Efficient Implementation of Radon Transform and Encryption Techniques for Cancelable Speaker Identification

1st Neven Hassan
Department of Electronics and Electrical Communications Engineering
Faculty of Electronic Engineering
Menoufia University:
Menouf, Egypt
eng.nevenhassan@yahoo.com

3th Adel S. EL-Fishawy
Department of Electronics and Electrical Communications Engineering
Faculty of Electronic Engineering
Menoufia University:
Menouf, Egypt
aelfishawy@hotmail.com

2nd Walid El-Shafai
Department of Electronics and Electrical Communications Engineering
Faculty of Electronic Engineering
Menoufia University:
Menouf, Egypt
Security Engineering Laboratory, Department of Computer Science, Prince Sultan University, Riyadh, Saudi Arabia
eng.waled.elshafai@gmail.com

4th Fathi E. Abd El-Samie
Department of Electronics and Electrical Communications Engineering
Faculty of Electronic Engineering
Menoufia University:
Menouf, Egypt
fathi_sayed@yahoo.com

***Abstract*— This paper introduces three cancelable speaker identification techniques based on the spectrogram estimation of speech signals subjected to either chaotic encryption process, or RSA algorithm in addition to Radon transform to produce cancelable templates instead of the original speech signals. The resulting transformed versions of the voice biometrics are stored in the server instead of the original biometrics. Therefore, the users' privacy can be protected well. It is evident from the obtained results that the proposed techniques are secure, reliable and practical. They have good encryption and ability to generate cancelable templates. These characteristics lead to good performance. The proposed cancelable speaker identification techniques are evaluated under the influence of Additive White Gaussian Noise (AWGN) with different strengths. This makes them more accurate in identifying the users and also more resistant to attack attempts. In addition, security is enhanced through maintaining the confidentiality of the processed data. In the experimental results, evaluation metrics such as Equal Error Rate (EER), False Rejection Rate (FRR), and False Acceptance Rate (FAR) are used to assess the performance of the proposed techniques. In addition, the genuine, impostor distributions, Receiver Operating Characteristic (ROC) curve and area under the ROC curve for the proposed techniques are estimated for better evaluation and comparison.**

Key words — Cancelable biometrics, Chaotic Baker map, Speaker identification, RSA algorithm, Radon transform, Spectrogram.

I. INTRODUCTION

Biometrics have been used for identifying persons' identities. Biometric authentication is now widely used in a lot of applications, such as border control, secure computer systems, secure banking services, mobile phones and credit cards. Hence, with biometrics, personal identification based on who he or she is instead of what he or she has (card - code - key) or what he or she knows (password) will be more secure. In addition, it is more complex to copy individuals' biometrics [1].

The ideal biometric information has some characteristics such as universality, which means that all individuals must be characterized by biometric information. In addition, this information must be as dissimilar as possible for two different individuals، and this indicates uniqueness, and permanency [2].

Speaker identification is the process of identifying a person from his speech signals. This is accomplished through training and testing operations [3] as shown in Fig.1. The training requires feature extraction, and hence speaker modeling through a certain classifier. On the other hand, testing requires a matching operation.

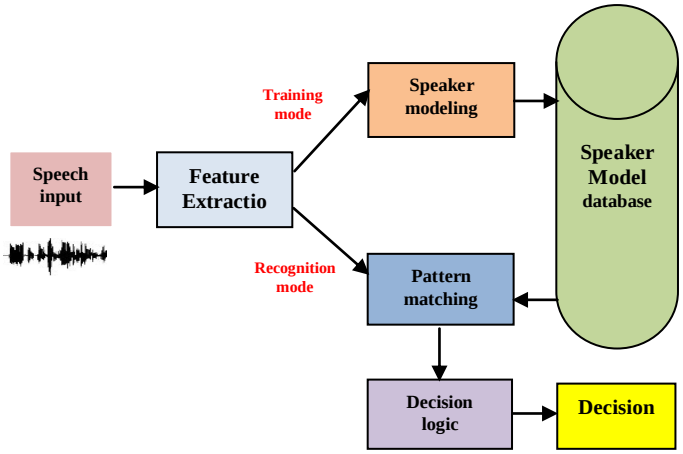


Fig .1 Training and testing modes of the speaker identification system [4].

Biometric systems can work in the verification or identification modes [5]. Unfortunately, these traditional systems are subject to several attacks as shown in Fig. 2. The solution for the problem of attacks is to adopt cancelable biometric systems.

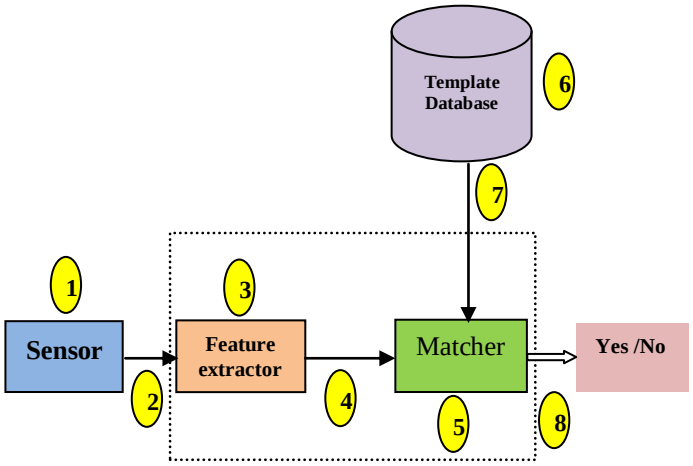


Fig .2 Probable attack points on a general biometric system [6]

Information security and user privacy are very important concerns in biometric-based systems. One of the main solutions to achieve these requirements is the cancelable biometrics as it is the protection mechanism used, even if the biometric system is breached. The original biometric template is intentionally distorted in order to be registered in the authentication system [7]. When a revocable biometric pattern is used, we really save a distorted model in the biometric database.

