Efficient Iris Recognition by Characterizing Key Local Variations

Li Ma, Tieniu Tan, *Fellow, IEEE*, Yunhong Wang, *Member, IEEE*, and Dexin Zhang

***Abstract*—Unlike other biometrics such as fingerprints and face, the distinct aspect of iris comes from randomly distributed features. This leads to its high reliability for personal identification, and at the same time, the difficulty in effectively representing such details in an image. This paper describes an efficient algorithm for iris recognition by characterizing key local variations. The basic idea is that local sharp variation points, denoting the appearing or vanishing of an important image structure, are utilized to represent the characteristics of the iris. The whole procedure of feature extraction includes two steps: 1) a set of one-dimensional intensity signals is constructed to effectively characterize the most important information of the original two-dimensional image; 2) using a particular class of wavelets, a position sequence of local sharp variation points in such signals is recorded as features. We also present a fast matching scheme based on exclusive OR operation to compute the similarity between a pair of position sequences. Experimental results on 2 255 iris images show that the performance of the proposed method is encouraging and comparable to the best iris recognition algorithm found in the current literature.**

***Index Terms*—Biometrics, iris recognition, local sharp variations, personal identification, transient signal analysis, wavelet transform.**

I. Introduction

With an increasing emphasis on security, automated personal identification based on biometrics has been receiving extensive attention over the past decade. Biometrics [1], [2] aims to accurately identify each individual using various physiological or behavioral characteristics, such as fingerprints, face, iris, retina, gait, palm-prints and hand geometry etc. Recently, iris recognition is becoming an active topic in biometrics due to its high reliability for personal identification [1]–[3], [9], [10], [12]. The human iris, an annular part between the pupil (generally appearing black in an image) and the white sclera as shown in Fig. 1, has an extraordinary structure and provides many interlacing minute characteristics such as freckles, coronas, stripes, furrows, crypts and so on. These visible characteristics, generally called the texture of the iris, are unique to each subject [5]–[12], [34]–[36]. The uniqueness of the iris pattern is the direct result of the individual differences that exist in the development of the anatomical structures in the body. Some research work [12], [34]–[36] has also stated that the iris is essentially stable over a person's life. Furthermore, since the iris is an internal organ as well as externally visible, iris-based personal identification systems can be noninvasive to their users [9]–[12], [35], [36], which is of great importance for practical applications. All these desirable properties (i.e., uniqueness, stability, and noninvasiveness) make iris recognition a particularly promising solution to security.

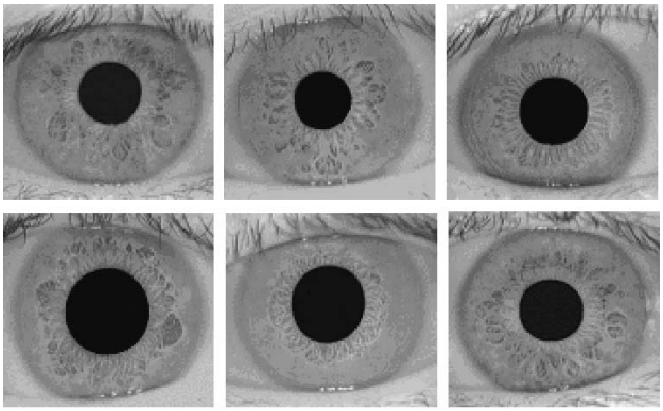

Fig. 1. Samples of iris images.

A. Related Work

Flom and Safir first proposed the concept of automated iris recognition in 1987 [34]. Since then, some researchers worked on iris representation and matching and have achieved great progress [7]–[23], [35], [36]. Daugman [8]–[10] made use of multiscale Gabor filters to demodulate texture phase structure information of the iris. Filtering an iris image with a family of filters resulted in 1024 complex-valued phasors which denote the phase structure of the iris at different scales. Each phasor was then quantized to one of the four quadrants in the complex plane. The resulting 2048-component iriscode was used to describe an iris. The difference between a pair of iriscodes was measured by their Hamming distance. Sanchez-Reillo *et al.* [16] provided a partial implementation of the algorithm by Daugman. Wildes *et al.* [11] represented the iris texture with a Laplacian pyramid constructed with four different resolution levels and used the normalized correlation to determine whether the input image and the model image are from the same class. Boles and Boashash [13] calculated a zero-crossing representation of one-dimensional (1-D) wavelet transform at various resolution levels of a concentric circle on an iris image

Manuscript received January 1, 2003; revised October 1, 2003. This work was supported by the National Science Foundation of China (NSFC) under Grants 69825105, 60121302, 60332010, and 60275003, and by the Chinese Academy of Sciences. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Philippe Salembier.

L. Ma is with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China (e-mail: lma@nlpr.ia.ac.cn).

T. Tan, Y. Wang, and D. Zhang are with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China (e-mail: tnt@nlpr.ia.ac.cn; wangyh@nlpr.ia.ac.cn; dxzhang@nlpr.ia.ac.cn).

Digital Object Identifier 10.1109/TIP.2004.827237

to characterize the texture of the iris. Iris matching was based on two dissimilarity functions. In [19], Sanchez-Avila *et al.* further developed the iris representation method by Boles *et al.* [13]. They made an attempt to use different similarity measures for matching, such as Euclidean distance and Hamming distance. Lim *et al.* [15] decomposed an iris image into four levels using 2-D Haar wavelet transform and quantized the fourth-level high-frequency information to form an 87-bit code. A modified competitive learning neural network (LVQ) was adopted for classification. Tisse *et al.* [20] analyzed the iris characteristics using the analytic image constructed by the original image and its Hilbert transform. Emergent frequency functions for feature extraction were in essence samples of the phase gradient fields of the analytic image's dominant components [25], [26]. Similar to the matching scheme of Daugman, they sampled binary emergent frequency functions to form a feature vector and used Hamming distance for matching. Park *et al.* [21] used a directional filter bank to decompose an iris image into eight directional subband outputs and extracted the normalized directional energy as features. Iris matching was performed by computing Euclidean distance between the input and the template feature vectors. Kumar *et al.* [22] utilized correlation filters to measure the consistency of iris images from the same eye. The correlation filter of each class was designed using the two-dimensional (2-D) Fourier transforms of training images. If the correlation output (the inverse Fourier transform of the product of the input image's Fourier transform and the correlation filter) exhibited a sharp peak, the input image was determined to be from an authorized subject, otherwise an imposter. Bae *et al.* [23] projected the iris signals onto a bank of basis vectors derived by independent component analysis and quantized the resulting projection coefficients as features.

Our earlier attempts to iris recognition developed the texture analysis-based methods [14], [17], [18] and a local intensity variation analysis-based method [37]. In [17], the global texture features of the iris were extracted by means of well-known Gabor filters at different scales and orientations. Based on the experimental results and analysis obtained in [17], we further constructed a bank of spatial filters [18], whose kernels are suitable for iris recognition, to represent the local texture features of the iris and thus achieved much better results. Different from the above two methods, we also developed a Gaussian–Hermite moments-based method [37]. This method is our preliminary work using local intensity variations of the iris as features. Gaussian–Hermite moments [24] which use Gaussian–Hermite polynomial functions as transform kernels belong to a class of orthogonal moments. This means that they produce minimal information redundancy. Gaussian–Hermite moments can well characterize local details of a signal since they construct orthogonal features from the signal's derivatives of different orders [24]. We decomposed an iris image into a set of 1-D intensity signals (see Section IV-A for more details) and represented local variations of the intensity signals using Gaussian–Hermite moments (order 1 to 4). To reduce computational cost and improve classification accuracy, we adopted Fisher linear discriminant to reduce the dimensionality of original features and the nearest center classifier for matching.

