a pixel are representation of histogram and various statistical moments; the transform coefficient features are properties of the feature vector using various mathematical transforms; the latter represent intrinsic attributions of an image. Any image can be considered as a matrix. Therefore, various algebraic transforms or matrix decompositions can be used for algebraic features extraction of the image. In the light of the types of the above-mentioned features, there are four approaches for extracting image features until now as follows:

### 4.1 Knowledge-Based Method

This approach depends on generic visual and statistical knowledge to extract features. So far there has been much literature concerning the problem of extracting these features [37, 41, 114]. In [113] facial features were extracted based on generic knowledge of facial components.

### 4.2 Mathematical Transform Method

It is well known that Fourier transform, Hadamard transform, Karhunen-Loeve transform, Singular Value Decomposition, Foley-Sammon transform, Discrete Cosine transform, multispace Karhunen-Loeve transform, etc. can be used to extract features of an image. Literature [81, 114–120] are correlative studies. In [115] Yilmaz and Gkmen proposed a new algorithm based on KLT to overcome edge problems due to illumination variation and pose change. [116] exploited the feature extraction capabilities of the discrete cosine transform (DCT) and invoked certain normalization techniques that increased system's robustness to variation in facial geometry and illumination. The Variance distribution for a selection of discrete transforms given a first-order Markov process of length $N = 16$ and $\rho = 0.9$ in [116] was shown in Figure 16. Data is shown for the following transforms: discrete cosine transform (DTC), discrete Fourier transform (DFT), slant transform (ST), discrete sine transform (type I) (DST-1), discrete sine transform (type II) (DST II), and Karhunen-Loeve transform (KLT). [118] introduced the multi-space KL to improve KLT when the data distribution was far from a multidimensional Gaussian and to better cope with large sets of pattern, which could cause a severe performance drop in KL. Lai et al. 2001 [120] presented a new method for holistic face representation, called spectro-face which combined the wavelet transform and the Fourier transform. Figure 17 in [120] showed the decomposition process by the 2D wavelet transform on a face image. Fourier invariant features could represent the facial features which were invariant for the rotation in the $z$-axis as Figure 18.

The other approach is the morphological transform method. This method extracts face features by the scale-space morphological techniques which is an alternative to linear techniques for generating an information pyramid. [121–123] are associated with it.

### 4.3 Neural networks or Fuzzy Extractor Method

Park and Bien 2000 [124] used a fuzzy observer as a means of extracting features of wrinkledness directly from a camera image of a human face. The fuzzy observer in [124] was shown in Figure 19.

### 4.4 Other Method

Literatures [125–127] are associated with this method. They combined generic geometry properties and mathematical transform to extract the image features.

## 5 RECOGNITION TECHNIQUES

After detection and feature extraction of face has been finished. Face recognition then is the last step of the bottom-up image processing approach. Research on face recognition technology has been studied for more than 20 years. It has become a new major research area in the last few years because of a number of potential applications ranging from security access control, personal identification to human-computer communication. A number of face methods have been proposed.

These methods can be divided into the following several categories which depends on the classifier selection while diverse classifier differs on the assumptions

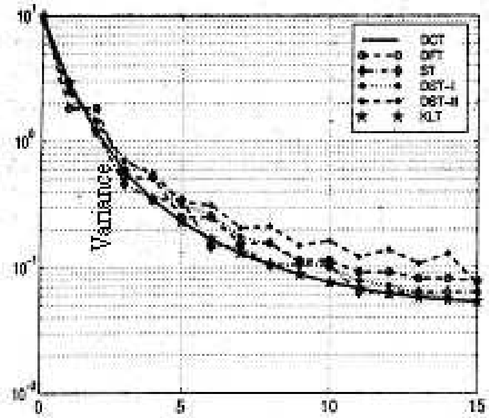


Fig. 16. Variance distribution for a selection of discrete transforms for $N = 16$ and $\rho = 0.9$

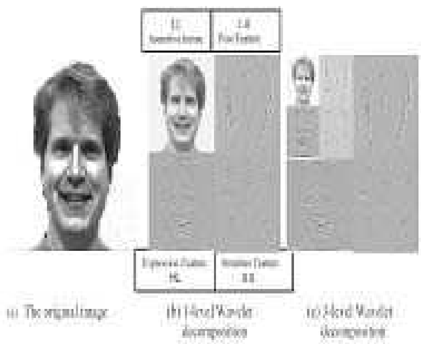

Fig. 17. 2D wavelet decomposition of a face

about classifier dependencies, type of classifier outputs, aggregation strategy (global or local), aggregation procedure (a function, a neural network, an algorithm), etc.

## 5.1 Statistical Approach

In this approach quantitative description of faces is characteristic, elementary numerical description — features — are used. The set of all possible patterns forms the pattern or feature space. The classes form clusters in the feature space, which can be separated by discrimination hyper-surface. The approach chiefly embraces geometrical parameterization method [39, 48, 50], eigneface method [114–120, 128, 129], Fisherface method [128, 130], evolutionary pursuit algorithm, etc.

