International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.8, No.7 (2015), pp.309-316
http://dx.doi.org/10.14257/ijsip.2015.8.7.30

The Key Extraction from Iris Features based on Wavelet Packet

Kun Yu and Juan Wei

Henan Mechanical and Electrical Engineering College, Xinxiang 453002, China
2315306@qq.com

Abstract

Key extraction from the biometric is a new direction of encryption, which solves the safety problem of key a certain extent. Iris has rich textures, and its generated signature is longer than other biometric, and has unique advantages in term of key extraction. The wavelet packet is used to decompose effective iris region into the three layers. The third diagonal high frequency coefficient is extracted as the iris feature, and the encryption key was randomly generated from iris feature to meet the encryption demand. The United States NIST encryption standard is adopted to test the extracted key, and the experimental results show that through the 7 seed test of NIST, the generated key meets the security of encryption algorithm needs.

Keywords: *Wavelet packet; Biometric; key; extraction; NIST test*

1. Introduction

Key number in encryption algorithms is longer, such as AES (Advanced Encryption Standard). Key encryption is generally random binary sequence of 128bit, which is hard to remember when the key will be stored in a file or device, or directly stored in the computer hard disk. The user releases the key through an easily memorial password or secret code. The security of encryption algorithm is based on the security of traditional password. The traditional password is stolen by standers through peeking or observing keystroke. Furthermore, the password has nonrandom (people often choose meaningful digital) and shorter length features, which is also easy to be cracked or through the exhaustive method, potential risk is always hidden for encryption algorithm [1].

The biological characteristics code extracted from the human body features is long for hundreds or even thousands of bits, which key can be generated from [2]. Therefore, the key from the man body itself, which is carried, without memory to solve an escrow problem of the key. In the term of extracting key from the biological characteristics, extracted key of Monors is based on keystroke behavior [3] and voice of speech [4]. Clancy et al. extracted the key from human fingerprint [5].

The key length of different biological characteristics is shown in Table 1. Because the iris has rich texture, the generated code is longer, which means that the key generation space is larger, and extracted key is even longer with higher encryption security. However, the key length generated by other biological characteristics is shorter, which cannot meet the requirements of the encryption algorithm to the length of key (at least 128bit). This becomes the unique advantages of key generated by iris feature. Based on iris feature, the earliest generation of key was Davida's private template scheme in 1998[6]. The iris feature of direct typical extraction regarded as the key, which was bound with error correcting codes. Feng Hao [7] generated a 140bit random key. Then through Reed-Solomon and Hadamard code, the key was expanded as 2048bit pseudo code for iris feature. Finally, XOR operation was done on 2048bit pseudo code with iris feature template to get the lock code, and was stored in the smart card. Boyen used the improved fuzzy extraction technique to generate a key based on iris feature [8]. In the existing

ISSN: 2005-4254 IJSIP

biometric key extraction method, iris feature extracting key has its unique advantages. In the algorithm reports, key security generated by iris is higher compared to other biological features, and therefore key extracted from the iris feature is a promising research direction for information security. However, whether the key extracted from the iris can meet the demand of the security of encryption algorithm needs standards to measure and test, and related work in the whole biometric key extraction research is less, which needs further study.

Table 1. Comparing Keys of Several Biometric Features

Biometric features	Key length
Keystroke[3]	12
Voice [4]	46
Fingerprint[5]	69
Iris[7]	140

2. Iris Recognition System Principle

Iris feature recognition goes through capturing a template, and then the feature extraction algorithm conserves template sample into biological templates. The template should provide standardized, effective and highly differential feature performance, which can be objectively compared with other templates to identify [9].

Iris recognition system includes authentication and identification model. Iris recognition system is mainly composed of four parts: iris image acquisition, iris location, iris feature extraction, match and recognition. The system principle is shown in Fig.1.

(1) Image acquisition. The professional high resolution acquisition equipment is used to gain the eye image containing the iris;
(2) Iris image preprocessing. It includes the iris inner, outer edge location, normalization steps. Effective area of iris is extracted, and the interference is removed;
(3) Feature extraction. The extraction algorithm is employed to converse the iris texture into feature code to be differentiated easily;
(4) Feature matching. The iris feature codes extracted and stored beforehand feature templates are compared to identity or distinguish.

