Multimodal Biometric System Using Dual Digital Watermarking

Abdulmawla Najih*, [1,4], Salem s.m Khalifa[2], Salem Enajeh[3], Nabila Albannai[4]

[1*,4]Department of computer engineering, The High Institute of Science & Technology Gharian
[2]Department of computer engineering College of Science & Technology Alriyayna
[3]Department of computer engineering, The High Institute of Science & Technology Tripoli
nabdulmawla@gmail.com

الملخص

يعتمد نظام القياسات الحيوية متعدد الوسائط على معرفات بيومترية متعددة لتحديد هوية الشخص الذي أصبح أكثر شيوعًا في المصادقة وتحديد الهوية. نظرًا لشعبية النظام متعدد الوسائط، يتم أخذ ثلاثة عوامل (الأمان والمصادقة والمتانة) في الاعتبار عند مصادقة الصورة والتعرف على الوجه وتحديد الهوية البيومترية. يعرض هذا البحث النظام البيومتري متعدد الوسائط من خلال تطبيق نظام العلامة المائية الرقمية المزدوجة. تقوم العلامة المائية المزدوجة المقترحة بدمج علامات مائية عمياء وشبه هشة وقوية في صورة الوجه. لتضمين علامة مائية رقمية عمياء شبه هشة، يُقترح(IWT تحويل المويجات الصحيحة) والإجراء العكسي، وبالنسبة للعلامات المائية الرقمية القوية، يُقترح DCT (تحويل كونتورليت منفصل) و QIM(تعديل مؤشر القياس الكمي). في هذا البحث، تم استخدام اثنين من القياسات الحيوية بما في ذلك الوجه والصوت لنظام القياسات الحيوية. بالنسبة لتضمين العلامة المائية، يتم استخراج الميزات الصوتية لـ MFCC معاملات التردد (Cepstral) وتضمينها في صورة الوجه. يتم تنفيذ عملية اختيار الوجه باستخدام خوارزمية ICP (النقطة التكرارية الأقرب) التي تعمل على أساس أوزان التعلم. يتم اتخاذ القرار النهائي باستخدام التعلم المعزز العميق المسمى-"Double Deep-Q Network".يتم استخدام مجموعتين مثلTIMIT (قاعدة بيانات الصوت) و ORL(قاعدة بيانات الوجه) لتقييم النظام واختبار الأداء. يُظهر نظام العلامات المائية المزدوجة المقترح أداءً أفضل من حيث الدقة، ونسبة كفاءة الطاقة EER ، وPSNR، وSSIM.

Abstract:

Multimodal biometric system relies on multiple biometric identifiers for person identification that becomes more popular in authentication and identification. Due to the popularity of multimodal system, three factors (security, authentication and robustness) are considered for image authentication, face recognition and biometric identification. This paper presents multimodal biometric system by applying double digital watermarking scheme. The proposed dual watermarking is embedding blind semi-fragile and robust watermarks into the facial image. To embed blind semi-fragile digital watermarking, IWT (Integer Wavelet Transform) and Reversible Procedure is proposed and for robust digital watermarks, DCT (Discrete Contourlet Transform) and QIM (Quantization Index Modulation) is proposed. In this, we have used two biometrics including face and voice for biometric system. For watermark embedding, MFCC (Mel Frequency Cepstral Coefficients) voice features are extracted and embedding into facial image. Face selection operation is performed using ICP (Iterative Closest Point) algorithm that works based on learning weights. Final decision making is performed using Deep Reinforcement Learning called “Double Deep-Q-Network”. Two corpuses such as TIMIT (voice) and ORL (Face) are used for system evaluation and performance testing. our proposed double watermarking scheme exhibits better performance in terms of accuracy, EER, PSNR, SSIM,

Keywords: Multimodal, Double watermarking, Face recognition, Discrete Contourlet Transformation (DCT), Quantization Index Modulation (QIM), Integer Wavelet Transform (IWT) and Reversible Procedure**.**

i. INTRODUCTION

The rapid growth in Information Technology (IT) has revised the need to protect sensitive and personal data from any unauthorized access. Many techniques have been proposed to protect these data, such as the knowledge-based method, where login credentials, such as passwords, Personal Identification Number (PIN) or patterns, are required from the users to access these data. However, the importance of protecting these data and the sensitivity of such methods to simple attacks, such as shoulder surfing, have imposed the need for more secure techniques [9]. Therefore, many methods have been proposed to protect these techniques from known attacks but the tendency of humans to use easy-to-remember credentials has limited the capabilities of such techniques, as easy-to-remember credentials are also easy-to-predicting [6].

