Computer Systems Science & Engineering
DOI:10.32604/csse.2022.022722
Article

Tech Science Press

Cancelable Speaker Identification System Based on Optical-Like Encryption Algorithms

Safaa El-Gazar[1], Walid El-Shafai[2,3,*], Ghada El-Banby[4], Hesham F. A. Hamed[1], Gerges M. Salama[1], Mohammed Abd-Elnaby[5] and Fathi E. Abd El-Samie[2,6]

[1]Department of Electrical Engineering, Faculty of Engineering, Egyptian - Russian University, Cairo, Egypt
[2]Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf, 32952, Egypt
[3]Security Engineering Laboratory, Department of Computer Science, Prince Sultan University, Riyadh, 11586, Saudi Arabia
[4]Department of Industrial Electronics and Control Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf, 32952, Egypt
[5]Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif, 21944, Saudi Arabia
[6]Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, 84428, Saudi Arabia
*Corresponding Author: Walid El-Shafai. Email: eng.waled.elshafai@gmail.com
Received: 17 August 2021; Accepted: 30 September 2021

Abstract: Biometric authentication is a rapidly growing trend that is gaining increasing attention in the last decades. It achieves safe access to systems using biometrics instead of the traditional passwords. The utilization of a biometric in its original format makes it usable only once. Therefore, a cancelable biometric template should be used, so that it can be replaced when it is attacked. Cancelable biometrics aims to enhance the security and privacy of biometric authentication. Digital encryption is an efficient technique to be used in order to generate cancelable biometric templates. In this paper, a highly-secure encryption algorithm is proposed to ensure secure biometric data in verification systems. The considered biometric in this paper is the speech signal. The speech signal is transformed into its spectrogram. Then, the spectrogram is encrypted using two cascaded optical encryption algorithms. The first algorithm is the Optical Scanning Holography (OSH) for its efficiency as an encryption tool. The OSH encrypted spectrogram is encrypted using Double Random Phase Encoding (DRPE) by implementing two Random Phase Masks (RPMs). After the two cascaded optical encryption algorithms, the cancelable template is obtained. The verification is implemented through correlation estimation between enrolled and test templates in their encrypted format. If the correlation value is larger than a threshold value, the user is authorized. The threshold value can be determined from the genuine and imposter correlation distribution curves as the midpoint between the two curves. The implementation of optical encryption is adopted using its software rather than the optical setup. The efficiency of the proposed cancelable biometric algorithm is illustrated by the simulation results. It can improve the biometric data security without deteriorating the recognition accuracy. Simulation results give close-to-zero values for the Equal Error Rate (EER) and close-to-one values for the Area under Receiver Operator Characteristic (AROC) curve.

Keywords: Cancelable biometrics; spectrogram; OSH; DRPE; EER; AROC

1 Introduction

Biometric is a Greek derivation word, where Bio means life and Metric means to measure. Biometrics represent the physical and behavioral characteristics that distinguish between individuals. A biometric can be physical such as fingerprint, iris, palm-print, etc. On the other hand, it can be behavioral, such as voice, gait, hand gesture, etc. [1]. The main thing that distinguishes biometrics is that it is impossible to have two people with the same physical or behavioral biometrics. Due to biometric uniqueness, it is efficient and practical for personal recognition. However, there is a limitation in using biometrics in privacy issues. When a biometric is compromised, it is no longer valuable and cannot be used again. Therefore, it is necessary to store biometrics in an intended destructive way, so that when a biometric is compromised, the hacked template can be replaced by another destructive one in a different way. These intentional distorted biometric versions are called cancelable biometrics [2]. Four characteristics must distinguish the cancelable biometrics:

1) Reusability. If a biometric template is compromised, it can be reissued.
2) Diversity. Each different application has its own cancelable biometric template.
3) Non-invertibility. Compromised biometric templates cannot be used to recover the original biometrics.
4) Performance. Cancelable biometrics should not degrade the recognition performance [3]. A cancelable biometric recognition system has two stages: (1) enrollment which means storage of the individual cancelable biometrics in the database, and (2) authentication, where the test biometric is converted to a cancelable template with the same method used in the enrollment stage. Then the test and stored templates are matched, and hence, recognition is carried out. There are different methods for cancelable biometrics such as random projection, cancelable biometric filters, hybrid methods etc. A survey of cancelable biometric methods is provided in [4].

The main motivation of this paper is to introduce a new algorithm for cancelable biometrics for the recognition systems based on voice-print. Speech signal is the biometric in this paper. Cancelable biometrics is investigated through two cascaded optical encryption algorithms applied on the spectrograms of the speech signals. The first optical encryption algorithm is the OSH followed by DRPE with two randomly generated RPMs. Then, to verify the user identity, the same steps are applied on the test speech signals. The correlation between the two encrypted templates is estimated and compared with the threshold value to determine if a user is authorized or not. The proposed algorithm can achieve two characteristics of cancelable biometrics, which are non-invertibility and high performance. The rest of the paper is organized as follows. Section 2 introduces some recent related studies. Section 3 gives the OSH algorithm. Section 4 introduces the DRPE algorithm. Section 5 provides the proposed algorithm. Section 6 gives a discussion of the simulation results. Finally, Section 7 summarizes the concluding remarks.

2 Related Works

Cancelable biometrics is based on repeatable intentional distortion of different biometrics to use deformed or transformed biometric versions in the verification process. Increasing user privacy and preserving a high performance of the recognition process are two conflicting requirements of cancelable biometric systems. Different algorithms have been proposed for cancelable biometrics. Some of them are unimodal biometrics, which use only one biometric for the verification process. Some others are multimodal biometrics that use more than one biometric for the verification process. In addition, several cancelable biometric methods depend on encryption algorithms.

One of the methods depends on encrypted multimodal biometrics was introduced in Tarek et al. [5]. They presented an image-based multimodal biometric authentication scheme that utilizes the DRPE. In the enrollment, they encrypt the fingerprint image using DRPE and a palm-print image as a secret key for the encryption algorithm. At the authentication, the user is considered to be authorized if the encrypted fingerprint can be successfully decrypted by the palm-print image from the same user. Then, the decrypted image is matched against a fresh fingerprint image. In addition, another scheme based on utilizing DRPE for cancelable face and iris recognition systems was presented by Soliman et al. [6]. They used the scale-invariant feature transform to extract the face image features, which are encrypted using the DRPE algorithm. For the iris recognition, the features from two iris images of the same person are extrcated. One of the two iris images is encrypted with the DRPE, and the other is used to generate the second phase mask used in the DRPE.

Also, Rachapalli et al. [7] presented a cancelable biometric cryptosystem using color QR codes. Their system depends on key generation, free registration, and works with a conventional matcher. They proposed a system that takes the multimodal biometric fused templates for texture, shape, and color classification as input to the cipher data conversion module to generate the color QR code image. Sudhakar et al. depended on steganography in cancelable iris recognition, where a cross-folded iris is embedded into a generated QR code. The verification process is investigated using deep learning by implementing the multilayer perceptron architecture for user recognition [8]. Another multimodal cancelable biometric system was proposed by Tarif et al. [9]. It depends on digital encryption and hiding techniques. It aims to achieve secure transmission of multimodal biometric information in the identification system. The method adopts face, iris, and fingerprint biometrics. It factorizes the fingerprint and iris images, separately, and then embeds them in the host Slantlet Transform Singular Values (SLT-SVs) of the face image.

