Iris Feature Extraction Using Independent Component Analysis

Kwanghyuk Bae, Seungin Noh, and Jaihie Kim

Department of Electrical and Electronic Engineering, Yonsei University
Biometrics Engineering Research Center, Seoul, Korea
paero@yonsei.ac.kr

Abstract. In this paper, we propose a new feature extraction algorithm based on Independent Component Analysis (ICA) for iris recognition. A conventional method based on Gabor wavelets should select the parameters (e.g., spatial location, orientation, and frequency) for fixed bases. We apply ICA to generating optimal basis vectors for the problem of extracting efficient feature vectors which represent iris signals. The basis vectors learned by ICA are localized in both space and frequency like Gabor wavelets. The coefficients of the ICA expansion are used as feature vector. Then, each iris feature vector is encoded into an iris code. Experimental results show that our proposed method has a similar Equal Error Rate (EER) to a conventional method based on Gabor wavelets and two advantages: first, the size of an iris code and the processing time of the feature extraction are significantly reduced; and second, it is possible to estimate the linear transform for feature extraction from the iris signals themselves.

1 Introduction

The iris recognition system has been in the limelight for high-security biometric applications since the method which uses unique patterns of the human iris makes it possible to recognize an individual without any contact or invasion at extremely high confidence levels[1][3]. In order to use the iris pattern for identification, it is important to define a representation that is well adapted for extracting the iris information content from images of the human eye[5].

The feature vector for the iris recognition system should satisfy the two basic requirements. First, feature vector must have distinctiveness for the personal identification. In general, the gray-level information of iris image is treated as the characteristic. It is required that iris image be effectively transformed into the distinctive feature vectors for recognition. Second, iris feature vector should be compact code. The iris recognition system has been used under an extensive database because of the unique nature and the extreme richness of the human iris. So, the size of the iris code should be compact for enrollment or recognition, and it must be accomplished at very high rates of speed to extract iris feature. In practice, iris recognition can also be used with card technologies (e.g., smart cards), where an iris code is stored on the card and then verified with the acquired iris code at the point of interaction[2]. But the limited memory size of cards is responsible for the relative high cost. Therefore, the size of iris code should be reduced.

Most works on the iris feature extraction have been done by linear transforms. In the early 1990s, J. Daugman developed the feature extraction algorithm using Gabor wavelets which showed promising results in [3]. R. Wildes used Laplacian pyramid to represent distinctive spatial characteristics of the iris, and W. Boles proposed the one-dimensional dyadic wavelet transform to obtain zero-crossing representation[4][5]. In general, the linear transforms such as Gabor and wavelets are employed to extract local characteristics from texture images. However, one disadvantage that the linear transforms inherently possess is that the basis vectors are fixed independently of any data. Therefore, conventional methods for iris feature extraction should select the parameters (e.g., spatial location, orientation, and frequency) for fixed bases.

In this paper, we introduce a new method to find the iris feature using the Independent Component Analysis (ICA), which satisfies two requirements mentioned above. Since the ICA is a good method to estimate the linear transform from the iris data themselves, the transform is ideally adapted to iris texture. The transformed coefficients from iris signal are used as the iris feature vector. For efficiently storing and comparing the iris feature vectors, the iris code is generated by quantizing the coefficients. Experimental results show that the size of an iris code and the processing time of the feature extraction are significantly reduced compared to Gabor wavelets

2 Independent Component Analysis

The ICA is an unsupervised learning algorithm using high order statistics. It would be most useful to estimate the linear transform from the data itself, in which case the transform could be ideally adapted to the kind of data that is being processed. The conventional transforms are fixed transforms, meaning that the basis vectors are fixed once and for all, independence of any data[6][7]. To define ICA, we can use a statistical latent variable model. We observe n random variables $x_1, \ldots, x_n$, which are modeled as linear combination of n random variables $s_1, \ldots, s_n$:

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \ldots + a_{in}s_n \tag{1}$$

where the a_{ij}, and i, $j = 1, \ldots, n$ are some real coefficients. By definition, the s_i are statistically mutually independent. This is the basic ICA model. The independent components s_i are latent variables, meaning that they cannot be directly observed. Also the mixing coefficients a_{ij} are assumed to be unknown. All we observe is the random variable x_i, and we must estimate both the mixing coefficients a_{ij} and the independent components s_i using x_i. This could be done under general assumptions. If we use vector-matrix notation, the mixing model is written as

$$\mathbf{x} = \mathbf{A}\mathbf{s} = \sum_{i=1}^{n} \mathbf{a}_i s_i \tag{2}$$

This model is considered as linear combination of basis vectors $\mathbf{a}_i$ in image processing. And the basis vectors are clearly localized in space, as well as in frequency and orientation[6]. Our goal is to find a good set of basis vectors to represent iris patterns effectively.

