Face Recognition Using Line Edge Map

Yongsheng Gao, *Member*, *IEEE*, and Maylor K.H. Leung, *Member*, *IEEE*

Abstract—The automatic recognition of human faces presents a significant challenge to the pattern recognition research community. Typically, human faces are very similar in structure with minor differences from person to person. They are actually within one class of "human face." Furthermore, lighting condition changes, facial expressions, and pose variations further complicate the face recognition task as one of the difficult problems in pattern analysis. This paper proposed a novel concept, "faces can be recognized using line edge map." A compact face feature, Line Edge Map (LEM), is generated for face coding and recognition. A thorough investigation on the proposed concept is conducted which covers all aspects on human face recognition, i.e., face recognition, under 1) controlled/ideal condition and size variation, 2) varying lighting condition, 3) varying facial expression, and 4) varying pose. The system performances are also compared with the eigenface method, one of the best face recognition techniques, and reported experimental results of other methods. A face prefiltering technique is proposed to speed up the searching process. It is a very encouraging finding that the proposed face recognition technique has performed superior to the eigenface method in most of the comparison experiments. This research demonstrates that LEM together with the proposed generic line segment Hausdorff distance measure provide a new way for face coding and recognition.

Index Terms—Face recognition, line edge map, line segment Hausdorff distance, structural information.

1 Introduction

Computerized human face recognition has been an active research area for the last 20 years. It has many practical applications, such as bankcard identification, access control, mug shots searching, security monitoring, and surveillance systems [1]. Face recognition is used to identify one or more persons from still images or a video image sequence of a scene by comparing input images with faces stored in a database. It is a biometric system that employs automated methods to verify or recognize the identity of a living person based on his/her physiological characteristic. In general, a biometric identification system makes use of either physiological characteristics (such as a fingerprint, iris pattern, or face) or behavior patterns (such as handwriting, voice, or key-stroke pattern) to identify a person. Because of human inherent protectiveness of his/her eyes, some people are reluctant to use eye identification systems. Face recognition has the benefit of being a passive, nonintrusive system to verify personal identity in a "natural" and friendly way.

The application of face recognition technology can be categorized into two main parts: law enforcement application and commercial application. Face recognition technology is primarily used in law enforcement applications, especially mug shot albums (static matching) and video surveillance (real-time matching by video image sequences). The commercial applications range from static matching of photographs on credit cards, ATM cards, passports, driver's licenses, and photo ID to real-time matching with still images or video image sequences for access control. Each presents different constraints in terms of processing requirement.

To the best of our knowledge, there is hardly any reported research work on face recognition using edge curves of faces except the outlines of face profiles. The only most related work was done by Takács [2] using binary image metrics. The face coding and matching techniques presented in this paper are different from [2]. Takács used edge map face coding and pixel-wise Hausdorff distance template matching techniques, while this paper proposed line-based face coding and line matching techniques to integrate geometrical and structural features in the template matching. A novel concept, "faces can be recognized using line edge map," is proposed here. A compact face feature, Line Edge Map (LEM), is extracted for face coding and recognition. A feasibility investigation and evaluation for face recognition based solely on face LEM is conducted-which covers all conditions of human face recognition, i.e., face recognition under controlled/ideal condition, varying lighting condition, varying facial expression, and varying pose. The system performances are compared with the eigenface method, one of the best face recognition techniques, and reported experimental results of other methods. It is a very encouraging finding that the proposed face recognition technique has performed consistently superior to (or equally well as) the eigenface method in all the comparison experiments except under large facial expression changes. A prefiltering scheme (two-stage identification) is proposed to speed up the searching using a 2D prefiltering vector derived from the face LEM. This research demonstrates that LEM together with the proposed generic Line Segment Hausdorff Distance measure provide a new way for face coding and recognition.

Some of the ideas presented in this paper were initially reported in [3]. In this paper, we report the full and new formulation and extensive experimental evaluation of our

• *The authors are with the School of Computer Engineering, Nanyang Technological University, Nanyang Ave., 639798, Singapore. E-mail: {asysgao, asmkleung}@ntu.edu.sg.*

Manuscript received 15 Dec. 2000; revised 17 Aug. 2001; accepted 28 Nov. 2001.
Recommended for acceptance by Z. Zhang.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 113312.

techniques. In the following, a literature review of face recognition techniques is given in Section 2. Most successful approaches to frontal face recognition, namely, eigenface, neural network, dynamic link architecture, hidden Markov model, geometrical feature matching, and template matching are discussed. Section 3 describes the concepts of face recognition using line edge map. A novel line segment Hausdorff distance measure for human face recognition is proposed in Section 4. The advantages of the proposed approach are discussed and compared with existing pixel wise Modified Hausdorff Distance method. In Section 5, the system is extensively examined on face recognition under controlled/ideal condition, size variation, varying lighting condition, varying expression, and varying pose. The storage and computational requirements are analyzed, and the system performance is compared with existing approaches. Section 6 presents a two-stage face identification schemewhich speeds up the searching process by filtering out part of the unlikely candidates. Finally, the paper concludes in Section 7.

2 Background

This section overviews the major human face recognition techniques that apply mostly to frontal faces. The methods considered are eigenface (eigenfeature), neural network, dynamic link architecture, hidden Markov model, geometrical feature matching, and template matching. The approaches are analyzed in terms of the facial representations they used.

Eigenface is one of the most thoroughly investigated approaches to face recognition. It is also known as Karhunen-Loève expansion, eigenpicture, eigenvector, and principal component. Sirovich and Kirby [5] and Kirby et al. [6] used principal component analysis to efficiently represent pictures of faces. They argued that any face images could be approximately reconstructed by a small collection of weights for each face and a standard face picture (eigenpicture). The weights describing each face are obtained by projecting the face image onto the eigenpicture. Turk and Pentland [7] used eigenfaces, which was motivated by the technique of Kirby and Sirovich, for face detection and identification. In mathematical terms, eigenfaces are the principal components of the distribution of faces, or the eigenvectors of the covariance matrix of the set of face images. The eigenvectors are ordered to represent different amounts of the variation, respectively, among the faces. Each face can be represented exactly by a linear combination of the eigenfaces. It can also be approximated using only the "best" eigenvectors with the largest eigenvalues. The best M eigenfaces construct an M dimensional space, i.e., the "face space." The authors reported 96 percent, 85 percent, and 64 percent correct classifications averaged over lighting, orientation, and size variations, respectively. Their database contained 2,500 images of 16 individuals. As the images include a large quantity of background area, the above results are influenced by background. The authors explained the robust performance of the system under different lighting conditions by significant correlation between images with changes in illumination. However, Grudin [26] showed that the correlation between images of the whole faces is not efficient for satisfactory recognition performance. An illumination normalization [6] is usually necessary for the eigenface approach. Zhao and Yang [4] proposed a new method to compute the covariance matrix using three images each taken in different lighting conditions to account for arbitrary illumination effects, if the object is Lambertian. Pentland et al. [8] extended their early work on eigenface to eigenfeatures corresponding to face components, such as eyes, nose, and mouth. They used a modular eigenspace which was composed of the above eigenfeatures (i.e., eigeneyes, eigennose, and eigenmouth). This method would be less sensitive to appearance changes than the standard eigenface method. The system achieved a recognition rate of 95 percent on the FERET database of 7,562 images of approximately 3,000 individuals. In summary, eigenface appears as a fast, simple, and practical method. However, in general, it does not provide invariance over changes in scale and lighting conditions.

The attractiveness of using neural network could be due to its nonlinearity in the network. Hence, the feature extraction step may be more efficient than the linear Karhunen-Loève methods. One of the first artificial neural network (ANN) techniques used for face recognition is a single layer adaptive network called WISARD which contains a separate network for each stored individual [9]. The way in constructing a neural network structure is crucial for successful recognition. It is very much dependent on the intended application. For face detection, multilayer perceptron [10] and convolutional neural network [11] have been applied. For face verification, Cresceptron [12] is a multiresolution pyramid structure. Lawrence et al. [11] proposed a hybrid neural network which combined local image sampling, a self-organizing map (SOM) neural network, and a convolutional neural network. The SOM provides a quantization of the image samples into a topological space where inputs that are nearby in the original space are also nearby in the output space, thereby providing dimension reduction and invariance to minor changes in the image sample. The convolutional network extracts successively larger features in a hierarchical set of layers and provides partial invariance to translation, rotation, scale, and deformation. The authors reported 96.2 percent correct recognition on ORL database of 400 images of 40 individuals. The classification time is less than 0.5 second, but the training time is as long as 4 hours. Lin et al. [13] used probabilistic decision-based neural network (PDBNN) which inherited the modular structure from its predecessor, a decision based neural network (DBNN) [14]. The PDBNN can be applied effectively to 1) *face detector:* which finds the location of a human face in a cluttered image, 2) *eye localizer*: which determines the positions of both eyes in order to generate meaningful feature vectors, and 3) *face recognizer.* A hierarchical neural network structure with nonlinear basis functions and a competitive credit-assignment scheme was adopted. PDBNN-based biometric identification system has the merits of both neural networks and statistical approaches, and its distributed computing principle is relatively easy to implement on parallel computer. In [13], it was reported that PDBNN face recognizer had the capability of recognizing up to 200 people and could achieve up to 96 percent correct recognition rate in approximately 1 second. However, when the number of persons increases, the computing expense will

become more demanding. In general, neural network approaches encounter problems when the number of classes (i.e., individuals) increases. Moreover, they are not suitable for a single model image recognition task because multiple model images per person are necessary in order for training the systems to "optimal" parameter setting.