In traditional biometric systems, most people are reluctant to present their vital features, because they are worried about their secrecy. Cancelable biometrics solve these secrecy issues, because they prevent the system from storing the users' original biometrics. Cancelable biometrics are generated through a set of intentional, systematic, and repeated distortions of biometric signals with the aim of protecting user information [8].

The main objective of cancelable biometric schemes is achieving diversity, where multiple cancelable patterns can be created from the same biometric for several applications. In addition, renewability/revocability means direct cancellation and re-issuance if the pattern is breached. Furthermore, non-invertability is adopted to stop fraud. It has to be difficult to get information about the original biometric features from the transformed forms. Finally, the recognition performance using a converted template should be high [9].

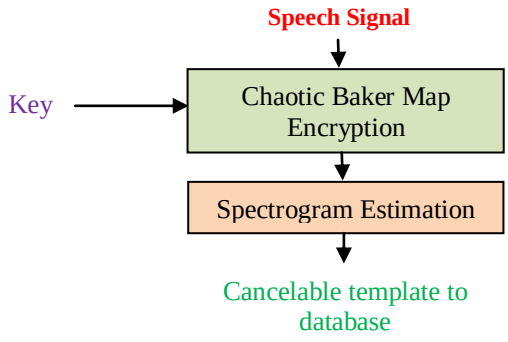
Cancelable biometrics achieve a high standard of privacy by allowing various versions to be related with the same biometric data. In each registration, several transformations can be performed to create the protected templates. This helps in enhancing the ability of generating different user biometrics for different databases. Transforms can be used at two levels to create cancelable templates: signal level and feature level. The transformations are used to make it hard to restore the original templates, when the transformed template is reversed, thus providing the required response against any potential attackers. In addition, the transformed templates have to maintain the discrimination ability between patterns [10-12].

II. PROPOSED CANCELABLE SPEAKER IDENTIFICATION TECHNIQUES

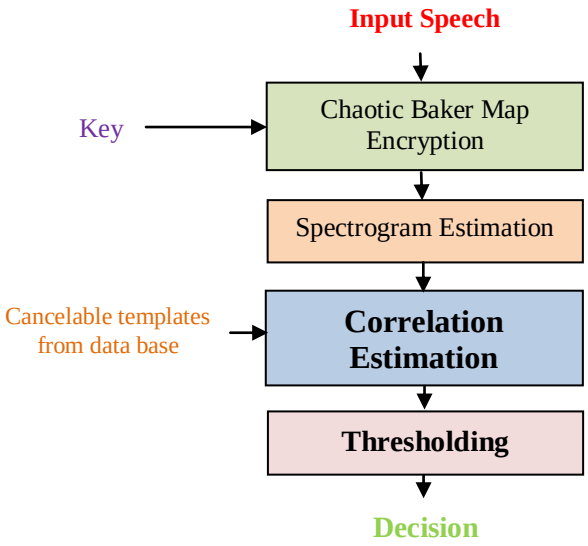
This section presents the three proposed techniques for cancelable speaker identification. All of them depend on spectrogram estimation preceded by encryption. The difference between them is in the encryption algorithm. The first one depends on chaotic Baker map for encryption. The second one adds a Radon transform step after encryption. The third one depends on RSA algorithm for encryption.

A. Proposed Technique Based on Chaotic Baker Map

This technique begins with chaotic Baker map encryption as a randomization tool for speech signals. Its block diagram is shown in Fig.3.



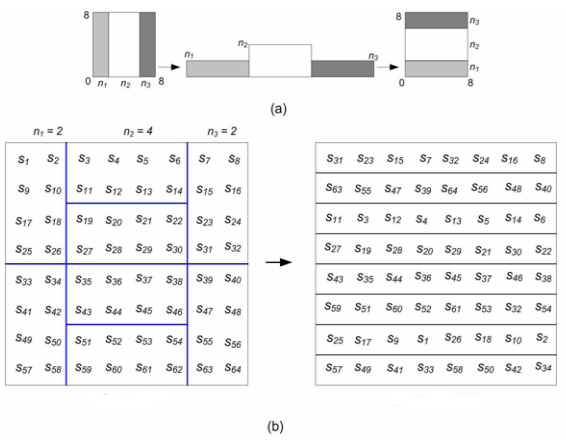
(a) Cancelable template generation



(b) Cancelable template verification

Fig .3 Block diagram of the proposed cancelable biometric technique based on chaotic Baker map.

Chaotic systems, in general, allow permutation of certain data in matrix format [13]. Chaotic Baker map is one of the most popular maps. It allows permutation of samples in a matrix into new positions in the same matrix as illustrated in Fig. 4.