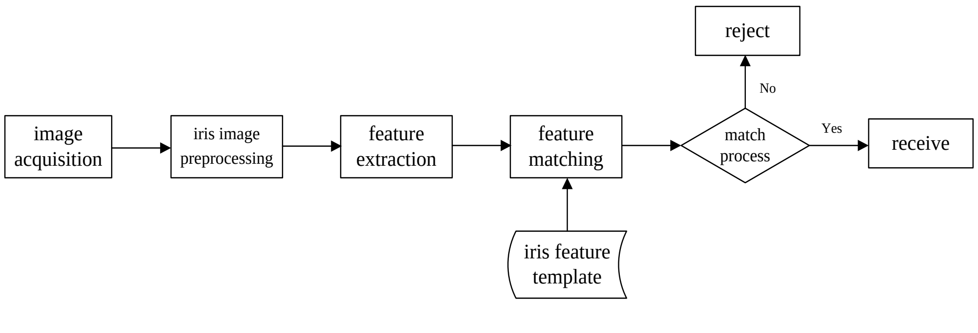


Figure 1. Iris Recognition System Principle

3. Wavelet Packet Decomposition Principle

Wavelet packet is regarded as an expansion of the wavelet concept. Wavelet packet analysis is a kind of the more sophisticated analysis method based on t in the multi-resolution analysis [10], which divides the band into multi levels, and according to the signal characteristics analyzed, adaptively selects the corresponding frequency band, which can be matched with the signal spectrum, to improve the time-frequency resolution ratio.

f (t) is set a decomposed signal; $p_j^t(t)$ presents a wavelet packet of j layer, called wavelet packet coefficient; G and H are the filters of the wavelet packet decomposition; H is associated with the scaling function and G is related to wavelet function. Therefore, Wavelet transform is defined as:

$$p_0^1(t) = f(t) \tag{1}$$

$$p_j^{2i-1} = \sum_k H(k-2t)p_{j-1}^i(t) \tag{2}$$

$$p_j^{2i} = \sum_k G(k-2t)p_{j-1}^i(t) \tag{3}$$

Among them, t=1,2,...,2^{J-i};i=1,2,…,2^j;J=$\log_2 N$ and N is the number of the layer of wavelet packet decomposition.

Figure 2 and Figure 3 are tree structures of the wavelet decomposition and wavelet packet decomposition. The left nodes of each sub-tree represent the low frequency part of signals, while the right nodes represent the high frequency part. (i, j) represent the j node of the i layer (scale degrees). The node of (0,0) represents the original signal; (1,0) represents low frequency coefficients W_{10} of the decomposition first layer. (1,1) represents decomposition high-frequency coefficient W_{11} of the first layer nodes, and the others are presented by analogy. It is seen that wavelet decomposition doesn't decompose the high frequency part of the signal, and therefore the frequency resolution of the high frequency part cannot be improved. Based on the wavelet analysis, the wavelet packet analysis further decomposes the high frequency part of wavelet multi-resolution decomposition analysis stopping the decomposition, which makes the decomposition have the same high frequency resolution in the frequency domain.

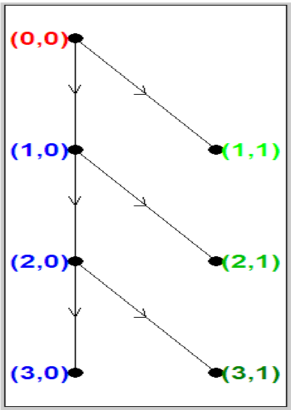


Figure 2. Tree Structure of the Wavelet Decomposition

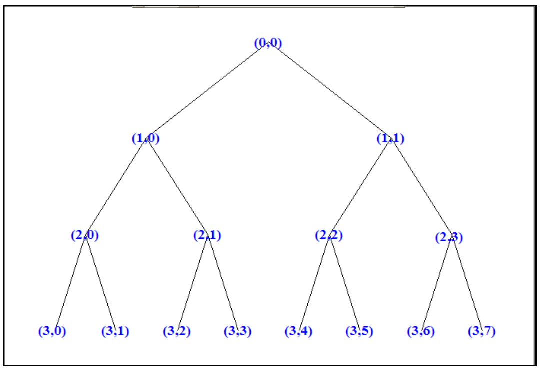
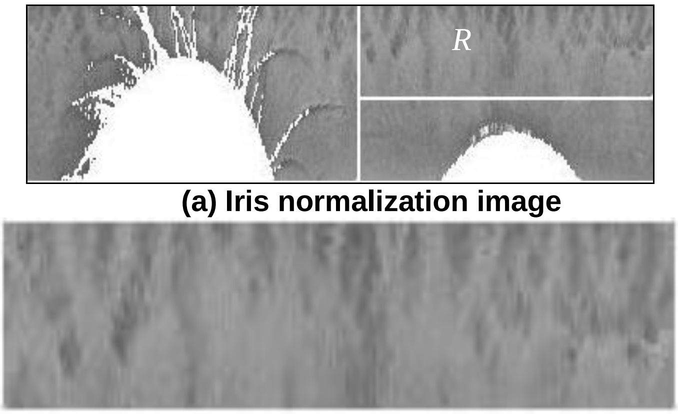