Yang et al. [10] worked on an access control system for a critical infrastructure. Fingerprint and face images are the considered biometrics. They introduced two cascaded encryption layers: the core and the expendable layers. The non-invertible transformation key and face feature set are merged together by utilizing the fuzzy commitment. Abouelazm et al. [11] presented a multimodal cancelable biometric scheme, which adopts optical encryption and the 3D jigsaw transform. Face and fingerprint are the considered biometrics. The algorithm adopts a single random phase mask with the Fractional Fourier Transform (FRFT). It depends on utilizing two stages of the 2D-FRFT separated with kernels in both dimensions and a random phase mask.

The comb filter has been used to generate cancelable speaker templates by Kareem et al. The multiple nulls in comb filter are used to induce intentional distortion in the speech signals [12]. Barrero et al. introduced a multi-biometric template protection scheme based on the Bloom filter. The introduced method estimates the main parameters of the different biometrics. Based on these estimations, a new weighted feature level fusion scheme of Bloom-filter-based templates was introduced in [13].

Moreover, Mostafa et al. [14] presented encrypted cancelable templates that can be checked based on correlation estimation. Ouda et al. used an extended bio-encoding algorithm to generate cancelable templates of iris images. This type of cryptosystem depends on key binding. The system treats two problems of cancelable biometrics: accuracy preservation, non-invertibility [15]. Another encryption-based algorithm for cancelable biometric generation was introduced by Alarifi et al. [16]. An a symmetric cryptography algorithm was implemented based on optical Phase Truncated Fourier Transform (PTFT). The ciphering key is completely different from the deciphering key. Each of them is obtained through two different random independent phase operations. The phase truncation in Fourier transform is used to obtain the cancelable biometrics.

Kim et al. implemented a cancelable ECG biometric algorithm for identity verification. The cancelable ECG templates are obtained by deriving the generalized likelihood ratio test from a composite hypothesis testing in the compressive sensing domain [17]. Kaur et al. used a random distance metric to generate the cancelable biometric templates. The extracted feature vector is represented as a point in the Cartesian coordinates. The matching process is investigated through the feature vector distance from a random point. If the distance between two feature vectors is small, then the feature vectors will belong to the same user [18]. Chee et al. investigated the projection method to obtain the cancelable templates. Random Binary Orthogonal Matrices Projection (RBOMP) hashing was the method used to project the biometric features to an ordinal space using binary orthogonal matrices. The authors also implemented Prime Factorization (PF) features to enhance security and privacy [19].

Another trend depends on using neural networks for face recognition and authentication. It was proposed by Abdellatef et al. [20] and Jang et al. [21]. Abdellatef et al. [20] extract the deep features from different facial regions by multiple Convolutional Neural Networks (CNNs). They adopted a CNN architecture that has the advantages of depth concatenation, batch normalization, and residual learning. They adopted a region-based technique that depends on detecting faces, eyes, nose, and mouth regions from the original face images. Deep features of each region are extracted using multiple CNNs. Then, a fusion network merges these features. Finally, bio-convolving encryption is implemented on the final facial descriptors to obtain the cancelable biometric templates. Jang et al. [21] also exploited a CNN for cancelable biometric generation. They proposed a Deep Table-based Hashing (DTH) algorithm, which is based on CNN-based feature encoding into binary codes using the index of the hashing table. To train the CNN, they proposed a segment-clustering loss and a pairwise Hamming loss with two classification losses. The final authentication results were obtained by voting on the outcome of the retrieval system.

Most of the previous work implemented biometric images such as fingerprints, palm print, and iris, while ignoring the voice-print which is a biometric that can be considered in system access. Nevertheless, the researches that considered the voice-print used traditional methods of encryption to obtain the cancelable templates. Reusability, diversity, non-invertibility and high performance are the four characteristics of cancelable biometrics algorithms. It is difficult for an algorithm to achieve the four characteristics at the same time. In this paper, a new and efficient encryption algorithm is implemented to encrypted the voice-prints and obtain the cancelable templates. Two cascaded optical encryption algorithms are implemented, OSH followed by DRPE, to encrypt voice-print spectrograms. Non-invertibility and high performance are two characteristics of the cancelable biometrics obtained with the proposed algorithm.

3 Optical Scanning Holography Algorithm

In 1948, light intensity and pahse have been recorded on films by Denis Gabor. This method is known as holography. The hologram illumination is a reproduction of the original 3-D wave field. In 1979, this hologram was recorded using a heterodyne system to be stored digitally on the computer. This process is called Optical Scanning Holography (OSH), and it was invented by Poon et al. [22].

The OSH physical architecture is shown in Fig. 1. It includes a photodetector, a heterodyne system, a lens, and a Band-Pass Filter (BPF). The OSH working theory depends on the object transparency scanning. The transparency carries all object information. An optical beam scans out the transparency. The photodetector receives all light and converts it to an electrical holography output, which can be displayed on the monitor or stored on the computer. The OSH transfers the outcome optical information into an electrical signal [23]. In this paper, speech signal spectrogram is encrypted using the OSH algorithm. The OSH algorithm is adopted using its software rather than its physical structure.

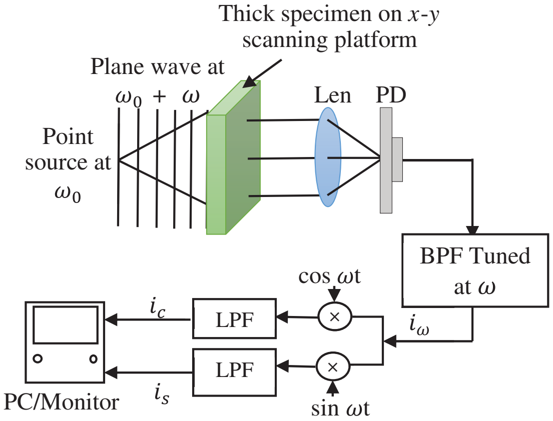

Thick specimen on x-y scanning platform
Plane wave at $\omega_0 + \omega$
Point source at ω_0
Len
PD
BPF Tuned at ω
$\cos \omega t$
i_c
LPF
i_ω
LPF
PC/Monitor
i_s
$\sin \omega t$

Figure 1: Architecture of the OSH [24]

Mathematically, the OSH can be explained as follows. At the beginning, we assume that the scanned optical beam complex field is represented by $\Phi_b(i, j)$ and the scanned object transparency is represented by $T_o(i', j')$ at the point (i', j'). Then, the complex field at the photodetector can be represented by $T_o(i', j')\Phi_b(i' - i, j' - j)$. Because of the transitional nature of transparency, there are shifts in the coordinates of the complex field as $\Phi_b(i' - i, j' - j)$. The electrical current output from the photodetector can be represented by the spatial integration of the light intensity over the active area R of the detector $|\Phi_b(i' - i, j' - j)T_o(i'.j')|^2$. Then, the resultant current can be stored or displayed in a 2D format as follows:

$$c(i, j) \propto \iint_R |T_o(i', j')\Phi_b(i' - i, j' - j)|^2 di' dj' \tag{1}$$

The scanning beam instantaneous position can be represented by $i(t)$ and $j(t)$, where the scanning beam velocity can be represented by $i(t) = j(t) = vt$. Eq. (1) can be denoted in 2D convolution as follows:

$$c(i, j) = |\Phi_b(i, j)|^2 \oplus |T_o(i, j)|^2 \tag{2}$$

The object $T_o(i, j)$ initially may be complex. The object intensity, $|T_o(i, j)|^2$ is processed by a real and non-negative quantity, $|\Phi_b(i, j)|^2$ [23].

