A data hiding approach for the self-security of iris recognition

Xueyi Ye (叶学义), Zhiwei He (何志伟), and Zhijing Zhao (赵知劲)

Laboratory of Pattern Recognition and Information Security, Hangzhou Dianzi University, Hangzhou 310018

Received July 19, 2007

Attacks to biometric data are the primary danger to the self-security of biometrics. To improve the iris feature template data security, a data hiding approach based on bit streams is proposed, in which an iris feature template is embedded into a face image. The proposed approach is applicable to present dominant techniques of iris recognition. With the low computation cost and the zero decoding-error-rate, this data hiding approach, embedding target biometric data into other biometric data for improving the security of target data in identity recognition, data storage and transmission, can deceive attackers more effectively. Furthermore, it does not degrade the iris recognition performances. Experimental results prove that the proposed approach can be used to protect iris feature templates and enhance the security of the iris recognition system itself.

OCIS codes: 100.0100, 110.2990, 200.3050.
doi: 10.3788/COL20080607.0487.

It is gradually being real possibility that traditional identification technologies are replaced by biometrics because of intrinsic characteristics of biometric features such as a fingerprint, iris and face. Actually, biometric technologies have been used in some occasions and countries[1]. But, because biometric features are one person's inherent behavior or physiological characteristics, unique and unchangeable, the self-security of biometrics has come to popular attention. If a legal registration user's biometric features were stolen, the consequence is more serious than what happens when a user of the traditional identification techniques loses his or her password, key, or smart card. Therefore, the requisite basis to use biometrics effectively in large scale practical applications is that the biometric features entered Biometrics network systems must come from legal registration users[2].

Generally, a typical biometrics network system comprises four parts: capture level, features level, decision level, and databases, as shown in Fig. 1. In the figure, arrows denote possible attacking behaviors to vulnerable sections of the system. All arrows in Fig. 1 can be separated into three categories. The first one is the impostor's attack (arrow 1) in which forged biometric samples are imported to the capture level to cheat the system. The second one is terminal attacks (arrow 9) in which attackers directly intercept the system export result or take over the end control equipment. The third category includes all others in which all attackers have a similar purpose — whether they attack the system transport channel (arrows 3, 5, and 7) or the system processing and memory unit (arrows 2, 4, 6, and 8) — to steal or tamper legal registration biometric data. So, the self-security of a biometrics network system depends on the security of biometric data much more[3].

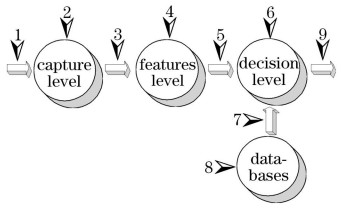


Fig. 1. Biometrics network system and its vulnerable sections under possible attacks.

In order to increase the security of the biometric data, encryption and data hiding technologies have been adopted. Encryption focuses on methods to make encrypted information meaningless to unauthorized parties[4], and data hiding is to hide critical information in unsuspected carrier data. Namely, data hiding is based on concealing the information itself. Data hiding techniques reduce the chance of biometric data intercepted by attackers, hence reducing the chance of illegal uses. Evolving from encryption and data hiding techniques, digital watermark techniques can be used to embed proprietary information, such as company logo, in host data to protect the intellectual property rights of the data[5,6].

There are only a few published papers on the security of the biometric data so far. Jain proposed a semi-unique key method based on local block average to detect tampering of host images, such as fingerprints and faces[7]. Ratha *et al.* described a data hiding method, which is applicable to fingerprint images compressed with wavelet scalar quantization (WSQ) wavelet-based scheme. The discrete wavelet transform coefficients are changed during WSQ encoding, by taking into consideration the possible image degradation[8]. A fragile watermarking method was used by Pankanti and Yeung for fingerprint image verification[9]. A spatial watermarking image was embedded in the spatial domain of a fingerprint image by utilizing a verification key. The method can localize any region of image that has been tampered and did not lead to a significant performance loss in fingerprint verification[9]. Gunsel *et al.* proposed a spatial domain watermarking method utilizing gradient orientation analysis in watermarking embedding for fingerprint images and meanwhile preserving the singular points so as not to affect the classification of the watermarked images[10].

At present, there are many attacking types opposed to the biometric data security. But if we just discuss the biometric feature template data, two modes should be

mentioned. One is fake in which attackers first reconstruct the biometric sample according to an intercepted feature template[3], then import it to cheat the biometric system and get a legal login or an intention export. The other is replacement in which attackers directly use an impostor identity data to replace a genuine one in a database and get a legal identity.