Graph matching is another approach to face recognition. Lades et al. [15] presented a dynamic link structure for distortion invariant object recognition which employed elastic graph matching to find the closest stored graph. Dynamic link architecture is an extension to classical artificial neural networks. Memorized objects are represented by sparse graphs, whose vertices are labeled with a multiresolution description in terms of a local power spectrum and whose edges are labeled with geometrical distance vectors. Object recognition can be formulated as elastic graph matching which is performed by stochastic optimization of a matching cost function. They reported good results on a database of 87 people and a small set of office items comprising different expressions with a rotation of 15 degrees. The matching process is computationally expensive, taking about 25 seconds to compare with 87 stored objects on a parallel machine with 23 transputers. Wiskott and von der Malsburg [16] extended the technique and matched human faces against a gallery of 112 neutral frontal view faces. Probe images were distorted due to rotation in depth and changing facial expression. Encouraging results on faces with large rotation angles were obtained. They reported recognition rates of 86.5 percent and 66.4 percent for the matching tests of 111 faces of 15 degree rotation and 110 faces of 30 degree rotation to a gallery of 112 neutral frontal views. In general, dynamic link architecture is superior to other face recognition techniques in terms of rotation invariant; however, the matching process is computationally expensive.

Stochastic modeling of nonstationary vector time series based on hidden Markov models (HMM) has been very successful for speech applications. Samaria and Fallside [27] applied this method to human face recognition. Faces were intuitively divided into regions such as the eyes, nose, mouth, etc., which can be associated with the states of a hidden Markov model. Since HMMs require a one-dimensional observation sequence and images are two-dimensional, the images should be converted into either 1D temporal sequences or 1D spatial sequences. In [28], a spatial observation sequence was extracted from a face image by using a band sampling technique. Each face image was represented by a 1D vector series of pixel observation. Each observation vector is a block of L lines and there is an M lines overlap between successive observations. An unknown test image is first sampled to an observation sequence. Then, it is matched against every HMMs in the model face database (each HMM represents a different subject). The match with the highest likelihood is considered the best match and the relevant model reveals the identity of the test face. The recognition rate of HMM approach is 87 percent using ORL database consisting of 400 images of 40 individuals. A pseudo 2D HMM [28] was reported to achieve a 95 percent recognition rate in their preliminary experiments. Its classification time and training time were not given (believed to be very expensive). The choice of parameters had been based on subjective intuition.

Geometrical feature matching techniques are based on the computation of a set of geometrical features from the picture of a face. The fact that face recognition is possible even at coarse resolution as low as 8×6 pixels [17] when the single facial features are hardly revealed in detail, implies that the overall geometrical configuration of the face features is sufficient for recognition. The overall configuration can be described by a vector representing the position and size of the main facial features, such as eyes and eyebrows, nose, mouth, and the shape of face outline. One of the pioneering works on automated face recognition by using geometrical features was done by Kanade [19] in 1973. Their system achieved a peak performance of 75 percent recognition rate on a database of 20 people using two images per person, one as the model and the other as the test image. Goldstein et al. [20] and Kaya and Kobayashi [18] showed that a face recognition program provided with features extracted manually could perform recognition apparently with satisfactory results. Bruneli and Poggio [21] automatically extracted a set of geometrical features from the picture of a face, such as nose width and length, mouth position, and chin shape. There were 35 features extracted to form a 35 dimensional vector. The recognition was then performed with a Bayes classifier. They reported a recognition rate of 90 percent on a database of 47 people. Cox et al. [22] introduced a mixture-distance technique which achieved 95 percent recognition rate on a query database of 685 individuals. Each face was represented by 30 manually extracted distances. Manjunath et al. [23] used Gabor wavelet decomposition to detect feature points for each face image which greatly reduced the storage requirement for the database. Typically, 35-45 feature points per face were generated. The matching process utilized the information presented in a topological graphic representation of the feature points. After compensating for different centroid location, two cost values, the topological cost, and similarity cost, were evaluated. The recognition accuracy in terms of the best match to the right person was 86 percent and 94 percent of the correct person's face was in the top three candidate matches. In summary, geometrical feature matching based on precisely measured distances between features may be most useful for finding possible matches in a large database such as a mug shot album. However, it will be dependent on the accuracy of the feature location algorithms. Current automated face feature location algorithms do not provide a high degree of accuracy and require considerable computational time.

A simple version of template matching is that a test image represented as a two-dimensional array of intensity values is compared using a suitable metric, such as the Euclidean distance, with a single template representing the whole face. There are several other more sophisticated versions of template matching on face recognition. One can use more than one face template from different viewpoints to represent an individual's face. A face from a single viewpoint can also be represented by a set of multiple distinctive smaller templates [24], [21]. The face image of gray levels may also

be properly processed before matching [25]. In [21], Bruneli and Poggio automatically selected a set of four features templates, i.e., the eyes, nose, mouth, and the whole face, for all of the available faces. They compared the performance of their geometrical matching algorithm and template matching algorithm on the same database of faces which contains 188 images of 47 individuals. The template matching was superior in recognition (100 percent recognition rate) to geometrical matching (90 percent recognition rate) and was also simpler. Since the principal components (also known as eigenfaces or eigenfeatures) are linear combinations of the templates in the data basis, the technique cannot achieve better results than correlation [21], but it may be less computationally expensive. One drawback of template matching is its computational complexity. Another problem lies in the description of these templates. Since the recognition system has to be tolerant to certain discrepancies between the template and the test image, this tolerance might average out the differences that make individual faces unique. In general, template-based approaches compared to feature matching are a more logical approach.

In summary, no existing technique is free from limitations. Further efforts are required to improve the performances of face recognition techniques, especially in the wide range of environments encountered in real world. Edge information is a useful object representation feature that is insensitive to illumination changes to certain extent. Though the edge map is widely used in various pattern recognition fields, it has been neglected in face recognition except in recent work reported in [2]. Face recognition employing the spatial information of edge map associated with local structural information remains an unexplored area. This paper proposed a novel approach that exploits such information. The eigenface technique provides a compact representation of the human face which is optimal for face reconstruction. It is one of the most thoroughly investigated approaches and has demonstrated excellent performance. Hence, this technique is used, in this study, as a baseline for recognition performance comparison.

3 Line Edge Map

Cognitive psychological studies [48], [49] indicated that human beings recognize line drawings as quickly and almost as accurately as gray-level pictures. These results might imply that edge images of objects could be used for object recognition and to achieve similar accuracy as gray-level images. Takács [2] made use of edge maps, which was motivated by the above finding, to measure the similarity of face images. The faces were encoded into binary edge maps using Sobel edge detection algorithm. The Hausdorff distance was chosen to measure the similarity of the two point sets, i.e., the edge maps of two faces, because the Hausdorff distance can be calculated without an explicit pairing of points in their respective data sets. The modified Hausdorff distance in the formulation of $h(A,B) = \frac{1}{N_a}\sum_{a\in A} \min_{b\in B}||a-b||$ was used, as it is less sensitive to noise than the maximum or kth ranked Hausdorff distance formulations. A 92 percent accuracy was achieved in their experiments. Takács argued that the process of face recognition might start at a much earlier stage and edge images can be used for the recognition of faces without the involvement of high-level cognitive functions. This is in accordance with the psychological reports of [48], [49]. However, the Hausdorff distance uses only the spatial information of an edge map without considering the inherent local structural characteristics inside such a map. Bruneli and Poggio [21] argued *that successful object recognition approaches might need to combine aspects of feature-based approaches with template matching method*. This is a valuable hint for us when proposing a Line Edge Map (LEM) approach which extracts lines from a face edge map as features. This approach can be considered as a combination of template matching and geometrical feature matching. The LEM approach not only possesses the advantages of feature-based approaches, such as invariant to illumination and low memory requirement, but also has the advantage of high recognition performance of template matching. The above three reasons together with the fact that edges are relatively insensitive to illumination changes motivated this research.

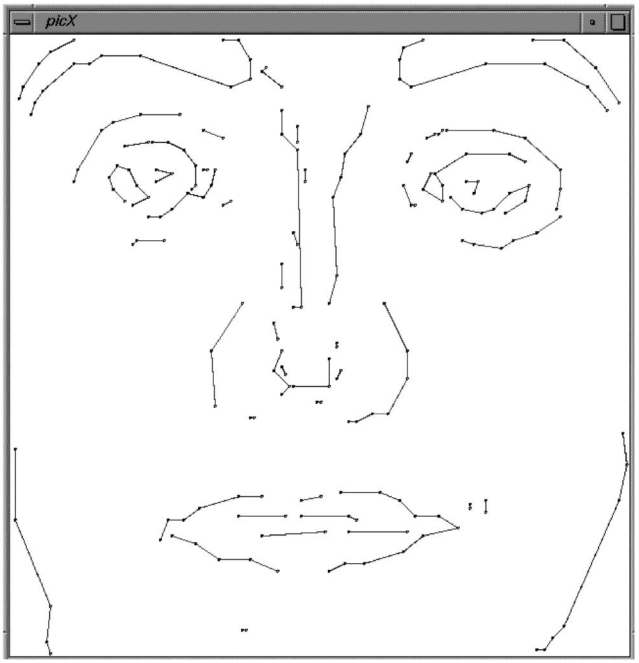

Fig. 1. An illustration of a face LEM.