In order to perform a 3-D imaging, phase information of light must not change. During 3D recording, it is not allowed to process any phase information. To solve this problem, the heterogeneous optical scanning is used. A time-dependent Fresnel Zone Plate (FZP), which depends on the superposition of a spherical wave and a plane wave with different temporal frequencies, is used. At the top of Fig. 1, the sent light is collected by the lens to be directed to the photodetector. We assume that there is a distance k between the spot of the focusing laser beam, which generates the spherical wave and the transparency of the object $T_o(i, \mathrm{j}; k)$. At the transparency, the scanning beam pattern can be defined as:

$$\Phi_b(i.j;\ k) = a\exp[I(\omega_0 + \omega)t] + \frac{Iz_o b}{2\pi k}\exp\left[\frac{-Iz_o}{2k}(i^2 + j^2)\right]\exp(I\omega_0 t) \tag{3}$$

where a is the plane wave amplitude on the film, b is the amplitude of the point source, $z_o = 2\pi/\lambda,\ \ \omega_o$ and $\omega_o + \omega$ are the spherical and the plane wave frequencies, respectively. Then, Eq. (1) can describe the generated current from the photodetector:

$$c(i, j) = \iint_R |T_o(i'.j';\ k)\Phi_b(i' - i, j' - j;\ k)|^2 di' \ \ dj' \tag{4}$$

By substituting Eq. (3) into Eq. (4) and keeping the heterodyne current by using a BPF tuned at the heterodyne frequency ω, we have

$$c_{\omega}(i,\ j) = ab\frac{z_o}{\pi k}\sin\left[\omega t + \frac{z_o}{2k}(i^2 + j^2)\right] \oplus |T_o(i.j;\ k)|^2 \tag{5}$$

From Eq. (5), the optical transfer function OTF can be obtained as follows:

$$OTF(i,\ j,\ k) = fft\{c_{\omega}(i,\ j,\ k)\}/fft\{|T_o(i,\ j,\ k)|^2\} \tag{6}$$

The heterodyne current frequency can represent the object information. This information can be obtained though an electronic mixture of sine and cosine functions at the heterodyne frequency to obtain the in-phase and the quadrature components of the heterodyne current, respectively.

$$c_{\omega}(x.y) \times \sin\omega t = \frac{abz_o}{2\pi k}\left\{\cos\left[\frac{z_o}{2k}(i^2 + j^2)\right] \oplus |T_o(i.j;\ k)|^2 - \cos\left[2\omega t + \frac{z_o}{2k}(i^2 + j^2)\right] \oplus |T_o(x.y;\ z)|^2\right\} \tag{7}$$

Applying an electronic Low Pass Filter (LPF), the hologram of the cosine FZP $G_{cos}(i, j)$ can be obtained:

$$c_s(i,\ j) \propto G_{cos}(i,\ j) = \frac{z_o}{2\pi k}\left\{\cos\left[\frac{z_o}{2k}(i^2 + j^2)\right] \oplus |T_o(i.j;\ k)|^2\right\} \tag{8}$$

Similarly, the sine FZP hologram after LPF $G_{cos}(x.y)$ can be obtained:

$$c_c(i,\ j) \propto G_{sin}(i,\ j) = \frac{z_o}{2\pi k}\left\{\sin\left[\frac{z_o}{2k}(i^2 + j^2)\right] \oplus |T_o(i,\ j;\ k)|^2\right\} \tag{9}$$

Then, holograms can be obtained as follows:

$$G_{cos}(i,\ \mathrm{j}) = \int \frac{z_o}{2\pi k}\left\{\cos\left[\frac{z_o}{2k}(i^2 + j^2)\right] \oplus |T_o(i,\ j;\ k)|^2\right\}dk \tag{10a}$$

$$G_{sin}(i,\ j) = \int \frac{z_o}{2\pi k}\left\{\sin\left[\frac{z_o}{2k}(i^2 + j^2)\right] \oplus |T_o(i,\ j;\ k)|^2\right\}dk \tag{10b}$$

Eq. (10) can be rewritten as follows giving the OTF:

$$G_{cos}(i,\ j) = Im\lfloor fft^{-1}\{ fft\{C(i,\ j)\}OTF_{osh}(i,\ j;\ k)\}\rfloor \tag{11a}$$

$$G_{sin}(i,\ j) = Re\lfloor fft^{-1}\{ fft\{C(i,\ j)\}OTF_{osh}(i,\ j;\ k)\}\rfloor \tag{11b}$$

The reconstructed complex FZP hologram $G_{compelx}(i, j)$ can be obtained by Eq. (12).

$$\begin{aligned} G_{complex}(i,\ j) &= G_{sin}(i,\ j) + jG_{cos}(i,\ j) \\ &= fft^{-1}\{ fft\{C(i,\ j)\}OTF_{osh}(i,\ j,\ k)\} \end{aligned} \tag{12}$$

The reconstructed image can be obtained by convolving any of the above holograms with the spatial impulse response $g(i, j, k)$ as in the following equation:

$$G(i,\ j) * g(i,\ j,\ k) \tag{13}$$

Finally, we adopt the inverse fft^{-1} to reconstruct the image as follows:

$$\begin{aligned} \text{Reconstructed real image } \alpha\ & fft^{-1}\{ fft\{G(i,\ j)\}G(i,\ j,\ k)\} \\ &= fft^{-1}\{ fft\{G(i,\ j)\}OTF_{osh}*(i,\ j,\ k)\} \end{aligned} \tag{14}$$

The OSH MATLAB program and its flowchart are indicated in detail in [25,26]. In this paper, the OSH is implemented using its software not the physical structure.

4 Double Random Phase Encoding (DRPE) Algorithm

The DRPE was presented first in 1995 by Refregier and Javidi. It is one of the most popular optical encryption schemes. This is due to its applicability using an optical system or a software. DRPE is based on the modification of the spectral distribution of the image. The optical system of the DRPE algorithm is illustrated in Fig. 2. It consists of two cascaded lenses with a Fourier Transform (FT) plane between them, two image planes, one for the input and the other for the output. They are all separated by focal length f to make the total separation distance as $4f$ as shown in Fig. 2. Hence, the DRPE system is known as a $4f$ system [27]. Two random phase masks are applied (RPM1 and RPM2). One is applied at the input plane, and the other is applied on the FT plane in order to increase the security level. An inverse FT is applied through the second lens to obtain the encrypted image in the spatial domain.