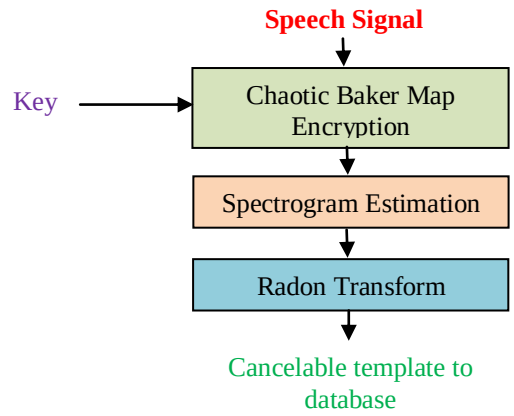


(a) Generalized map. (b) Discretized map

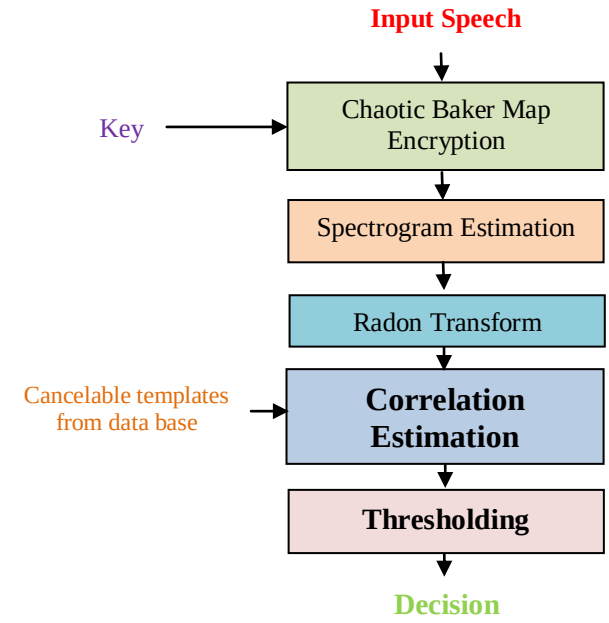
Fig .4 Baker map operation [14].

B. Proposed Technique Based on Radon Transform

This technique depends on using Radon transform after Baker map encryption as shown in the block diagram in Fig.5.



(a) Cancelable template generation



(b) Cancelable template verification

Fig .5 Block diagram of the proposed cancelable biometric technique based on chaotic baker map and Radon transform.

Radon transformation is an integral transformation, which aggregates spectrogram values. Fig.6 illustrates three projections of a matrix, *M*, as an example. Hence, Fig.6 shows three various information sectors of *M* with respect to the three angles. Obviously, the more the projections utilized are, the more the information obtained from the image. Note that this leads to better performance, but it is very time-consuming. To address the aforementioned problem, we first choose eight projections. An overall study has been done on this choice. The second solution for this problem is the application of eight expectations of the shrinking research space. This greatly reduces the cost of calculations [15-17].

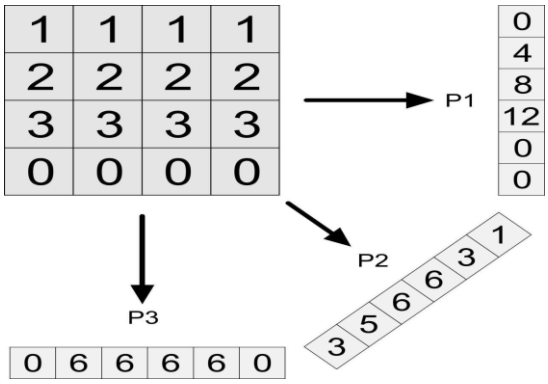
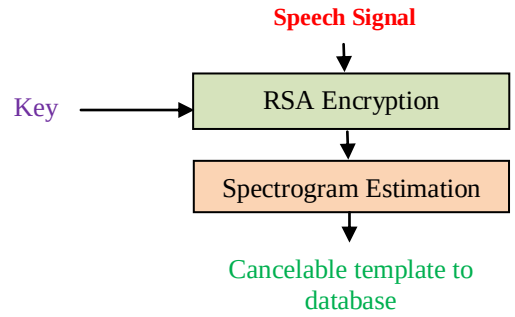


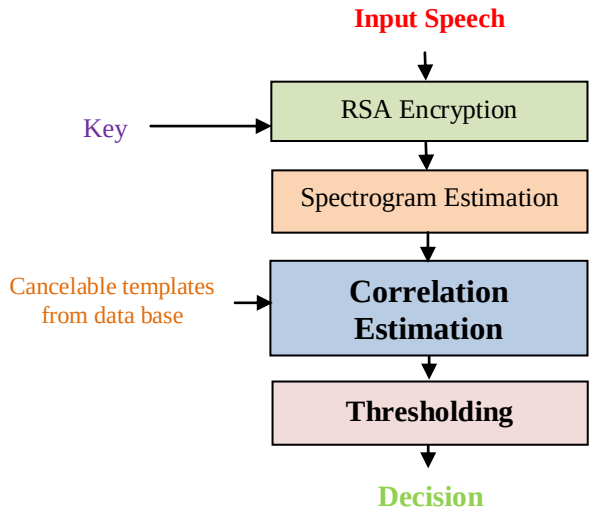
Fig .6 An example of three projections of Radon transform on a matrix [18].

C. *Proposed Technique Based on RSA Algorithm*

This technique uses the RSA algorithm to encrypt the speech signals as shown in the block diagram Fig.7.



(a) Cancelable template generation



(b) Cancelable template authentication

Fig .7 Block diagram of the proposed cancelable biometric technique based on RSA Algorithm.

The Rivest-Shamir-Adleman (RSA) is an algorithm that modern computers use to encrypt and decrypt messages. It is an asymmetrical encryption algorithm. Asymmetrical indicates that there are two different keys. This is called public key cryptography, because one of the keys can be presented to any person. Another key has to be kept secret. The algorithm was constructed based on the fact that finding the factors of a huge complex number is hard. When the factors are primary numbers, the problem is called initial analysis [19].