The use of geometrical parameterization, i.e., distances and angles between points such as eye corners, mouth extremities, nostrils, and chin top is one method of characterizing the face. A simple distance measure is used to check for similarity

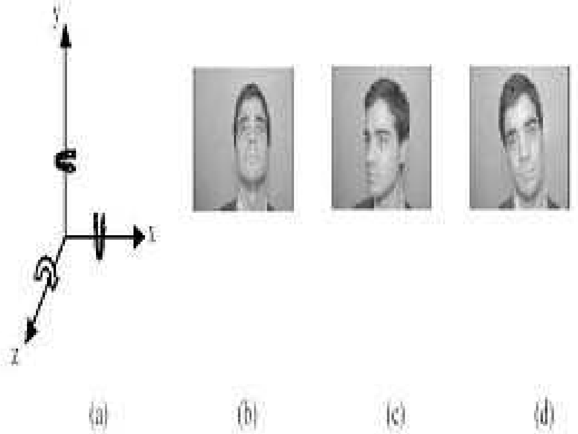


Fig. 18. (a) The direction of rotation; (b) rotation in the $x$-axis; (c) rotation in the $y$-axis; (d) rotation in the $z$-axis

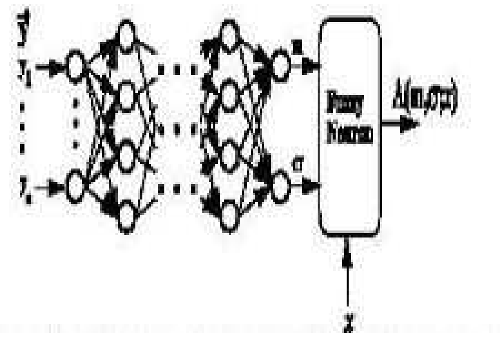


Fig. 19. Artificial neural network structure of fuzzy observer

between an image of the test set and the image in the reference set. Matching accuracies depends on the geometrical feature parameters extracted. The references denote their experimental results are not comparatively satisfactory. The deformable template method in [48], where an energy function is employed to adjust the geometrical parameter configuration also belongs to this category. Though it is a reformative algorithm, there still exist two disadvantages in practice. On the one hand, the matching process is sensitive to the values in the energy function. Hence for the input images under different conditions the values of fields are rather different, and consequently affect the matching process and result. What's more, there is another weakness of more time-consuming and expensive computation due to using the matching process. The deformable face model in [48] was a 3D polygon mesh, shown smoothly shaded in Figure 20(a), and wireframe in 20(b) in its default configuration.

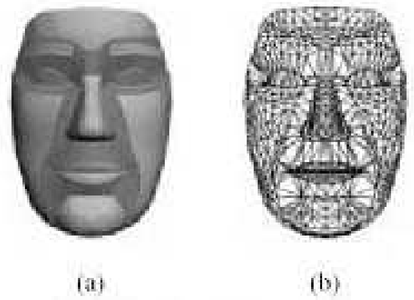


Fig. 20. The deformable face model

| method | # axes | top 1 region rate | top 3 rate |
|---|---|---|---|
| Eigenface | 26 | 87.26 % | 95.66 % |
| Eigenface | 30 | 88.62 % | 95.93 % |
| Fisherfaces | 26 | 86.45 % | 93.77 % |
| Fisherfaces | 30 | 88.08 % | 95.39 % |
| Evolutionary Pursuit | 26 | 92.14 % | 97.02 % |

Table 1. The comparative testing performance for eigenfaces, fisherfaces, and evolutionary pursuit (30 images)

The eigenface method is a successful holistic approach to face recognition using the Karhunen-Loeve transform (KLT). This transform produces an expansion of an input image in terms of a set of basis images or the so-called "eigenimages". Turk and Pentland (1991) [129] proposed a face recognition system based on the KLT in which only a few KLT coefficients were used to represent faces in what they termed "face space". Each set of KLT coefficients representing a face formed a point in this high-dimensional space. The system performed well for frontal mug shot images. Whereas the KLT does not achieve adequate robustness against variations in face orientation, position, and illumination. Furthermore, in practice the two main drawbacks of KLT in multi-space, including linearity and scalability problem, exist — namely, the data distribution cannot be model with a multidimensional Gaussian and a linear mapping for feature extraction demonstrates its weakness; the feature power progressively vanishes since the pattern features tend to become very smooth and the training time can become daunting. Attempting to overcome these limitations, a multi-space KL as a new approach is introduced in [118]. Other literatures are referred to [114–117, 119, 120].

The fisherface method based on Fisher Linear Discriminant (FLD), following Principal Component Analysis (PCA) and operating then in a compressed subspace, seeks for disciminatory features by taking into account within- and between-class scatter as the relevant information for pattern classification. It overcomes one of PCA's drawbacks as it can distinguish within- and between-class scatters. Its weakness is that it requires large sample sizes for good generalization. The fisherface space is superior to the eigenface space for face recognition only when the training images are representative of the range of face (class) variations; otherwise, the performance difference between the both is not significant.

Evolutionary Pursuit (EP) is a novel and adaptive representation method for image encoding and classification. In analogy to projection pursuit methods, EP seeks to learn an optimal basis for the dual purpose of data compression and pattern classification. It implements strategies characteristic of genetic algorithms (GAs) and is similar to random search techniques for nonlinear optimization and variable selection. Experimental results show that EP improves on face recognition performance when compared against PCA ("Eigenface") and displays better generalization abilities than FLD ("Fisherfaces"). The comparative testing performance for eigenfaces, fisherfaces, and evolutionary pursuit (30 images) in [131] was shown in Table 1. And the 30 eigenfaces derived by the Eigenface method and optimal basis derived by the EP in [132] were displayed in Figure 21 and Figure 22, respectively. All details refer to [131, 132].

