Proceedings of the 2012 International Conference on Machine Learning and Cybernetics, Xian, 15-17 July, 2012

A TWO-STEP HYBRID APPROACH FOR VOICEPRINT-BIOMETRIC TEMPLATE PROTECTION

HUA-HONG ZHU[1,2], QIAN-HUA HE[1], YAN-XIONG LI[1]

[1] School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China
[2] Data Communication Department, Guangdong Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China
E-MAIL:zhuhuah@gsta.com

Abstract:

Biometric template protection is a crucial issue to be addressed for widespread deployment of biometrics-based recognition systems in real life application. Although a number of biometric template protection methods have been reported, it is still a challenging task to devise a scheme to satisfy both security and performance. In this paper, a two-step hybrid approach is proposed to generate a cancelable voiceprint template utilizing the advantages of both Template Transformation and Biometric Cryptosystem. The original voiceprint is transformed with random matrix based on the similarity-preserving of random projection. Chaff points are added to the codebook and matching is also performed in the transformed domain. Random projection improves the cancelability while chaff points conceal the genuine codeword to enhance the security. Binary indexes help identify the genuine codeword accurately. The effectiveness of the proposed method is well supported by detailed analysis. The experimental results demonstrate that the recognition performance is well-kept as the original template does.

Keywords:

Voiceprint; Template protection; Random projection; Fuzzy vault

1. Introduction

With the recent advances of speech recognition and speaker recognition, voiceprint-biometric system has raised an important research concern. Applications such as Siri on iPhone 4S become popular. At the same time, "voice cloud" service [1] has been brought about with the development of cloud computing in the mobile Internet. Voice technology will become one of the most important human-computer interfaces for mobile and Internet devices. However, voiceprint-biometric systems also face great challenges in the field of data security and privacy protection under the open environment. It's permanent if voiceprint is lost, and voiceprint always connects to the identity of one user. Thus, the security of voiceprint templates is an important and valuable issue.

Many biometric template protection methods and algorithms have been proposed. Typical solutions based software can be classified two main categories [2]: Template Transformation and Biometric Cryptosystem. Template transformation applies passwords or functions to the input and query biometrics. By changing the key, it's easy to change the template. However, the security mainly relies on the protection of the key. And it's difficult to find suitable one-way functions to obtain non invertible template. Biometric Cryptosystems make use of biometric features to secure a cryptographic key or directly generate a cryptographic key. There are three main modes: key release, key binding and key generation. Generally, the biometric features need to be quantified into the binary bit strings using some sort of user-specific information known as helper data. It mostly relies on error correction coding theory. In order to improve the performance, extracting invariant and discriminative features is necessary. However, it is difficult due to intrinsic intra-class variation of biometric features.

Since the two solutions have their own advantages and limitations, we present a two-step hybrid approach to enhance the security strength. In the enrollment phase, the original voiceprint is projected into another space with random matrix which preserves the Euclidean distances between vectors before and after the transformation. Then codebook is generated which can be thought as a template for voiceprint-biometric system which provides a cancelable parametric representation for cryptosystem. We add chaff vectors to genuine vectors and store them as a whole codebook to database. In order to distinguish the genuine vectors, binary indexes are generated randomly for every vector. In the verification phase, the query voiceprint is also transformed by the same random matrix. Matching is conducted by measuring the Euclidian distances between the query voiceprint and encrypted codebook without disclosing the genuine vectors. At last, the decision

978-1-4673-1487-9/12/$31.00 ©2012 IEEE

calculates the average distortion measurement by separating the genuine vectors and makes decision.

The remainder of this paper is organized as follows. In Section 2, we review the related works. Section 3 introduces the proposed approach and the new scheme. Security analysis and experimental results are presented in Section 4. Finally, the conclusion is drawn in Section 5.

2. Related Works

The design of a privacy-preserving template protection method critically depends on the characteristics of the biometric data and features. All kinds of tentative solutions have been proposed in the literature using various biometrics. Main software solutions are Template Transformation and Biometric Cryptosystem.