The RSA includes a public key and a private key. The public key can be popular to everybody. It is utilized to encrypt messages. Messages, which are encrypted by the public key, can only be decrypted by the private key. The private key has to be preserved as a secret. Calculating the private key from the public key is very severe.

The RSA is essentially an authentication system suitable for the Internet. This algorithm was introduced by its inventors in 1977. It is one of the most prevalent asymmetric key cryptosystems included as part of Netscape and Microsoft Web browsers. First, two huge prime numbers are selected and multiplied by this algorithm to generate the public and the private key pair for encryption and decryption operations [20].

The creation of the key in the RSA algorithm depends on the selection of two integers' p and q that are prime in nature in order to determine a modulus n as follows,

$$n = p \square q \tag{1}$$

Then, we estimate the Euler function $\emptyset$ as follows:

$$\emptyset(n) = (p-1)(q-1) \tag{2}$$

We select an integer e as follows:

$$\gcd(\emptyset(n), e) = 1 \tag{3}$$

i.e. $\emptyset(n)$ and e are co-prime, where e represents the public exponent.

As well, we calculate an integer number d by the formula:

$$d = -1 mod \quad \emptyset(n) \tag{4}$$

where d is a private exponent.

Encryption and decryption are implemented with number pairs (n,e) and (n,d) [21-23] as follows:

$$C = M^e \ mod \ n \tag{5}$$

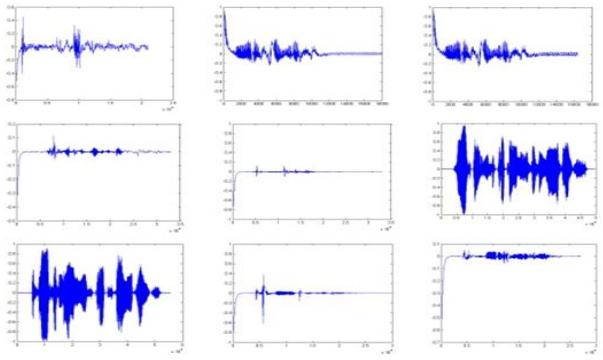
and

$$M = C^d \ mod \ n \tag{6}$$

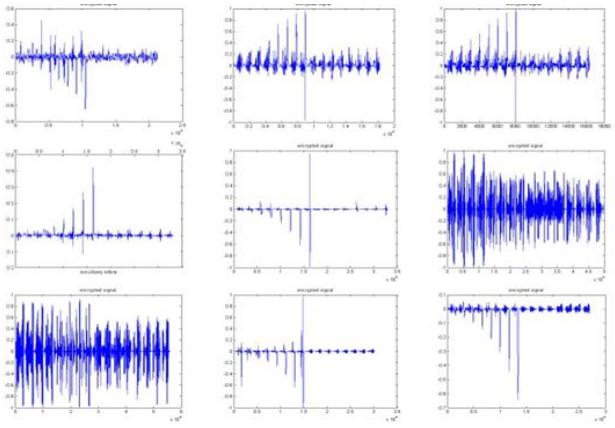
III. RESULTS AND PERFORMANCE ANALYSIS

This section presents the simulation results for all proposed cancelable speaker identification techniques. A unified scenario for simulation is adopted. Twenty speech signals are used and encrypted first. Spectrograms of these signals are estimated and stored in the database. All simulation experiments on all techniques depend on genuine and imposter tests. Correlation values for genuine as well as

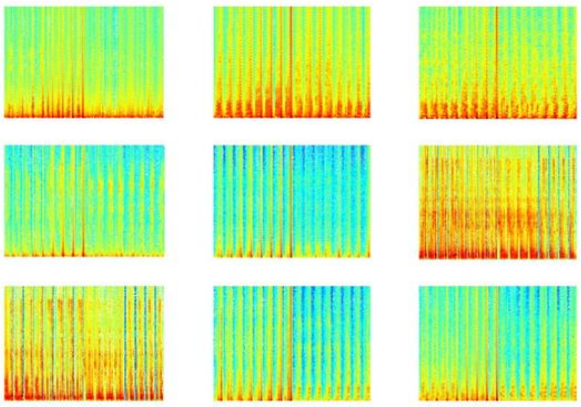
imposter records are estimated, and hence the PDFs of correlation values for genuine and imposter tests are obtained. The intersection points for these PDFs are taken as the threshold values. Based on these values, new records are classified as either genuine or imposter records. Simulation results for the first technique based on chaotic Baker map are given in Fig.8. The PDFs of genuine and imposter tests and the ROC curves in the presence of noise for the first proposed technique using chaotic Baker map are given in Fig.9.



a) Original Signals

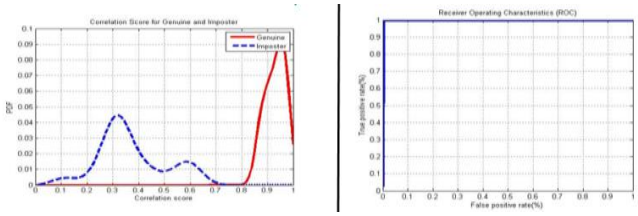


b) Encrypted signals

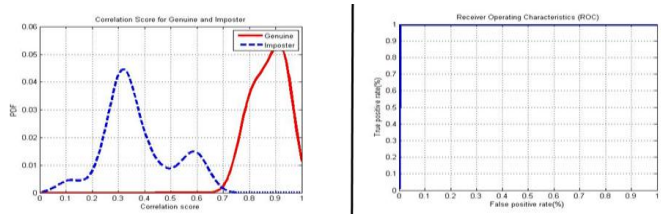


c) Spectrogram of encrypted signals

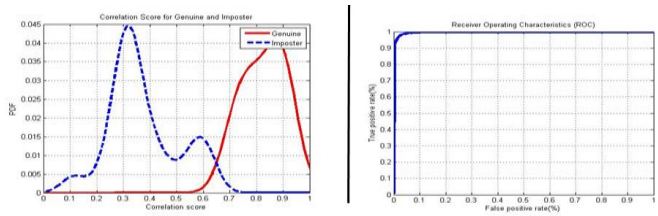
Fig .8 The tested nine speech samples of the first proposed technique using chaotic Baker map.