### 5.2 Feature Matching Approach

This method stores feature points detected using the Gabor wavelet decomposition or multi-scale morphological dilation-erosion into data files for each image. Its identification process utilizes the information present in a topological graphic representation of the feature points. After compensating for differing centroid locations, two cost values are evaluated. One is the topological cost and the other a similarity cost. Kotropoulos et al. [121] applied morphological elastic graph matching to frontal face authentication on databases ranging from small to large multimedia ones collected under either well-controlled or real-world conditions. [122] proposed a novel morphological dynamic link architecture (MDLA) based on multi-scale morphological dilation-erosion instead of Gabor wavelets for frontal face authentication. Figure 23 in [122] depicted the grids formed in the matching procedure of a test person with himself and with another person for a pair of test persons extracted from the Multi-modal Verification Techniques for Tele-services and Security Applications (M2VTS) database. In [123] (a) Model grid for person BP; (b) best grid for test person BP after elastic graph matching with the model grid; (c) best grid for test person BS after elastic graph matching with the model grid for person BP; (d) model grid for person BS; (e) best grid for test person BP after elastic graph matching with the model grid for person BS; and (f) best grid for test person BS after elastic graph matching with the model grid. Tefas et al. [123] used support vector machines to enhance the performance of elastic graph matching for frontal face authentication. It is relatively insensitive to variations in lighting, face position, and expression and its database is easily expanded, whereas it may require more computational effort than the eigenface.

Fig. 21. The 30 eigenfaces derived by the Eigenface method

### 5.3 Neural Networks Approach

Neural networks can be viewed as massively parallel computing systems consisting of an extremely large number of simple processors with many interconnections. Neural networks attempt to use some organizational principles (such as learning, generalization, adaptability, fault tolerance and distributed representation, and computation) in a network of weighted directed graphs in which the nodes are artificial neurons and

Fig. 22. Optimal basis derived by the EP

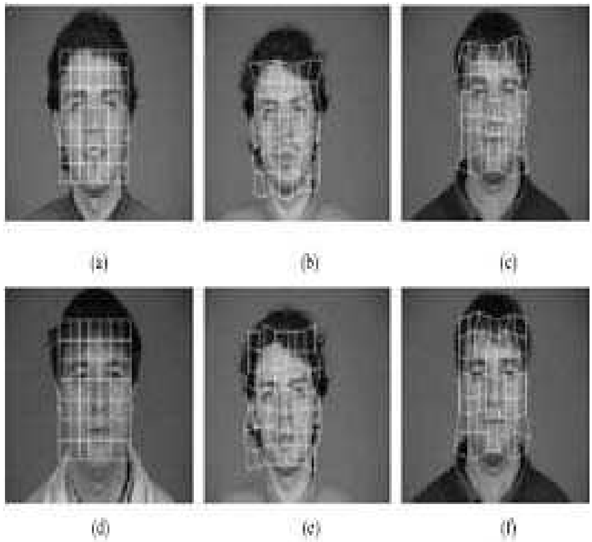


Fig. 23. Grid matching procedure in MDLA

directed edges (with weights) are connections between neuron outputs and neuron inputs. The main characteristics of neural networks are that they have the ability to learn complex nonlinear input-output relationships, use sequential training procedures, and adapt themselves to the data. To date, Hopfield Neural Networks (HNN), Self-Organizing Map (SOM), Back-Propagation (BP) have been used for face recognition. [113, 104–110] are detailed to describe them.

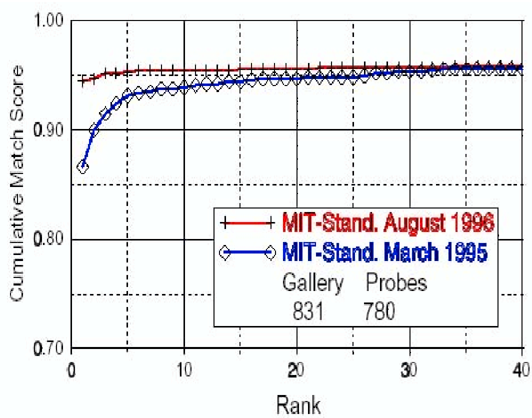


Fig. 24. Cumulative recognition rates with standard eigenface matching (bottom) and the newer Bayesian similarity metric (top)

### 5.4 Fuzzy Theory Approach

This approach uses fuzzy theory to representing diverse, non-exact, uncertain, and inaccurate knowledge or information. And information carried in individual fuzzy set is combined to make a decision. Processes of composition and de-fuzzification form the basis of fuzzy reasoning. Fuzzy reasoning is performed to recognize face in the context of a fuzzy system model that consists of control, solution, and working data variables; fuzzy sets; hedges; fuzzy; and a control mechanism. Many paradigms are appeared in [87, 124].

## 6 OTHER APPROACHES

Beside the above-mentioned approach to face recognition, some researchers also used other methods to perform the studies on face recognition, i.e., the rules of the shape and albedo of a face under all possible illumination conditions, Bayesian decision, etc. Georghiades et al. 2001 [133] presented a generative appearance-based method for recognition human face under variation in lighting and viewpoint, exploiting the fact that the set of images of an object in fixed pose, but under all possible illumination conditions, is a convex cone in the space of images. In [134] Moghaddam et al. 2000 utilized Bayesian decision for the purpose of face recognition and image retrieval. They arrived at the conclusion that Bayesian method gained better results than standard eigenface method and effectively halved the error rate of eigenface matching. Figure 24 in [134] highlighted the performance difference between standard eigenfaces and the Bayesian method from a small test set of 800+ individuals.

## 7 SUMMARY AND CONCLUSION

In this paper, we have presented an extensive review of recent research development on face recognition. Also have we focused on face recognition systems, detection and localization, feature extraction, and recognition aspects of the face recognition problem. Here we give below a concise summary followed by conclusions in the same order as the topics appear in the paper.

A crucial step in face recognition system is the evaluation and benchmarking of numerous algorithms. Several important face databases and their associated evaluation methods are reviewed. The availability of these protocols which include the FERET protocol and the XM2VTS protocol has had a significant impact on progress in the development of face recognition algorithms.