It should be noted that all these algorithms are based on gray images, and color information is not used. The main reason is that the most important information for recognition (i.e., texture variations of the iris) is the same in both gray and color images. From the methods described above, we can conclude that there are four main approaches to iris representation: phase-based methods [8]–[10], zero-crossing representation [13], [19], texture analysis [11], [14], [15], [17], [18], [21], and intensity variation analysis [23], [37]. However, the question of which approach is most suitable for extracting iris features has never been answered.

B. Outline

In this paper, we first present our intuitive observations about the characteristics of the iris based on the appearance of numerous iris images, and then introduce a new algorithm for iris recognition inspired by such observations. Finally, we perform a series of experiments to evaluate the proposed algorithm. Moreover, in order to answer the question of which approach is most suitable for extracting iris features, we carry out extensive quantitative comparison among some existing methods and provide detailed discussions on the overall experimental results. To the best of our knowledge, this is the first attempt in comparing the existing algorithms on a reasonably sized database.

The remainder of this paper is organized as follows. Section II provides an overview of our method based on an intuitive understanding for iris features. Detailed descriptions of image preprocessing, feature extraction and matching are given in Section III and Section IV respectively. Experimental results and discussions are reported in Section V. Section VI concludes this paper.

II. Overview of Our Approach

Six iris samples captured by our home-made digital optical sensor are shown in Fig. 1. We can see from these images that the iris consists of many irregular small blocks, such as freckles, coronas, stripes, furrows, crypts, and so on. Furthermore, the distribution of these blocks in the iris is also random. Such randomly distributed and irregular blocks constitute the most distinguishing characteristics of the iris.

Intuitively, if we can precisely locate each of these blocks in the image and recognize the corresponding shape as well, then we will obtain a high performance algorithm. But it is almost impossible to realize such an idea. Unlike fingerprint verification, where feature extraction can rely on ridge following, it is difficult to well segment and locate such small blocks in gray images. Moreover, classifying and recognizing the shape of such blocks is unpractical due to their great irregularity. From the viewpoint of signal processing, however, we can regard these irregular blocks as a kind of transient signals. Therefore, iris recognition can be solved using some approaches to transient signal analysis. As we know, local sharp variations denote the most important properties of a signal. In our framework, we thus record the position of local sharp variation points as features instead of locating and recognizing those small blocks. Fig. 2 illustrates the main steps of our method.

First, the background in the iris image is removed by localizing the iris. In order to achieve invariance to translation and scale, the annular iris region is normalized to a rectangular block

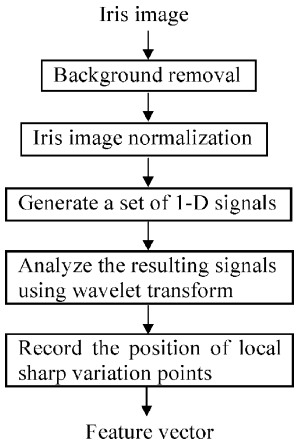

Fig. 2. Diagram of our approach.

of a fixed size. After lighting correction and image enhancement, we construct a set of 1-D intensity signals containing the main intensity variations of the original iris for subsequent feature extraction. Using wavelet analysis, we record the position of local sharp variation points in each intensity signal as features. Directly matching a pair of position sequences is also very time-consuming. Here, we adopt a fast matching scheme based on the exclusive OR operation to solve this problem. The proposed method is detailed in the following sections.

III. Iris Image Preprocessing

Our preprocessing operates in three steps. First, the iris is localized and the irrelevant parts (e.g. eyelid, pupil etc.) are removed from the original image. Then, the localized iris is unwrapped to a rectangular block of a fixed size in order to reduce the deformation caused by variations of the pupil and obtain approximate scale invariance. Finally, lighting correction and contrast improvement are applied to compensate for differences of imaging conditions.

A. Localization

The iris is an annular portion between the pupil (inner boundary) and the sclera (outer boundary). Both the inner boundary and the outer boundary of a typical iris can approximately be taken as circles. However, the two circles are usually not concentric [8]. We first roughly determine the iris region in the original image, and then use edge detection and Hough transform to exactly compute the parameters of the two circles in the determined region. The detailed steps are as follows.

1) Project the image in the vertical and horizontal direction to approximately estimate the center coordinates (X_p, Y_p) of the pupil. Since the pupil is generally darker than its surroundings, the coordinates corresponding to the minima of the two projection profiles are considered as the center coordinates of the pupil

$$X_p = \arg\min_x \left(\sum_y I(x,y) \right)$$
$$Y_p = \arg\min_y \left(\sum_x I(x,y) \right) \quad (1)$$

where X_p and Y_p denote the center coordinates of the pupil in the original image $I(x,y)$.

2) Compute a more accurate estimate of the center coordinates of the pupil. We binarize a 120×120 region centered at the point (X_p, Y_p) by adaptively selecting a reasonable threshold using the gray level histogram of this region. The centroid of the resulting binary region is considered as a new estimate of the pupil coordinates. Note that one can improve accuracy for estimating the center coordinates of the pupil by repeating this step since the coordinates (X_p, Y_p) estimated by image projection described in the first step are sometimes slightly far from the real center coordinates of the pupil.
3) Calculate the exact parameters of these two circles using edge detection (Canny operator [27] in our experiments) and Hough transform [28] in a certain region determined by the center of the pupil (X_p, Y_p).

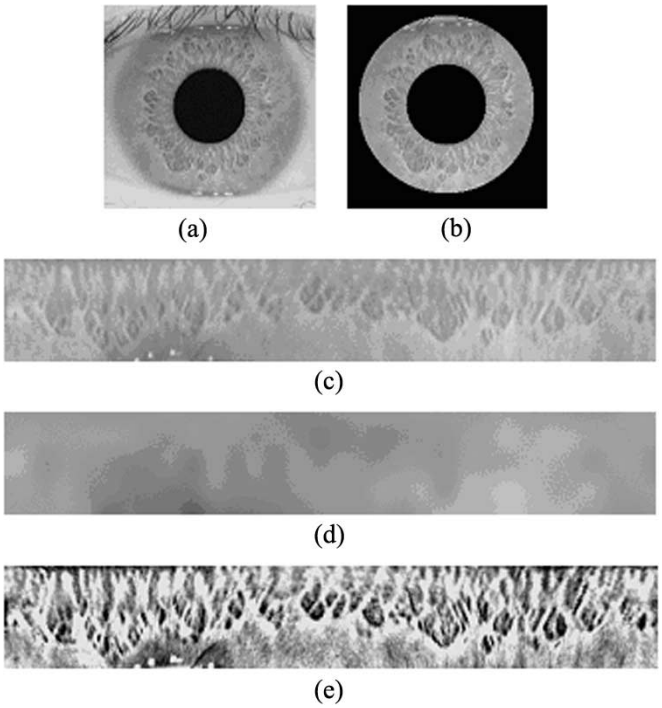

Fig. 3. Iris image preprocessing: (a) original image; (b) localized image; (c) normalized image; (d) estimated local average intensity; and (e) enhanced image.

In the experiments, we perform the second step twice for a reasonably accurate estimate. Compared with the localization method by Wildes *et al.* [11] where the combination of edge detection and Hough transform is also adopted, our method approximates the iris region before edge detection and Hough transform. This will reduce the region for edge detection and the search space of Hough transform, and thus result in lower computational cost. An example of iris localization is shown in Fig. 3(b).