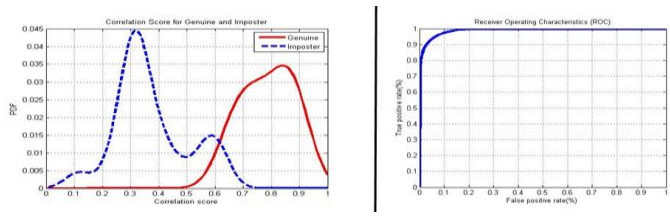
(a) 0.01 noise variance



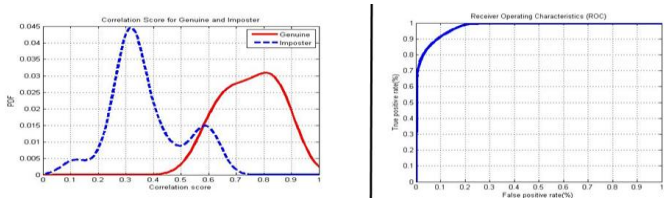
(b) 0.02 noise variance



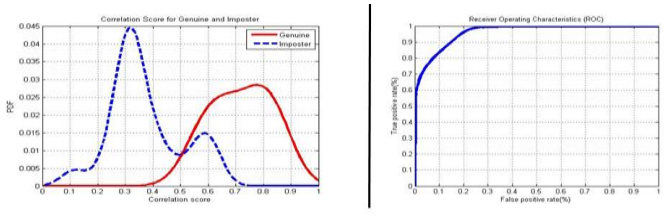
(c) 0.03 noise variance



(d) 0.04 noise variance

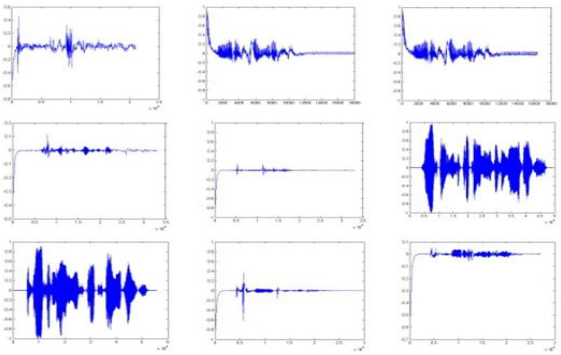


(e) 0.05 noise variance

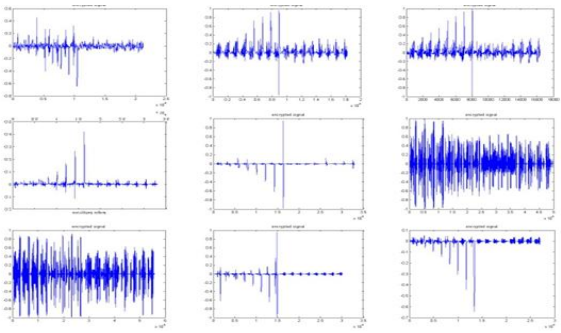


(f) 0.06 noise variance

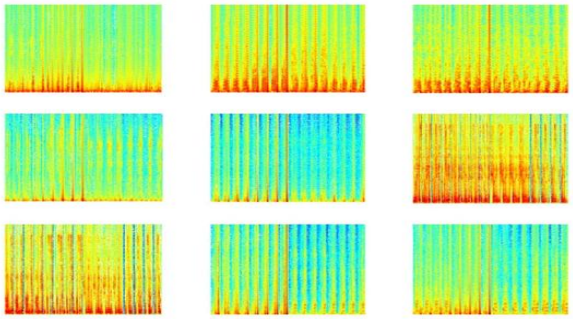
Fig .9 The PDFs of genuine and imposter tests and the ROC curves in the presence of noise for the first proposed technique using chaotic Baker map.



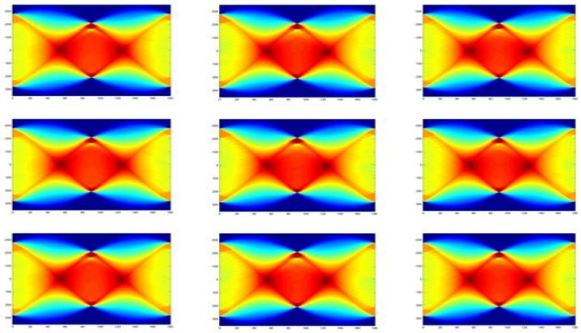
a) Original Speech Signals



b) Encrypted signals

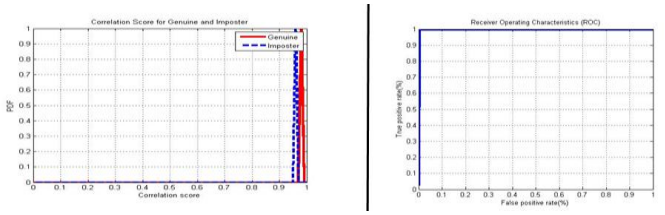


c) Spectrogram of encrypted signals

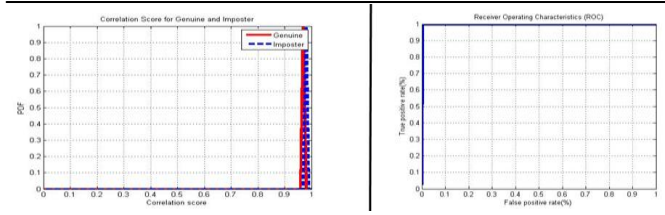


d) Radon transforms of spectrograms of encrypted signals using chaotic Baker map.