Figure 3. Tree Structure of the Wavelet Packet Decomposition

4. Iris Feature Extraction based on Wavelet Packet Decomposition

4.1 Iris Feature Extraction

Before iris feature extraction, the normalized operation is done [11]. Iris textures without interference are selected and area *R* is extracted with size of 40×200, which is shown in Figure 4.



(a) Iris normalization image

(b) Region of feature extraction

Figure 4. Extraction Region of the Iris Texture Feature

The *Haar* wavelet packet is used to decompose region *R* into three layers, and decomposed sub-band image is shown in Figure 5. Information distribution of each node frequency is shown in Figure 6. After the decomposition of the second layer, each sub-images are further decomposed, which represent low-frequency information of *LL* image, horizontal high frequency information of *HL*, vertical high frequency information of *LH*, diagonal high frequency information of *HH*, and the subscript represents node's layer number. In Fig.5, the big red box represents the frequency domain information of each direction of the second layer. Each node of the second layer is decomposed by the wavelet packet to generate the frequency domain information of each direction of a third layer. The small red box shows the *HH* diagonal high frequency information of the third layer. After the wavelet packet decomposition, the size of each sub node of the third layer is 125. Because the iris texture information is detailed information of grayscale change, it mainly focuses on the high frequency coefficients. To represent the iris texture information of each direction better and eliminate noise effects, *HH* diagonal high-frequency information of the third layer is selected as the feature object extracted, with a total of 16 nodes, and 2000 size.

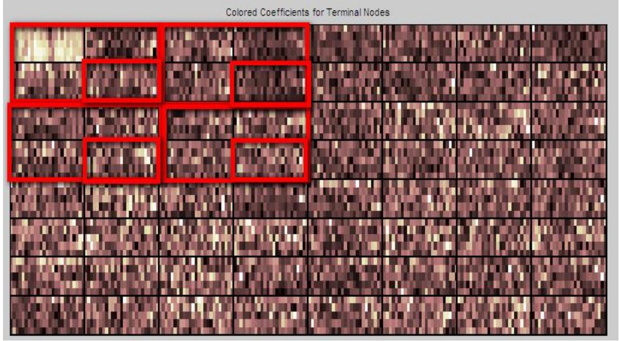


Figure 5. Three Layers Results of Iris Texture Region Decomposition

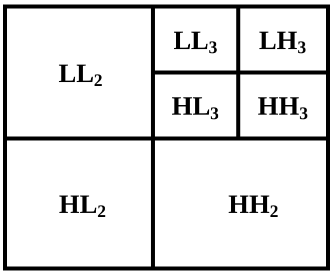


Figure 6. The Frequency Distribution of Each Sub-band Image

The wavelet coefficients represent similarity degree between the wavelet and signal [12], and the similarity between positive wavelet coefficients and negative wavelet coefficients has a larger disparity, and therefore the *T* threshold is used for the binary code. It elements of the feature space C(*i*) obey encoding rules as follows:

$$\begin{cases} C(i)=0 \quad if \;\; C(i)<T \\ C(i)=1 \quad if \;\; C(i)\geq T \end{cases} \tag{4}$$

Here, Donoho sets *T*[13]:

$$T=\sigma\sqrt{2\ln N} \tag{5}$$

σ is the variance of noise and *N* is signal length. In the real application, the intensity of noise is unknown, which can be estimated through using the diagonal high frequency wavelet coefficient:

$$\sigma=(\text{median}\left|HH_1,\cdots,HH_j,\cdots,HH_n\right|)/0.6475, 1\leq j\leq 16 \tag{6}$$

HH_j is the node of diagonal high frequency coefficients of the third layer, *N*=2000. Through the above code, 2000bits iris feature codes are formed.

4.2 Key Generation

AES encryption key is generally random binary sequence of 128bit, and therefore it is necessary to extract from a 2000 bit 0-1 sequence to meet the demand of encryption. C(*i*) is the extracted code, while the key extracted is *K*(*j*), 1≤*j*≤128. The real key is selected from 2000 randomly selecting 128 integers *z*.

$$\begin{cases} K(j)=C(z),\ 1\leq j\leq 128 \\ z=\text{random}(2000) \end{cases} \tag{7}$$

Random represents random number generating function, and a 128 bit binary sequence is generated.