B. Normalization

Irises from different people may be captured in different size, and even for irises from the same eye, the size may change due to illumination variations and changes of the camera-to-eye dis-

tance. Such elastic deformation in iris texture will affect the matching results. For the purpose of achieving more accurate recognition results, it is necessary to compensate for such deformation. Wildes *et al.* [11] solved this problem by registering the input image with the model image. Daugman [8]–[10] represented the iris using a fixed parameter interval in a doubly dimensionless pseudo polar coordinate system, whereas our previous method [14], [17], [18] normalized the iris into an image of a fixed size. The normalization schemes described in [15], [16], [20], [21], and [23] are similar to our approach. These existing methods are essentially the same except the method by Wildes *et al.* [11]. In experiments, we counter-clockwise unwrap the annular iris to a rectangular texture block with a fixed size. The normalization not only reduces to a certain extent the distortion of the iris caused by pupil movement but also simplifies subsequent processing.

C. Enhancement

The normalized iris image has low contrast and may have nonuniform brightness caused by the position of light sources. All these may affect the subsequent processing in feature extraction and matching. In order to obtain a more well-distributed texture image, we first approximate intensity variations across the whole image. The mean of each 16×16 small block constitutes a coarse estimate of the background illumination. This estimate is further expanded to the same size as the normalized image by bicubic interpolation. The estimated background illumination as shown in Fig. 3(d) is subtracted from the normalized image to compensate for a variety of lighting conditions. Then we enhance the lighting corrected image by means of histogram equalization in each 32×32 region. Such processing compensates for the nonuniform illumination, as well as improves the contrast of the image. Fig. 3(e) shows the preprocessing result of an iris image, from which we can see that finer texture characteristics of the iris become clearer than those in Fig. 3(c).

IV. Feature Extraction and Matching

As mentioned earlier, the characteristics of the iris can be considered as a sort of transient signals. Local sharp variations are generally used to characterize the important structures of transient signals. We thus construct a set of 1-D intensity signals which are capable of retaining most sharp variations in the original iris image. Wavelet transform is a particularly popular approach to signal analysis and has been widely used in image processing [29]–[33]. In this paper, a special class of 1-D wavelets (the wavelet function is a quadratic spline of a finite support) is adopted to represent the resulting 1-D intensity signals. The position of local sharp variation points is recorded as features.

A. Generation of 1-D Intensity Signals

Local details of the iris generally spread along the radial direction in the original image corresponding to the vertical direction in the normalized image [see Fig. 3(e)]. Therefore, information density in the angular direction corresponding to the horizontal direction in the normalized image is much higher than that in other directions [9], [18]; i.e., it may suffice only to capture local sharp variations along the horizontal direction in the normalized image to characterize an iris. In addition, since our basic idea is to represent the randomly distributed blocks of the iris by characterizing local sharp variations of the iris, it is unnecessary to capture local sharp variation points in every line of the iris image for recognition. Bearing these two aspects in mind, we decompose the 2-D normalized image I into a set of 1-D intensity signals S according to the following equation:

$$S_i = \frac{1}{M}\sum_{j=1}^{M} I_{(i-1)^* M+j} \quad i = 1, 2, \ldots N$$

$$I = \begin{pmatrix} I_1 \\ \vdots \\ I_x \\ \vdots \\ I_K \end{pmatrix} = \left(I_1^T, \cdots I_x^T, \cdots I_K^T\right)^T \tag{2}$$

where I is the normalized image of $K \times L$ (64×512 in our experiments), I_x denotes gray values of the xth row in the image I, M is the total number of rows used to form a signal S_i, N is the total number of 1-D signals. In essence, each intensity signal is a combination of M successive horizontal scan lines which reflect local variations of an object along the horizontal direction. A set of such signals contains the majority of the local sharp variations of the iris. This is confirmed by the experimental results reported in Section V. Moreover, such processing reduces the computational cost required for subsequent feature representation. In experiments, we find that the iris regions close to the sclera contain few texture characteristics and are easy to be occluded by eyelids and eyelashes. Therefore, we extract features only in the top-most 78% section (corresponding to the regions closer to the pupil) of the normalized image. The relation between the total row number K of the normalized image, the total number N of 1-D signals and the number M of rows used to form a 1-D signal is denoted as $K \times 78\% = N \times M$. Since the total row number K of the normalized image is fixed, the product of the total number N of 1-D signals and the number M of rows used to form a 1-D signal is a constant in experiments. The recognition rate of the proposed algorithm can be regulated by changing the parameter M. A small M leads to a large set of signals which results in characterizing the iris details more completely, and thus increases recognition accuracy. A large M, however, implies a lower recognition rate with a higher computational efficiency. This way, we can trade off between speed and accuracy. In experiments, we choose $M = 5$ and $N = 10$.

B. Feature Vector

As a well-known multiresolution analysis approach, the dyadic wavelet transform has been widely used in various applications, such as texture analysis, edge detection, image enhancement and data compression [29]–[33]. It can decompose a signal into detail components appearing at different scales. The scale parameter of the dyadic wavelets varies only along the dyadic sequence $(2^j)_{j \in z}$. Here, our purpose is to precisely locate the position of local sharp variations which generally indicate the appearing or vanishing of an important

image structure. The dyadic wavelets satisfy such requirements as well as incur lower computational cost, and are thus adopted in our experiments. The dyadic wavelet transform of a signal $S(x)$ at scale 2^j is defined as follows (in convolution form):

$$WT_{2^j}S(x) = \frac{1}{2^j}\int S(X)\psi\left(\frac{x-X}{2^j}\right)dX \tag{3}$$

where $\psi(x/2^j)$ is the wavelet function at scale 2^j. In our algorithm, the function $\psi(x)$ is a quadratic spline which has a compact support and one vanishing moment [29], [31], [33]. This means that local extremum points of the wavelet transform correspond to sharp variation points of the original signal. Therefore, using such a transform, we can easily locate the iris sharp variation points by local extremum detection. Mallat [33] has proved that the dyadic wavelet transform based on the above wavelet function could be calculated with a fast filter bank algorithm. The detailed implementation may be referred to [33]. As (3) shows, the wavelet transform of a signal includes a family of signals providing detail components at different scales. There is an underlying relationship between information at consecutive scales, and the signals at finer scales are easily contaminated by noise. Considering these two points (information redundancy at consecutive scales and the effect of noise on signals at finer scales), we only use two scales to characterize differences among 1-D intensity signals. As we know, a local extremum is either a local minimum or a local maximum. Iris images shown in Fig. 3 illustrate that the irregular blocks of the iris are slightly darker than their surroundings. Therefore, it is reasonable to consider that a local minimum of the wavelet transform described above denotes the appearing of an irregular block and a local maximum denotes the vanishing of an irregular block. A pair of adjacent local extremum points (a minimum point and a maximum point) indicates that a small block may exist between them. However, there are a few adjacent local extremum points between which the amplitude difference is very small. Such local extremum points may correspond to relatively faint characteristics in the iris image (i.e., local slow variations in the 1-D intensity signals) and are less stable and reliable for recognition. A threshold-based scheme is used to suppress them. If the amplitude difference between a pair of adjacent local extrema is less than a predetermined threshold, such two local extremum points are considered from faint iris characteristics and not used as discriminating features. That is, we only utilize distinct iris characteristics (hence local sharp variations) for accurate recognition. For each intensity signal S_i, the position sequences at two scales are concatenated to form the corresponding features:

$$f_i = \{d_1, d_2, \cdots d_i, \cdots d_m; d_{m+1}, d_{m+2}, \cdots d_{m+n}; p_1, p_2\} \tag{4}$$

where the first m components are from the first scale, the next n components from the other scale, d_i denotes the position of a local sharp variation point in the intensity signal, p_1 and p_2, respectively, represent the property of the first local sharp variation point at two scales. If the first local sharp variation point d_1 (or d_{m+1}) is a local minimum of the wavelet transform, p_1 (or p_2) is set to 1, otherwise -1. Features from different 1-D intensity signals are concatenated to constitute an ordered feature vector