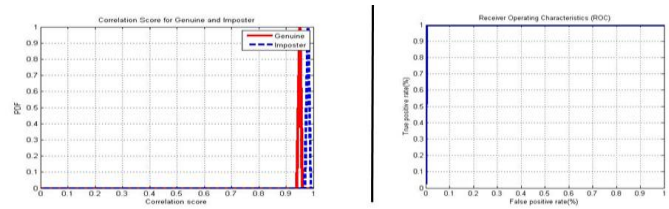
Fig. 10 The tested nine speech samples of the second proposed technique using Radon transform.



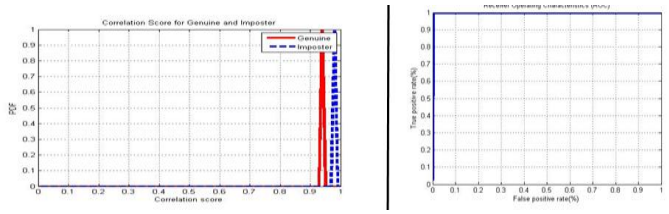
(a) 0.01 noise variance



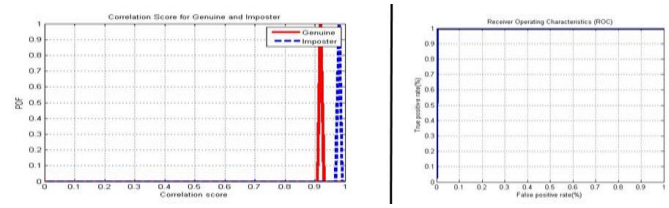
(b) 0.02 noise variance



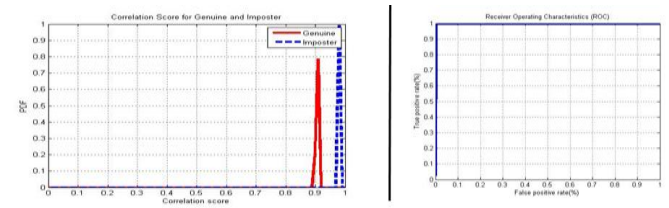
(c) 0.03 noise variance



(d) 0.04 noise variance

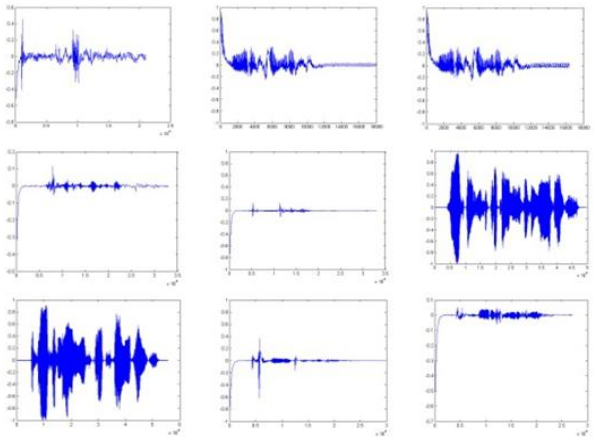


(e) 0.05 noise variance

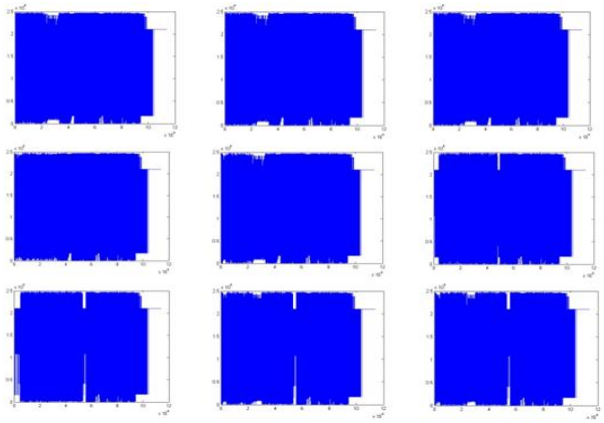


(f) 0.06 noise variance

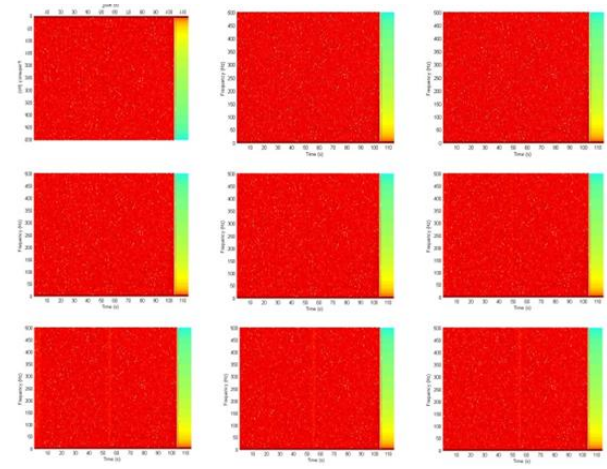
Fig .11 The PDFs of genuine and imposter tests and the ROC curves in the presence of noise for the second proposed technique using Radon transform.