We herein present a comprehensive and critical survey of face detection algorithms. Face detection is a necessary first-step in face recognition systems, with the purpose of localizing and extracting the face region from the background. The algorithms presented in this paper are classified as either feature-based or holistic and are discussed in terms of their technical approach and performance. In the light of the above detailed analysis and outline, it is clear that, at present, the approaches on face detection are chiefly concentrated on feature-based and holistic aspects. And it is also at least no doubt to come to the following conclusions: (1) The present face detection techniques are developing from the traditional 2D image detection to 3D image detection. (2) In a recent study of face detection systems, most of them have adopted hybrid processing algorithms to obtain better results for detection. It deeply indicates that the future in face detection will lie in the research of hybrid detection algorithms and systems. (3) The interest and research activities on the application of neural networks and fuzzy method to face detection will continue to gain attention and increasing significance in the future. (4) More commercial applications of FRT will emerge, such as face verification based on ATM and access control, while law enforcement applications include video surveillance.

In terms of image features, the extraction techniques are classified into four types: knowledge-based, mathematical transform based, neural networks or fuzzy theory based, and other. Feature extraction is an old problem in the field of pattern recognition; however, it has been the most fundamental and important problem. Whether feature extraction is effective is always the key to solving the problem or completing the task of image recognition. It is believed that better methods for extraction will be improved to represent intrinsic and more attributions of images, reduce the information redundancy, and increase simultaneously the entropy, which will merit further recognition for better results.

A multitude of techniques are available in literature [128–134] for recognition. They include statistical method, feature matching method, neural networks, fuzzy theoretical method, and other. The intensity-based approaches, Eigenface, Fisherface, Support Vector Machines (SVM's), Morphological Elastic Graph Matching (MEGM), Neural Networks (NN), fuzzy systems, etc. are typical. In the meantime, MEGM, NN, and fuzzy systems are the most promising methods which will be further progressed. The eigenface is sensitive to face position and lighting variations in the images, but MEGM is insensitive and its database expansion is easier, whereas it may require more computational effort. The performance of the auto-association and classification nets is upper bounded by that of the eigenface, but is more difficult to implement in practice. Consequently, for commercial/law enforcement applications to date, only intensity-based approaches may be pursued. Methods for integrating variety of methods above should be developed, combining the latest achievements. And it would also be of interest to investigate ways to make the MEGM, NN, fuzzy systems more practical and efficient.

## REFERENCES


[1] Proceedings of the International Conferences on Automatic Face and Gesture Recognition. 1996, 1998, 2000.

[2] Proceedings of the International Conferences on Audio- and Video-Based Persin Authentication. 1997, 1999.

[3] Proceedings of the International Conferences on Computer Vision. 1995, 1998, 1999, 2001.

[4] Proceedings of the International Conferences on Pattern Recognition. 1998, 2000.

[5] Proceedings of the International Conferences on Computer Vision and Pattern Recognition. 1997–2000.

[6] Proceedings of the IEEE Workshops on Applications of Computer Vision. 1996, 1998, 2000.

[7] Proceedings of SIBGRAPI International Symposiums on Computer Graphics, Image Processing, and Vision. 1998, 1999, 2000.

[8] Proceedings of the IEEE Workshop on Computer Vision Beyond the Visible Spectrum Methods and Applications. 2000.

[9] Proceedings of the IEEE Colloquium on Computer Vision for Human Modeling. 1998.

[10] Proceedings of the IEEE Signal Society Workshops on Neural Networks. 1997–2001.

[11] Proceedings of the International Conferences on Image Processing. 1997–2001.

[12] Phillips, P. J.—Moon, H.—Rauss, P.—Rizvi, S. A.: The FERET Methodology for Face Recognition Algorithm. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 1997, pp. 137–143.

[13] Phillips, P. J.—Wechsler, H.—Huang, J.—Rauss, P. J.: The FERET Database and Evaluation Procedure for Face-Recognition Algorithm. Image and Vision Computing, Vol. 16, 1998, No. 5, pp. 295–306.

[14] Phillips, P. J.—Moon, H.—Rauss, P.—Rizvi, S. A.: The FERET Methodology for Face Recognition Algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, 2000, No. 10, pp. 1090–1103.

[15] Rizvi, S. A—Phillips, P. J.—Moon, H.: A Verification Protocol and Statistical Performance Analysis for Face Recognition Algorithm. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 1998, pp. 833–838.

[16] Rizvi, S. A.—Phillips, P. J.—Moon, H.: The FERET Verification Testing Protocol for Face Recognition Algorithm. Proceedings of the International Conference on Automatic Face and Gesture Recognition, 1998, pp. 48–53.

[17] Jain, A. K.—Robert, P. W.—Jianchang, M.: Statistical Pattern Recognition: A Review? IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, 2000, No. 1, pp. 4–37.

[18] Hancock, P. J. B.—Bruce, V.—Burton, M. A.: Recognition of Unfamiliar Faces Trends in Cognitive Science. Vol. 4, 2000, No. 9, pp. 330–337.

[19] Gauthier, I.—Nelson, C. A.: The Development of Face Expertise. Cognitive Neuroscience, 2001, No. 11, pp. 219–224.

[20] Phillips, P. J.—Moon, H.—Rizvi, S. A.—Rauss, P.: The FERET Testing Protocol of Face Recognition: from Theory to Application. (H. Vechsler, P. J. Phillips, V. Bruce, F. F. Soulie, and T. S. Huang), pp. 244–261, Springer-Verlag, Berlin, 1998.

[21] Messer, K.—Matas, J.—Kittler, J.—Luettin, J.—Maitre, G.: M2VTSDB: The Extended XM2VTS Database. Proceedings of the Interna-

tional Conference on Audio- and Video-Based Persin Authentication, 1999, pp. 72–79.

[22] Matas, J.—Hamous, M.—Jonsson, K. et al.: Comparison of Face Verification Results on the XM2VTS Database. Proceedings of the International Conference on Pattern Recognition, 2000, pp. 858–863.