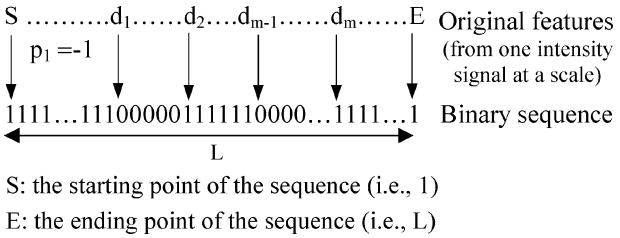

Fig. 4. Illustration of feature transform.

$$f = \{f_1, f_2, \cdots f_i, \cdots f_N\} \tag{5}$$

where f_i denotes the features from the ith intensity signal, and N is the total number of 1-D intensity signals. Note that since the number of local sharp variation points is distinct for different irises, the dimensionality of the feature vector f is not a constant. In our experiments, the feature vector generally consists of about 660 components. In addition, at each scale, we replace each feature component d_i denoting the actual position of local sharp variation points by its difference d_i' with the previous component. This will save memory since the difference d_i' (i.e., the interval between two consecutive local sharp variation points) is less than 256 and can be represented in a byte.

C. Matching

We determine whether two irises are from the same class by comparing the similarity between their corresponding feature vectors. Directly computing the distance between a pair of position sequences is inconvenient. Inspired by the matching scheme of Daugman [8]–[10], a two-step approach is proposed to solve this problem.

1) The original feature vector is expanded into a binary feature vector (called feature transform, in our algorithm).
2) The similarity between a pair of expanded feature vectors is calculated using the exclusive OR operation.

Fig. 4 illustrates how to transform original features derived from one intensity signal at a scale into a sequence of 0's and 1's (hereinafter, called a binary sequence). The length L of the binary sequence at a scale is the same as the length of the 1-D intensity signal defined in (2). At each position denoted by the original feature components, the binary sequence changes from 1 to 0 or from 0 to 1. In the algorithm, we utilize the component p of the original features to set the values of the first $(d_1 - 1)$ elements of the binary sequence. We define that the values of the first $(d_1 - 1)$ elements of the binary sequence are set to 1 if p_1 is equal to -1 (i.e., d_1 is a local maximum point of the wavelet transform), otherwise 0. By such processing, a position sequence as shown in Fig. 4 can be expanded into a binary sequence. Using the same scheme, the original features of an iris defined in (5) can be written as

$$Ef = \{Ef_{(1,1)}, Ef_{(1,2)}, \cdots Ef_{(i,1)}, Ef_{(i,2)} \cdots Ef_{(N,1)}, Ef_{(N,2)}\} \tag{6}$$

where $Ef_{(i,1)}$ and $Ef_{(i,2)}$ are the binary sequences from the ith 1-D intensity signal at the first and the second scale, re-

TABLE I
CASIA IRIS DATABASE

Subjects		Age and Gender		Time Interval	Environment
Chinese	95.3%	Age < 25	41%	At lease one month between two capture stages	Normal office conditions (indoor)
		25 < = Age < 50	55%		
Others	4.7%	Age >=50	4%		
		Male : Female	7 : 3		

spectively. As described above, the length of a binary sequence $Ef_{(i,1)}$ is L. Therefore, the expanded binary feature vector Ef defined in (6) contains $2L \times N$ components whose values are 0 or 1. For such feature vectors, a very fast exclusive OR operation can be utilized for matching. The similarity function is defined in the following:

$$D = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2L}\sum_{j=1}^{2}\left(Ef^1_{(i,j)} \oplus Ef^2_{(i,j)}\right) \tag{7}$$

where Ef^1 and Ef^2 denote two different expanded binary feature vectors, $\oplus$ is the exclusive OR operator, L is the length of the binary sequence at one scale, and N is the total number of 1-D intensity signals.

It is desirable to obtain an iris representation invariant to translation, scale, and rotation. By simple operations, the proposed iris representation can be translation, scale and rotation invariant. Translation invariance is inherent in our algorithm because the original image is localized before feature extraction. To achieve approximate scale invariance, we normalize irises of different size to the same size. Rotation invariance is important for an iris representation since changes of head orientation and binocular vergence may cause eye rotation. In our algorithm, the annular iris is unwrapped into a rectangular image. Therefore, rotation in the original image corresponds to translation in the normalized image. Fortunately, translation invariance can easily be achieved for the intensity signals derived from the normalized image. Since an intensity signal of length L in essence originates from a closed ring, it can be regards as a complete period of an infinite signal with period L. Accordingly, both the wavelet coefficients and features (including the original features and the expanded binary features) are also considered as signals of period L. Because the binary sequence at each scale [such as $Ef_{(i,1)}$, described in (6)] can be regarded as a periodic signal, we obtain translation invariant matching by circular shift. With a pair of binary sequences from two different images of the same iris as an example, if there is a rotation between two original images, there is a horizontal shift between such a pair of sequences. To register the two sequences, we first circularly shift one of the binary sequences, and then compute the similarity between them. After several circular shifts, the minimum matching score is considered from the situation where the two sequences are best registered and is taken as the final matching score. That is, the proposed algorithm will be rotation invariant only by the circular shift-based matching. The method for solving rotation invariance is similar to the brute force search scheme by Daugman [8]–[10] and our previous scheme [17], [18].

V. EXPERIMENTS AND RESULTS

To evaluate the performance of the proposed method, we collected a large number of iris images using a homemade sensor to form a database named *CASIA Iris Database*. The database includes 2255 iris images from 306 different eyes (hence, 306 different classes) of 213 subjects. The images are acquired during different sessions and the time interval between two collections is at least one month, which provides a challenge to our algorithm. To the best of our knowledge, this is the largest iris database available in the public domain. The profile of the database is shown in Table I. The subjects consist of 203 members of the CAS Institute of Automation and ten visiting students from Europe.

Most existing methods for iris recognition generally used small image sets for performance evaluation, and only the method by Daugman has been tested on a larger image set involving over 200 subjects [3], [9], [10]. As we mentioned earlier, there is no detailed comparison among these methods [9]–[23], despite of the great importance of such comparative studies (especially from the practical point of view). We thus conducted a comparative study of these methods on the *CASIA Iris Database*. The experiments were completed in two modes: verification (one-to-one matching) and identification (one-to-many matching). In verification mode, the receiver operating characteristic (ROC) curve and equal error rate (EER) are used to evaluate the performance of the proposed method. The ROC curve is a false match rate (FMR) versus false non-match rate (FNMR) curve [3], [4], which measures the accuracy of matching process and shows the overall performance of an algorithm. The FMR is the probability of accepting an imposter as an authorized subject and the FNMR is the probability of an authorized subject being incorrectly rejected. Points on this curve denote all possible system operating states in different tradeoffs. The ideal FMR versus FNMR curve is a horizontally straight line with zero false nonmatch rate. The EER is the point where the false match rate and the false nonmatch rate are equal in value. The smaller the EER is, the better the algorithm. In identification mode, the algorithm is measured by correct recognition rate (CRR), the ratio of the number of samples being correctly classified to the total number of test samples.