Template Transformation uses parameters to transform the features extracted to the template. The template is stored and matching is performed in the transformed domain. Biohashing scheme [3] is a popular method that combines orthogonal random vectors with user specific biometric data. A token is a sequence of numbers or symbols that are used as a seed to generate random vectors. Each user can have its own private token or every person can use the public token. Jin et al. [4] proposed a two factor authenticator based on iterated inner products between token pseudo-random number and the user specific fingerprint feature, which generated from the integrated wavelet and Fourier–Mellin transform. The performance is as good as that before transformation by elimination of false accept rates (FAR) without suffering from increased occurrence of false reject rates (FRR). BioHashing technique was presented in [5] using the random projection (RP) theory. The Johnson–Lindenstrauss (JL) [6] result stated that Euclidean distances are retained well in RP. Ling et al. [7] applied appropriate dimensionality reduction first and then running RP on the reduced space for face recognition. The methods used in this paper are based on orthogonal transforms including principal component analysis (PCA), fisher linear discriminant (FLD), wavelet transform (WT), wavelet transform with PCA (WT_PCA), and wavelet transform with fourier–mellin transform (WT_FMT). Teoh and Yuang [8] have proposed a multispace RP (MRP) method, which applies user-specific RP on dimensionality-reduced feature vectors without the quantization procedure of BioHashing. Wang and Plataniotis [9] presented a systematic analysis of a random-projection based method for addressing cancelability and privacy protection. RP on both high-dimensional image vectors and dimensionality-reduced feature vectors was discussed and compared by detailed theoretical analysis. Extensive experimentations on a face-based biometric verification problem showed the effectiveness of the method.

Biometric Cryptosystem integrates the advantages of both biometrics and cryptography to enhance the overall security and privacy of a biometric system. Juels and Wattenberg [10] proposed a fuzzy commitment scheme based an error-correction which binds binary representation of biometric features with randomly generated keys by an XOR operation. Hao et al. [11] designed an iris-based encryption method using a two-level error correction mechanism base on fuzzy commitment scheme and got the good performance with FAR=0 and FRR=0.47%. Juels and Sudan [12] proposed a fuzzy vault scheme which works with unordered set of features, such as the minutia points in fingerprints. The security of the fuzzy vault scheme is based on the infeasibility of the polynomial reconstruction problem. Nyang and Lee [13] incorporated well-known face authentication schemes such as PCA and LDA into the fuzzy vault scheme for cancelable face templates. Most of face authentication algorithms are dependent upon weighted features, which are incompatible with the original fuzzy vault scheme. To reflect the level of importance of individual features from feature set, they introduce another layer between captured feature set and points in the polynomial to be interpolated. Liu et al. [14] pointed out the fuzzy vault scheme is not applicable to biometric feature data, because the scheme is designed based on set differences metric while in biometric systems Euclidean distance metric is often used to measure the similarity between two feature vectors. They proposed a multidimensional fuzzy vault scheme, in which every point is a vector, and data matching using Euclidian distance measurement could be performed in fuzzy vault space. Dodis et al. [15] presented secure sketch and fuzzy extractor, trying to generation of cryptographic keys from noisy biometric data using error correction code and hash functions. Different constructions for three metric spaces, namely, hamming distance, set difference, and edit distance, are introduced. A detailed review of different biometric cryptosystems has been presented in [2].

3. Proposed Hybrid Approach

Fuzzy vault scheme is a good idea for features represented as an unordered set like minutiae points because it utilizes the advantages of both cryptography and biometrics. Since voiceprint can be looked as an unordered set, Xu et al. [16] presented fuzzy vault for voiceprint template protection. In their paper, chaff points are added directly to MFCCs (Mel Frequency Cepstral Coefficients) which generally have hundreds of vectors. However, the

number of chaff points influences the security of the vault. The more number of chaff points the more noise to conceal genuine points from attack. So it is necessary that the number of chaff points is far