[23] Ashok, S.—Prasana, A.—Iyengar: Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey. Pattern Recognition, Vol. 25, 1992, No. 1, pp. 65–77.

[24] Valentin, D.—Abdi, H.—Coole, A. J.—Cottrell, G. W.: Connectionist Models for Face Processing: A survey. Pattern Recognition, Vol. 27, 1994, No. 11, pp. 1209–1230.

[25] Chellappa, R.—Wilson, C. L.—Sirohey, S.: Human and Machine Recognition of Faces: A Survey? Proceedings of the IEEE, Vol. 83, 1995, No. 5, pp. 704–740.

[26] Fromherz, T.—Stucki, P.—Bichsel, M.: A Survey of Face Recognition. MML Technical Report, No. 97.01, Dept. Of Computer Secience, University of Zurich, Zurich 1997.

[27] Jie, Z.—Chunyu, L.—Changshui, Z. et al.: A Survey of Automatic Human Face Recognition. Acta Electronica Sinica, Vol. 28, 2000, No. 4, pp. 102–106.

[28] Fell, N.—Dempsey, P.: Whitehall Dithers on Airport Security. Ee Times, London, Jan 14, 2002.

[29] Blue Bell, Pa. and Littleton, Mass.: Firms Forge Alliance to Enhance. Businessworld, Manila, Dec 28, 2001.

[30] Pope, C.: Ways to Keep the Terrorist Grounded. Professional Engineering, Vol. 14, 2001, No. 19, pp. 20–21.

[31] Gorsuch, J.: Improved Security Through Technology. Overhaul & Maintenance, Washington, Nov 20, 2001.

[32] Candela, G. T.—Chellappa, R.: Comparative Performance of Classification Methods for Fingerprints. Technical Report, National Institute of Standards and Technology, 1993.

[33] Grother, P. J.—Candela, G. T.: Comparison of Hand-printed Digit Classifiers. Technical Report, National Institute of Standards and Technology, 1993.

[34] Zhao, W.—Chellappa, R.—Krishnaswamy, A.: Discriminant Analysis of Principal Components for Face Recognition. Proceedings of the International Conference on Automatic Face and Gesture Recognition, 1998, pp. 336–341.

[35] Moghaddam, B.—Nastar, C.—Pentland, A.: A Bayesian Similarity Measure for Direct Image Matching. Proceedings of the International Conference on Pattern Recognition, 1996, pp. 88–95.

[36] Wiskott, L.—Fellous, J. M.—von der Malsburg, C.: Face Recognition by Elastic Bunch Graph Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, 1997, No. 7, pp. 775–779.

[37] Jeng, S. H.—Liao, H. Y. M—Hua, C. C. et al.: Facial Feature Detection Using Geometrical Face Model: An Efficient Approach. Pattern Recognition. Vol. 31, 1998, No. 3, pp. 273–282.

[38] Bernögger, S.—Yin, L. J. et al.: Eye Tracking and Animation for MPEG-4 Coding. Proceedings of the International Conference on Pattern Recognition, 1998, pp. 1281–1284.

[39] Feng, G. C.—Yuen, P. C.: Variance Projection Function and Its Application to Eye Detection for Human Face Recognition. Pattern Recognition Letters, Vol. 19, 1998, No. 7, pp. 899–906.

[40] Ohta, H.—Saji, H.—Nakatani, H.: Recognition of Facial Expressions Using Muscle-Based Feature Methods. Proceedings of the International Conference on Pattern Recognition, 1998.

[41] Jingru, W.—Min, W.—Guangzheng, Y.: Knowledge-Based Edge Detection and Feature Extraction of Human-Face Organs. Pattern Recognition and Artificial Intelligence, Vol. 12, 1999, No. 3, pp. 340–346.

[42] Reisfeld, D.—Teshurun, Y.: Preprocessing of Face Images: Detection of Features and Pose Normalization. Computer Vision and Image Understanding, Vol. 71, 1998, No. 3, pp. 413–430.

[43] Jie, Z.—Chunyu, L.—Changshui, Z. et al.: Human Face Location Based on Directional Symmetry Transform. Chinese Acta Electronica Sinica, Vol. 27, 1999, No. 8, pp. 12–15.

[44] Lv, X.—Jie, Z.—Zhang, C.: A Novel Algorithm for Rotated Human Face Detection. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 2000.

[45] Kwon, Y. H.—Lobo, N. D.: Age Classification from Facial Images. Computer Vision and Image Understanding, Vol. 74, 1999, No. 1, pp. 1–21.

[46] Hai, T.—Huang, T. S.: Explanation-Based Facial Motion Tracking Using a Piecewise Bzier Volume Deformation Model. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 1999.

[47] Ying-LI, T.—Kanade, T.—Cohn, J. F.: Recognizing Upper Face Action Units for Facial Expression Analysis. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 1999.

[48] Decarlo, D.—Metaxas, D.: Optical Flow Constraints on Deformable Models with Applications to Face Tracking. International Journal of Computer Vision, Vol. 38, 2000, No. 2, pp. 99–127.

[49] Jianguo, W.—Tieniu, T.: A New Face Detection Method Based on Shape Information. Pattern Recognition Letters, Vol. 21, 2000, No. 3, pp. 463–471.

[50] Roach, M. J.—Brand, J. D.—Mason, J. S. D.: Acoustic and Facial Features for Speaker Recognition. Proceedings of the International Conference on Pattern Recognition, 2000.

[51] Lin, C.—Kuo-Chin, F.: Human Face Detection Using Geometric Triangle Relationship. Proceedings of the International Conference on Pattern Recognition, 2000.