A. Performance Evaluation of the Proposed Method

To assess the accuracy of the proposed algorithm, each iris image in the database is compared with all the other irises in the database. For the *CASIA Iris Database*, the total number of comparisons is 2 304 242, where the total number of intra-class comparisons is 7223 and that of inter-class comparisons is 2 297 019.

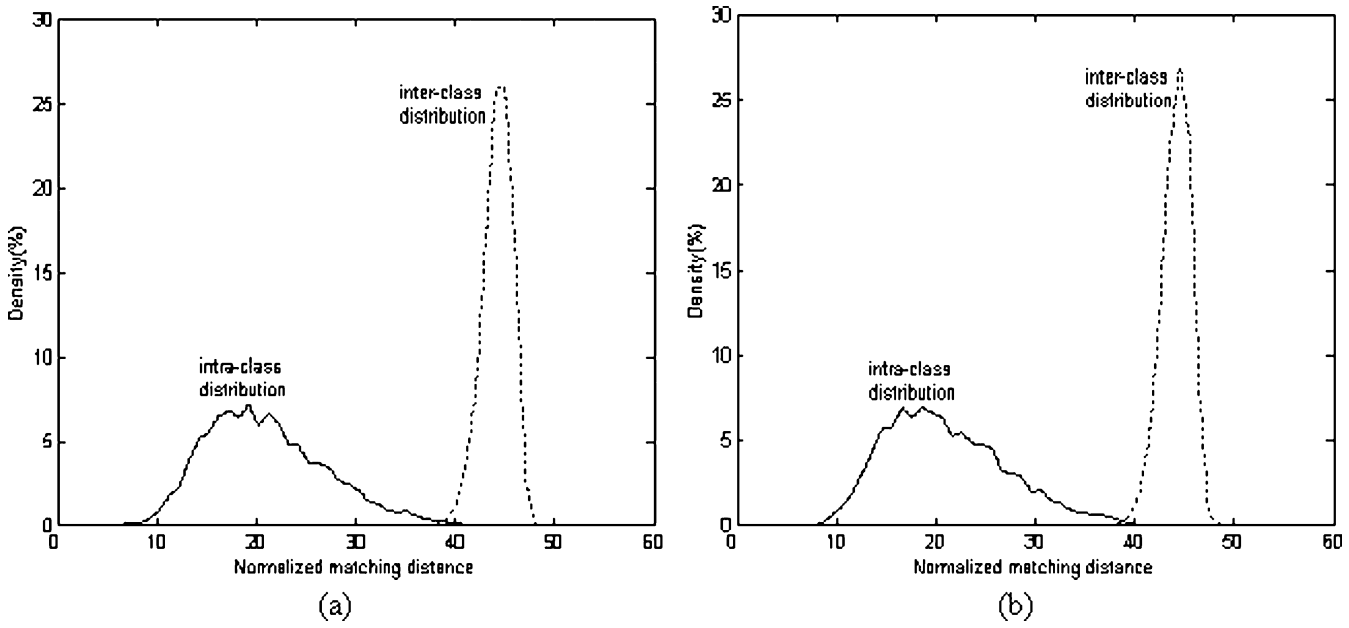

Fig. 5. Distributions of intra-class and inter-class distances. (a) The results from comparing images taken at the same session. (b) The results from comparing images taken at different sessions.

As we know, the time lag between the date when images used for building the templates are captured and the date when the test images are taken has an effect on intra-class matching distances since there may be great variations between images of the same iris taken at different time. To explore the impact of the time lag on the proposed method, we respectively analyze the experimental results based on comparisons between images taken at the same session and those based on comparisons between images taken at different sessions. In our experiments, the time lag between different capture sessions is at least one month. For the results from comparing images taken at the same session, the distribution of the intra-class matching distance is estimated with 3512 comparisons and the inter-class distribution is estimated with 1 165 164 comparisons. For the results from comparing images taken at different sessions, the distribution of the intra-class matching distance is estimated with 3711 comparisons and the inter-class distribution is estimated with 1 131 855 comparisons. Fig. 5 shows distributions of intra-class and inter-class matching distances in the two cases.

For a satisfying biometrics algorithm, intra-class distances should hardly vary with time. From the results shown in Fig. 5, we can see that the intra-class distance distribution derived from comparing images of the same iris taken at the same session and that derived from comparing images of the same iris taken at different sessions are very close. This demonstrates the high stability of the proposed iris features. It should be noted that the same iris sensor is used to capture images at different sessions. If we make use of different iris sensors at different image capture sessions, the differences between the above two intra-class distance distributions may increase. However, this needs to be further investigated. Fig. 5 also reveals that the distance between the intra-class and the inter-class distribution is large, indicating the good discriminability of the extracted features. This is verified by the following verification results. Fig. 6 shows the ROC curves of the proposed method, from which two observations can be made. First, the ROC curve based on different session comparisons interlaces with that based on the same session comparisons. That is, the performance change caused by the time-lag is extremely small for the proposed method. Second, the performance of our algorithm is very high and the EER is only 0.09% for different session comparisons. In particular, if one and only one false match occurs in 1 000 000 trails, false nonmatch rate is less than 1.60%. The above experimental results are highly encouraging. This also demonstrates that our iris representation and matching schemes are very effective and the 1-D intensity signals defined in (2) well capture the most discriminating information of the iris.

Experiments were carried out to investigate the cause of a few large intra-class distances. Such two pairs of iris images are listed in Figs. 7 and 8, from which two main reasons can be identified.

1) Eyelids and eyelashes may occlude the effective regions of the iris for feature extraction, and the failure of iris localization (i.e., large localization errors) may cause false nonmatching. Such an example is shown in Fig. 7. In our experiments, we found that 57.7% false nonmatches are incurred by the occlusion of eyelids and eyelashes and 21.4% false nonmatches come from the inaccurate localization. In addition, the inaccurate localization usually occurs in the occluded images since the eyelids and eyelashes bring about some edge noises and decrease localization accuracy. In order to reduce such false nonmatches, we are working on detecting eyelids and eyelashes so that feature extraction is only performed in the regions of no occlusion as well as localizing the iris more exactly.
2) As shown in Fig. 8, the difference of the pupil size between these two original images of the same eye is very significant. One has a pupil of normal size, and the other a considerably dilated pupil. Currently, we can recover

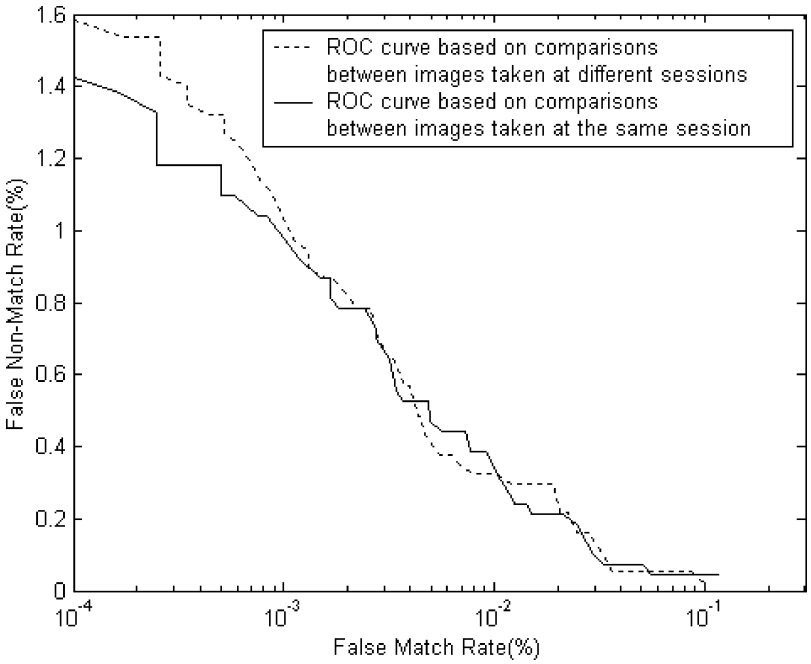

Fig. 6. Verification results.