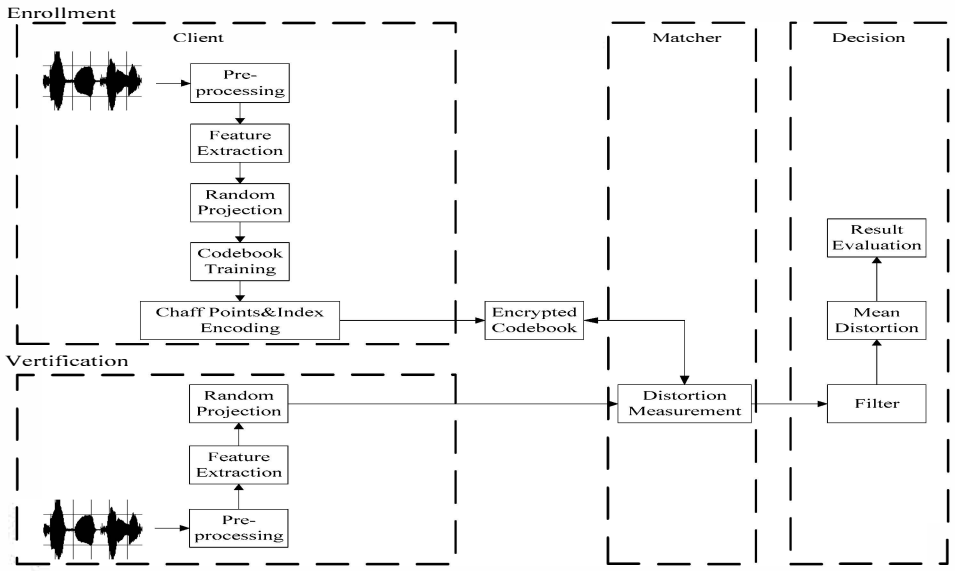

Figure 1. Block diagram of the proposed two-step hybrid approach

larger than the number of genuine points. Chaff points added directly to MFCCs results in an easy way to distinguish the genuine points from the chaff points [17]. What's more, the plain template is stored in database and matching is performed in the original space which is vulnerable to attacks. And it is not easily changed if the template is compromised. So we present a two-step hybrid approach to achieve better cancelability and security. The scheme is as shown in Figure 1.

The proposed method combines RP with fuzzy vault of voiceprint vectors to protect voiceprint template. In order to focus on voiceprint template protection, we assume that components of servers are malicious but they do not collude. Additionally, the client is always honest. In the enrollment phase, the original MFCCs are projected into another space with random orthogonal matrix which preserves the Euclidean distances between vectors before and after the transformation. In this way, we can provide revocability and diversity. When the biometric template in one application is compromised, a new one can be reissued by simply using a new matrix. But this step is not one-way function, so we need further protection by adding chaff points. Difference with [16], chaff points are not added directly to MFCCs but to the codebook trained in the transformed domain by Linde-Buzo-Gray (LBG) so that the number of genuine points is small. The encrypted codebook is stored as both voiceprint template and reference model for pattern matching in the verification phase. In the verification phase, the query voiceprint is also transformed by the same random matrix. The matcher calculates the Euclidian distances between every query vector and encrypted codebook. In paper [14],

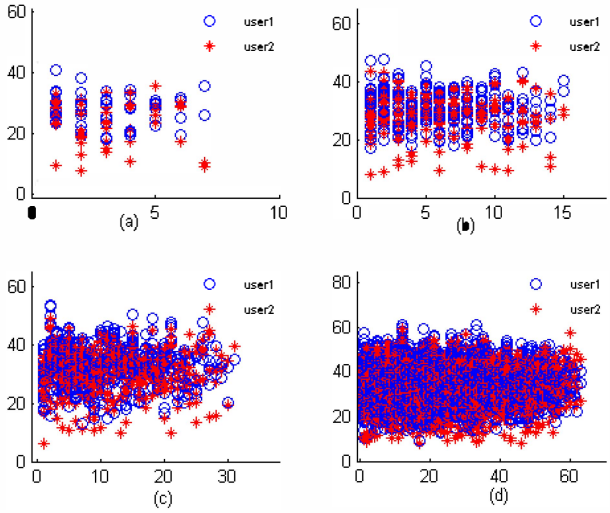

Figure 2. Distances between codewords in two different codebooks: (a) distances with the codebook size of 8; (b) distances with the codebook size of 16; (c) distances with the codebook size of 32; (d) distances with the codebook size of 64.

Euclidian distance is used to identify genuine points. For

given query points, the point in vault which has the minimum distance is considered as the genuine point. It seems suitable for voiceprint recognition since Euclidean distance metric is also used to measure the similarity between two feature vectors by VQ (Vector Quantization) technology. So we can generate chaff points whose distance is larger than the threshold. But as shown in Figure 2, Some Euclidian distances between codewords in two different codebooks are smaller than those in the same codebook so that FAR will increase if the attacker uses the another user's features may find the genuine points by measuring the minimum Euclidian distance. In order to keep the performance of system, we add binary indexes to bind genuine vectors which are also stored with codebook in database. For example, given binary bit strings I_1, I_2 as indexes of the genuine points, we can calculate the key by OR operator and filter I_1, I_2 by AND operator. Decision filters the genuine points by the key which is shared by Client and calculates the average distortion measurement of genuine vectors to make decision. Binary indexes are more efficient than prime number in [16] and easier to be protected by encryption algorithm. The detail procedure is described as follows.