The encryption process with the DRPE algorithm can be represented mathematically as follows:

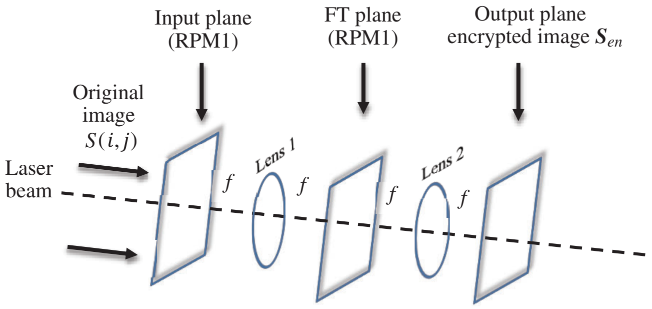


Figure 2: Optical DRPE system [27]

$$S_{en} = \{S(i,\, j)\ \varphi_1(i,\, j)\} * fft^{-1}\{\varphi_2(i,\, j)\} \tag{15}$$

where $S(i,\, j)$ and S_{en} are the original image and the encrypted image in the spatial domain, respectively. $\varphi_1(i,\, j)$ and $\varphi_2(i,\, j)$ are the two random phase masks RPM1 and RPM2, respectively.

RPM1 and RPM2 are 2D matrices of the same size as the original image $S(i,\, j)$. Their values are distributed uniformly between 0 and 2π. RPM1 and RPM2 are generated using one of the different random distributions such as uniform and Gaussian distributions [28] for efficient encryption of images.

5 Proposed Algorithm

This paper introduces a proposed algorithm for cancelable biometric recognition. Speech signal is considered to be the used biometric in this paper. The proposed algorithm depends on speech spectrogram encryption, and two cascaded optical encryption algorithms, which are OSH and DRPE. The spectrogram is a 3D representation of the signal amplitude with time and frequency. Spectrogram can be obtained by implementing the Short-Time Fourier Transform (STFT) of the signal. Then, the signal is segmented into fixed-length frames. After that, a small overlapped window is applied. The spectrogram is usually obtained as an image with color or brightness representing the varying amplitude with frequencies on the vertical axis and time instants on the horizontal axis [29].

This paper introduces two optical encryption algorithms with software and mathematics explained in Sections 3 and 4 rather than the optical setup. First, the spectrogram image is encrypted by the OSH software. Then, the OSH encrypted spectrogram is encrypted again by the DRPE software with its two randomly generated RPMs to increase the security level. The proposed algorithm has two phases: the enrollment phase, and the verification phase as shown in Fig. 3. The enrollment phase can be explained as follows:

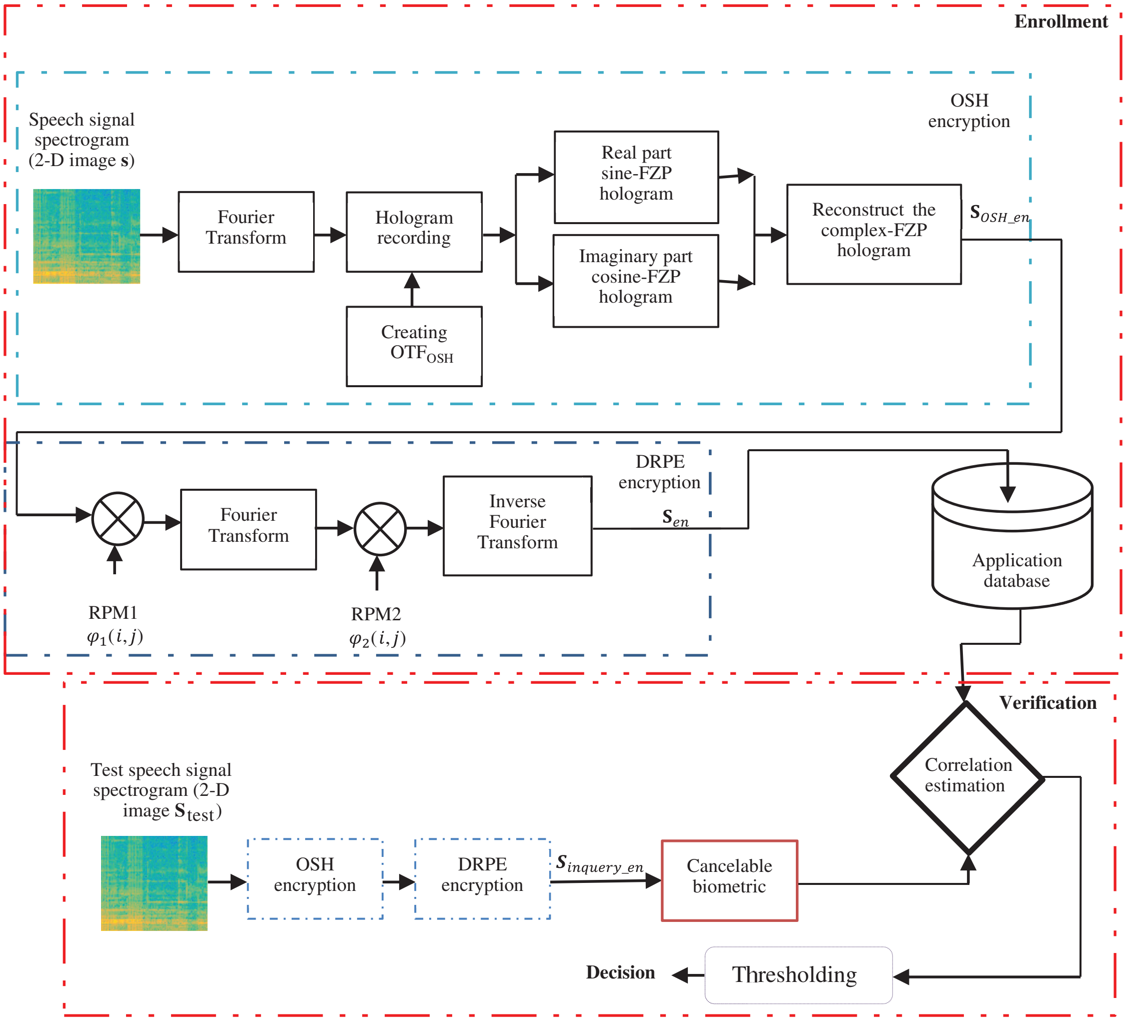


Figure 3: Cancelable verification system

1) The speech signal is transformed into spectrogram image s.
2) The spectrogram is encrypted using OSH software that begins with creating the OTF$_{\text{OSH}}$ through the equation:

$$OTF_{OSH} = \exp[-j\frac{z}{2k_0}(k_i^2 + k_j^2)] \quad (16)$$

where k_i and k_j are the spatial frequencies.