[52] Zhong, J.—Mariani, R.: Glasses Detection and Extraction by Deformable Contour. Proceedings of the International Conference on Pattern Recognition, 2000.

[53] Jiang-Gang, W.—Sung, E.: Pose Determination of Human Faces by Using Vanishing Points. Pattern Recognition, Vol. 34, 2001, No. 12, pp. 2427–2445.

[54] Ying, L.—Xiang-Lin, Q.—Yun-Jiu, W.: Eye Detection by Using Fuzzy Template Matching and Feature-Parameter-Based Judgement. Pattern Recognition Letters, Vol. 22, 2001, No. 10, pp. 1111–1124.

[55] Sclaroff, S.—Lifeng, L.: Deformable Shape Detection and Description via Model-Based Region Grouping. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, 2001, No. 5, pp. 475–489.

[56] Shinn-Ying, H.—Hui-Ling, H.: Facial Modeling from an Uncalibrated Face Image Using a Coarse-to-Fine Genetic Algorithm. Pattern Recognition, Vol. 34, 2001, No. 9, pp. 1015–1031.

[57] Wong, K. W.—Lam, K. M.—Siu, W.-C.: An Efficient Algorithm for Human Face Detection and Facial Feature Extraction under Different Conditions. Pattern Recognition, Vol. 34, 2001, No. 10, pp. 1993–2004.

[58] Amit, Y.—Geman, D.: A Computational Model for Visual Selection. Neural Computation, 1999, No. 11, pp. 1691–1715.

[59] Fleuret, F.: Coarse-to-Fine Face Detection. International Journal of Computer Vision, Vol. 41, 2001, Nos. 1–2, pp. 85–107.

[60] Terrillon, J.-C.—David, M.—Akamatsu, S.: Detection of Human Faces in Complex Scene Images by Use of a Skin Color Model and of Invariant Fourier-Mellion Moments. Proceedings of the International Conference on Pattern Recognition, 1998.

[61] Terrillon, J.-C.—David, M.—Akamatsu, S.: Detection of Human Faces in Complex Scene Images by Use of a Skin Color Model and of Invariant Fourier-Mellion Moments. Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998.

[62] Yang, M.-H.—Ahuja, N.: Detection of Human Faces in Color Images. Proceedings of International Conference on Image Processing, 1998.

[63] Bakiá, V.—Stockman, G.: Real-Time Tracking of Face Features and Gaze Determination. Proceedings Fourth IEEE Workshop on Applications of Computer Vision, 1998.

[64] Albol, A.—Charles, A.—Edward, J. D.: Face Detection for Pseudo-Semantic Labeling in Video Databases. Proceedings of International Conference on Image Processing, 1999.

[65] Garcia, C.—Tziritas, G.: Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis. IEEE Transactions on Multimedia, Vol. 1, 1999, No. 3, pp. 264–277.

[66] Chunyu, L.—Changshui, Z.—Fang, W. et al.: Regional Feature Based Fast Human Face Detection. Journal of Tsinghua University, Vol. 39, 1999, No. 1, pp. 101–105.

[67] Tao, X.—Yun, H.—Chongxi, F.: Automatic Facial Feature Detection and Its Application in Image Coding. Journal of Tsinghua University, Vol. 39, 1999, No. 5, pp. 81–84.

[68] Li, Y.—Goshtasby, A.—Garcia, O.: Detecting and Tracking Human Faces in Videos. Proceedings of the International Conference on Pattern Recognition, 2000.

[69] Ying, W.—Huang, T. S.: Color Tracking by Transductive Learning. Proceedings of the International Conference on Pattern Recognition, 2000.

[70] Gang, W.—Dongge, L. V.—Sethi, I. K.: Detection of Side-View Faces in Color Images. Proceedings, Fourth IEEE Workshop on Applications of Computer Vision, 2000.

[71] Störring, M.—Andersen, H. J.—Granum, E.: Estimation of the Illuminant Color from Human Skin Color. Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition, 2000.

[72] Schwerdt, K.—Crowley, J. L.: Robust Face Tracking Using Color. Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition, 2000.

[73] Lu-Hong, L.—Hai-Zhou, A.—Ke-Zhong, H. et al.: Single Rotated Face Location Based on Affine Template Matching. Chinese Journal of Computers, Vol. 23, 2000, No. 6, pp. 640–645.

[74] Dass, S. C.—Jain, A. K.: Markov Face Models. Proceedings of the International Conference on Computer Vision, 2001.

[75] Jun, Z.—Jie, G.—Guilin, Z. et al.: Fast Hierarchical Knowledge-Based Approach for Human Face Detection in Color Images. Proceedings of the International Society for Optical Engineering in Wuhan, 2001.

[76] Duta, N.—Jain, A. K.: Learning the Human Face Concept in Black and White Images. Proceedings of the International Conference on Pattern Recognition, 1998.

[77] Cascia, M. L.—Sclaoff, S.: Fast, Reliable Head Tracking under Varying Illumination. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 1999.

[78] Descombes, X.—Sigelle, M.—Preteux, F.: Estimating Gaussian Markov Random Field Parameters in a Nonstationary Framework: Application to Remote Sensing Imaging. IEEE Transactions on Image Processing, Vol. 8, 1999, No. 3, pp. 490–503.

[79] Fan, L.—Sung, K. K.: A Combined Feature-Texture Similarity Measure for Face Alignment under Varying Pose. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 2000.

[80] Li, Y.—Peng, J.: A New Estimation of Hurst Parameter for Texture Analysis. Proceedings of the International Society for Optical Engineering in Wuhan, 2001.

[81] Hui, J.—Wen, G.: The Human Facial Combined Expression Recognition System. Chinese Journal of Computers, Vol. 23, 2000, No. 6, pp. 602–608.