Enrollment:

1) Extract *d*-dimensional voiceprint feature vectors $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_T)$ from the input voice.
2) Transform the vector $\boldsymbol{x}_t \in \boldsymbol{X}$ $(t = (1,2,...,T))$ by an orthogonal random matrix $\boldsymbol{R} \in \Re^{d*n}$ $(d = n)$, the result is a *n*-dimensional vector $\boldsymbol{y}_t$:
$$\boldsymbol{y}_t = \boldsymbol{R}^{\mathrm{T}} \boldsymbol{x}_t \quad (1)$$
3) Train the transformed features to get the codebook $\boldsymbol{C} = (\boldsymbol{c}_1, \boldsymbol{c}_2, ..., \boldsymbol{c}_m)$, $\boldsymbol{c}_i$ is a *n*-dimensional vector.
4) Generate an encrypted template $\boldsymbol{C}' = (\boldsymbol{c}_1', \boldsymbol{c}_2', ..., \boldsymbol{c}_r')$ by adding chaff points randomly to the codebook $\boldsymbol{C}$.
5) Bind the binary index I_i for every vector $\boldsymbol{c}_i'$ $(i = (1,2,...,r))$ and calculate the accumulator of genuine indexes I_i by OR operator as a key k.
6) Store the encrypted template with binary indexes in database and send the key k to decision.

Verification:

1) Extract *d*-dimensional voiceprint feature vectors $\boldsymbol{X}' = (\boldsymbol{x}_1', \boldsymbol{x}_2', ..., \boldsymbol{x}_{T'}')$ from the query voice.
2) Transform the vector $\boldsymbol{x}_t' \in \boldsymbol{X}'$ $(t = (1,2,...,T'))$ by the same matrix $\boldsymbol{R} \in \Re^{d*n}$ $(d = n)$, the result is a *n*-dimensional vector $\boldsymbol{y}_t'$:
$$\boldsymbol{y}_t' = \boldsymbol{R}^{\mathrm{T}} \boldsymbol{x}_t' \quad (2)$$
3) Match by computing the Euclidean distance between $\boldsymbol{y}_t'$ and codeword $\boldsymbol{c}_i'$, defined as
$$d(\boldsymbol{y}_t', \boldsymbol{c}_i') = \left\| \boldsymbol{y}_t' - \boldsymbol{c}_i' \right\| \quad (3)$$
4) Compare the binary indexes I_i with the key k stored in decision to filter the genuine vectors $\boldsymbol{c}_i$ and calculate the average distortion measurement:
$$\mathrm{T} = \frac{1}{T'} \sum_{t=1}^{T'} \min_{1 \le i \le m} (d(\boldsymbol{y}_t', \boldsymbol{c}_i)) \quad (4)$$
5) Evaluate whether the authentication is successful according to the threshold.

Of course, we can also use the traditional fuzzy fault scheme by constructing a polynomial to encrypt and decrypt the vault.

4. Experiment and Analysis

4.1. Cancelability analysis

To satisfy the cancelability, different applications need different templates to ensure security. Thus the voiceprint templates of an individual in different applications will be different. In our approach, the voiceprint template can be changed by simply varying the random matrix if template is compromised. To ensure strong cancelability, chaff points can be different for the codebook in different applications from the same user. According formula (5), random orthogonal matrix can preserve the similarity measure in the transformed domain. If genuine points can be identified, the performance of system can be kept well. In this way, we can get two-factor changeability while accuracy of recognition can be assured. The experimental results also support above analysis.

$$\begin{aligned} \left\| \boldsymbol{y}_i - \boldsymbol{y}_j \right\|^2 &= \left(\boldsymbol{R}^T \boldsymbol{x}_i - \boldsymbol{R}^T \boldsymbol{x}_j \right)^T \left(\boldsymbol{R}^T \boldsymbol{x}_i - \boldsymbol{R}^T \boldsymbol{x}_j \right) \\ &= \left(\boldsymbol{x}_i - \boldsymbol{x}_j \right)^T \boldsymbol{R} \boldsymbol{R}^T \left(\boldsymbol{x}_i - \boldsymbol{x}_j \right) \\ &= \left(\boldsymbol{x}_i - \boldsymbol{x}_j \right)^T \left(\boldsymbol{x}_i - \boldsymbol{x}_j \right) \\ &= \left\| \boldsymbol{x}_i - \boldsymbol{x}_j \right\|^2 \end{aligned} \quad (5)$$