A novel face feature representation, Line Edge Map (LEM), is proposed here to integrate the structural information with spatial information of a face image by grouping pixels of face edge map to line segments. After thinning the edge map, a polygonal line fitting process [42] is applied to generate the LEM of a face. An example of a human frontal face LEM is illustrated in Fig. 1. The LEM representation, which records only the end points of line segments on curves, further reduces the storage requirement. Efficient coding of faces is a very important aspect in a face recognition system. LEM is also expected to be less sensitive to illumination changes due to the fact that it is an intermediate-level image representation derived from low-level edge map representation. The basic unit of LEM is the line segment grouped from pixels of edge map. In this study, we explore the information of LEM and investigate the feasibility and efficiency of human face recognition using LEM. A novel Line Segment Hausdorff Distance

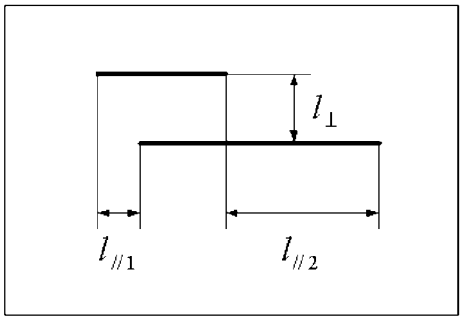

Fig. 2. Line displacement measures.

(LHD) measure is then proposed to match LEMs of faces. LHD has better distinctive power because it can make use of the additional structural attributes of line orientation, line-point association, and number disparity in LEM, i.e., it is not encouraged to match two lines with large orientation difference, and all the points on one line have to match to points on another line only.

4 Line Segment Hausdorff Distance

A novel disparity measure, Line Segment Hausdorff Distance (LHD), is designed to measure the similarity of face LEMs. The LHD is a shape comparison measure based on LEMs. It is a distance defined between two line sets. Unlike most shape comparison methods that build a one-to-one correspondence between a model and a test image, LHD can be calculated without explicit line correspondence to deal with the broken line problem caused by segmentation error. The LHD for LEM matching is more tolerant to perturbations in the locations of lines than correlation techniques since it measures proximity rather than exact superposition.

Given two LEMs $M^l = \{m_1^l, m_2^l, \ldots, m_p^l\}$ (representing a model LEM in the database) and $T^l = \{t_1^l, t_2^l, \ldots, t_q^l\}$ (representing an input LEM), LHD is built on the vector $\vec{d}(m_i^l, t_j^l)$ that represents the distance between two line segments m_i^l and t_j^l. The vector is defined as

$$\vec{d}(m_i^l, t_j^l) = \begin{bmatrix} d_\theta\left(m_i^l, t_j^l\right) \\ d_{//}\left(m_i^l, t_j^l\right) \\ d_\perp\left(m_i^l, t_j^l\right) \end{bmatrix},$$

where $d_\theta(m_i^l, t_j^l)$, $d_{//}(m_i^l, t_j^l)$, $d_\perp(m_i^l, t_j^l)$ are the *orientation distance*, *parallel distance*, and *perpendicular distance*, respectively. All these three entries are independent and defined as

$$d_\theta\left(m_i^l, t_j^l\right) = f\left(\theta\left(m_i^l, t_j^l\right)\right). \quad (1)$$

$$d_{//}\left(m_i^l, t_j^l\right) = \min\left(l_{//1}, l_{//2}\right). \quad (2)$$

$$d_\perp\left(m_i^l, t_j^l\right) = l_\perp. \quad (3)$$

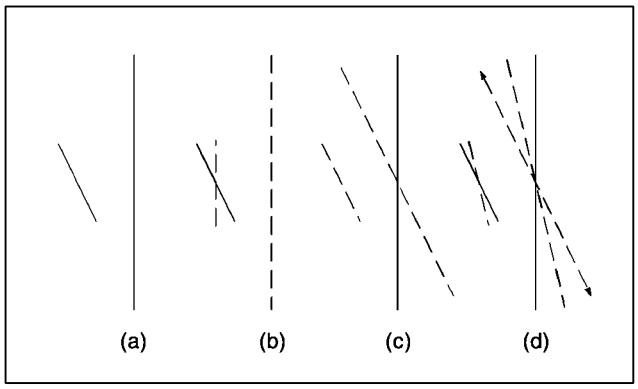

Fig. 3. Choices of rotation. (a) Two lines to be measured. (b) Rotate the shorter line. (c) Rotate the longer line. (d) Rotate both lines half of their angle difference in opposite direction. Solid lines represent lines before rotation. Dashed lines represent lines after rotation. The line with arrows illustrates the angle difference of the two segments.

$\theta(m_i^l, t_j^l)$ computes the smallest intersecting angle between lines m_i^l and t_j^l. $f()$ is a nonlinear penalty function to map an angle to a scalar. It is desirable to ignore small angle variation but penalize heavily on large deviation. In this study, the quadratic function $f(x) = x^2/W$ is used, where W is the weight to be determined by a training process. The designs of the parallel and perpendicular displacements can be illustrated with a simplified example of two parallel lines, m_i^l and t_j^l, as shown in Fig. 2. $d_{//}(m_i^l, t_j^l)$ is defined as the minimum displacement to align either the left end points or the right end points of the lines. $d_\perp(m_i^l, t_j^l)$ is simply the vertical distance between the two lines. In general, m_i^l and t_j^l would not be parallel, but one can rotate the shorter line with its midpoint as rotation center to the desirable orientation before computing $d_{//}(m_i^l, t_j^l)$ and $d_\perp(m_i^l, t_j^l)$. The shorter line is selected to rotate because this would cause less distortion to the original line pair, as illustrated in Fig. 3. In order to cater for the effect of broken lines caused by segmentation error and alleviate the effect of adding, missing, and shifting of feature points (i.e., end points of line segments) caused by inconsistency of feature point detection, the parallel shifts $l_{//1}$ and $l_{//2}$ are reset to zero if one line is within the range of the other, as shown in Fig. 4. Finally, the distance between two line segments m_i^l and t_j^l is defined as

$$d\left(m_i^l, t_j^l\right) = \sqrt{d_\theta^2\left(m_i^l, t_j^l\right) + d_{//}^2\left(m_i^l, t_j^l\right) + d_\perp^2\left(m_i^l, t_j^l\right)}. \quad (4)$$

A primary line segment Hausdorff distance (pLHD) [3] is defined as

$$H_{pLHD}\left(M^l, T^l\right) = \max\left(h\left(M^l, T^l\right), h\left(T^l, M^l\right)\right), \quad (5)$$

where

$$h\left(M^l, T^l\right) = \frac{1}{\sum_{m_i^l \in M^l} l_{m_i^l}} \sum_{m_i^l \in M^l} l_{m_i^l} \cdot \min_{t_j^l \in T^l} d\left(m_i^l, t_j^l\right) \quad (6)$$

and $l_{m_i^l}$ is the length of line segment m_i^l. In the calculation of (6), the distance contribution from each line is weighted by

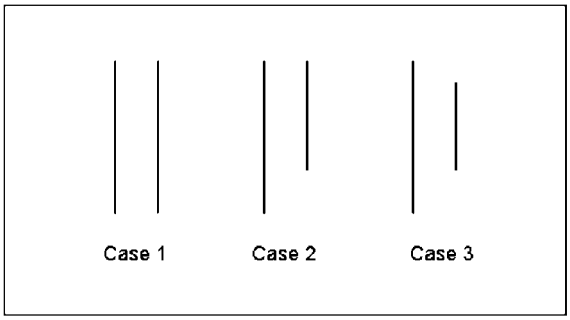

Fig. 4. All cases with $d_{//}(m,t)=0$.

its length. In the definition of (2), it can be found that the displacement distance (Fig. 2) depends on the smaller distance between the left/right end points of the two line segments to be matched, which means the measure only reflects the smallest shift of the two line end points. If one of the line end points is consistently detected and the other one shifts, the displacement distance is almost zero no matter how far the other end point shifts. This helps to alleviate the problem of shifting feature points.

However, the design of pLHD still has the following weakness. Suppose T is the LEM of a test image to be matched and t_i is a line segment in T, M_c is the corresponding identical model of T in the database, and M_n is a nonidentical model of T in the database. If the corresponding line of t_i in M_c is missing because of segmentation error, pLHD will take the nearest line, $m_{cn} \in M_c$, as the corresponding line of t_i. And $d(m_{cn}, t_i) = \min_{m_c \in M_c} d(m_c, t_i)$ is used in the calculation of pLHD. Similarly, the nearest line ($m_{nn} \in M_n$) of t_i in M_n is considered as the correspondent line of t_i, though M_n and T are different objects. It is possible to have $d(m_{cn}, t_i) >> d(m_{nn}, t_i)$ for the matching of complicated and similar objects such as faces. This kind of missing lines can cause larger disparity between T and M_c than that between T and M_n, though both m_{cn} and m_{nn} are actually not the corresponding line segment of t_i and both $d(m_{cn}, t_i)$ and $d(m_{nn}, t_i)$ are relatively large. This may cause mismatch.

The number of corresponding line pairs between the input and the model is another measure of similarity. The number of corresponding line pairs between two identical images should be larger than that between two images of different objects. Hence, the problem mentioned above can be alleviated by introducing this number information into pLHD measure.