3) The FT is applied to image s.

$$S = fft(s) \tag{17}$$

4) The hologram FH is recorded in the frequency domain by the equation:

$$FH = S * OTF_{OSH} \tag{18}$$

5) The real part sine-FZP hologram and the imaginary part cosine-FZP hologram are obtained.

$$G_{sin}(i,\ j) = Re[\,fft^{-1}\{\,fft\{S(i,\ j)\}OTF_{OSH}(k_i,\ k_j;\ k_0)\}] \tag{19}$$

$$G_{cos}(i,\ j) = Im[\,fft^{-1}\{\,fft\{S(i,\ j)\}OTF_{OSH}(k_i,\ k_j;\ k_0)\}] \tag{20}$$

6) The reconstructed encrypted image S_{OSH_en} is obtained by reconstructing the complex FZP hologram $G_c(i,\ j)$.

$$G_c(i,\ j) = G_{sin}(i,\ j) + jG_{\cos}(i,\ j) \tag{21}$$

$$S_{OSH\_en} = G_c(i,\ j) = fft^{-1}\{\,fft\{S(i,\ j)\}OTF_{OSH}(i,\ j;\ k_0)\} \tag{22}$$

7) The OSH encrypted spectrogram S_{OSH_en} is encrypted again by DRPE software as follows.
8) The two random phase RPM1 and RPM2 known as $\varphi_1(i,\ j)$ and $\varphi_2(i,\ j)$ are generated. At the end, the encrypted spectrogram can be given by the equation:

$$S_{en} = \{S_{OSH_{en}}(i,\ j)\ \varphi_1(i,\ j)\} * fft^{-1}\{\varphi_2(i,\ j)\} \tag{23}$$

where $\varphi_1(i,\ j)$ and $\varphi_2(i,\ j)$ are 2D matrices of the same size as that of the OSH encrypted spectrogram S_{OSH_en}.

In the verification phase, the cancelable biometric template of the test user is obtained using the same steps in the enrollment. Then, the stored and test templates are 1:1 matched. After the matching process, the resultant correlation value between the two templates is compared with a threshold value to give the decision. The verification stage is, in fact, a thresholding task. The threshold is determined at the intersection point of correlation distributions for genuine and imposter users. First of all, several genuine tests are performed, and the correlation score is treated as a random variable. The Probability Distribution Function (PDF) of the genuine score is estimated. Similarly, several tests are performed for imposter users, and the correlation scores are recorded. The PDFs of these correlation scores of imposter test are estimated. The intersection point between both PDFs is estimated, and hence the threshold value is evaluated. Correlation value can be estimated through the following equation:

$$c_r = \frac{\sum_{\boldsymbol{l}=1}^{\boldsymbol{m}}(x_l - E(\mathbf{x}))(y_l - E(\mathbf{y}))}{\sqrt{\left(\sum_{l=1}^{m} x_l - E(\mathbf{x})\right)^2}\ \sqrt{\left(\sum_{l=1}^{m} y_l - E(\mathbf{y})\right)^2}} \tag{24}$$

where c_r is the correlation value, x_l and y_l are the values of lth pixel intensity of the stored and test templates, respectively. $E(\mathbf{x})$ and $E(\mathbf{y})$ are the values of stored and test templates mean intensity [30].

6 Experimental Results

This section displays the proposed algorithm results with some discussions. An Intel 2.5 GHz processor that has a 6.00 GB RAM has been used in simulation with MATLAB R2016b. Sample speech signals have been captured from the MIT dataset to be used in the simulation experiments.

The proposed algorithm depends on two levels of security to generate cancelable templates from speech signals. The speech signal spectrogram is encrypted with two cascaded optical encryption algorithms. The OSH and the DRPE are the two implemented optical encryption algorithms. The verification process is performed on the encrypted cancelable templates, which means that the test biometric is checked, while it is encrypted. The verification process depends on the correlation value between the enrolled and test biometrics. If the correlation value is larger than a threshold, the user is considered as an authorized user. The threshold value is determined from correlation distributions of genuine and impostor curves as the intermediate point between the two curves.

Several evaluation metrics are adopted to evaluate the proposed system. The Receiver Operator Characteristic (ROC) curve and the distributions of genuine and impostor tests are obtained from the simulation experiments. In addition, the Area under the ROC curve (AROC), Equal Error Rate (EER), False Accept Rate (FAR), and False Reject Rate (FRR) are calculated to evaluate the performance efficiency.

The ROC curve is a plot of False Positive Rate (FPR) *versus* True Positive Rate (TPR). The FPR is the number of false positives to the total number of negatives, and TPR is the number of true positives to the total number of positives. The point at which FAR and FRR are equal is the EER value. Both FAR and FRR represent the probability of wrong decisions, where FAR defines imposter cases classified as genuine cases and FRR defines genuine cases classified as imposter cases. The closer the EER value to zero is, the higher the accuracy of the security system. The closer the value of AROC to one is, the higher the system efficiency [31–33].

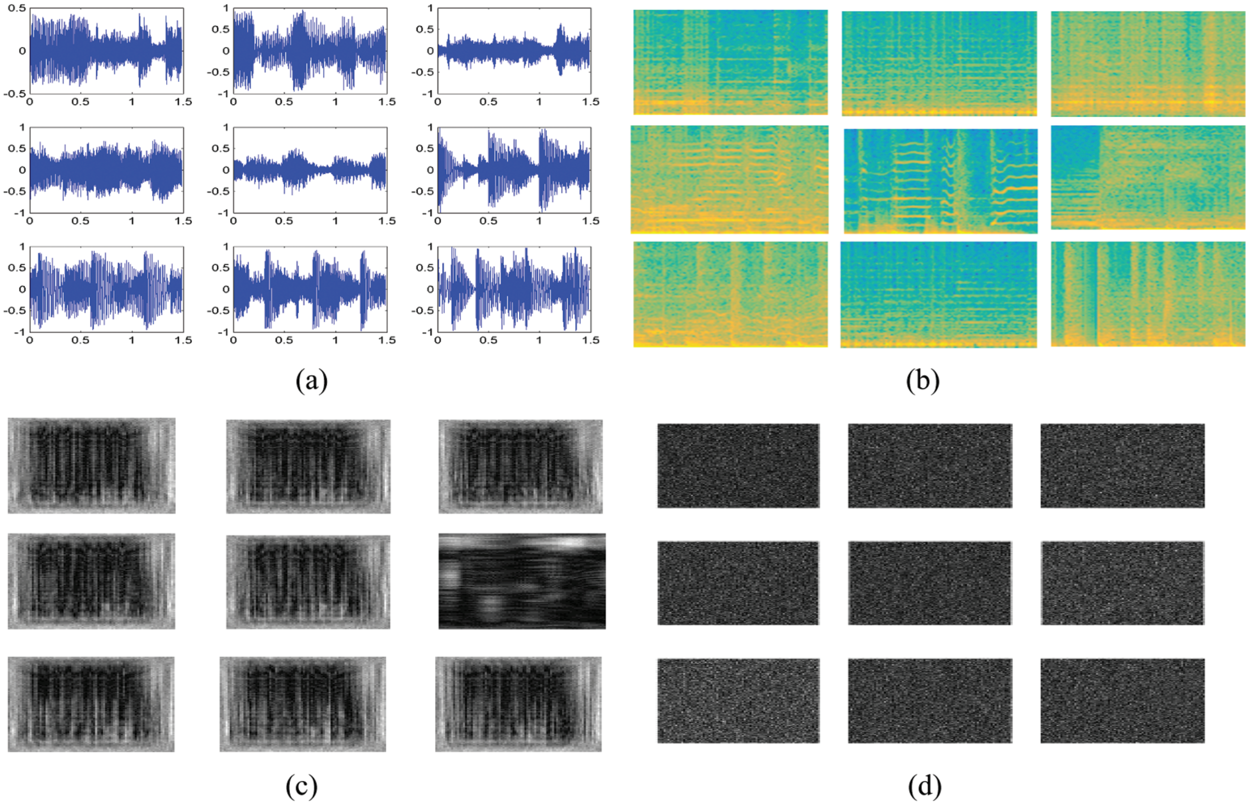
Displayed results start with the numerical results of the evaluation metrics, which are given in Tab. 1 in noise scenarios with variances of 0.01, 0.02, 0.03, and 0.04. The table gives the EER, AROC, FAR, FRR, and processing time values. The table compares the two security levels. The first depends on the OSH encryption algorithm, and the second depends on a cascaded structure of OSH and DRPE encryption algorithms. It is clear from the table that the cascaded encryption of two stages gives better results than the those of the OSH, only. For the cascaded encryption, EER values are close to zero indicating a good performance of the proposed algorithm. From the table, the AROC value is one in the case of cascaded stages, which indicates a high accuracy of the cancelable biometric recognition algorithm to distinguish between genuine and imposter cases.