4.2. Security analysis

Traditional fuzzy vault is vulnerable to cross-matching attack. The adversary can filter genuine points by comparing the different vaults of the same user if he can get them from different applications, as shown in Figure 3. We consider the scenario where there are two applications using the same voiceprint as template. Suppose the adversary gets all protected templates in both database and tries to find the

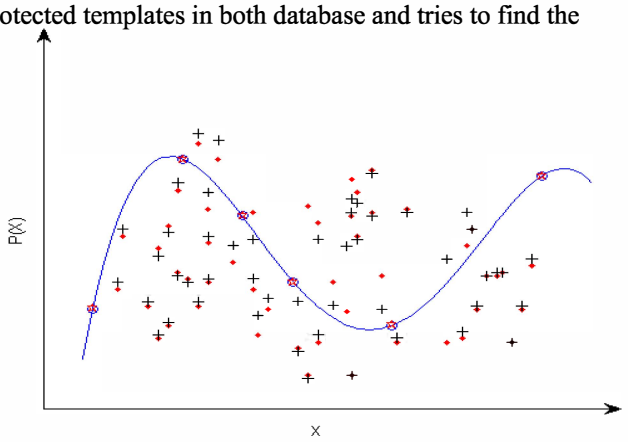

Figure 3. Cross-matching in traditional fuzzy vault

genuine codeword. If the number of chaff points is 300 and the number of genuine points is 32 in each application, the total number of distance comparisons is 332*332=110224. In this way, the adversary can decide the genuine points belong to the same user. However, the hybrid method overcomes the limitation by mapping the original voiceprint to different space and randomly adding chaff points. Since the generation of chaff points does not rely on the requirement of the threshold and it is not possible to estimate the original features using only the chaff points.

On the other hand, RP is vulnerable to masquerade attack. Assuming the worst case that the matrix is compromised, the adversary can estimate the original voiceprint. However, chaff points can conceal the genuine points and increase the complexity of decryption. Suppose the adversary attempts a brute-force attack. In above scenario, if the binary indexes are secret, in order to find the genuine points, the total number of possible combinations is $\binom{332}{32} \approx 3.8*10^{44}$. This corresponds to a computational time of 10^{36} years based on our current implementation. If constructing a polynomial is used to decrypt the fault and the degree of the encoding polynomial which can secure a key of size 128 bit is 8. Among these combinations, $\binom{32}{9} \approx 2.8*10^{7}$ combinations will successfully decode the secret. The probability that a combination of points decodes the secret is $((2.8*10^{7})/(3.8*10^{44})) \approx 7.3*10^{-38}$ and the expected number of combinations that need to be evaluated is $1.4*10^{37}$. This corresponds to a computational time of 10^{29} years based on our current implementation.

4.3. Experimental results

We use Mandarin continuous speech recognition training database as experimental data. We select 1600 sentences from 40 speakers in which 400 sentences are for training and other 1200 sentences for testing. Each utterance length is between 4s to 10s after removing the mute. All data are 16kHz, 16bit and mono channel WAV formats. The 24-order MFCCs are used as the feature parameters. The frame length is 32ms, and the frame shift is 16ms. We choose the size of codebook with 32 and 64. The results in table I show that FRR and FAR after encryption are the same as that of the original features. Generally , the performance is better with the increasing of codebook size. Although FRR of our method is similar to that of multidimensional fuzzy vault in [14], FAR of our method is far better than that of multidimensional fuzzy vault in [14]. This has also been proved by above analysis. So we can safely come to the conclusion that the performance after our encryption is unchanged which is related to the accuracy of VQ.

TABLE 1. EXPERIMENTAL RESULTS

Method	Performance Comparison		
	Codebook size	*FRR*	*FAR*
Original feature	32	5.12%	0.08%
Our Method	32	5.12%	0.08%
Multidimensional Fuzzy vault [14]	32	5.28%	26.43%
Original feature	64	4.37%	0.07%
Our Method	64	4.37%	0.07%
Multidimensional Fuzzy vault [14]	64	4.51%	28.46%

5. Conclusions

In this paper, we present a two-step hybrid approach for addressing the challenging problem of template cancelability and privacy protection in voiceprint verification systems. The proposed method is based on random projection in conjunction with the improved fuzzy vault scheme. This method overcomes the limitation of random projection and the fuzzy vault scheme. Security and cancelability analysis has been given. Random projection improves the

cancelability while chaff points conceal the genuine codewords to enhance the security. Experimental results demonstrate the recognition accuracy is unchanged as the original features do.