In terms of the above analysis and by emerging the standards of the biometric template data, the existing technologies of encryption, and data hiding[1,10,11], especially digital watermark techniques, we propose a data hiding method based on bit streams to increase the security of iris feature template itself. In this method, an iris feature template is embedded in a face image to protect the template data.

In biometric data hidden process (i.e., encoding) using digital watermark techniques, the biometric data are called the watermark, while the carrier data are called the host data (usually is an image). Consequently, the contrary process in which the biometric data are picked up from the host data is called decoding. The amplitude modulation based technique[12] is one of the techniques widely used in digital watermark. An improved version of the technique was presented in Ref. [4], in which the fingerprint minutiae data as the watermark are hidden into a host image. If $P(i,j)$ denotes the (i,j)th pixel's value in the host image, the encoding calculation is briefly described as

$$P_{\mathrm{WM}}(i,j) = P(i,j) + (2s-1)\cdot P_{\mathrm{AV}}(i,j)\cdot q \cdot(1+\frac{P_{\mathrm{SD}}(i,j)}{A})\cdot(1+\frac{P_{\mathrm{GM}}(i,j)}{B})\cdot\beta(i,j), \quad (1)$$

where $P_{\mathrm{WM}}(i,j)$ is the value of the watermarked pixel corresponding to the original pixel $P(i,j)$ at the location (i,j) of the host image. The value of watermark bit is denoted as s, $s\in[0,1]$. $P_{\mathrm{AV}}(i,j)$ and $P_{\mathrm{SD}}(i,j)$ denote the average and the standard deviations of pixel values in the neighborhood of pixel (i,j) and $P_{\mathrm{GM}}(i,j)$ denotes the gradient magnitude at location (i,j). The parameters A and B are weights for the standard deviation $P_{\mathrm{SD}}(i,j)$ and the gradient magnitude $P_{\mathrm{GM}}(i,j)$, respectively. The watermark embedding strength (or the degree of the host image changed by watermark) is denoted as q ($q>0$). $\beta(i,j)$ denotes the modulated (or watermarked) symbol. If the pixel (i,j) is a marked pixel, $\beta(i,j)$ takes the value 0, otherwise, the value 1. Modulation (or encoding) details are given in Ref. [12].

Extracting the watermark data from the host image, which is decoding, starts with finding the data embedding locations in the host image by the secret key used during the encoding stage. For every bit embedding location (i,j), if its value during decoding is estimated as the linear combination of pixel values in a 5×5 cross-shaped neighborhood of the watermarked pixels, the (i,j)th estimated value $\hat{P}(i,j)$ is calculated according to

$$\hat{P}(i,j) = \frac{1}{8}\Big(\sum_{k=-2}^{2} P_{\mathrm{WM}}(i+k,j) + \sum_{k=-2}^{2} P_{\mathrm{WM}}(i,j+k) - 2P_{\mathrm{WM}}(i,j)\Big). \quad (2)$$

The differences δ between the estimated and watermarked pixel values are calculated by

$$\delta = P_{\mathrm{WM}}(i,j) - \hat{P}(i,j). \quad (3)$$

These differences are averaged over all the embedding locations associated with the same bit, to yield $\bar{\delta}$.

Finally, the watermark bit value $\hat{s}$ is estimated by

$$\hat{s} = \begin{cases} 1, & \text{if } \bar{\delta} > \frac{\bar{\delta}_{\mathrm{R0}}+\bar{\delta}_{\mathrm{R1}}}{2} \\ 0, & \text{else} \end{cases}. \quad (4)$$

To find an adaptive threshold, these averages are calculated separately for the reference bits, 0 and 1, as $\bar{\delta}_{\mathrm{R0}}$ and $\bar{\delta}_{\mathrm{R1}}$, respectively.

Consequently, it is possible to bring the estimation error in the decoding procedure as given in Ref. [12], even if the watermarked image is not tampered. Furthermore, the cost of the encoding and decoding calculation is huge. For example, if the watermark data comprise K bits and the host image comprises S pixels, the redundancy coefficient α will be S/K, and $\alpha>1$. According to these original conditions, the encoding calculation contains more than $95K\alpha$ additions and $62K\alpha$ multiplications, and the decoding calculation contains more than $12K\alpha$ additions and $6K\alpha$ multiplications. Totally, a whole encoding and decoding computation contains at least $107K\alpha$ additions and $68K\alpha$ multiplications.