3 Representation of Iris Patterns

We assume that the observed iris signals are generated by a linear combination of unknown set of statistically independent sources. We estimate the linear transform that generates the iris signal using ICA. The coefficients of a linear transform of iris signal are the iris feature vector.

3.1 Selection of Iris Signals

A set of regions is taken with 100 $\times$ 16 pixels at random locations of each stretched iris image from database, and used as training data. For the testing data, we select a set of 14 regions. We excluded regions at the top and bottom of iris image which are partially occluded by eyelid and illumination. These training or testing data are first converted to one-dimensional signals with length 100. This reason we have chosen one-dimensional signals is that the patterns of iris in the direction of circumference are more various than those in the radial direction (from center to outside). Although we model the iris inner boundary as a circle, we can not completely describe it. There are some errors for modeling its boundary. So, we multiply training or testing iris data by Gaussian function to minimize errors, and make one-dimensional iris signals, which are shown in Fig. 1.

3.2 ICA Training

We choose the FastICA algorithm to estimate the ICA basis vectors[6][7]. This algorithm provides rapid convergence and estimates the independent components by maximizing a measure of independence among the estimated original components. The idea of FastICA is that independent components are the maximally nongaussian components. The FastICA is based on a fixed-point iteration scheme for finding a maximum of the nongaussianity of $y = \mathbf{w}^T\mathbf{z}$. In this algorithm the weight vector $\mathbf{w}$ is updated as follows:

$$\mathbf{w} \Leftarrow E\left\{\mathbf{z}g\left(\mathbf{w}^T\mathbf{z}\right)\right\} - E\left\{g'\left(\mathbf{w}^T\mathbf{z}\right)\right\}\mathbf{w} \tag{3}$$