Assume that, for each line t_j in the test LEM T, its corresponding line m_i in model M of the identical face should locate near t_j because the test image and the model have been aligned and scale normalized by preprocessing before matching. Therefore, a position neighborhood N_p and an angle neighborhood N_a are introduced. *Similarity neighborhood* N_s is a combination of N_p and N_a as:

$$N_s = N_p \cap N_a.$$

If at least one line in model M locates within the similarity neighborhood of a line t_j in test LEM T (that is, locates in a given position neighborhood of t_j and their angle difference is also within the given angle neighborhood), it is most likely for t_j to find a correct corresponding line among those lines inside the similarity neighborhood. This line (t_j) is named as a *high confident line*. A *high confident line ratio* (R) of an image is defined as the ratio of the number of high confident line (N_{hc}) to the total line number in the LEM (N_{total}).

$$R = \frac{N_{hc}}{N_{total}}. \tag{7}$$

Hence, a complete version of LHD integrated with number disparity is defined as (8) by taking the effect of similarity neighborhood into account.

$$H_{LHD}(M,T) = \sqrt{H_{pLHD}^2(M,T) + (W_n D_n)^2}, \tag{8}$$

where $H_{pLHD}(M,T)$ is the pLHD defined in (5) and W_n is the weight of number disparity D_n. The number disparity is defined as the average ratio, the number of lines located outside the similarity neighborhood to the total line number, of the two LEMs to be compared as

$$D_n = 1 - \frac{R_M + R_T}{2} = \frac{(1-R_M)+(1-R_T)}{2}, \tag{9}$$

where R_M and R_T are the high-confident line ratio of the model and the test LEMs, respectively.

One way to determine the parameters (W, W_n, N_p, N_a) in the LEM face recognition system using LHD is to select the values with the smallest error rate of face matching using a typical database. We use simulated annealing [29], [30] to perform the global minimization of the error rate of face identification. Simulated annealing is a well-known stochastic optimization technique where, during the initial stages of the search procedure, moves can be accepted which increase the objective function. The objective is to do enough exploration of the search space before resorting to greedy moves in order to avoid local minima. Candidate moves are accepted according to probability p as

$$p = e^{-\frac{\Delta Err}{t}}, \tag{10}$$

where Err is the error rate of face identification and t is the temperature parameter which is adjusted according to certain cooling schedule. Exponential cooling schedule [31] is adopted in this work. The face database [33] from the University of Bern is used as training data and $W = 30, W_n = 5, N_p = 6, N_a = 30$ are obtained. Since the parameter selection process is conducted offline, it is worth spending a substantial effort to design the best LHD format (in the sense of minimizing error rate). This can result in optimal performance of the proposed face recognition system.

5 Experimental Results

A thorough system performance investigation, which covers all conditions of human face recognition, has been conducted. They are face recognition under

1. controlled condition and size variation,
2. varying lighting condition,
3. varying facial expression, and
4. varying pose.

The system performances are compared with the eigenface method [7], the edge map approach [2], and reported experimental results of other methods studied in [32].

In this study, three face databases were tested. The database from the University of Bern [33] (Fig. 5) was used

Fig. 5. An example pair of faces from Bern University [33].

Fig. 6. An example pair of faces from the AR face database [34]. The two faces were taken with a two week interval.

TABLE 1
Face Recognition Results of Edge Map (EM) [2], Eigenface (20-Eigenvectors), and LEM

	Bern database			AR database		
Method	EM	Eigenface	LEM	EM	Eigenface	LEM
Recognition rate	96.7%	100%	100%	88.4%	55.4%	96.4%

to examine the system performances under controlled/ideal condition and head pose variations. The database contains frontal views of 30 people. Each person has 10 gray-level images with different head pose variations (Two fronto-parallel pose, two looking to the right, two looking to the left, two looking downwards, and two looking upwards). The AR face database [36] (Fig. 6) from Purdue University was used to evaluate the system performances under controlled/ideal condition, size variation, varying lighting condition, and facial expression. The database contains color images corresponding to 126 people's faces (70 men and 56 women). However, some images were found lost or corrupted after downloading through Internet. There were 112 sets of usable images (61 men and 51 women). No restrictions on wear (clothes, glasses, etc.), make-up, hairstyle, etc., were imposed to the participants. The Yale face database [35] was also tested in this work in order to compare the proposed approach to methods studied in [32]. The database, constructed at the Yale Center for Computational Vision and Control, is composed of 165 images of 15 subjects. The images demonstrate variations in lighting condition (left-light, center-light, right-light), facial expression (normal, happy, sad, sleepy, surprised, and wink), and with/without glasses. In all the experiments, preprocessing to locate the faces was applied. Original images were normalized (in scale and orientation) such that the two eyes were aligned roughly at the same position with a distance of 80 pixels. Then, the facial areas were cropped into the final images for matching. Sample cropped images can be found in Fig. 8.

5.1 Face Recognition under Controlled/Ideal Condition and Size Variation

The face images under controlled condition in the database of Bern University and AR database were used to evaluate the performance of the proposed approach. Two example pairs of the face images in the two databases are illustrated in Figs. 5 and 6. The recognition results are summarized in Table 1. It is found that the LEM approach performed better than the edge map and the eigenface methods. LEM and Eigenface achieved 100 percent accuracy for identifying faces in the database of Bern University. However, LEM significantly outperformed Eigenface on the AR face database. Detailed eigenface data were tabulated in Table 2.

TABLE 2
Performance Comparison on the AR Database

Method	Recognition rate
LEM	96.43%
Eigenface (20-eigenvectors)	55.36%
Eigenface (60-eigenvectors)	71.43%
Eigenface (112-eigenvectors)	78.57%

The performance of the eigenface method depends on the number of eigenvectors m. If this number is too small, important information about the identity is likely to be lost. If it is too high, the weights corresponding to small eigenvalues might be noises. The number of eigenvectors m is limited by the rank of the training set matrix. One hundred and twelve is the upper bound of m in the experiment of the AR database and, thus, 78.57 percent is the best performance that eigenface can achieve here. One way to interpret this is that the eigenface approach will work well as long as the test image is "similar" to the ensemble of images used in the calculation of eigenface [37]. And the training set should include multiple images for each person with some variations [7] to obtain a better performance. Here, only one image per person was used for training and the other is used for testing. Another reason is the difference between two identical faces is larger in the AR database than that in the database of Bern University. In particular, the illuminations of the input and the model are slightly different (Fig. 6). This might be the major negative impact to the eigenface approach as compared with LEM since LEM is relatively insensitive to illumination changes.

In the LEM method, the proposed LHD dissimilarity measure is also evaluated and compared with the pLHD definition. The experimental results are summarized in Fig. 7 with three classification conditions (top1, top3, and top5). In the top1 classification, the correct match is only counted when the best matched face from models is the identical face of the input. In the top3 or top5 classification, the correct match is counted when the identical face of the input is among the best 3 or 5 matched faces from models, respectively. It was found that for the top 1 match, LEM with pLHD performed better than edge map with MHD [2] by 5.36 percent, and LEM with LHD could further improve the performance of pLHD by 2.68 percent, i.e., it correctly identified 96.43 percent of the input faces.

A sensitivity analysis to size variation was conducted using the AR database. The size variation was generated by applying a random scaling factor, which was uniformly distributed within [1-10 percent, 1+10 percent], to the test

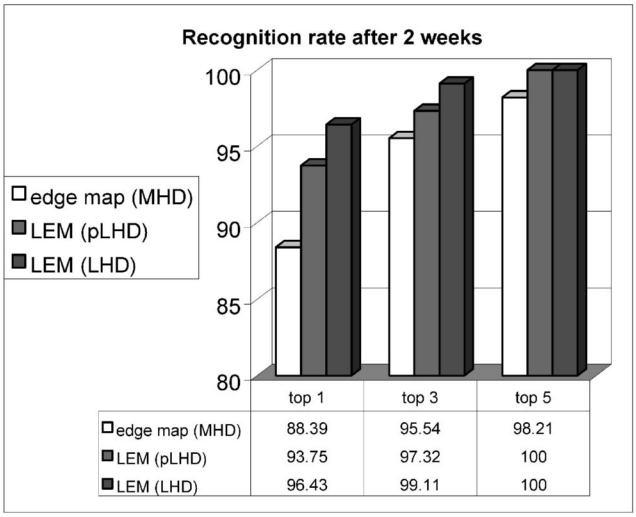

Fig. 7. Recognition results on the AR face database.

images. Four faces with different sizes were generated from each test image. Thus, we had 448 test faces in total with size variations ranging from -10 percent to +10 percent. The sizes of the 112 model images were not changed.

The experimental results are tabulated in Table 3. The results show that the edge map approach is slightly more sensitive to size variation than the eigenface approach. The LEM with pLHD performs better than the eigenface approach in the top 1 match. The proposed LEM with LHD, which outperformed the eigenface approach by an accuracy increment of 21.6 percent, is much more robust to size variations than all the others in the experiments. This is a very attractive property that can alleviate the difficulty of precisely locating faces in the previous face detection stage.

5.2 Face Recognition under Varying Lighting Conditions

Ideally, an object representation employed for recognition should be invariant to lighting variations. It has been shown theoretically that, for the general case, a function invariant to illumination does not exist [38]. Edge maps can serve as robust representations to illumination changes for some classes of objects with sharp edges, such as computers, tables, etc. However, for other objects, such as faces, part of the edges can not be obtained consistently. It can be shown theoretically that edges on a smooth surface are not stable with changes in the lighting direction [39]. The LEM is an intermediate-level image representation derived from low-level edge map representation. The basic unit of LEM is the line segment grouped from pixels of the edge map. It remains an open question whether LEM, edge map, and other possible

Fig. 8. Sample cropped images of model (leftmost) and test faces (under varying lighting and expression).

TABLE 3
Recognition Results with Size Variations

	Top 1	Top 5	Top 10
Edge map	43.3%	56.0%	64.7%
Eigenface (112-eigenvectors)	44.9%	68.8%	75.9%
LEM (pLHD)	53.8%	67.6%	71.9%
LEM (LHD)	66.5%	75.9%	79.7%

representations provide an illumination-insensitive representation for face recognition.