a) Original Speech Signals

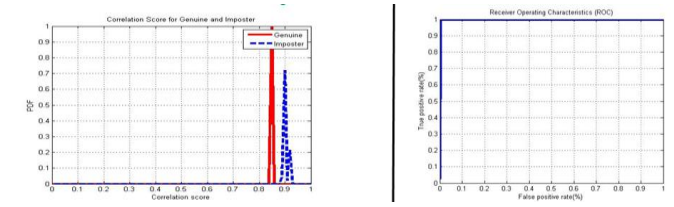


b) Encrypted signals

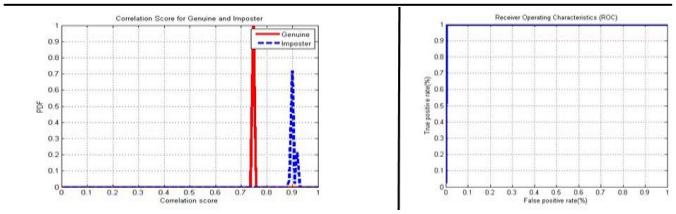


c) Spectrogram of encrypted signals

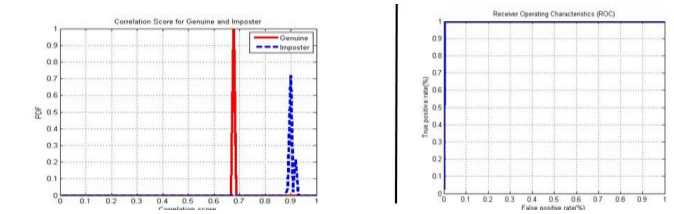
Fig. 12 The tested nine speech samples of the third proposed technique using RSA algorithm.



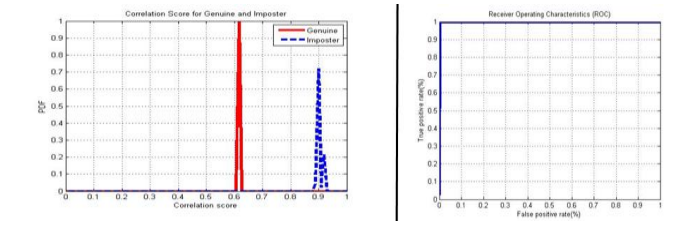
(a) 0.01 noise variance



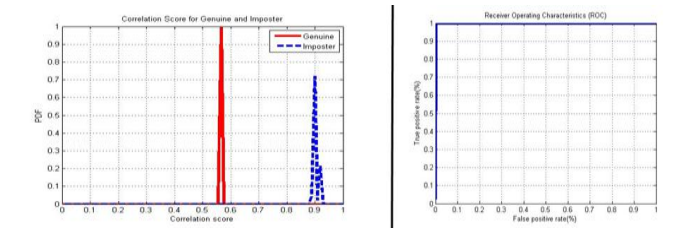
(b) 0.02 noise variance



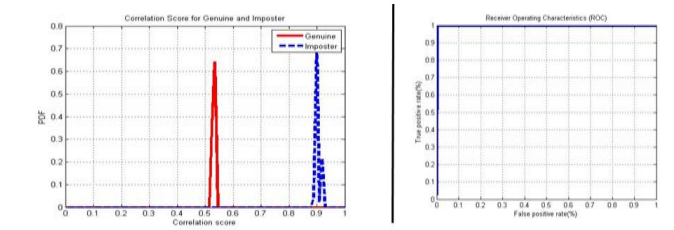
(c) 0.03 noise variance



(d) 0.04 noise variance



(e) 0.05 noise variance



(f) 0.06 noise variance

Fig.13 The PDFs of genuine and imposter tests and the ROC curves in the presence of noise for the third proposed technique using RSA algorithm.

A summary of all evaluation metrics is given in Table 1 at different levels of noise. It is clear that the performance of all techniques is slightly affected by noise with variance up to 0.06.

Table 1 EER, FAR, FRR, and AROC values of the proposed Cancelable biometric systems

Proposed algorithm	EER	FAR	FRR	AROC	Noise σ^2
Chaotic map	**0.0131**	**0**	**1**	**1**	0.01
Radon transform	**0.5000**	**0**	**1**	**1**	
RSA algorithm	**0**	**1**	**0**	**1**	
Chaotic map	**0.0058**	**0**	**1**	**0.9999**	0.02
Radon transform	**0**	**0**	**1**	**1**	
RSA algorithm	**0**	**1**	**0**	**1**	
Chaotic map	**0.0031**	**0**	**1**	**0.9981**	0.03
Radon transform	**0**	**0**	**0**	**1**	
RSA algorithm	**0**	**1**	**0**	**1**	
Chaotic map	**0.0019**	**0**	**1**	**0.9908**	0.04
Radon transform	**0**	**0**	**0**	**1**	
RSA algorithm	**0**	**1**	**0**	**1**	
Chaotic map	**0.0013**	**0**	**1**	**0.9781**	0.05
Radon transform	**0**	**0**	**0**	**1**	
RSA algorithm	**0**	1	**0**	**1**	
Chaotic map	**0**	**0**	**1**	**0.9625**	0.06
Radon transform	**0**	**0**	**0**	**1**	
RSA algorithm	**0**	**1**	**0**	**1**	