$$\mathbf{w} \Leftarrow \mathbf{w}/\|\mathbf{w}\| \tag{4}$$

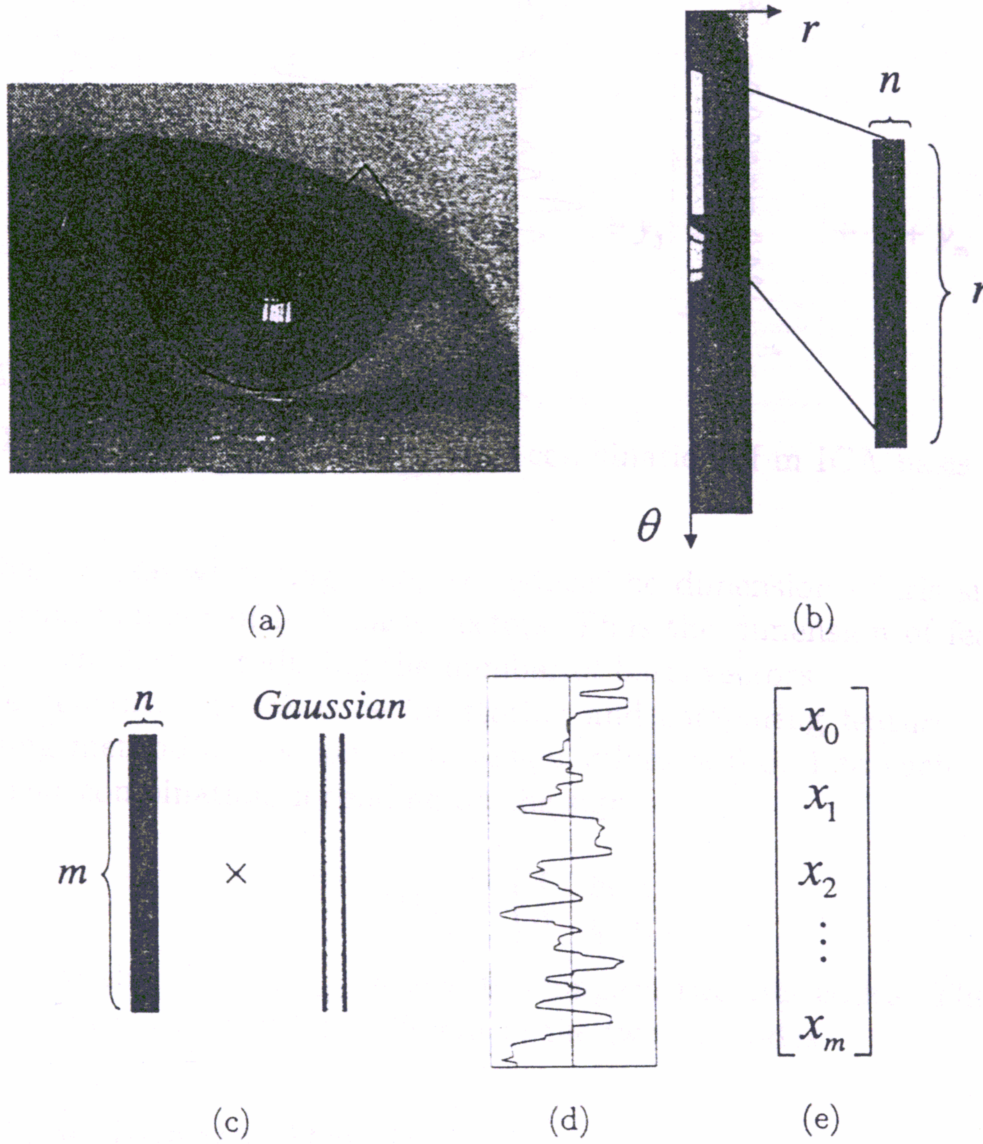

Fig. 1. Generating an iris signal: (a) An eye image after iris localization. (b) Extraction of an iris region from a stretching iris image. (c) Multiplying an iris region by gaussian function. (d) An iris signal. (e) An iris Vector

where the $\mathbf{z}$ denotes the data that has been prewhitened, and g is a suitable nonlinearity (e.g., the tanh function). After convergence, $\mathbf{y} = \mathbf{w}^T\mathbf{z}$ gives an estimate of one of the independent components.

3.3 Extraction of Iris Feature Vectors

Our approach considers that the iris signal is described using a set of basis vectors, which are estimated using FastICA. The components of the iris feature vector are the coefficients of the linear combination of the ICA basis vectors. Fig. 2 shows that iris signal can be expressed by these linear combination. The dimension of iris feature space is the same number of ICA basis vectors. A common preprocessing step in ICA is whitening of the input data, which removes of the second order statistic effects and filters some additive noise as much as

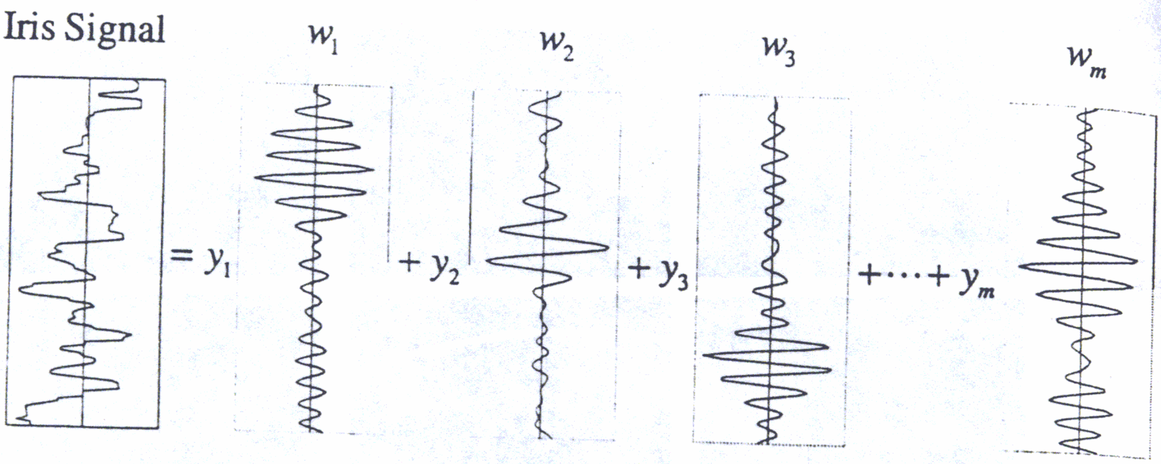

Fig. 2. An iris signal as a linear combination of m ICA basis vectors

possible. In the whitening step, we reduce the dimension of iris signal that is the actual number of ICA basis vectors. Thus the dimension of feature vector becomes smaller by reducing the number of basis vectors.

We generate the iris code for storing and comparing feature vectors. The encoding method of iris code is to assign values of 0 or 1 to each coefficient of the linear combination depending on the sign.

$$q(y_i) = \begin{cases} 1 & y_i \geq 0 \\ 0 & y_i < 0 \end{cases} \tag{5}$$

We use the Hamming distance to compare two iris codes. The Hamming distance is the count of bits different in the two iris codes.