[82] Bobick, A. F.—Davis, J. W.: The Recognition of Human Movement Using Temporal Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, 2001, No. 3, pp. 257–267.

[83] Ming, X.—Akatsuka, T.: Multi-Model Method for Detection of a Human Face from Complex Backgrounds. Proceedings of the International Society for Optical Engineering, 1998, pp. 793–802.

[84] Wee, S.—Ji, S.—Yoon, C. et al.: Face Detection Using Pattern Information and Deformable Template in Motion Images. Proceedings of the Fifith International Conference on Soft Computing and Information / Intelligence Systems, 1998, pp. 213–216.

[85] Aggarwal, J.—Cai, Q.: Human Motion Analysis: A Review. Computer Vision and Image Understanding, Vol. 73, 1999, No. 3, pp. 428–440.

[86] Moeslund, T. B.—Granum, E.: Multiple Cues used in Model-Based Human Motion Capture. Proceedings of Field Gun, France, 2000, No. 3.

[87] Mirhosseini, A. R.—Hong, Y.: Human Face Recognition: An Evidence Aggregation Approach. Computer Vision and Image Understanding, Vol. 71, 1998, No. 2, pp. 213–230.

[88] Yokoyama, T.—Yagi, Y.—Yachida, M.: Active Contour Model for Extracting Human Faces. Proceedings of the International Conference on Pattern Recognition, 1998.

[89] Jebara, T. S.—Pentland, A.: Parameterized Structure from Motion for 3D Adaptive Feedback Tracking of Faces. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 1998.

[90] Sun, Q. B.—Huang, W. M.—Wu, J. K.: Face Detection Based on Color and Local Symmetry Information. Proceedings of First IEEE International Conference on Automatic Face and Gesture Recognition, 1998.

[91] Romdhani, S.—Shaogang, G.—Psarrou, A.: A Generic Face Appearance Model of Shape and Texture under Very Large Pose Variations from Profile to Profile Views. Proceedings of the International Conference on Pattern Recognition, 2000.

[92] Yong, Y.—Challapali, K.: A System for the Automatic Extraction of 3-D Facial Feature Points for Face Model Calibration. Proceedings of International Conference on Image Processing, 2000.

[93] Weimin, H.—Qinib, S.—Lam, C.-P. et al.: A Robust Approach to Face and Eyes Detection from Images with Cluttered Background. Proceedings of the International Conference on Pattern Recognition, 1998.

[94] Weimin, H.—Mariani, R.: Face Detection and Precise Eyes Location. Proceedings of the International Conference on Pattern Recognition, 2000.

[95] Reiter, M.—Matas, J.: Object-Detection with a Varying Number of Eigenspace Projections. Proceedings of the International Conference on Pattern Recognition, 1998.

[96] Sung, K. K.—POGGIO, T.: Example-Based Learning for View-Based Human Face Detection. IEEE Transactions on Pattern analysis and Machine Intelligence, Vol. 20, 1998, No. 1, pp. 39–51, 1998.

[97] Ying, Z.—Schwartz, S.—Orchard, M.: Fast Face Detection Using Subspace Discriminant Wavelet Features. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 2000.

[98] Papageorgiou, C. P.—Oren, M.—Poggio, T.: A General Framework for Object Detection. Proceedings of the International Conference on Computer Vision, 1998, pp. 555–562.

[99] Pontil, M.—Verri, A.: Support Vector Machine for 3D Object Recognition. IEEE Transactions on Pattern analysis and Machine Intelligence, Vol. 20, 1998, No. 6, pp. 637–646.

[100] Roobaert, D.—van Hulle, M. M.: View-Based 3D Object Recognition with Support Vector Machines. Proceedings of IEEE International Workshop Neural Networks for Signal Processing VIII, 1998.

[101] Terrillon, J.-C.—Shirazi, M. N.—Sadek, M.: Invariant Face Detection with Support Vector Machines. Proceedings of the International Conference on Pattern Recognition, 2000.

[102] Chao, Y.—Chang-Shui, Z.: Multiple Template Matching for Frontal-View Face Detection. Acta Electronica Sinica, Vol. 38, 2000, No. 3, pp. 95–98.

[103] Keren, D.—Osadchy, M.—Gotsman, C.: Antiface: A Novel, Fast Method for Image Detection. IEEE Transactions on Pattern analysis and Machine Intelligence, Vol. 23, 2001, No. 7, pp. 747–761.

[104] Fang, W.—Jie, Z.—Changshui, Z.: LLM Netural Networks and Compensation for Light Condition Based Color Face Detection. Journal of Tsinghua University, Vol. 39, 1999, No. 9, pp. 37–40.

[105] Rowley, H. A.—Baluja, S.—Kanade, T.: Neural Network-Based Face Detection. IEEE Transactions on Pattern analysis and Machine Intelligence, Vol. 20, 1998, No. 1, pp. 23–30.

[106] Hui, P.—Changshui, Z.—Zhaoqi, B.: A fully Automated Face Recognition System under Different Conditions. Proceedings of the International Conference on Pattern Recognition, 1998.

[107] Rowley, H. A.—Baluja, S.—Kanade, T.: Rotation Invariant Neural Network-Based Face Detection. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 1998.

[108] Qian, G.—Li, S. Z.: Combining Feature Optimization into Neural Network Based Face Detection. Proceedings of the International Conference on Pattern Recognition, 2000.

[109] Terrillon, J.-C.—McReynolds, D.—Sadek, M.: Invariant Neural-Network Based Face Detection with Orthogonal Fourier-Mellin Moments. Proceedings of the International Conference on Pattern Recognition, 2000.

[110] Fraud, R.—Bernier, O.-J.—Viallet, J.-E. et al.: A Fast and Accurate Face Detector Based on Neural Networks. IEEE Transactions on Pattern analysis and Machine Intelligence, Vol. 23, 2001, No. 1, pp. 42–53.