Acknowledgements

This paper is supported by the National Natural Science Foundation of China (No. 60972132, No. 61101160), the Natural Science Foundation of Guangdong province, China (No. 9351064101000003, No. 10451064101004651), and the Fundamental Research Funds for the Central Universities, South China University of Technology, China (No. 2011ZM0029).

References

[1] Tao Jiang, "Voice cloud in mobile Internet", http://cloud.csdn.net/a/20110520/298342.html, 2011.

[2] Anil. K. J., Karthik. N., and Abhishek. N., "Biometric template security", EURASIP Journal on Advances in Signal Processing, Vol. 2008, pp. 1-17, Jan. 2008.

[3] Alessandra. L., and Loris. N., "An improved BioHashing for human authentication", Pattern Recognition, Vol.40, pp. 1057–1065, Mar. 2007.

[4] Andrew. B. J. Teoh., David. C. L. Ngo., and Alwyn. G., "Biohashing: Two factor authentication featuring fingerprint data and tokenised random number", Pattern Recognition, Vol.37, pp. 2245–2255, Apr. 2004.

[5] Andrew. T., Beng. J., Tee. C., David. N., and Chek. L., "Remarks on BioHash and its mathematical foundation", Information Processing Letters, Vol. 100, No. 4, pp. 145–150, Nov. 2006.

[6] Ella. B., and Heikki. M., "Random projection in dimensionality reduction: Applications to image and text data", Proceeding of KDD2001 Conference, San Francisco, pp.245–250, August 2001.

[7] David. C. L. Ngo., Andrew. B. J. Teoh., and Alwyn. G., "Biometric Hash: High-Confidence face recognition", IEEE Transactions on Circuits and Systems for Video Technology, Vol.16, No.6, pp. 771–775, Jun. 2006.

[8] Andrew. B. J. Teoh. and Chong. T. Yuang., "Cancelable biometrics realization with multispace random projections", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 37, No. 5, pp. 1096–1106, Oct. 2007.

[9] Yongjin. W., and Konstantinos. N. P., "An analysis of random projection for changeable and privacy preserving biometric verification", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 40, No. 5, pp. 1280–1293, Oct. 2010.

[10] Ari. J. and Martin. W., "A fuzzy commitment scheme", Proceeding of CCS1999 Conference, Singapore, pp. 28–36, November 1999.

[11] Feng. H., Ross. A., and John. D., "Combining crypto with biometric effectively", IEEE Transactions on Computers, Vol. 55, No. 9, pp. 1081–1088, Sep. 2006.

[12] Ari. J. and Madhu. S., "A fuzzy vault scheme", Designs, Codes and Cryptography, Vol. 38, No. 2, pp. 237-257, Feb. 2006.

[13] DaeHun. N. and KyungHee. L., "Fuzzy Face Vault. How to implement fuzzy vault with weighted features", Proceeding of UAHCI2007 Conference, Beijing, pp. 491–496, July 2007.

[14] Hailun. L., Dongmei. S., Ke. X., and Zhengding. Q., "Is Fuzzy Vault Scheme very Effective for Key Binding in Biometric Cryptosystems?", Proceeding of CyberC2011 Conference, Beijing, pp. 279-284, October 2011.

[15] Yevgeniy. D., Leonid. R., and Adam. S., "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data", Proceeding of EUROCRYPT2004 Conference, Switzerland, pp. 523–540, May 2004.

[16] Wenhua. X., and Minying. C., "Cancelable voiceprint template based on chaff-points-mixture method", Proceeding of CIS2008 Conference, Suzhou, Volume II, pp. 263-266, December 2008.

[17] Ee-Chien. C., Ren. S, Francis. W. T., "Finding the original points set hidden among chaff", Proceeding of ASIACCS 2006 Conference, Taipei, pp. 182-188, March 2006.

[18] Karthik. N., Anil. K. J., and Sharath. P., "Fingerprint-Based fuzzy vault: Implementation and performance", IEEE Transactions on Information Forensics and Security, Vol. 2, No. 4, pp. 744–757,Dec. 2007.