4 Experimental Results

Our algorithm was tested on the iris image database which consists of 990 images, with 10 images for each person. Our database includes the various iris images obscured by eyelids, eyelashes, and specular reflections from the cornea or from eyeglasses. Some examples are shown in Fig. 3. Two experiments are performed to evaluate our proposed method; one shows that the number of ICA basis vectors influences Equal Error Rate (EER) of the system, and the other compares our method with the Gabor wavelets method.

Fig. 4 shows the influence of the number of ICA basis vectors on the EER in our iris database. The reason for the increase of EER in a large number of ICA basis vectors is that small eigenvalues are included in the whitening step, and this leads to an unstable ICA estimation. The explanation for this behavior is that the small eigenvalues correspond to high-frequency components and usually encode noise. A small number of ICA basis vectors lead to increasing EER as well, since the whitened data fail to capture enough information on the original data. From Fig. 4, the EER corresponding to the whitened data of 28 dimensions, that is the number of ICA basis vectors, has the lowest value. So we set the dimension of the whitened data as 28.

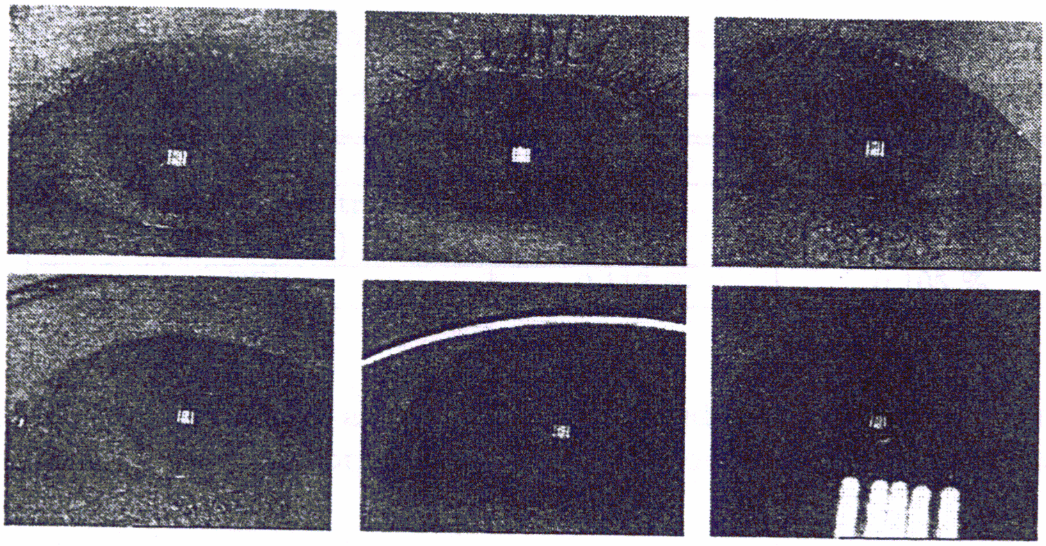

Fig. 3. Examples of iris images in our database

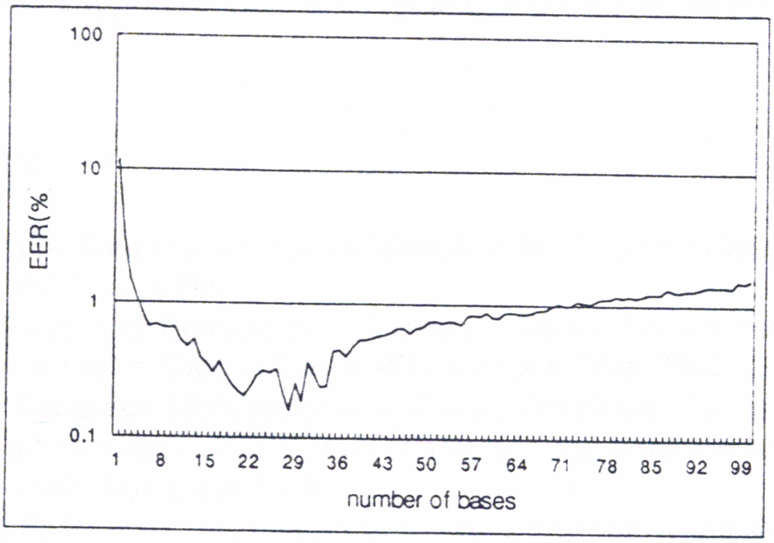

Fig. 4. EER with increasing dimensionality of the ICA bases

We compare the ICA method with the Gabor wavelets method in EER, size of iris code, and processing time. Results are shown in Table 1. This shows that our proposed method enhanced the performance of iris feature extraction. The EER has similar results. But our method reduces tremendously the size of iris code. Also, due to its low computational complexity, ICA method in iris feature extraction leads to decrease of processing time.