The issue addressed in this section is whether the LEM representation is sufficient or how well it performs for recognizing faces under varying lighting condition. To answer this question, an empirical study was performed to evaluate the sensitivity of the LEM, and the edge map and the eigenface representations to this appearance change through the performances of face recognition systems. The experiment was designed using face images taken under different lighting conditions from Purdue University (Fig. 8). The faces in neutral expression with background illumination (the leftmost image in Fig. 8) were used as single models of the subjects. The images under three different lighting conditions were used as test images. There are a total of 112 models (representing 112 individuals) and 336 test images. Note that the experiment was based on a single model view.

The experimental results with three different lighting conditions are illustrated in Table 4. These experiments reveal a number of interesting points:

1. In all the three experiments, the LEM consistently performed better than the edge map approach with an improvement of 10.72-19.65 percent in recognition rate. The eigenface approach performed very badly in these conditions. For the eigenface method, it has been suggested that the first three principal components are the primary components responding sensibly to lighting variation. Thus, the system error rate can be reduced by discarding these three most significant principal components [32]. Though the accuracy of the eigenface approach increased without using the first three eigenvectors, the LEM still significantly outperformed it.
2. The variations of lighting condition did affect the system performance. Nevertheless, the recognition rates of the LEM approach were still high and acceptable. Note that the recognition rates of LEM, when only one light was on, stayed as high as 92.86 and 91.07 percent, respectively. In other words, the effect on recognition rates when one light was on created only 3.57 to 5.36 percent decreases in recognition

TABLE 4
Recognition Results under Varying Lighting

Testing faces	Eigenface		Edge map	LEM
Left light on	20-eigenvectors	6.25%	82.14%	92.86%
	60-eigenvectors	9.82%		
	112-eigenvectors	9.82%		
	112-eigenvectors w/o 1st 3	26.79%		
Right light on	20-eigenvectors	4.46%	73.21%	91.07%
	60-eigenvectors	7.14%		
	112-eigenvectors	7.14%		
	112-eigenvectors w/o 1st 3	49.11%		
Both lights on	20-eigenvectors	1.79%	54.46%	74.11%
	60-eigenvectors	2.68%		
	112-eigenvectors	2.68%		
	112-eigenvectors w/o 1st 3	64.29%		

Conditions "w/o 1st 3" stands for without the first three eigenvectors.

TABLE 5
Recognition Results under Different Facial Expressions

Testing faces	Eigenface		EM	LEM
Smiling expression	20-eigenvectors	87.85%	52.68%	78.57%
	60-eigenvectors	94.64%		
	112-eigenvectors	93.97%		
	112-eigenvectors w/o 1st 3	82.04%		
Angry expression	20-eigenvectors	78.57%	81.25%	92.86%
	60-eigenvectors	84.82%		
	112-eigenvectors	87.50%		
	112-eigenvectors w/o 1st 3	73.21%		
Screaming expression	20-eigenvectors	34.82%	20.54%	31.25%
	60-eigenvectors	41.96%		
	112-eigenvectors	45.54%		
	112-eigenvectors w/o 1st 3	32.14%		

accuracy. The large number of subjects, compared with the database used in [32], [40], made it especially interesting. These results indicate that the proposed LEM together with LHD provide a novel face recognition solution which is insensitive to varying lighting condition to certain extent.

3. It was found that the recognition rates with left light on were always higher than that with right light on. This could be due to the fact that the illumination on faces from the right light was slightly stronger than that from the left light.
4. When both lights were on, the error rates became much higher than that of only one light on. This evidence shows that LEM (and edge map) would still be affected by extreme lighting condition variations, such as overillumination, though it is insensitive to certain extent. The overillumination would cause strong specular reflection on the face skin (it is no longer a Lambertian surface). Therefore, the shape information on faces would have been suppressed or lost which could result in the increase of the error rate of classification.

5.3 Face Recognition under Facial Expression Changes

Similar experiments were conducted to evaluate the effects of different facial expressions (smile, anger, and scream) on the system performances. The face images of different facial expressions from Purdue University were used in the experiments (Fig. 8). The faces in neutral expression (the leftmost image in Fig. 8) were used as single models of the subjects. Totally, there were 112 models (representing 112 individuals) and 336 test images.

The experimental results on faces with smile, anger and scream expressions were summarized in Table 5. The smile expression caused the recognition rate to drop by 17.86 percent as compared to neutral expression in Table 1, while the anger expression caused only a 3.57 percent drop of the rate. This is because the anger expression had produced less physical variation from neutral expression than the expression of smile. The scream expression could be the extreme case of deformation among various human facial expressions, i.e., most facial features had been distorted. The LEM performed better than the edge map approach in all three experiments. However, the eigenface approach is found the least sensitive to facial expression changes. Without using the first three significant eigenvectors, the method to reduce the negative effect of lighting condition variation, degraded the accuracy under facial expression changes.

5.4 View-Based Identification Experiment and Comparison with Existing Approaches

In [32], Belhumeur et al. designed tests to determine how different face recognition methods compared under a different range of conditions and presented the error rates of five different face recognition approaches (eigenfaces, eigenfaces without the three most significant principal components, correlation, linear subspace, and fisherface) using the Yale face database [35]. Experiments were performed using a "leaving-one-out" strategy: The models were constructed with all images of a subject except one image that was used as input.

The same strategy and database were adopted here in order to compare the LEM and the edge map approaches to methods studied in [32]. For comparison purposes, the images were similarly processed as stated by Belhumeur et al. to generate closely cropped images. Fig. 9 displays samples of 11 cropped images of one subject from the database.

The experimental results together with the results conducted by Belhumeur et al. were summarized in Table 6. A rough comparison of the experimental results shows that LEM is superior to all the methods except the Fisherface method (Table 6). It is worth highlighting that, though the performance of the edge map approach is the worst among all the methods in Table 6, LEM has greatly improved the system performance to ranks 2. Note that LEM is a more compact

Fig. 9. Sample cropped faces used in our experiment.

(less storage and computational requirement) representation than edge map. The results indicate that LEM performed much better than the eigenface method, the widely used baseline, and slightly better than Eigenface without the first three eigenvectors. Fisherface is specifically designed and only valid for applications of multiple models per person. By complicated computation, it maximizes the difference of the between-person variation and the within-person variation. The test and database (leave-one-out test on 15 individuals and 11 images/person) are "ideal" for Fisherface, whereas all the other six methods that can be applied on single model recognition do not get any favor.

5.5 Face Recognition under Varying Poses

The face database from Bern University is used to evaluate the system performance on face images of different poses. One fronto-parallel face per person was used as the model. The system was tested using the eight poses looking to the right, left, up, and down for each person. There are 240 test images in total. The recognition results are summarized in Table 7. It can be observed that pose variations degrade the recognition rate of all the three investigated methods, but LEM is the most robust to pose variation, while the edge map approach performed the worst. Note that 30 is the maximum number of eigenvectors that can be used for the eigenface approach in this experiment. Thus, 65.12 percent is the best average performance that the eigenface approach can achieved here.

5.6 Storage and Computational Complexity

The storage requirement of LEM is analyzed and compared with edge map based on the frontal face database from Bern University. The data sizes in the experiment are listed in Table 8. On average, a face can be represented by an LEM with 293 feature points. Compared with the edge map, LEM requires only 28.5 percent of its storage space.

The computational complexities of LHD and MHD [2] are of the order $O(k_l m_l t_l)$ and $O(k_p m_p t_p)$, respectively. m_l and t_l are the line segment numbers of model and test LEMs, while m_p and t_p are pixel numbers of model and test edge maps. k_l

TABLE 6
"Leave-One-Out" Test of Yale Face Database

Method	Error Rate
Edge map	**26.06%**
Eigenface*	24.4%
Correlation*	23.9%
Linear Subspace*	21.6%
Eigenface w/o 1st 3*	15.3%
LEM	**14.55%**
Fisherface*	7.3%

*The values with * are from [32]*

TABLE 7
Face Recognition Results under Pose Different Variations

	Recognition rate			
Method	Edge map	Eigenface (20-eigenvectors)	Eigenface (30-eigenvectors)	LEM
Looks left/right	50.00%	70.00%	75.00%	74.17%
Looks up	65.00%	51.67%	56.67%	70.00%
Looks down	67.67%	45.00%	55.00%	70.00%
Average	**58.17%**	**59.17%**	**65.12%**	**72.09%**

TABLE 8
Average Storage Requirement of Faces in the Experiments

	Edge map	LEM
Average number of (feature) point	1027	293

and k_p are the time to compute $d(m_i^l, t_j^l)$ in LHD and the Euclidean norm $||m_i^p - t_j^p||$ in MHD. Table 9 shows the real average computational time of MHD and LHD on faces of [33]. The experiments were conducted on a SGI Octane workstation with 300MHz CPU and 512MB RAM. The computational time for LHD is less than 50 percent of MHD. Since the calculation of $d(m_i^l, t_j^l)$ spends much more time than a simple Euclidean distance calculation in MHD (that is, k_l is larger than k_p),

$$\frac{O(k_l m_l t_l)}{O(k_p m_p t_p)} > \frac{O(m_l t_l)}{O(m_p t_p)}.$$

With some acceleration techniques (such as hardware acceleration, look up table), the LHD computational time can be further reduced by minimizing k_l. When $k_l \approx k_p$, the LHD computational time could be reduced to only about 10 percent of MHD

$$\left(\frac{m_l}{m_p} \cdot \frac{t_l}{t_p} < (28.5\%)^2 < 10\%\right).$$

This is the ideal upper bound of computational time decrement for LHD with respect to MHD.