5. Results and Analysis of Experiment

The security of the encryption algorithm depends on the security of the key, and the randomness of the key is considered as the most important security standard. The key generation algorithm proposed in this study is adopted to extract 600 keys and security was tested by NIST key safety standards [14]. Key length is 128 bits. In the NIST test, FT, FBT, RT, ST, AET, LROBT and CST 7 seeds are selected to test, and each seed test gets a evaluation value *P_Value* which is used to evaluate the randomness of key in certain aspects. If *P_value* > α (α is significance level set), the key sequence selected passes the test. If the last test pass rate is more than *T*, and then 0-1 sequence key is randomly generated. *T* is generated based on the following formula

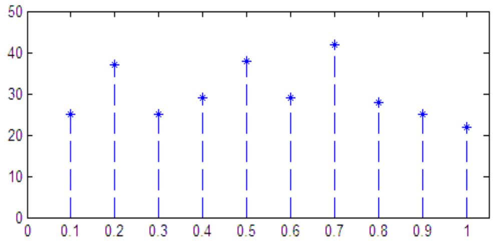
$$T = \hat{p} - 3\sqrt{\frac{\hat{p}(1-\hat{p})}{m}} \times 100\% \tag{8}$$

m is the number of test samples, where *m* is 10^2, $\alpha=1/10^2=0.01$, $\hat{p}=1-\alpha=0.99$. The known value is put into the above formula. The sub-test pass rate is shown in Table 2, which is more than *T*, and the extracted key goes through NIST standard test.

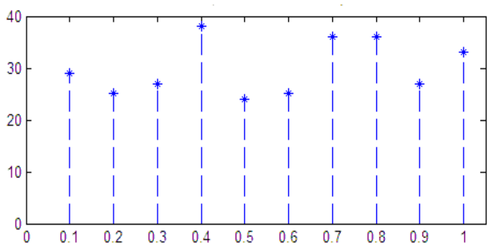
Table 2. Pass Rate by Seven Kinds of Randomness Test

Test	Pass rate /%
FT	98.45
FBT	99.11
RT	99.22
AET	99.13
LROBT	97.43
ST	98.27
CST*	98.67

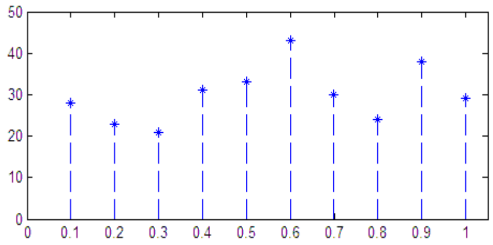
P_value distribution of seven tests is shown in Figure7.