IV. CONCLUSION

Three cancelable speaker identification techniques based on the spectrogram estimation of the encrypted signal using chaotic encryption process, Radon transform algorithm and RSA algorithm have been presented. A lot of simulations have been presented to verify the efficiency of the proposed encryption algorithms. Performance comparison has been made between these techniques to determine the most accurate one. In the simulation study, 20 different samples of voice signals for men and women have been used. First, the original signals are encrypted using the proposed encryption algorithms. Then, the spectrograms of the encrypted signals are estimated and stored in the database instead of the original ones. In the experimental results, the values of EER, FRR, FAR, and AROC have been estimated for each proposed work. The ROC curve and the genuine and impostor distributions are also estimated. The proposed speaker identification techniques were also evaluated using cancelable features under the influence of different noise levels. When comparing the results of all techniques, it was found that the first one using chaotic encryption is clearly affected by noise variance variation. The second one using Radon transformation shows better results at the expense of having much execution time. The RSA algorithm shows the most accurate results with the shortest execution time. This makes the technique more accurate in recognizing the user and also more powerful to resist attack attempts. It also becomes more secure through maintaining the confidentiality of the data.

REFERENCES

[1] A. Mostafa, N. Soliman, M. Abdullah and F. E. Abd El-samie, "Speech encryption using two dimensional chaotic maps," 11th International Computer Engineering Conference (ICENCO), 2015.

[2] S. Guennouni, A. Mansouri and A. Ahaitouf, "Biometric Systems and Their Applications," Submitted: October 19th 2018Reviewed: January 30th 2019Published: March 1st 2019 DOI: 10.5772/intechopen.84845

[3] L. Muda, M. Begam and I. Elamvazuthi, "Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques," Journal of Computing, Vol. 2, ISSN 2151-9617, March 2010.

[4] Naglaa F. Soliman, Zhraa Mostafa, Fathi E. Abd ElSamie and Mahmoud I. Abdalla, "Performance enhancement of speaker identification systems using speech encryption and cancelable features," International Journal of Speech Technology, 2017

[5] D. Ambika and V. Radha, "Secure speech Review", *International Journal of Engineering Research and application (IJERA)*, vol. 2, no. 5, pp. 1044-1049, 2012.

[6] Mohammad El-Abed and Christophe Charrier, "Evaluation of Biometric Systems," New Trends and Developments in Biometrics, pp. 149 - 169, 2012, ff10.5772/52084ff. ffhal-00990617

[7] S. Rane et al., "Secure Biometrics: Concepts Authentication Architectures and Challenges", *IEEE Signal Processing*, vol. 30, no. 5, pp. 51-64, 2013.

[8] Christian Rathgeb and Andreas Uhl, "A survey on biometric cryptosystems and cancellable biometrics", proc. EURASIP Journal on Information Security, March 2011.

[9] R. Jain and C. Kant, "Attacks on Biometric Systems: An Overview," International Journal of Advances in Scientific Research; 1(07): 283-288, 2015.

[10] A. Nagar, K. Nandakumar and A. K. Jain, \Biometric template transformation: A security analysis," in Proc. SPIE 7541, Media Forensics and Security II, 7541 (2010).

[11] B. Choudhury, P. Then, B.Issac, V. Raman and M. K. Haldar, "A Survey on Biometrics and Cancelable Biometrics Systems. International Journal of Image and Graphics, 18(01), 1850006 10.1142/S0219467818500067,(2018).

[12] R. Aparna and PL. Chithra, " Role of Windowing Techniques in Speech Signal Processing For Enhanced Signal Cryptography," Advanced Engineering Research and Applications, Chapter 28,Volume V, pp. 446-458, (2017).

[13] Manisha and N. Kumar, "Cancelable Biometrics: a comprehensive survey," Artificial Intelligence Review | 10.1007/s10462-019-09767-8, 2019.

[14] S. Davies. Touching Big Brother, "How Biometric Technology Will Fuse Flesh and Machine," *Information Technology and People*, 7(4), 1994

[15] A. Matsunaga, K. Koga and M. Ohkawa, "An analog speech scrambling system using the FFT technique with high level security", *IEEE J. Select. Areas Common*, vol. 7, pp. 540-547, 1989.

[16] Pointcheval, D. "How to Encrypt Properly with RSA." CryptoBytes, Winter/Spring 2002.

[17] P. Kuchment, "The Radon transform and medical imaging", *SIAM*, 2013.

[18] A. Khatami, M. Babaie, A. Khosravi, H. R. Tizhoosh, S. M. Salaken, & S. Nahavandi, (2017). A deep-structural medical image classification

for a Radon-based image retrieval. 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE). doi:10.1109/ccece.2017.7946756
[19] Fathi E. Abd El-Samie, “Information Security for Automatic Speaker Identification,” Springer, pp 3-9, 2011
[20] N. K. Ratha, J. H. Connell, and R. M. Bolle, “Enhancing security and privacy in biometrics-based authentication systems,” IBM systems Journal, vol. 40, no. 3, pp. 614–634, 2001.
[21] R. Mohammed, "An Enhancement of Some Chaotic Maps for Speech Encryption", *Ph.D. Thesis*, 2014
[22] Yin, S.-C., Rose, R. and Kenny, P. (2007) “A Joint factor Analysis Approach to Progressive Model Adaptation in Text-Independent Speaker Verification,” IEEE Trans. On Audio, Speech, and Language Process., 15, 7, pp. 1999-2010.
[23] J. Godfrey, D. Graff and A. Martin, "Public Databases for Speaker Recognition and Verification," ESCA Workshop on Automatic Speaker Recognition, Identification and Verification, pp. 39-42, 1994.