Table 1: Different evaluation metric values and the processing times at different values of the noise variance and zero mean for the proposed algorithm

Performance metrics/Algorithm		AWGN variance			
		0.01	0.02	0.03	0.04
EER	First level of security (OSH)	3.23×10^{-07}	3.23×10^{-07}	3.23×10^{-07}	3.23×10^{-07}
	Second level of security (OSH and DRPE)	7.9×10^{-020}	1.63×10^{-30}	3.66×10^{-43}	2.78×10^{-55}
FAR	First level of security (OSH)	6.468×10^{-07}	6.468×10^{-07}	6.46×10^{-07}	6.46×10^{-07}
	Second level of security (OSH and DRPE)	0.0	0.0	0.0	0.0
FRR	First level of security (OSH)	1.00	1.00	1.00	1.00
	Second level of security (OSH and DRPE)	1.00	1.00	1.00	1.0000
AROC	First level of security (OSH)	0.3261	0.0501	0.0500	0.0500
	Second level of security (OSH and DRPE)	1	1	1	1
T (Sec)	First level of security (OSH)	2.467	2.447	2.446	2.465
	Second level of security (OSH and DRPE)	3.483	3.540	3.444	3.230

Fig. 4a shows nine random versions of speech signals existing in the database. Fig. 4b shows the speech signal spectrograms. Figs. 4c and 4d show the OSH and cascaded structure encrypted spectrograms, respectively. The genuine and impostor distributions are shown in Fig. 5. It is obvious from the figure that the proposed algorithm has a good performance, where there is an enough distance between the distribution curves to distinguish the genuine and imposter ones. The ROC curves in the presence of different noise variance values are shown in Fig. 6. From these curves, it is obvious that the performance of the proposed algorithm is good with AROC values close to one.

Tab. 2 gives the average EER values for the proposed algorithm compared to those of the algorithms given in Refs. [8,10,11,16,19]. It is clear from the table that the proposed algorithm gives low EER values, which means a good performance.



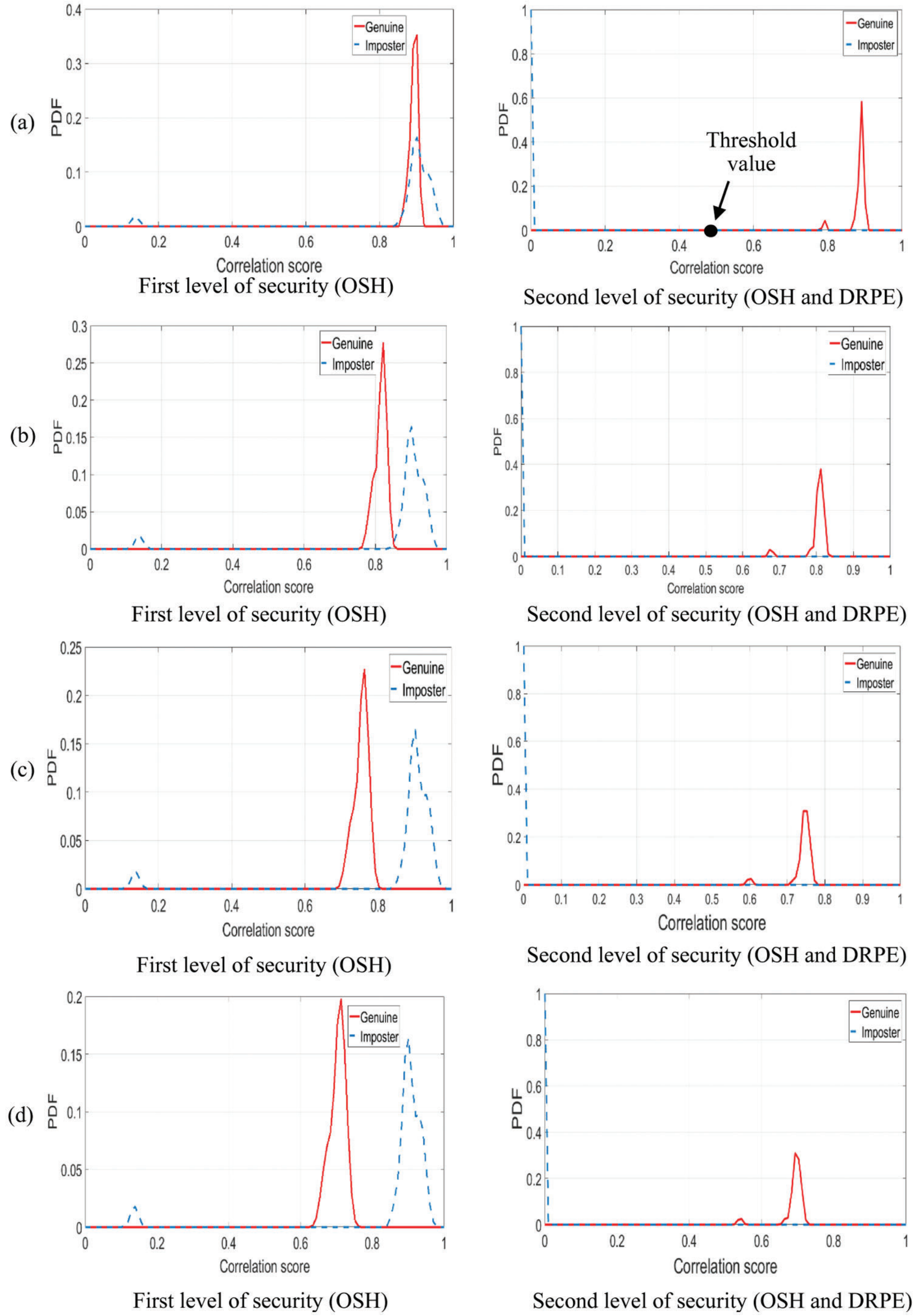
(a)

(b)

(c)

(d)

Figure 4: Speech biometrics, spectrograms, and encrypted versions. (a) Nine random original speech waveforms from the used databases. (b) Spectrograms. (c) OSH encrypted spectrograms. (d) Encryption spectrograms using the cascaded encryption algorithm



First level of security (OSH)