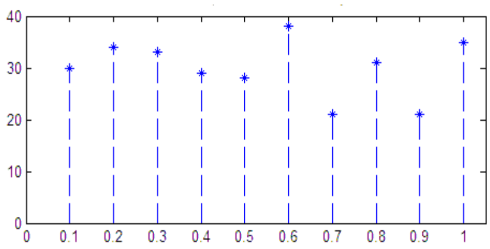
FT



FBT



RT



AET

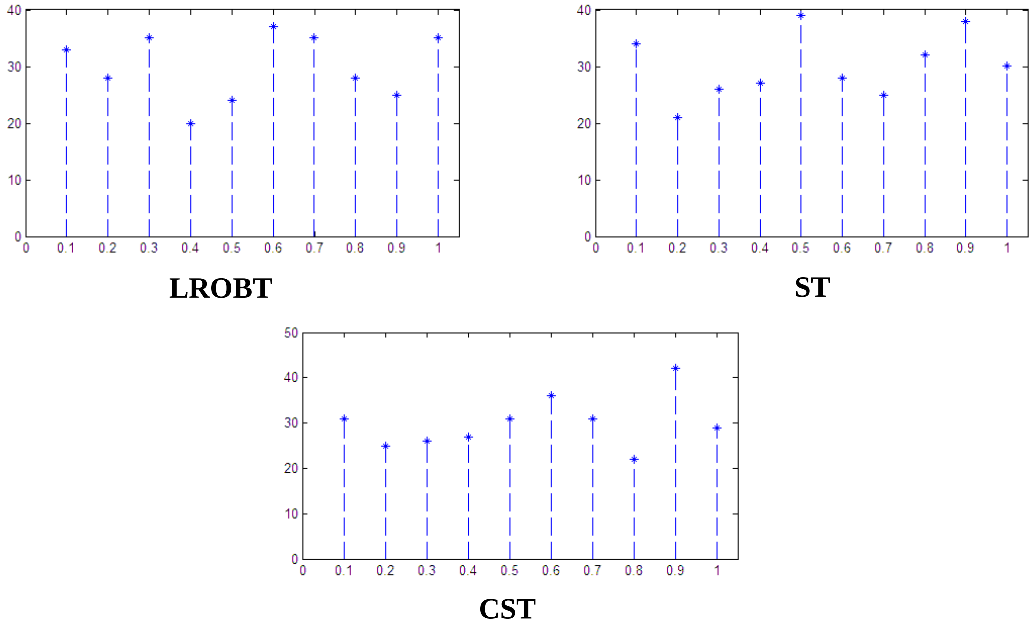


Figure 7. P_value Distribution of Seven Tests

6. Conclusion

An approach is proposed based on wavelet packet decomposition to extract key of iris feature. The wavelet packet is used to decompose effective iris region into the three layers. The third diagonal high frequency coefficient is extracted as the iris feature, and the encryption key was randomly generated from iris feature to meet the encryption demand. The United States NIST encryption standard is adopted to test the extracted key, and the experimental results show that pass through the NIST test, the generated key meets the security of encryption algorithm needs.

References

[1] Q. Tian, “Review of biological feature recognition”, Application research of computers, vol. 26, no. 12, (**2009**), pp. 4401-4410.

[2] X. Y. Yuan, “A study of iris localization, deformation and feature extraction”, PhD thesis, (**2010**), Shanghai Jiao Tong University.

[3] F. Monrose, M. K. Reiter and S. Wetzel, “Password hardening based on keystroke dynamics”, Proceedings of the 6th ACM Conference on Computer and Communications Security, (**1999**),Singapore, pp. 73-82.

[4] F. Monrose, M. K. Reiter and Q. Li, “Cryptographic key generation from voice”, IEEE Symposium on Security and Privacy, (**2001**), Oakland, CA, USA, pp. 202- 213.

[5] T. C. Clancy, N. Kiyavash and D. J. Lin, “Secure Smart Card-Based Fingerprint Authentication”, Proceedings of 2003 ACM SIGMM Workshop Biometrics Methods and Application, (**2003**), New York, USA, pp. 45-52.

[6] G. Davida, Y. Frankel and B. Matt, “On enabling secure applications through offline biometric Identification, Proceedings of IEEE Symposium on Security and Privacy, (1998), Oakland, California, pp. 148–157.

[7] F. Hao, R. Anderson and J. Daugman, “Combining Cryptography with Biometrics Effectively”, IEEE Transactions on Computers, vol. 55, no. 9, (**2006**), pp. 1081–1088.

[8] X. Boyen, “Reusable Cryptographic Fuzzy Extractors”, proceedings of the 11th ACM Conference on Computer and Communications Security, (**2004**),Washington DC, pp. 82-91.

[9] Y. H. Wang, Y. Zhu and T. N. Tan, “Identification based on iris recognition”, Journal of automation chemistry, vol. 28, no. 1, (**2010**), pp. 3-4.

[10] J. G. Yang, “Wavelet analysis and its engineering application”, Mechanical Industry Press, Beijing, **(2009).**

[11] J. Daugman, “The importance of being random: Statistical principles of iris recognition”, Pattern Recognition, vol. 36, (**2009**), pp. 279–291.

[12] C. Y. Lin, “A research on iris recognition algorithm based on Haar wavelet”, PhD thesis, (**2009**), Electronic Technology University.

[13] D. L. Donoho, “De-noising by soft-thresholding”, IEEE Transactions on Information Theory, vol. 42, no. 3, (**2009**), pp. 613 – 627.

[14] US National Institute of Standards and Technology, NIST Special Publication 800-22 [EB/OL],[2008-04-14], http：//csrc.nist.gov/rng/rng2.html.

Authors

Kun Yu, received his Master Degree from Wuhan University of Technology in 2010. He is now a lecturer of Henan Mechanical and Electrical Engineering College in China.

His research interests include software engineering, Network technology and information processing.

He has published more than 6 papers in journals and conferences.

Jun Wei, she received her Master Degree from Wuhan University of Technology i n 2010.

She is now a lecturer of Henan Mechaninal and Electrical Enginee ring College in China.her research interests include Information techn ology and database applications.

He has published more than 10 papers on in journals and conferen ces.