According to the data hiding method represented above, there are three aspects which could be improved: the redundancy $\alpha>1$ is necessary; the decoding-error-rate is not equal to 0; high computation cost. Two main advantages of the previous iris recognition techniques are the comparatively low recognition error rate and the extraordinary searching efficiency[13]. If the above method is used to iris template data hiding, it would result in two negative effects. Firstly, because the decoding-error-rate is not 0, the equal-error-rate (EER) of iris recognition will increase. Secondly, because of the high computation cost, when the search of the 1:N matching or the m:N list matching[14] is executed, especially for a big value of N, the searching efficiency will decrease observably. It is sure that the decoding-error-rate will decrease when the watermark embedding strength q is enlarged, but the strategy will weaken the hidden effect, consequently.

The data hiding method should not only improve the security of the biometric data, but also prevent the degradation in performance of biometrics. Therefore, we propose a new data hiding method, based on bit stream, to embed an iris feature template into a face image. The proposed method has the following advantages. The decoding-error-rate is zero, the redundancy coefficient is not necessarily to be larger than or equal to 1, and the higher computation efficiency can be obtained. Meanwhile, by regarding other biometric features such as the host data, the fraudulence of the data hiding method is better. The details of the proposed method are described as follows.

An iris feature template is denoted as $I(k,l)$ which is a binary matrix depicted in Fig. 2, comprising K rows and L columns as the embedded data (or the watermark data), where an element of the matrix means a bit.

Let $F(m,n)$ denote a face gray image (e.g., intersection of gray levels 0 and 255) comprising M rows and N

columns as the host data, in which every byte element of the matrix means a pixel in the image. $I(k,l)$ is regarded as the binary code stream and (k,l) denotes the location of an element in $I(k,l)$ ($k \in [1,K]$, $l \in [1,L]$), in which K and L are the numbers of rows and columns in the matrix $I(k,l)$, respectively. (m,n) denotes the location of an element in $F(m,n)$ ($m \in [1,M]$, $n \in [1,N]$), in which M and N are the numbers of rows and columns in the matrix, respectively. It is necessary that $(M \times N) \geq (K \times L)$ (i.e., the redundancy coefficient $\alpha \geq 1$).

As the embedded data, the original bit can discretionarily choose an element of the iris feature template $I(k,l)$, while the information is recorded in the secret key. Then bits in all rows in $I(k,l)$ as shown in Fig. 2 are connected end to end to become a roll. Beginning from the original bit, all bits are embedded into the host image clockwise (or anticlockwise). For example, in Fig. 2, the bit framed with a small rectangle is chosen as the original bit and its value is denoted as $S(i)$, $S(i) \in \{0,1\}$ (where i denotes the bit's location in the sequence and the value of i is equal to $(k-1)L+l$). Likewise, as the host data, the original byte can discretionarily choose a pixel of the face image $F(m,n)$ too, and certainly the information is recorded in the secret key. As shown in Fig. 3, the pixel framed with a white rectangle is the original byte and is denoted as $P(j)$, where j is the pixel's location in the image and $j = (m-1)N + n$. In the gray image, the value of $P(j)$ can be represented by a byte (e.g., the gray scale is $0-255$) as the rectangle frame pointed by the arrow in the right part of Fig. 3. In the rectangle frame, all 8 bits of a byte are shown from the low bit to the high bit ranking from the numbers 0 to 7 under the rectangle, respectively. Each bit of the watermark data is serially embedded in a pixel of the host data (all pixels of $F(m,n)$ are linked end to end, too). Note that $F(m,n)$ can be the whole or the part of the host image.

As shown in the right side of Fig. 3, all bits in the byte are denoted as $P^{(0)}(j), P^{(1)}(j), \cdots, P^{(7)}(j)$, where the superscripts in brackets mean the location in the byte, $P^{(0)}(j)$ denotes the flag bit and the reference bit can be selected from other bits (e.g., $P^{(1)}(j)$ denotes the reference bit). When $S(i)$ is embedded into $P(j)$, the value of $P^{(0)}(j)$ is replaced by $P^{(0)}_{\mathrm{WM}}(j)$ according to

$$P^{(0)}_{\mathrm{WM}}(j) = \begin{cases} 0, & \text{if } S(i) = P^{(1)}(j) \\ 1, & \text{if } S(i) \neq P^{(1)}(j) \end{cases} . \tag{5}$$

Consequently, the pixel $P(j)$ of the host image $F(m,n)$ is denoted as $P_{\mathrm{WM}}(j)$. As shown in Fig. 3, after the encoding calculation, all other bits in $P(j)$ and $P_{\mathrm{WM}}(j)$ are unchangeable except that the value of $P^{(0)}(j)$ changes into the value of $P^{(0)}_{\mathrm{WM}}(j)$. The next bit of $S(i)$ in $I(k,l)$ is embedded into the next byte of $P(j)$ in $F(m,n)$ till each bit of the iris feature template data is orderly hidden in each pixel of the face image one time.