Second level of security (OSH and DRPE)

First level of security (OSH)

Second level of security (OSH and DRPE)

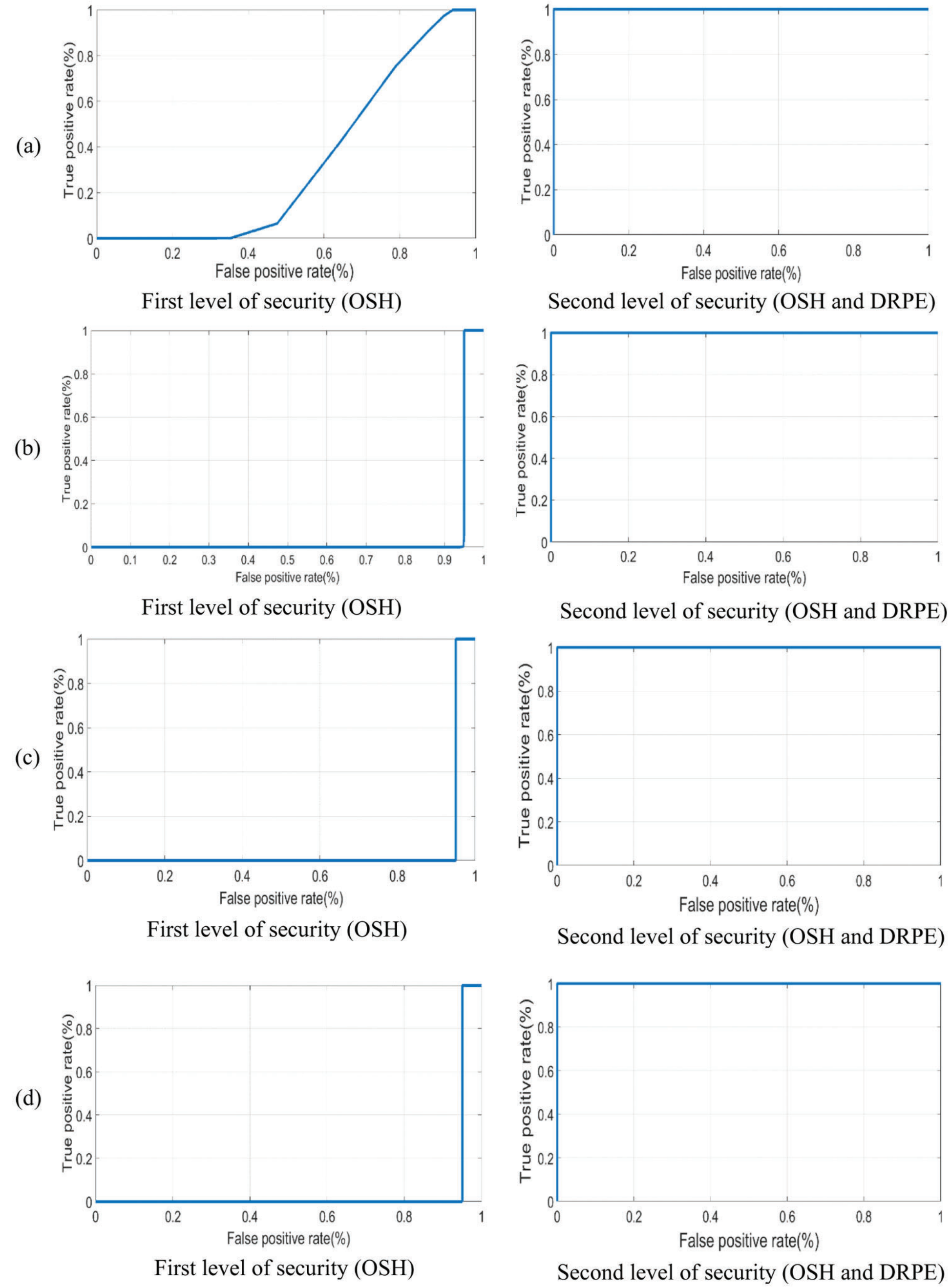
First level of security (OSH)

Second level of security (OSH and DRPE)

First level of security (OSH)

Second level of security (OSH and DRPE)

Figure 5: Genuine and impostor distributions in noisy scenarios with zero mean and different variance values. (a) 0.01, (b) 0.02, (c) 0.03 and (d) 0.04



First level of security (OSH)

Second level of security (OSH and DRPE)

First level of security (OSH)

Second level of security (OSH and DRPE)

First level of security (OSH)

Second level of security (OSH and DRPE)

First level of security (OSH)

Second level of security (OSH and DRPE)

Figure 6: ROC curves in noisy scenarios with zero mean and different variance values. (a) 0.01, (b) 0.02, (c) 0.03 and (d) 0.04

Table 2: Average EER values of the proposed cancelable algorithm and the previous cancelable biometric algorithms in [8,10,11,16,19]

Cancelable algorithm	EER
Proposed	1.9750×10^{-20}
Ref. [8]	0.04
Ref. [10]	0.002
Ref. [11]	9.3997×10^{-15}
Ref. [16]	0.0019
Ref. [19]	0.0016

7 Conclusions and Suggestions for Future Work

This paper presented a robust proposed algorithm for cancelable biometric verification. The proposed algorithm adopts two cascaded encryption stages. The speech signal is the proposed algorithm input. First, the speech signal is transformed into its spectrogram. Next, the spectrogram is encrypted using OSH. Finally, the OSH spectrogram is encrypted again using DRPE with two randomly generated RPMs to get the cancelable templates. To verify the authorized users access, the correlation between the two cancelable templates of the stored and test users is estimated and compared with a threshold value. The two encryption stages are adopted using their software implementation in MATLAB. The proposed algorithm can generate non-invertible cancelable templates and high efficiency even with noise. This is clear from the evaluation metric values. The proposed algorithm achieves a trade-off between the two requirements of cancelable biometric systems with a high level of security and high verification performance. In the future work, we will work on iris images to generate RPMs for DRPE encryption. Deep learning can be exploited in feature extraction from spectrograms, training, and testing although it needs a large dataset and resources with high capabilities.

Acknowledgement: The authors would like to acknowledge the support received from Taif University Researchers Supporting Project Number (TURSP-2020/147), Taif University, Taif, Saudi Arabia.

Funding Statement: This work was funded and supported by the Taif University Researchers Supporting Project Number (TURSP-2020/147), Taif University, Taif, Saudi Arabia.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] K. Jain, A. Ross and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 14, no. 1, pp. 4–20, 2004.

[2] Y. Chung, D. Moon, K. Moon and S. Pan, "Hiding biometric data for secure transmission," in *Proc. of the Int. Conf. on Knowledge-Based and Intelligent Information and Engineering Systems*, Springer, Berlin, Heidelberg, pp. 1049–1057, 2005.

[3] S. Rane, Y. Wang, S. Draper and P. Ishwar, "Secure biometrics: Concepts authentication architectures and challenges," *IEEE Signal Processing Magazine,* vol. 30, no. 5, pp. 51–64, 2013.

[4] V. Patel and R. Chellappa, "Sparse representations compressive sensing and dictionaries for pattern recognition," in *Proc. of the IEEE First Asian Conf. on Pattern Recognition*, Beijing, China, pp. 325–329, 2011.