The storage requirement for the eigenface approach is $Nm + dm$ (Table 10), where N is the number of faces, m is the number of eigenvectors, and d is the dimensionality of the image. The storage demand of LEM is Nn, where n is the average number of feature points. n is content-based, whereas m is fixed. Usually, LEM demands more storage space than the eigenface approach for large N as n is larger than m (not always). In this experiment, $n = 293$ and $m = 112$. Each feature point is represented by two 8-bit integers and each eigenvalue is a 16-bit float number. But, LEM does not need to store any vector of size d, whereas the eigenface approach must store the projection matrix (m vectors of dimension d). d is $160 \times 160 = 25,600$ in this experiment.

The computational complexity includes three aspects, i.e., matching time, training time, and updating time. The matching time is the most important one for large database searching. LEM requires more matching time than the eigenface approach as $n > m$. However, the eigenface approach demands a substantial amount of training time

TABLE 9
Average Computational Time of MHD and LHD on [33]

	MHD	LHD
Computational time	1.146s	0.503s

TABLE 10
Comparison of Storage Requirements

	LEM	Eigenface
Storage space	$O(Nn)$	$O(Nm + dm)$

in the order of $O(N^3 + N^2 d)$ to obtain a satisfactory projection matrix. When a new individual is added into the database, the projection matrix must be recomputed. This incremental update retraining of $O(N^3 + N^2 d)$ is another expensive operation that the LEM approach does not need. This retraining can be avoided by assuming that the new images do not have a significant impact on the eigenfaces, that is, just calculate the weights for the new images using the old projection matrix. This is only valid if the system was initially "ideally" trained.

6 Face Prefiltering

In a face identification system, searching is the most computationally expensive operation due to the large number of images available in the database. Therefore, efficient search algorithms are a prerequisite of identification systems. In most systems, face features are extracted/coded offline from the original images, and stored in the face feature database. In identification, the same features are extracted from the input face and the features of the input images are compared with the features of each model image in the database. Apart from adopting a fast face matching algorithm, a prefiltering operation can further speed up the search by reducing the number of candidates and the actual face feature matching, LEM matching in this work, should only be carried out on a subset of target images. Due to the high similarity of human faces, fast prefiltering based on image features is a very difficult task and is usually neglected or avoided. This study is believed to be the first piece of work on face image prefiltering.

Prefiltering feature selection is critical because it significantly affects the remaining aspects of the system design and greatly determines the filtering capability of the system. For indexing visual features, a common approach is to obtain numerical values of n features and then represent the image or object as a point in the n-dimensional space. In this work, a 2D vector derived from the LEM of a face is proposed as the face prefiltering signature. This signature is proven to be able to filter out certain percent of candidates while preserving a very low false negative rate though it is a very low dimensional vector and human faces are highly similar.

Experimental results show that 22.02 percent of the candidates can be filtered out while the false rejection rate was 0.89 percent using the AR face database of Purdue University. The experiments on the face database of Bern University, which contains fewer variations between the two identical faces, demonstrated much better filtering rate. The system correctly filtered out 51.95 percent of candidates, while the false rejection rate was zero.

6.1 Prefiltering Signature

The LEM of a human face contains rich identity information. In order to speed up the search process, a prefiltering signature, $\vec{S}$, is generated using the length and weighted orientation features of a face LEM as

$$\vec{S} = \begin{bmatrix} \Gamma \\ \Theta \end{bmatrix}, \tag{11}$$

where Γ is the total length of the face LEM and Θ is the weighted orientation defined as

$$\Theta = \frac{1}{\sum_i l_i} \sum_i \vartheta_i l_i, \tag{12}$$

where l_i is the length of the ith line segment in the face LEM and ϑ_i is the angle of the ith line segment if it is less than 90 degrees. The reason that only those line segments whose $\vartheta_i < 90^\circ$ are taken into account of the Θ calculation is given as follows.

We assume that the human face is symmetric, thus its LEM is also symmetric. Ideally, any line segment with ϑ_i can find its symmetric correspondent line segment with $\vartheta_j = 180^\circ - \vartheta_i$. Therefore, the overall average of line segment orientation tends to be 180 degrees for all faces because the distinctive information of ϑ_i is counteracted by ϑ_j. To prevent this compensation effect, only half of the line segments $(\vartheta_i < 90^\circ)$ are used to compute Θ. As a whole, the prefiltering signature $\vec{S}$ represents the richness of a face's edges and the orientation of the face LEM.

6.2 Prefiltering Criteria

Consider two prefiltering signatures $\vec{S}_i$ and $\vec{S}_m$, where $\vec{S}_i$ is a signature vector of input face for inquiry and $\vec{S}_m$ is a signature vector of model face in the database. The process needs to decide whether the model face can be selected as one of the possible candidates corresponding to the input. The error vector $(\Delta\vec{S})$ between the input signature $(\vec{S}_i)$ and the model signature $(\vec{S}_m)$ is considered as an observable value.

$$\Delta\vec{S} = \vec{S}_i - \vec{S}_m = \begin{bmatrix} \Gamma_i - \Gamma_m \\ \Theta_i - \Theta_m \end{bmatrix} = \begin{bmatrix} \Delta\Gamma \\ \Delta\Theta \end{bmatrix}. \tag{13}$$

Error vectors are often assumed to be normally distributed [41]. Here, we assume the errors $\Delta\Gamma$ and $\Delta\Theta$ between signatures of identical faces obey normal distribution of zero mean. Hence, the error vector $\Delta\vec{S}$ is a bivariate normal distribution variable

$$\Delta\vec{S} \sim N_2\left(\vec{\mu}, \vec{\Sigma}\right),$$

where

$$\vec{\mu} = \begin{bmatrix} \mu_l \\ \mu_\theta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \vec{\Sigma} = \begin{bmatrix} \sigma_l^2 & \sigma_{l\theta} \\ \sigma_{\theta l} & \sigma_\theta^2 \end{bmatrix} = \begin{bmatrix} \sigma_l^2 & \sigma_l\sigma_\theta\rho \\ \sigma_\theta\sigma_l\rho & \sigma_\theta^2 \end{bmatrix},$$

and the correlation coefficient

$$\rho = \frac{\sigma_{l\theta}}{\sigma_l\sigma_\theta}.$$

Then, the density function of the error vector can be represented as

$$f(\Delta\vec{S}) = \frac{1}{2\pi|\vec{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}\left(\Delta\vec{S} - \vec{\mu}\right)^T \vec{\Sigma}^{-1}\left(\Delta\vec{S} - \vec{\mu}\right)\right\}, \Delta\vec{S} \in \Re^2. \tag{14}$$

Since $|\vec{\Sigma}| = \sigma_l^2\sigma_\theta^2(1-\rho^2)$, the inverse of $\vec{\Sigma}$ exists if and only if $|\rho| < 1$. Straightforward calculation shows that

$$\vec{\Sigma}^{-1} = \frac{1}{\sigma_l^2\sigma_\theta^2(1-\rho^2)} \begin{bmatrix} \sigma_l^2 & -\sigma_l\sigma_\theta\rho \\ -\sigma_\theta\sigma_l\rho & \sigma_\theta^2 \end{bmatrix}. \tag{15}$$

Thus, the density function of $\Delta\vec{S}$ becomes

$$f(\Delta\vec{S}) = \frac{1}{2\pi\sigma_l\sigma_\theta\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{\Delta\Gamma-\mu_l}{\sigma_l}\right)^2 - 2\rho\left(\frac{\Delta\Gamma-\mu_l}{\sigma_l}\right)\left(\frac{\Delta\Theta-\mu_\theta}{\sigma_\theta}\right) + \left(\frac{\Delta\Theta-\mu_\theta}{\sigma_\theta}\right)^2\right]\right\}, \Delta\vec{S} \in \Re^2. \tag{16}$$

The constant density contours for a bivariate normal are a series of ellipses with different values of d as shown in the following equation:

$$\left(\Delta\vec{S} - \vec{\mu}\right)^T \vec{\Sigma}^{-1}\left(\Delta\vec{S} - \vec{\mu}\right) = d^2$$

or

$$\left(\frac{\Delta\Gamma-\mu_l}{\sigma_l}\right)^2 - 2\rho\left(\frac{\Delta\Gamma-\mu_l}{\sigma_l}\right)\left(\frac{\Delta\Theta-\mu_\theta}{\sigma_\theta}\right) + \left(\frac{\Delta\Theta-\mu_\theta}{\sigma_\theta}\right)^2 = d^2(1-\rho^2). \tag{17}$$

The probability that $\Delta\vec{S}$ falls in the elliptic region Ω of parameter d is given by

$$\begin{aligned} F(d) &= \Pr\left(\Delta\vec{S} \in \Omega\right) = \iint_\Omega f\left(\Delta\vec{S}\right) d(\Delta\Gamma)d(\Delta\Theta) \\ &= \iint_\Omega \frac{1}{2\pi\sigma_l\sigma_\theta\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{\Delta\Gamma-\mu_l}{\sigma_l}\right)^2 - 2\rho\left(\frac{\Delta\Gamma-\mu_l}{\sigma_l}\right)\left(\frac{\Delta\Theta-\mu_\theta}{\sigma_\theta}\right) + \left(\frac{\Delta\Theta-\mu_\theta}{\sigma_\theta}\right)^2\right]\right\} d(\Delta\Gamma)d(\Delta\Theta). \end{aligned} \tag{18}$$

Let

$$u = \frac{\Delta\Gamma-\mu_l}{\sigma_l}, \qquad v = \frac{\Delta\Theta-\mu_\theta}{\sigma_\theta}. \tag{19}$$