[111] Haiuan, W.—Qian, Ch.—Yachida, M.: Face Detection from Color Image Using a Fuzzy Pattern Matching Method. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, 1999, No. 6, pp. 557–563.

[112] Lier, J. J.-J.—Kanade, T.—Cohn, J. F. et al.: Subtly Different Facial Expression Recognition and Expression Intensity Estimation. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 1998.

[113] Yoon, K. S.—Ham, Y. K.—Park, R. H.: Hybrid Approaches to Frontal View Face Recognition Using the Hidden Markov Model and Neural Network. Pattern Recognition, Vol. 31, 1998, No. 3, pp. 283–293.

[114] Leonardis, A.—Bischof, H.: Robust Recognition Using Eigenimages. Computer Vision and Image Understanding, Vol. 78, 2000, No. 1, pp. 99–118.

[115] Yilmaz, A.—Gökmen, M.: Eigenhill vs. Eigenface and Eigenedge. Pattern Recognition, Vol. 34, 2001, No. 1, pp. 181–184.

[116] Hafed, Z. M.—Levine, M.: Face Recognition Using the Discrete Cosine Transform. International Journal of Computer Vision, Vol. 43, 2001, No. 3, pp. 167–188.

[117] Zhong, J.—Jing-Yu, Y.—Zhong-Shan, H. et al.: Face Recognition Based on the Uncorrelated Discriminant Transform. Pattern Recognition, Vol. 34, No. 4, pp. 1405–1416.

[118] Cappelli, R.—Miao, D.—Maltoni, D.: Multispace KL for Pattern Representation and Classification. IEEE on Pattern Analysis and Machine Intellegence, Vol. 23, 2001, No. 9, pp. 977–996.

[119] Ramasubramanian, D.—Venkatesh, Y.V.: Encoding and Recognition of Faces Based on the Human Visual Model and DCT. Pattern Recognition, Vol. 34, 2001, No. 8, pp. 2447–2458.

[120] Lai, J. H.—Yuen, P. C.—FENG, G. C.: Face Recognition Using Holistic Fourier Invariant Features. Pattern Recognition, Vol. 34, 2001, No. 1, pp. 95–109.

[121] Kotropoulos, C.—Tefas, A.—Pitas, I.: Morphological Elastic Graph Matching Applied to Frontal Face Authentication under Well-Controlled and Real Conditions. Pattern Recognition, Vol. 33, 2000, No. 9, pp. 1936–1947.

[122] Kotropoulos, C.—Tefas, A.—Pitas, I.: Frontal Face Authentication Using Morphological Elastic Graph Matching. IEEE Transactions on Image Processing, Vol. 9, 2000, No. 4, pp. 555–560.

[123] Tefas, A.—Kotropoulos, C.—Pitas, I.: Using Support Vector Machines to Enhance the Performance of Elastic Graph Matching for Frontal Face Authentication. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, 2001, No. 7, pp. 735–746.

[124] Gyu-Tae, P.—Bien, Z.: Neural Network-Based Fuzzy Observer with Application to Facial Analysis. Pattern Recognition Letters, Vol. 21, 2000, No. 1, pp. 93–105.

[125] Ryu, Y.-S.—Oh, S.-Y.: Automatic Extraction of Eye and Mouth Fields from a Face Image Using Eigenfeatures and Multilayer Perceptirons. Pattern Recognition, Vol. 34, 2001, No. 8, pp. 2459–2466.

[126] Ying-Li, T.—Kanade, T.—Cohn, J. F.: Recognizing Action Units for Facial Expression Analysis. IEEE on Pattern Analysis and Machine Intelligence, Vol. 23, 2001, No. 2, pp. 97–115.

[127] Nikolaidis, A.—Pitas, I.: Facial Feature Extraction and Pose Determination. Pattern Recognition, Vol. 33, 2000, No. 5, pp. 1783–1791.

[128] Martinez, A. M.—Kak, A. C.: PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, 2001, No. 2, pp. 228–233.

[129] Turk, M.—Pentland, A.: Face Recognition Using Eigenfaces. Proceedings of International Conference on Computer Vision and Pattern Recogntion, 1991, pp. 586–591.

[130] Liu, C.—Wechsler, H.: Enhanced Fisher Linear Disciminant Models for Face Recognition. Proceedings of International Conference on Pattern recognition, 1998.

[131] Liu, C.—Wechsler, H.: Face Recognition Using Evolutionary Pursuit. Proceedings of European Conference on Computer Vision, 1998.

[132] Chnegjun, L.—Wechsler, H.: Evolutionary Pursuit and Its application to Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, 2000, No. 6, pp. 570–582.

[133] Georghiades, A. S.—Belhumeur, P. N.—Kriegman, D. J.: From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, 2001, No. 6, pp. 643–660.

[134] Moghaddam, B.—Jebara, T.—Pentland, A.: Bayesian Faec Recognition. Pattern Recognition, Vol. 33, 2000, No. 7, pp. 1771–1782.

**Yongzhong Lu** is a lecturer in the School of Computer Science and Technology in Huazhong University of Science and Technology in China. His current interests include image processing and analysis, computer vision and pattern recognition, CAD/CAM, engineering optimization, dynamical modeling and simulation, system control engineering.

**Jingli Zhou** is a doctoral supervisor, professor in Department of Computer Science and Technology in Huazhong University of Science and Technology in China, and member of IEEE. Her research interests include computer storage system and multimedia technology.

**Shengsheng Yu**is a doctoral supervisor, professor in Department of Computer Science and Technology in Huazhong University of Science and Technology in China, and member of IEEE. His research interests include computer architecture and networking communication technology.