[5] E. Tarek, O. Ouda and A. Atwan, “Image-based multimodal biometric authentication using double random phase encoding,” *International Journal of Network Security,* vol. 20*,* no. 6*,* pp. 1163–1174, 2018.

[6] R. Soliman, G. El Banby, A. Algarni, M. Elsheikh, N. Soliman *et al.,* “Double random phase encoding for cancelable face and iris recognition,” *Applied Optics,* vol. 57*,* no. 2*,* pp. 10305–10316, 2018.

[7] D. Rachapalli and H. K. Kalluri, “Multimodal biometric template protection using color QR code,” *International Journal of Recent Technology and Engineering (IJRTE),* vol. 7*,* pp. 7–11, 2019.

[8] T. Sudhakar and M. Gavrilova, “Cancelable biometrics using deep learning as a cloud service,” *in IEEE Access,* vol. 8*,* pp. 112932–112943, 2020.

[9] E. Tarif, S. Wibowo, S. Wasimi and A. Tareef, “A hybrid encryption/hiding method for secure transmission of biometric data in multimodal authentication system,” *Multimedia Tools and Applications,* vol. 77*,* no. 5*,* pp. 2485–2503, 2018.

[10] W. Yang, S. Wang, G. Zheng, J. Chaudhry and C. Valli, “ECB4CI: An enhanced cancelable biometric system for securing critical infrastructures,” *Journal of Supercomputing,* vol. 74*,* no. 4*,* pp. 4893–4909, 2018.

[11] L. Abouelazm, S. Ibrahim, M. Egila, H. Shawky, M. Elsaid *et al.,* “Cancelable face and fingerprint recognition based on the 3d jigsaw transform and optical encryption,” *Multimedia Tools and Applications,* vol. 79*,* no. 9*,* pp. 14053–14078, 2020.

[12] M. Kareem, A. Saleeb, S. M. El-Dolil, A. El-Fishawy, F. E. Abd El-Samie *et al.,* “Efficient comb-based filter for cancelable speaker identification system,” in *Proc. Int. Conf. on Electronic Engineering (ICEEM)*, Menouf, Egypt, pp. 1–7, 2021.

[13] M. G. Barrero, C. Rathgeb, G. Li, R. Ramachandra, J. Galbally *et al.*, “Multi-biometric template protection based on bloom filters,” *Information Fusion,* vol. 42*,* pp. 37–50, 2018.

[14] A. Mostafa, N. F. Soliman, M. Abdalluh and F. E. Abd El-Samie, “Effect of voice features cancellation in speaker identification system,” in *Proc. Fourth Int. Japan-Egypt Conf. on Electronics, Communications and Computers (JEC-ECC)*, Alexandria, Egypt, pp. 139–142, 2016.

[15] O. Ouda, K. Nandakumar and A. Ross, “Cancelable biometrics vault: a secure key-binding biometric cryptosystem based on chaffing and winnowing,” in *25th Int. Conf. on Pattern Recognition (ICPR)*, Cairo, Egypt, pp. 8735–8742, 2021.

[16] A. Alarifi, M. Amoon, M. H. Aly and W. El-Shafai, “Optical PTFT asymmetric cryptosystem-based secure and efficient cancelable biometric recognition system,” *IEEE Access,* vol. 8*,* pp. 221246–221268, 2020.

[17] H. Kim and S. Y. Chun, “Cancelable ECG biometrics using compressive sensing-generalized likelihood ratio test,” *IEEE Access,* vol. 4, pp. 9232–9242, 2019.

[18] H. Kaur and P. Khanna, “Privacy preserving remote multi-server biometric authentication using cancelable biometrics and secret sharing,” *Future Generation Computer Systems,* vol. 102*,* pp. 30–41, 2020.

[19] K. K. Chee, Z. Jin, D. Cai, M. Li, W. S. Yap *et al.*, “Cancellable speech template via random binary orthogonal matrices projection hashing,” *Pattern Recognition,* vol. 76*,* pp. 273–287, 2018.

[20] E. Abdellatef, N. Ismail, S. Abd Elrahman, K. Ismail, M. Rihan *et al.,* “Cancelable multi-biometric recognition system based on deep learning,” *Journal of Visual Computer,* vol. 36*,* no. 7*,* pp. 1–13, 2019.

[21] Y. Jang and N. Cho, “Deep face image retrieval for cancelable biometric authentication,” in *Proc. IEEE Int. Conf. on Advanced Video and Signal Based Surveillance (AVSS)*, Taipei, Taiwan, pp. 1–8, 2019.

[22] T. Poon and T. Chung, “Digital holography and three-dimensional display: Principles and applications,” *New York: Springer Science-Business Media,* vol. 3*,* no. 1*,* pp. 1–14, 2006.

[23] L. Z. Zhang, X. Zhou, D. Wang, N. N. Li, X. Bai *et al.,* “Multiple-image encryption based on optical scanning holography using orthogonal compressive sensing and random phase mask,” *Opt. Eng.,* vol. 59*,* no. 10*,* pp. 102411, 2020.

[24] T. Poon, “On the fundamentals of optical scanning holography,”*American Journal of Physics,* vol. 76*,* no. 2*,* pp. 738–745, 2008.

[25] T. Poon, “*Optical Scanning Holography with MATLAB,*” New York: Springer, 2007.

[26] M. Al-Bermani and W. Al Aaraje, “Optical scanning holography (OSH),” *Journal of Kufa-Physics,* vol. 4*,* no. 3*,* pp. 37–47, 2012.

[27] R. Soliman, G. El Banby, A. Algarni, M. Elsheikh, N. Soliman *et al.,* “Double random phase encoding for cancelable face and iris recognition,” *Applied Optics,* vol. 57*,* pp. 10305–10316, 2018.

[28] A. Elshamy, F. Abd El-Samie, O. Faragallah, E. Elshamy, H. El-sayed *et al.,* “Optical image cryptosystem using double random phase encoding and arnold's cat map,” *Optical and Quantum Electronics,* vol. 212*,* no. 1*,* pp. 1–15, 2016.

[29] S. Fulop and K. Fitz, “Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications,” *The Journal of Acoustical Society of America,* vol. 119*,* no. 2*,* pp. 360–371, 2006.

[30] S. Wang, A. Sekey and A. Gersho, “An objective measure for predicting subjective quality of speech coders,” *IEEE Journal on Selected Areas in Communications,* vol. 10*,* pp. 819–829, 1992.

[31] S. Sree and N. Radha, “Cancelable multimodal biometric user authentication system with fuzzy vault,” in *Proc. IEEE Int. Conf. on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, pp. 1–6, 2016.

[32] T. Dang, Q. Truong, T. Le and H. Truong, “Cancelable fuzzy vault with periodic transformation for biometric template protection,” *IET Biometrics,* vol. 5*,* no. 8*,* pp. 229–35, 2016.

[33] P. Kumar, J. Joseph and K. Singh, “Optical image encryption using a jigsaw transform for silhouette removal in interference-based methods and decryption with a single spatial light modulator,” *Applied Optics,* vol. 50*,* no. 2*,* pp. 1805–1811, 2011.