5 Conclusion

In this paper, we proposed a new method for iris feature extraction. We made use of independent component analysis which aims to find a linear transform for the iris signal using a set of bases as statistically independent as possible. In experiments compared to conventional method for iris recognition, we found two advantages of our method: first, the size of an iris code and the processing time

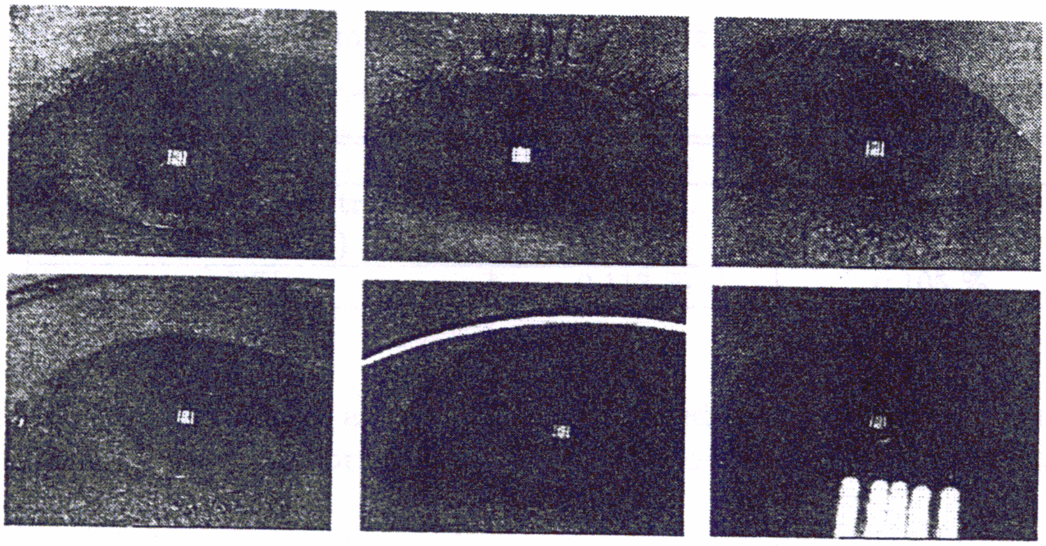

Fig. 3. Examples of iris images in our database

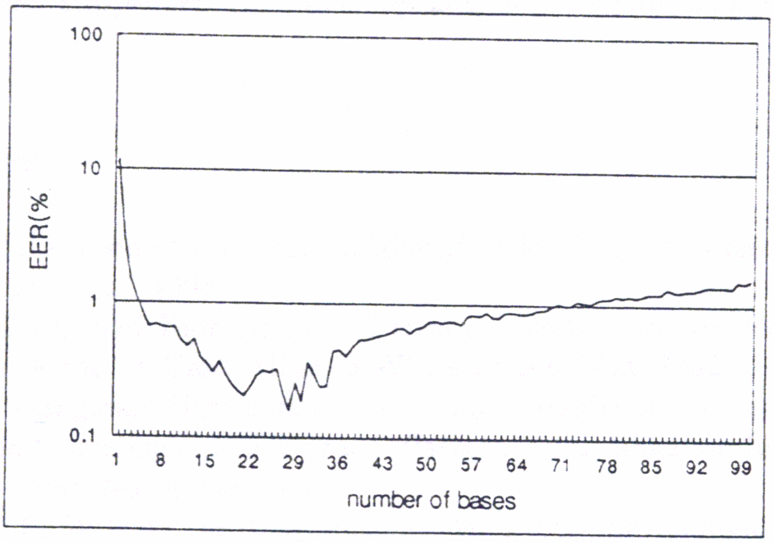

Fig. 4. EER with increasing dimensionality of the ICA bases

We compare the ICA method with the Gabor wavelets method in EER, size of iris code, and processing time. Results are shown in Table 1. This shows that our proposed method enhanced the performance of iris feature extraction. The EER has similar results. But our method reduces tremendously the size of iris code. Also, due to its low computational complexity, ICA method in iris feature extraction leads to decrease of processing time.

5 Conclusion

In this paper, we proposed a new method for iris feature extraction. We made use of independent component analysis which aims to find a linear transform for the iris signal using a set of bases as statistically independent as possible. In experiments compared to conventional method for iris recognition, we found two advantages of our method: first, the size of an iris code and the processing time