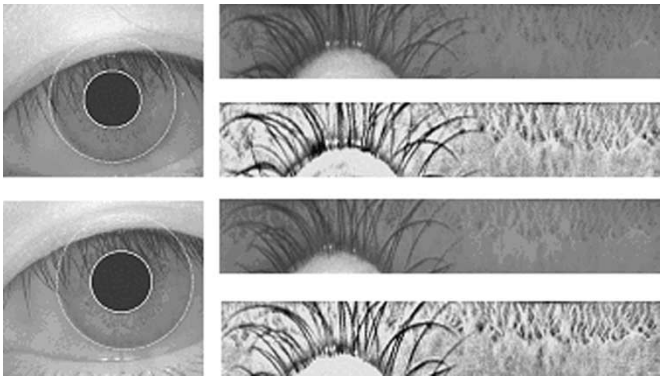

Fig. 7. An example of false nonmatch due to eyelid/eyelash occlusion.

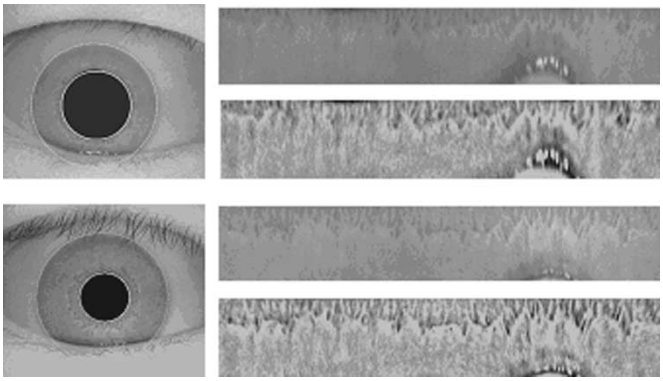

Fig. 8. An example of false nonmatch due to excessive pupil dilation.

the deformed iris caused by the dilation or constriction of the pupil to the normal iris by normalization. However, under the extreme conditions (namely the iris texture is excessively compressed by the pupil), the iris after normalization still has many differences with its normal state (i.e., the iris has a pupil of normal size). Therefore, the matching distance between such a pair of iris images is very large. In our experiments, 10.7% false nonmatches result from the pupil changes. This is a common problem in all iris recognition methods. Iris normalization is thus an important research issue in the future.

Besides the above reasons, other factors that can result in false nonmatching include the specular reflection from the cornea or eyeglasses, poorly focused images and motion-blurred images. Since these factors are relevant to noninvasive iris imaging, it is difficult to completely avoid the corresponding false nonmatches. However, with the improvement of iris imaging, such cases can be reduced.

B. Comparison With Existing Methods

The methods proposed by Daugman [8]–[10], Wildes *et al.* [11], Boles and Boashash [13] are the best known among existing schemes for iris recognition. Furthermore, they characterize local details of the iris based on phase, texture analysis and zero-crossing representation. Therefore, we choose to compare our algorithm with theirs. For the purpose of comparison, we implemented the three methods according to the related papers [8]–[13], [35], [36] (in our current implementations of the methods by Daugman [9] and Wildes *et al.* [11], we did not carry out their schemes for eyelid and eyelash detection. Nevertheless, this should not invalidate our comparison experiments and the resulting conclusions. This is because if we performed eyelid and eyelash detection with all these methods, their respective performance will be slightly improved.). The method by Wildes *et al.* only works in verification mode [11], so we do not test its performance in identification mode. Some methods such as our previous method based on local intensity variation analysis [37] and those presented in [11], [18] need more than one sample for each class for training. Therefore, for

TABLE II
COMPARISON OF CRRS AND EERS

Methods	Correct recognition rate (%)	Equal error rate (%)
Boles [13]	92.64	8.13
Daugman [9]	100	0.08
Previous*	99.60	0.29
Proposed	100	0.07
Tan [18]	99.19	0.57
Wildes [11]	-	1.76

each iris class, we choose three samples from images taken at the first session for training and all samples captured at other sessions serve as test samples. This is also consistent with the widely accepted standard for biometrics algorithm testing [3], [4] (training images and testing images should be respectively captured at different sessions). For the *CASIA Iris Database*, there are 918 images for training and 1237 images for testing. (To satisfy the above requirement, 100 images taken at the first session are not used in the experiments.) When matching the input feature vector with the three templates of a class, the average of the three scores is taken as the final matching distance. Table II and Fig. 9 describe the experimental results conducted on the *CASIA Iris Database* in two modes (verification and identification), where $*$ denotes our previous method based on local intensity variation analysis [37].

From the results shown in Table II and Fig. 9, we can find that Daugman's method and the proposed method have the best performance, followed by our previous method based on local intensity variation analysis [37], our previous method [18], the methods by Wildes *et al.* [11] and Boles *et al.* [13]. The proposed method has an encouraging performance and its EER is only 0.07%. Wildes *et al.* [11] decomposed the iris texture into four different frequency bands and used the normalized correlation for verification. Image registration, which generally brings about high computational cost, is an important step in this method. Although combining the block correlation values by the median operation leads certain robustness against misregistration, misregistration is still a main reason for false nonmatching. That is, misregistration may affect the verification accuracy of this method. Our previous method [18] adopted a bank of spatial filters to represent local texture information of the iris. The disadvantage of this method is that it cannot exactly capture the fine spatial changes of the iris. The theoretical basis of the method [13] comes from the signal reconstruction theory based on the wavelet transform [30], [33]. However, good reconstruction does not necessarily mean accurate recognition. This is because that information used for reconstruction includes some "individual" features of an image which do not exist in all samples from the same class and may reduce recognition accuracy. Furthermore, this method only employed extremely little information along a concentric circle on the iris to represent the whole iris. These factors result in a relatively low accuracy as shown in Fig. 9. More recently, Sanchez-Avila and Sanchez-Reillo [19] further developed the method of Boles *et al.* by using different similarity measures for matching. When the similarity measure was Hamming distance, this method captured a small amount of local variations of the iris and thus achieved 97.9% correct recognition rate on a data set of 200 images from 10 subjects. However, the proposed method constructs a set of intensity signals to contain the most important details of the iris and makes use of stable and reliable local variations of the intensity signals as features. This leads to the high performance of the proposed method. Another method from our group was based on local intensity variation analysis [37]. Similar to the proposed method, it characterized local variations of a set of 1-D intensity signals using Gaussian–Hermite moments and obtained satisfying results. But averaging the adjacent feature components for dimensionality reduction overlooks the effect of the most discriminating features on recognition accuracy. Both the proposed algorithm and the method by Daugman achieve the best results. Daugman projected each small local region onto a bank of Gabor filters, and then quantized the resulting phasor denoted by a complex-valued coefficient to one of the four quadrants in the complex plane. In essence, this scheme analyzed local variations of the iris by comparing and quantizing the similarity between Gabor filters and the local regions. To achieve high accuracy, the local region for feature extraction must be small enough. This results in a high dimensional feature vector (2048 components). Compared with Daugman's method (which is the most exploited commercially), our current method contains about 660 components. This is because that our method only records the position of local sharp variations as features and contains less redundant information. Daugman skillfully applied Gabor filters to represent local shape of the iris, while we used quadratic spline wavelets to characterize local sharp variations of the iris. That is, the two methods aim to capture local variations of the iris as discriminating features. Therefore, they achieve quite close performance. Since the proposed iris representation is based on 1-D filtering and feature matching adopts the exclusive OR operation, our method is very efficient. Table III illustrates the computational cost of the methods described in [9], [11], [13], [18] and the current algorithm, including the CPU time for feature extraction (from an input image to a feature vector) and matching (from a pair of feature vectors to the matching result).