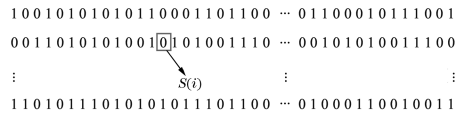


Fig. 2. An iris feature template.

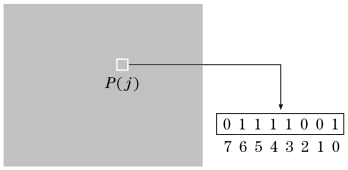


Fig. 3. Host image.

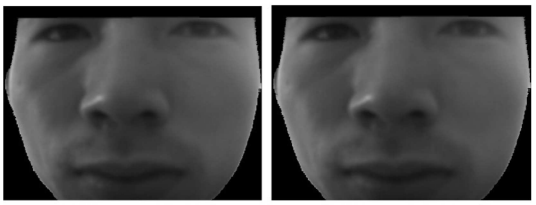
Decoding (i.e., extracting the iris template data $I(k,l)$ from the face image $F(m,n)$) starts with finding the beginning location j ($j \in [1, M \times N]$) of the data hiding (i.e., the original byte $P_{\mathrm{WM}}(j)$ in the host image) by the secret key used during the encoding stage. Furthermore, via the secret key, we can get the beginning location i ($i \in [1, K \times L]$) (i.e., the original bit). For every bit $S(i)$ embedded by $P_{\mathrm{WM}}(j)$, its value during decoding is calculated as

$$S(i) = \begin{cases} P^{(1)}_{\mathrm{WM}}(j) & \text{if } P^{(0)}_{\mathrm{WM}}(j) = 0 \\ P^{(1)}_{\mathrm{WM}}(j) & \text{if } P^{(0)}_{\mathrm{WM}}(j) = 1 \end{cases} , \tag{6}$$

where the reverse of $P^{(1)}_{\mathrm{WM}}(j)$ is denoted as $\overline{P^{(1)}_{\mathrm{WM}}(j)}$, and the bit $P^{(0)}_{\mathrm{WM}}(j)$ and $P^{(1)}_{\mathrm{WM}}(j)$ denote the flag bit and reference bit as same as above, respectively. Finally, every bit $S(i)$ is orderly filled in the corresponding location of the iris template $I(k,l)$.

According to the proposed method based on bit streams, an iris template is directly hidden into a face image. Not only the encoding calculation but also the decoding calculation is just the bit's comparison or reversion, and the computation cost of encoding and decoding depends on the bit length of the iris template data. This processing mode is very suitable to digital computers for computation, so the high calculation efficiency can be obtained. Because the main existing iris recognition techniques are based on the Hamming distance of two iris templates to complete the matching, after getting parameters of the bit length of the iris template and original pixel's locations of host images, two host images embedded in iris templates can be used to calculate the Hamming distance at the same time of the decoding (i.e., it is not necessary to recover the iris template from the face image firstly). Furthermore, the decoding-error-rate of the method itself is zero from the perspective of decoding calculation.

The proposed method is not lossless to the host image because the gray value of the pixel hidden data in the host image is changed. But the change just happens to the lowest bit of a hidden pixel in the image in which the changed amplitude and the probability of the pixel gray value are 1 and 0.25, respectively. In this way, even in the worst situation (e.g., a pixel gray value is added by 1, while gray values of all neighbor pixels are subtracted by 1), there is no difference in vision between the original



(a) (b)

Fig. 4. Comparison of images before and after data hiding. (a) Original image; (b) image hidden data.

host image and the hidden data image as shown in Fig. 4. In Fig. 4, the host image is randomly selected from the Iris and Face Database of the Intelligent Information Processing Laboratory in University of Science and Technology of China and likewise half pixels selected from the image are changed as the worst situation. By the visual observation, it is almost impossible to judge whether the face image is hidden by an iris template.

As the template data security of the iris recognition increases, the proposed bit stream based data hiding method has more practical application values. Firstly, even if the host image is intercepted by attackers, it is not certain for attackers whether the intercepted image is hidden by an iris template. Secondly, even if attackers know that an iris template is embedded into the intercepted image, it is difficult to get the correct iris template data if they do not obtain the secret key used during the encoding. Therefore, the proposed method can effectively protect the security of the iris template and consequently improve the security of the iris recognition system. Moreover, a face image as the host embedding with an iris template using the proposed method can be directly used in the existing face recognition techniques[15], and it is not necessary to recover the original face image by the decoding processing.