The equation of constant density contour can be rewritten as

$$u^2 - 2\rho uv + v^2 = d^2(1-\rho^2). \tag{20}$$

If $\rho > 0$, the major axis of the ellipse would be located along the 45 degree line passing through the origin. The lengths of the major axis and minor axis are $d\sqrt{1+\rho}$ and $d\sqrt{1-\rho}$, respectively. If $\rho < 0$, the major axis of the ellipse would locate along the 135 degree line passing through the origin. The lengths of the major axis and minor axis are $d\sqrt{1-\rho}$ and $d\sqrt{1+\rho}$, respectively. Let

$$u' = u\cos 45^\circ + v\sin 45^\circ, v' = -u\sin 45^\circ + v\cos 45^\circ,$$

or

$$\begin{aligned} u &= u'\cos(-45^\circ) + v'\sin(-45^\circ), \\ v &= -u'\sin(-45^\circ) + v'\cos(-45^\circ), \end{aligned} \tag{21}$$

i.e., rotate the ellipse of (20) such that its major axis is coincident with either of the axes of the Cartesian coordinate. The Jacobian of the above transformation is

$$|J| = \left|\frac{\partial(u,v)}{\partial(u',v')}\right| = \begin{vmatrix} \cos(-45^\circ) & \sin(-45^\circ) \\ -\sin(-45^\circ) & \cos(-45^\circ) \end{vmatrix} = 1 \tag{22}$$

and one can change the integration coordinate system of (18) by substituting (19), (20), (21), and (22) into (18) as [Substitute (19) into (18)]

$$F(d) = \iint_\Omega \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[u^2 - 2\rho uv + v^2\right]\right\} dudv. \tag{23}$$

[Substitute (21) and (22) into (23)]

$$\begin{aligned} F(d) &= \iint_\Omega \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2}\left(\frac{u'^2}{1+\rho} + \frac{v'^2}{1-\rho}\right)\right\} |J| du'dv' \\ &= \iint_\Omega \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2}\left(\frac{u'^2}{1+\rho} + \frac{v'^2}{1-\rho}\right)\right\} du'dv'. \end{aligned} \tag{24}$$

Let $a = \sqrt{1+\rho}$, $b = \sqrt{1-\rho}$, thus, we get

$$F(d) = \iint_\Omega \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2}\left(\frac{u'^2}{a^2} + \frac{v'^2}{b^2}\right)\right\} du'dv'. \tag{25}$$

To further change the integration coordinate system, let

$$u' = ra\cos\theta, \quad v' = rb\sin\theta, \tag{26}$$

where r and θ represent the new polar system. The Jacobian of the transformation is

$$|J| = \begin{vmatrix} \frac{\partial u'}{\partial r} & \frac{\partial v'}{\partial r} \\ \frac{\partial u'}{\partial \theta} & \frac{\partial v'}{\partial \theta} \end{vmatrix} = \begin{vmatrix} a\cos\theta & b\sin\theta \\ -ra\sin\theta & rb\cos\theta \end{vmatrix} = rab. \tag{27}$$

By substituting (26) and (27) into (25), we have

$$\begin{aligned} F(d) &= \int_0^d \int_0^{2\pi} \frac{1}{2\pi ab} \exp\left\{-\frac{1}{2}r^2\right\} |J| dr d\theta \\ &= 1 - e^{-\frac{1}{2}d^2}. \end{aligned} \tag{28}$$

TABLE 11
AR Face Database Training Results

ρ	c_l	c_θ	μ_l	μ_θ
0.02	145.36	4.33	26.42	0.27

Finally, (28) can be rewritten for d in terms of $F(d)$, which is the desired true acceptance rate of the prefiltering.

$$d = \sqrt{-2\ln[1 - F(d)]}. \tag{29}$$

To summarize, given a desired probability $F(d)$, one can obtain a constant density ellipse in the form of (17) whose d value can be calculated by (29). The prefiltering criteria can be described as follows: If the error vector $\Delta\bar{S}$ between the input and the model satisfies the equation

$$\left(\frac{\Delta\Gamma}{\sigma_l}\right)^2 - 2\rho\left(\frac{\Delta\Gamma}{\sigma_l}\right)\left(\frac{\Delta\Theta}{\sigma_\theta}\right) + \left(\frac{\Delta\Theta}{\sigma_\theta}\right)^2 < d^2\left(1-\rho^2\right), \tag{30}$$

the model is classified as the potential identical face of the input, and selected as the candidate for the following actual face matching. Otherwise, the model is not considered.

6.3 Experimental Results

Experiments were conducted using two face databases from Purdue University (AR face database) and Bern University. Each person contains two images. One was used as model, the other was used as input. The AR face database was used for training to obtain the necessary values of ρ, σ_l, and σ_θ in the prefiltering equation (30). Then, the prefiltering system was tested using the above two face databases by varying the desired true acceptance rate $F(d)$.

The training results are summarized in Table 11. The near zero value of ρ, the correlation coefficient between the two features, indicates that $\Delta\Gamma$ and $\Delta\Theta$ are nearly independent. Since two independent features provide the maximum classification capability, the proposed signature is thus the optimal selection in terms of the relationship between the two components. The actual mean vector of the error bivariate signature obtained from the training database is a nonzero vector, though it is assumed zero. This bias could be caused by the environmental changes between acquiring the model images and the inputs. All the inputs were taken in the second session, which is 14 days after obtaining the model images in the first session. The error vectors were computed as input signatures minus model signatures. Therefore, the consistent environmental changes between the two sessions would produce a bias on the mean vector. In this study, we argue that the mean vector of the error signature should be a zero vector and any shifting from zero vector can be considered as the effect of environmental changes but not the face signature's features themselves. The prefiltering criteria of (30) with a zero bivariate mean value was adopted accordingly.

Two evaluation tests were conducted to investigate the performance of the proposed prefiltering method. The results are tabulated in Tables 12 and 13. The real true acceptance rates in Table 12 were similar to, but slightly lower than, the desired true acceptance rates ($F(d)$). This can be caused by the effect of mean vector bias in the

TABLE 12
Prefiltering Results on AR Face Database

$F(d)$	d^2	True acceptance rate	Filter out rate
90%	4.61	88.39%	50.31%
95%	5.99	92.86%	41.37%
99%	9.21	97.32%	26.58%
99.5%	10.60	99.11%	22.02%
99.7%	12.43	100%	17.06%

TABLE 13
Prefiltering Results on Bern University Face Database

$F(d)$	d^2	True acceptance rate	Filter out rate
90%	4.61	96.67%	61.55%
95%	5.99	96.67%	53.91%
96%	6.44	100%	51.95%

database. The system filtered out 22.02 percent of non-identical faces while keeping 99.11 percent of identical faces in the experiment on AR face database. A much better result was obtained from the experiment on the face database of Bern University. The true acceptance rates in Table 13 were higher than the desired true acceptance rates ($F(d)$). The filter out rate was as high as 51.95 percent while preserving 100 percent of the identical faces. The underlying reason is that the identical face variation (i.e., the standard deviation) in this database is smaller than that in the training data. Considering the high similarity of human faces and the extremely low dimensions (only 2 dimensions) of the signature, the results are satisfactory.

The prefiltering algorithm is one of the preprocesses for the LEM matching. To obtain a quick response to inquiries in huge face databases, the prefiltering speed is strongly required to be higher in comparison with the LEM matching speed. Nearly 1 million matches/sec (49,952 match operations in 0.051 seconds) prefiltering speed on a SGI Octane workstation with 300MHz and 512MB of RAM has been achieved. This makes it possible to construct high performance face identification systems.

7 Conclusion

The proposed LEM is a novel compact face feature representation generated from face edge map. It is less sensitive to illumination changes and only requires less than 30 percent storage space of face edge map. A novel Line Segment Hausdorff Distance (LHD), which incorporates spatial information, structural information of line orientation and line-point association, and number disparity, is proposed for face dissimilarity measuring. It is a very encouraging finding that the proposed LEM face recognition approach can achieve higher recognition accuracy than the edge map approach with much less storage requirement and computational time. Experiments on frontal faces under controlled/ideal conditions indicate that the proposed LEM is consistently superior to edge map. LEM correctly identify 100 percent and 96.43 percent of the input frontal faces on face databases [33] and [34], respectively. Compared with the eigenface method, the most widely used baseline for face recognition, LEM performed equally well as the eigenface method for faces under ideal conditions and significantly superior to the eigenface method for faces with slight appearance variations. Moreover, The LEM approach is much more robust to size variation than the eigenface method and the edge map approach.

Beside the recognition under controlled/ideal condition, this research also covers the three difficult issues in face recognition field, i.e., recognizing a face under 1) varying lighting condition, 2) varying facial expression, and 3) varying pose. The sensitivity investigation on the proposed LEM to lighting condition and facial expression shows that the LEM and the edge map are relatively insensitive to lighting changes to certain extent though the effect does exist. The proposed LEM approach is more robust to lighting condition variations than the edge map approach. The effect on recognition rates when one light is on creates only 3.57 to 5.36 percent decrease for the LEM method but 6.25 to 15.18 percent decrease for the edge map method. It is shown that the LEM approach performed significantly superior to the eigenface approach for identifying faces under varying lighting condition. The LEM approach is also less sensitive to pose variations than the eigenface method but more sensitive to large facial expression changes. In the identification experiment under multiple appearance changes, the proposed LEM approach performed better than the eigenface, eigenface without the first three eigenvectors, correlation, and linear subspace methods, while the performance of the edge map approach was the worst.