The above experiments used 200 different iris images and were performed in Matlab 6.0 on a 500-MHz PC with 128 M RAM. Table III shows that the current method, our previous method [37] and the method of Boles *et al.* [13] consume less time than others [9], [18] for feature extraction. The reason is that they are based on 1-D signal analysis and the other methods involve 2-D mathematical operation. The method by Wildes *et al.* [11] only takes about 210 ms to build a four-level Laplacian pyramid representation of an image, whereas the piecewise correlation based matching generally needs high computational expense as shown in Table III. Since Daugman's method and our method can compute the distance between a pair of feature vectors by the exclusive OR operation, they implement matching faster than others. If the exclusive OR operation is carried out using some optimization methods in C/C++, the running time for matching may be further reduced (as Daugman reported in [9], the matching speed of his method can be only 0.01 ms). Note that the method by Wildes *et al.* [11] and our previous method [37] require extra cost for image registration and feature reduction, respectively.

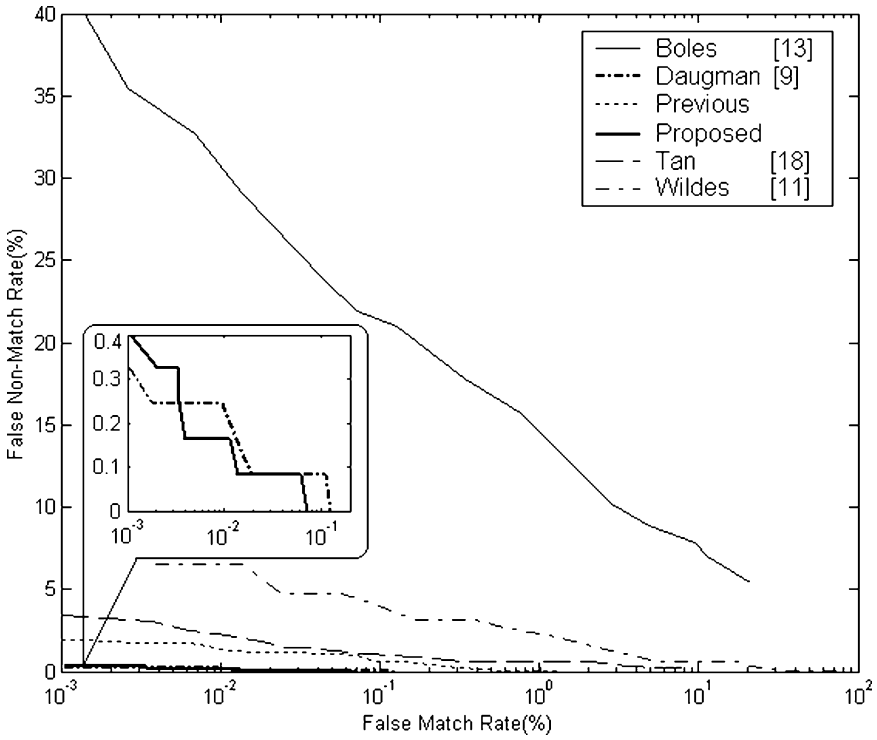

Fig. 9. Comparison of ROCs.

TABLE III
COMPARISON OF THE COMPUTATIONAL COMPLEXITY

Methods	Feature extraction (ms)	Matching (ms)	Feature extraction + Matching (ms)	Others
Daugman [9]	682.5	4.3	686.8	-
Wildes [11]	210.0	401.0	611.0	Registration
Boles [13]	170.3	11.0	181.3	-
Tan [18]	426.8	13.1	439.9	-
Previous*	260.2	8.7	268.9	Feature reduction
Proposed	244.2	6.5	250.7	-

* denotes our previous method based on local intensity variation analysis [37].

C. Discussions

From the above analysis and comparison, we can draw the following conclusions:

1) The proposed method and the methods described in [9], [13] can be considered as the local variation analysis based schemes. They all make some attempts to analyze local intensity variations of the iris using similar ideas but different approaches, namely they utilize local variations to reflect shape information of the iris characteristics. In recently published work on iris recognition, the methods described in [16], [19], [20], [23] also adopt different approaches to represent local variations of the iris, whereas Park *et al.* [21] uses texture features (normalized directional energy features) for recognition. Therefore, existing methods for iris recognition should be essentially classified into two main categories: local variation analysis based methods [8]–[10], [13], [16], [19], [20], [23] and texture analysis based methods [11], [15], [17], [18], [21].
2) In general, the methods based on local variation analysis (Daugman's method [9] and our previous one [37]) have better performance than the texture analysis based methods (see Table II and Fig. 9). The main reason is that texture features are incapable of precisely capturing local fine changes of the iris since texture is by nature a regional image property. In other words, features which can effectively represent local variations are more efficient than texture features for iris recognition.
3) All experimental results have demonstrated that the proposed method achieves high performance in both speed and accuracy. This confirms that local sharp variations truly play an important role in iris recognition and our intuitive observation and understanding are reasonable.
4) The current features are a kind of "minutiae" features or local features, so it may be easily affected by iris localization, noise and the iris deformation caused by pupil movement. To improve robustness, we can add some global features, which are complementary to local features, into the current method. A combination of local and global fea-

tures is expected to further improve the performance of our method. Texture features are well-known global information. In the next stage, we will associate the current features with global texture information for more efficient and robust iris recognition.

VI. Conclusions

In this paper, we have presented an efficient algorithm for iris recognition which is invariant to translation, scale and rotation. This method regards the texture of the iris as a kind of transient signals and uses the wavelet transform to process such signals. The local sharp variation points, good indicators of important image structures, are extracted from a set of intensity signals to form discriminating features. Experimental results have illustrated the encouraging performance of the current method in both accuracy and speed. In particular, a comparative study of existing methods for iris recognition has been conducted. Such performance evaluation and comparison not only verify the validity of our observation and understanding for the characteristics of the iris but also will provide help for further research.

Acknowledgment

The authors would like to thank Dr. J. Daugman (Cambridge University, U.K.) and Dr. R. Wildes (York University, Canada) for their helpful comments and discussions. They also thank the anonymous referees for their thorough reviews and constructive comments.