In order to assess the influence of the proposed method to the performance of the iris recognition, receiver operating characteristics (ROC) curves for original iris feature templates and templates that are recovered after the data hiding decoding are computed. A total of 1200 iris images and 1200 face images from the above-mentioned database are used in our experiments. These images come from 30 users, with 40 impressions of different gestures, irises, and faces, respectively, of each user, captured by the local multi-biometric collector. Certainly, 1200 original iris feature templates are calculated according to the method presented in Ref. [13], and then these templates are respectively embedded into 1200 face images using the proposed method one by one. To match with Hamming distance as given in Ref. [13], computations of the data hiding decoding and matching can be synchronously finished using the proposed method like the above representation. In both original iris feature templates and templates after decoding, the total number of template matching is $C_{1200}^2 = 719400$, with $30 \cdot C_{40}^2 = 23400$ same identity matches and $C_{1200}^2 - 30 \cdot C_{40}^2 = 696000$ different identity matches. Two ROC curves are shown in Fig. 5, where the horizontal coordinate means false accept rate (FAR) and the vertical coordinate means genuine accept rate (GAR). Experimental results prove that the proposed method do not bring negative effects to the performance of iris recognition.

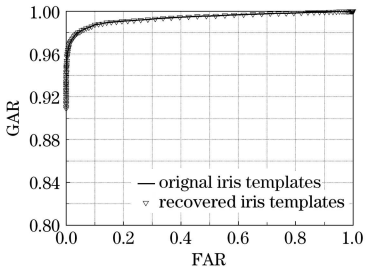


Fig. 5. ROC performance comparison.

In this paper, we analyze the security of a biometrics system and discuss typical vulnerable sections in a biometrics network system. Attacks to iris feature template data are regarded as the primary menace to the security of iris recognition itself. To improve iris feature template data security, a bit stream based data hiding approach is proposed, in which an iris feature template is embedded into a face image to improve the template data security. The proposed approach is applicable to present dominant techniques of iris recognition. With the low computation cost and the zero decoding-error-rate, this approach, embedding an iris template into another biometric data for verification, storage, and transmission, can deceive attackers more effectively. Moreover, those excellent performances of iris recognition are not degraded. Experimental results have shown that the proposed approach can be used to protect iris feature template data and enhance the self-security of the iris recognition system.

This work was supported by the Science and Technology Key Programs of Zhejiang Province, P.R.China under Grant No. 2008C21092. X. Ye's e-mail address is xueyiye@hdu.edu.cn or xueyi_ye@ustc.edu.

References

1. C. A. Shoniregun and S. Crosier, *Securing Biometrics Applications* (Springer, New York, 2007).
2. B. Schneier, Commun. ACM **42,** (8) 136 (1999).
3. A. Adler, in *Proceedings of IEEE CCECE 2003* **2,** 1163 (2003).
4. A. K. Jain and U. Uludag, IEEE Trans. Pattern Anal. Machine Intell. **25,** 1494 (2003).
5. F. Hartung and M. Kutter, Proc. IEEE **87,** 1079 (1999).
6. X. Peng, W. Bai, and J. Tian, Acta Opt. Sin. (in Chinese) **27,** 1011 (2007).
7. S. Jain, in *Proc. Indian Conf. Computer Vision, Graphics and Image Processing* 139 (2000).
8. N. K. Ratha, J. H. Connell, and R. M. Bolle, in *Proc. ACM Multimedia Workshops* 127 (2000).
9. S. Pankanti and M. M. Yeung, Proc. SPIE **3657,** 66 (1999).
10. B. Gunsel, U. Uludag, and A. M. Teklap, Pattern Recogn. **35,** 2739 (2002).
11. W. Bender, D. Gruhl, N. Morimoto, and A. Lu, IBM Systems Journal **35,** 313 (1996).
12. M. Kutter, F. D. Jordan, and F. Bossen, Proc. SPIE **3022,** 518 (1997).
13. J. Daugman, IEEE Trans. Circ. Syst. Video Technol. **14,** 21 (2004).
14. A. K. Jain, S. Pankanti, S. Prabhahar, L. Hong, A. Ross, and J. L. Wayman, in *Proceedings of the 17th International Conference on Pattern Recognition* 935 (2004).
15. X. Ye, "Iris and face based multi-biometrics identification and fusion algorithm", Ph.D Thesis (University of Science and Technology of China, 2006).