We have proposed a face prefiltering algorithm that can be used as a preprocess of LEM matching in face identification applications. The prefiltering operation can speed up the search by reducing the number of candidates and the actual face (LEM) matching is only carried out on a subset of remaining models. Nearly 1 million matches/sec prefiltering speed on a SGI Octane workstation with 300MHz and 512MB of RAM is achieved. Because of the high similarity of human faces, fast prefiltering based on image features is a difficult task and is usually neglected or avoided. This work has demonstrated that face image prefiltering is indeed a practical solution to speed up the image searching by carefully selecting/generating proper representation features. This is believed to be the first piece of work on face image prefiltering. Investigations on exploring more distinctive features and efficient representation forms are interesting future work to fine tune the approach.

In summary, this work has proven the proposed new concept: "Faces can be recognized using Line Edge Map." It provides a new way for human face coding and recognition, which is robust to lighting condition changes and size variations. It is a very attractive finding that the proposed face recognition technique has performed superior to the well-known eigenface method in most of the comparison experiments.

References

[1] B. Miller, "Vital Signs of Identity," *IEEE Spectrum*, pp. 22-30, Feb. 1994.
[2] B. Takács, "Comparing Face Images Using the Modified Hausdorff Distance," *Pattern Recognition*, vol. 31, pp. 1873-1881, 1998.

[3] Y. Gao and M.K. Leung, "Human Face Recognition Using Line Edge Maps," *Proc. IEEE Second Workshop Automatic Identification Advanced Technologies,* pp. 173-176, Oct. 1999.
[4] L. Zhao and Y.H. Yang, "Theoretical Analysis of Illumination in PCA-Based Vision Systems," *Pattern Recognition,* vol. 32, pp. 547-564, 1999.
[5] L. Sirovich and M. Kirby, "Low-Dimensional Procedure for the Characterisation of Human Faces," *J. Optical Soc. of Am.,* vol. 4, pp. 519-524, 1987.
[6] M. Kirby and L. Sirovich, "Application of the Karhunen-Loève Procedure for the Characterisation of Human Faces," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, pp. 831-835, Dec. 1990.
[7] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience,* vol. 3, pp. 71-86, 1991.
[8] A. Pentland, B. Moghaddam, and T. Starner, "View-Based and Modular Eigenspaces for Face Recognition," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,* pp. 84-91, 1994.
[9] T.J. Stonham, "Practical Face Recognition and Verification with WISARD," *Aspects of Face Processing,* pp. 426-441, 1984.
[10] K.K. Sung and T. Poggio, "Learning Human Face Detection in Cluttered Scenes," *Computer Analysis of Image and Patterns,* pp. 432-439, 1995.
[11] S. Lawrence, C.L. Giles, A.C. Tsoi, and A.D. Back, "Face Recognition: A Convolutional Neural-Network Approach," *IEEE Trans. Neural Networks,* vol. 8, pp. 98-113, 1997.
[12] J. Weng, J.S. Huang, and N. Ahuja, "Learning Recognition and Segmentation of 3D objects from 2D images," *Proc. IEEE Int'l Conf. Computer Vision,* pp. 121-128, 1993.
[13] S.H. Lin, S.Y. Kung, and L.J. Lin, "Face Recognition/Detection by Probabilistic Decision-Based Neural Network," *IEEE Trans. Neural Networks,* vol. 8, pp. 114-132, 1997.
[14] S.Y. Kung and J.S. Taur, "Decision-Based Neural Networks with Signal/Image Classification Applications," *IEEE Trans. Neural Networks,* vol. 6, pp. 170-181, 1995.
[15] M. Lades, J.C. Vorbrüggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Würtz, and M. Konen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture," *IEEE Trans. Computers,* vol. 42, pp. 300-311, 1993.
[16] L. Wiskott and C. von der Malsburg, "Recognizing Faces by Dynamic Link Matching," *Neuroimage 4,* pp. S14-S18, 1996.
[17] S. Tamura, H. Kawa, and H. Mitsumoto, "Male/Female Identification from 8×6 Very Low Resolution Face Images by Neural Network," *Pattern Recognition,* vol. 29, pp. 331-335, 1996.
[18] Y. Kaya and K. Kobayashi, "A Basic Study on Human Face Recognition," *Frontiers of Pattern Recognition,* S. Watanabe, ed., p. 265, 1972.
[19] T. Kanade, "Picture Processing by Computer Complex and Recognition of Human Faces," technical report, Dept. Information Science, Kyoto Univ., 1973.
[20] A.J. Goldstein, L.D. Harmon, and A.B. Lesk, "Identification of Human Faces," *Proc. IEEE,* vol. 59, p. 748, 1971.
[21] R. Bruneli and T. Poggio, "Face Recognition: Features versus Templates," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 15, pp. 1042-1052, 1993.
[22] I.J. Cox, J. Ghosn, and P.N. Yianios, "Feature-Based Face Recognition Using Mixture-Distance," *Computer Vision and Pattern Recognition,* 1996.
[23] B.S. Manjunath, R. Chellappa, and C. von der Malsburg, "A Feature Based Approach to Face Recognition," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,* pp. 373-378, 1992.
[24] R.J. Baron, "Mechanism of Human Facial Recognition," *Int'l J. Man Machine Studies,* vol. 15, pp. 137-178, 1981.
[25] M. Bichsel, "Strategies of Robust Object Recognition for Identification of Human Faces," PhD thesis, Eidgenossischen Technischen Hochschule, Zurich, 1991.
[26] M.A. Grudin, "A Compact Multi-Level Model for the Recognition of Facial Images," PhD thesis, Liverpool John Moores Univ., 1997.
[27] F. Samaria and F. Fallside, "Face Identification and Feature Extraction Using Hidden Markov Models," *Image Processing: Theory and Application,* G. Vernazza, ed., Elsevier, 1993.
[28] F. Samaria and A.C. Harter, "Parameterisation of a Stochastic Model for Human Face Identification," *Proc. Second IEEE Workshop Applications of Computer Vision,* 1994.
[29] P.J.M. van Laarhoven and E.H.L. Aarts, *Simulated Annealing: Theory and Applications.* Kluwer Academic Publishers, 1987.
[30] R.H.J.M. Otten and L.P.P.P. van Ginneken, *The Annealing Algorithm.* Kluwer Academic Publishers, 1989.
[31] S. Kirkpatrick Jr., C.D. Gelatt, and M.P. Vecchi, "Optimization by Simulated Annealing," *Science,* vol. 220, pp. 671-680, 1993.
[32] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, pp. 711-720, 1997.
[33] Bern Univ. Face Database, ftp://iamftp.unibe.ch/pub/Images/FaceImages/, 2002.
[34] Purdue Univ. Face Database, http://rvl1.ecn.purdue.edu/~aleix/aleix_face_DB.html, 2002.
[35] Yale Univ. Face Database, http://cvc.yale.edu/projects/yalefaces/yalefaces.html, 2002.
[36] A.M. Martinez and R. Benavente, "The AR Face Database," CVC Technical Report no. 24, June 1998.
[37] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proc. IEEE,* vol. 83, pp. 705-740, 1995.
[38] Y. Moses and S. Ullman, "Limitation of Non-Model-Based Recognition Schemes," *Proc. European Conf. Computer Vision,* pp. 820-828, 1992.
[39] Y. Moses, "Face Recognition: Generalization to Novel Images," PhD thesis, Weizman Inst. of Science, 1993.
[40] Y. Adini, Y. Moses, and S. Ullman, "Face Recognition: The Problem of Compensation for Changes in Illumination Direction," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, pp. 721-732, 1997.
[41] Y.L. Tong, "The Multivariate Normal Distribution," *Springer Series in Statistics.* Springer-Verlag, 1990.
[42] M.K.H. Leung and Y.H. Yang, "Dynamic Two-Strip Algorithm in Curve Fitting," *Pattern Recognition,* vol. 23, pp. 69-79, 1990.
[43] B. Duc, S. Fischer, and J. Bigun, "Face Authentication with Gabor Information on Deformable Graphs," *IEEE Trans. Image Processing,* vol. 8, pp. 504-516, 1999.
[44] M.J. Lyons, J. Budynek, and S. Akamatsu, "Automatic Classification of Single Facial Images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, pp. 1357-1362, 1999.
[45] I. Craw, N. Costen, T. Kato, and S. Akamatsu, "How Should We Represent Faces for Automatic Recognition?" *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, pp. 725-736, 1999.
[46] C. Kotropoulos, A. Tefas, and I. Pitas, "Frontal Face Authentication Using Morphological Elastic Graph Matching," *IEEE Trans. Image Processing,* vol. 9, pp. 555-560, 2000.
[47] A. Pentland and T. Choudhury, "Face Recognition for Smart Environments," *Computer,* vol. 33, pp. 50-55, 2000.
[48] I. Biederman and J. Gu, "Surface versus Edge-Based Determinants of Visual Recognition," *Cognitive Psychology,* vol. 20, pp. 38-64, 1988.
[49] V. Bruce et al., "The Importance of "Mass" in Line Drawings of Faces," *Applied Cognitive Psychology,* vol. 6, pp. 619-628, 1992.

Yongsheng Gao received the BSc and MSc degrees in electronic engineering from Zhejiang University, China, in 1985 and 1988 respectively, and the PhD degree in computer engineering from Nanyang Technological University, Singapore. Currently, He is an assistant professor at Nanyang Technological University, Singapore. His research interests include face recognition, biometrics, image retrieval, computer vision, and pattern recognition. He is a member of the IEEE.

Maylor K.H. Leung received the BSc degree in physics from the National Taiwan University in 1979 and the BSc, MSc, and PhD degrees in computer science from the University of Saskatchewan, Canada, in 1983, 1985, and 1992, respectively. Currently, Dr. Leung is an associate professor at Nanyang Technological University, Singapore. His research interest is in the area of computer vision, pattern recognition, and image processing. He is particularly interested in improving security using visual information. He is a member of the IEEE.