References

[1] A. Jain, R. Bolle, and S. Pankanti, Eds., *Biometrics: Personal Identification in a Networked Society*. Norwell, MA: Kluwer, 1999.
[2] D. Zhang, *Automated Biometrics: Technologies and Systems*. Norwell, MA: Kluwer, 2000.
[3] T. Mansfield, G. Kelly, D. Chandler, and J. Kane, "Biometric product testing final report," Nat. Physical Lab., Middlesex, U.K., 2001.
[4] A. Mansfield and J. Wayman, "Best practice standards for testing and reporting on biometric device performance," Nat. Physical Lab., Middlesex, U.K., 2002.
[5] F. Adler, *Physiology of the Eye: Clinical Application*, 4th ed. London, U.K.: C.V. Mosby Co., 1965.
[6] H. Davision, *The Eye*. London, U.K.: Academic, 1962.
[7] R. Johnson, "Can iris patterns be used to identify people?," Chemical and Laser Sciences Division LA-12 331-PR, Los Alamos Nat. Lab., Los Alamos, CA, 1991.
[8] J. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Analy. Machine Intell.*, vol. 15, pp. 1148–1161, Nov. 1993.
[9] ——, "Statistical richness of visual phase information: update on recognizing persons by iris patterns," *Int. J. Comput. Vis.*, vol. 45, no. 1, pp. 25–38, 2001.
[10] ——, "Demodulation by complex-valued wavelets for stochastic pattern recognition," *Int. J. Wavelets, Multi-Res. and Info. Processing*, vol. 1, no. 1, pp. 1–17, 2003.
[11] R. Wildes, J. Asmuth, G. Green, S. Hsu, R. Kolczynski, J. Matey, and S. McBride, "A machine-vision system for iris recognition," *Mach. Vis. Applic.*, vol. 9, pp. 1–8, 1996.
[12] R. Wildes, "Iris recognition: an emerging biometric technology," *Proc. IEEE*, vol. 85, pp. 1348–1363, Sept. 1997.
[13] W. Boles and B. Boashash, "A human identification technique using images of the iris and wavelet transform," *IEEE Trans. Signal Processing*, vol. 46, pp. 1185–1188, Apr. 1998.
[14] Y. Zhu, T. Tan, and Y. Wang, "Biometric personal identification based on iris patterns," in *Proc. Int. Conf. Pattern Recognition*, vol. II, 2000, pp. 805–808.
[15] S. Lim, K. Lee, O. Byeon, and T. Kim, "Efficient iris recognition through improvement of feature vector and classifier," *ETRI J.*, vol. 23, no. 2, pp. 1–70, 2001.
[16] R. Sanchez-Reillo and C. Sanchez-Avila, "Iris recognition with low template size," in *Proc. Int. Conf. Audio and Video-Based Biometric Person Authentication*, 2001, pp. 324–329.
[17] L. Ma, Y. Wang, and T. Tan, "Iris recognition based on multichannel Gabor filtering," in *Proc. 5th Asian Conf. Computer Vision*, vol. I, 2002, pp. 279–283.
[18] ——, "Iris recognition using circular symmetric filters," in *Proc. 16th Int. Conf. Pattern Recognition*, vol. II, 2002, pp. 414–417.
[19] C. Sanchez-Avila and R. Sanchez-Reillo, "Iris-based biometric recognition using dyadic wavelet transform," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 17, pp. 3–6, Oct. 2002.
[20] C. Tisse, L. Martin, L. Torres, and M. Robert, "Person identification technique using human iris recognition," in *Proc. Vision Interface*, 2002, pp. 294–299.
[21] C. Park, J. Lee, M. Smith, and K. Park, "Iris-based personal authentication using a normalized directional energy feature," in *Proc. 4th Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 2003, pp. 224–232.
[22] B. Kumar, C. Xie, and J. Thornton, "Iris verification using correlation filters," in *Proc. 4th Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 2003, pp. 697–705.
[23] K. Bae, S. Noh, and J. Kim, "Iris feature extraction using independent component analysis," in *Proc. 4th Int. Conf. Audio- and Video-Based Biometric Person Authentication*, 2003, pp. 838–844.
[24] J. Shen, W. Shen, and D. Shen, "On geometric and orthogonal moments," *Int. J. Pattern Recognit. Artific. Intell.*, vol. 14, no. 7, pp. 875–894, 2000.
[25] T. Tangsukson and J. Havlicek, "AM-FM image segmentation," in *Proc. EEE Int. Conf. Image Processing*, 2000, pp. 104–107.
[26] J. Havlicek, D. Harding, and A. Bovik, "The mutli-component AM-FM image representation," *IEEE Trans. Image Processing*, vol. 5, pp. 1094–1100, June 1996.
[27] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 679–698, Nov. 1986.
[28] D. Ballard, "Generalized Hough transform to detect arbitrary patterns," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-13, pp. 111–122, 1981.
[29] S. Mallat and W. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, pp. 617–643, Mar. 1992.
[30] S. Mallat, "Zero-crossings of a wavelet transform," *IEEE Trans. on Inform. Theory*, vol. 37, pp. 1019–1033, July 1992.
[31] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 710–732, July 1992.
[32] Q. Tieng and W. Boles, "Recognition of 2D object contours using the wavelet transform zero-crossing representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 910–916, Aug. 1997.
[33] S. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic, 1999.
[34] L. Flom and A. Safir, "Iris Recognition system," U.S. Patent 4 641 394, 1987.
[35] J. Daugman, "Biometric personal identification system based on iris analysis," U.S. Patent 5 291 560, 1994.
[36] R. Wildes, J. Asmuth, S. Hsu, R. Kolczynski, J. Matey, and S. Mcbride, "Automated, noninvasive iris recognition system and method," U.S. Patent 5 572 596, 1996.
[37] L. Ma, "Personal identification based on iris recognition," Ph.D dissertation, Inst. Automation, Chinese Academy of Sciences, Beijing, China, June 2003.

Li Ma received the B.Sc. and M.Sc. degrees in automation engineering from Southeast University, Nanjing, China, in 1997 and 2000, respectively, and the Ph.D. degree in computer science from the National Laboratory of Pattern Recognition, Chinese Academy of Sciences, Beijing, China, in 2003.

Currently, he is a Research Member of the IBM China Research Laboratory, Beijing, China. His research interests include image processing, pattern recognition, biometrics, Web mining, and knowledge management.

Tieniu Tan (M'92–SM'97–F'04) received the B.Sc. degree in electronic engineering from Xi'an Jiaotong University, Xi'an, China, in 1984 and the M.Sc., DIC, and Ph.D. degrees in electronic engineering from Imperial College of Science, Technology and Medicine, London, U.K., in 1986, 1986, and 1989, respectively.

He joined the Computational Vision Group at the Department of Computer Science, The University of Reading, Reading, U.K., in October 1989, where he worked as Research Fellow, Senior Research Fellow, and Lecturer. In January 1998, he returned to China to join the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China. He is currently Professor and Director of the National Laboratory of Pattern Recognition, as well as President of the Institute of Automation. He has published widely on image processing, computer vision, and pattern recognition. His current research interests include speech and image processing, machine and computer vision, pattern recognition, multimedia, and robotics.

Dr. Tan serves as referee for many major national and international journals and conferences. He is an Associate Editor of *Pattern Recognition* and IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, and the Asia Editor of Image and Vision Computing. He was an elected member of the Executive Committee of the British Machine Vision Association and Society for Pattern Recognition (1996–1997) and is a Founding Co-Chair of the IEEE International Workshop on Visual Surveillance.

Yunhong Wang (M'03) received the B.Sc. degree in electronic engineering from Northwestern Polytechnical University, Xi'an, China, and the M.S. and Ph.D. degrees in electronic engineering from Nanjing University of Science and Technology, Nanjing, China, in 1995 and 1998, respectively.

She joined the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 1998, where she has been an Associate Professor since 2000. Her research interests include biometrics, pattern recognition, and image processing.

Dexin Zhang received the B.Sc. degree in automation engineering from Tsinghua University, Beijing, China, in 2000. He is currently pursuing his Master's degree at the National Laboratory of Pattern Recognition, Chinese Academy of Sciences, Beijing.

His research interests include biometrics, image processing, and pattern recognition.