To protect such data, and according to the limited security that Knowledge-based techniques provide, many techniques have been proposed based on biometric information. This information is collected from the user upon authentication and compared to the information of the legitimate users who are allowed to access the system. The user of such information has shown better resistance against attacks that rely on identifying the information used for authentication, as it is more complex to replicate biometric features, than the traditional methods, such as passwords or patterns, in Knowledge-based authentication [10 ,13 ,16].

One of the widely used techniques to protect the authenticity of information communicated between different parts of the authentication system is digital watermarking, where biometric information is added to the captured biometric image that the biometric features are extracted from. Watermarking techniques are normally used for one of two reasons, which are to prove ownership of the biometric image or to detect any tampering with it. Moreover, some of the watermarking techniques add visible watermarks to the biometric image, while others hide the watermark inside the biometric image, so that, it is not visible unless it is extracted. For tamper detection, hidden watermarks are added to the biometric images, so that, the absence of the watermark or any distortion in the extracted watermark indicates tampering with the original biometric image. Moreover, the watermarking techniques used for tamper detection is fragile, so that, the extracted watermark is highly affected when any attack is executed against the watermarked biometric image [5].

In biometric systems, researchers [4, 14, 15, 17,18, and 19] have proposed digital watermarking algorithms against geometrical attacks and other attacks. Watermarking is referred to as hidden information and protect against unauthorized persons. To hidden data, certain information is considered and to be embedded on original host image using any water- marking techniques. In this case, original image content could not affect. Some of the researchers [1, 2, 3, 7, 11, 28, 29, 21, 24, 25, 26, and 27] have proposed frequency domain methods for watermarking while others proposed spatial domain methods [12 and 31]. Recently, researchers have concentrated on double watermarking methods. For each watermarking, they have used separate techniques. Table 1 demonstrates that recently used watermarking techniques. To solve those limitations, in this paper we propose a hybrid double watermarking method to build an online multimodal system by embedding MFCC voice features into face images. The main objective of our multimodal biometric watermarking system is to provide security to biometric data without conceding the quality of both biometric host image and watermark data. The main contributions of our work are as follows:

- DGA-SA (Dynamic Genetic Algorithm with Simulated Annealing). In watermarking, spatial location and best embedding point are required for efficient watermarking. So that DGA- SA is proposed DGA to compute the spatial location of the coefficients and then this suboptimal solution is forwarded to Simulated Annealing for best embedding selection.
- ICP (Iterative Closest Point) for face selection. Face selection is based on the patches. In this face image is split into patches (eye, nose, mouth). In this ICP algorithm the face selection in which closest points selected that will be used for classification.
- Deep Reinforcement Learning algorithm named “Double Deep-Q-Network” which is the recent network, this reduces the complexity and it will be used for face recognition.

The paper is organized as follows: Section 2 is a literature review; Section 3 describes the proposed methodology in three parts; biometric data processing, double digital watermarking and face identification; Section 4 presented database description and experimental results; and Section 5 concludes on the research study carried out.

ii. LITERATURE REVIEW

There is wide collection of approaches have been proposed for watermarking. Almost of them adopt the idea of digital watermarking using transformation methods DWT (Discrete Wavelet Transform), DCT (Discrete Cosine Transform), and DFT

(Discrete Fourier Transform). Some of these transformation methods are briefly reviewed in the following:

The authors of [14] proposed single biometric identifier/unimodal identifier for face recognition using PCA (Principle Component Analysis) and DCT (Discrete Contourlet Transformation). This combined system is used to ensure the image authenticity and security against robust attacks. Experimentation is proved that the system is successful in security, accuracy, robustness and watermarking techniques. The papers [20, 22, and 23] proposed an approach for multimodal biometric authentication system using watermarking technique. Three biometric traits are postulates in this paper are fingerprint, face (physical traits) and signature (behavioral trait). Initially, metamorphose is applied on biometric traits using Discrete Cosine and Discrete Wavelet Transformation. After this watermarking is achieved using singular value decomposition scheme. However, the security of biometric data is preeminent, therefore utmost importance to provide security for individual biometric data. The authors of [3] proposed Phase Congruency Model and DCT for personal identification system. Phase Congruency model is employed the low frequency on DCT coefficients of face image and normalization correlation is based on robust property and human perceptivity. Here authors claims that improve the quality, robustness and recognition accuracy against different types of image processing attacks. The authors of [7] described robust and secure watermarking for biometric data protection. This scheme is proposed for the biometric template protection. Communication channel is used for biometric system and the sparse information of watermark biometric data is generated based on the compressive sensing and wavelet coefficients. The proposed scheme is attaining the best results in accuracy and robustness. The authors of [8] introduced face detection methods for digital image authentication. Biometric watermarking technology is authenticating digital images automatically, in such processing authors proposed face detection methods such as principal component analysis, and eigen feature regularization. The objective of this paper is to find the relationship between the original data and extracted biometric data using neural networks. Authors obtained very promising results in experimentation. The authors of [3] proposed feature based 3-level RDWT [Redundant Discrete Wavelet Transform] for multimodal biometric system. Phase congruency model is used to compute the embedding locations which preserves the facial features from being watermarked and ensures that the face recognition accuracy. In order to improve the performance of proposed watermarking algorithm, an author uses adaptive user-specific watermarking parameters. 3-level wavelet decomposition of a face image is divided to four subbands such that the size of each subband is equal to the original image, the RDWT redundant space provides additional locations for watermark embedding. To address the issues of 3-level RDWT. The authors of [11] presented multimodal biometric identification system using L-level RDWT decomposition. Two biometric traits of the user i.e. the facial and iris features are embedded independently into the wavelet sub-bands. While using the fused score for evaluation, the accuracy was increased. The robustness of the system has been analyzed against various attacks and the verification accuracy evaluated based on false acceptance rate, area under curve and false rejection rate. The wide range of existing approaches uses single watermarking for authentication and face recognition. Single watermarking does not meet security requirements. To mitigate this, current researchers concentrated on dual/double watermarking [30, 31]. these approaches are not efficient for multimodal biometric system since its aim to satisfy the security requirements. To solve those problems, we designed effective multimodal biometric system which meets those requirements to all extent.

iii. System Overview:

To overcome the aforesaid problems and the best of our knowledge, Double Digital Watermarking for Online Multimodal Biometric System is proposed. The proposed double digital watermarking algorithm is a combination of blind semi fragile and robust watermarks. In blind semi-fragile digital watermarking, Integer Wavelet Transform (IWT) and Reversible procedure is performed. For blind robust digital watermarking, Discrete Contourlet Transform (DCT) and Quantization Index Modulation (QIM) are used. The proposed scheme consists of seven phases: Preprocessing, voice feature (MFCC) extraction, face selection, watermark embedding, extraction, authentication and identification. In first, user face and voice biometric traits are preprocessed like image refinement and normalization. In DCT, image is decomposed with five levels 1, 2, 4,8,16 under the directional band. To identify the best embedding point, we proposed Nature Inspired Algorithms called

DGA-SA (Dynamic Genetic Algorithm with Simulated Annealing). In voice data, Voice information segmented and then MFCC (Mel Frequency Cepstral Coefficients) features extracted and applied for watermark embedding. Before that, face selection is performed using ICP (Iterative Closest Point) Algorithm. In this learning weights computed and less weight will be discarded for online face recognition system. After watermark embedding, information transfer to Multimodal biometric system through communication channel. In this, watermark is extracted using Deep Reinforcement Learning algorithm named “Double Deep-Q-Network”.

This algorithm reduces time and improves the system efficiency. TIMIT and ORL corpuses will be used for evaluating the developed systems. Similarity matching is calculated for each training sample with testing sample and the Euclidean Distance formula is used to calculate the distance between training and testing samples.

A. Preprocessing

The first step of the proposed multimodal biometric system is pre-processing. This process makes the input images better suitable for the subsequent steps. The most important processes such as refinement and normalization are carried out under preprocessing. Also, this process is helpful for the feature extraction. Here, the facial images from the database are refined and normalized. Once the input images are acquired from the database, it undergoes the refinement process after refinement we move to normalization steps in which input image is converting into the range of pixels. In parallel, user voice input can be processed in the system. Voice features are used as a watermark to embed in face image. Segmenting voice sample is a crucial task in recognition. In this we segment voice to embed in the facial image.

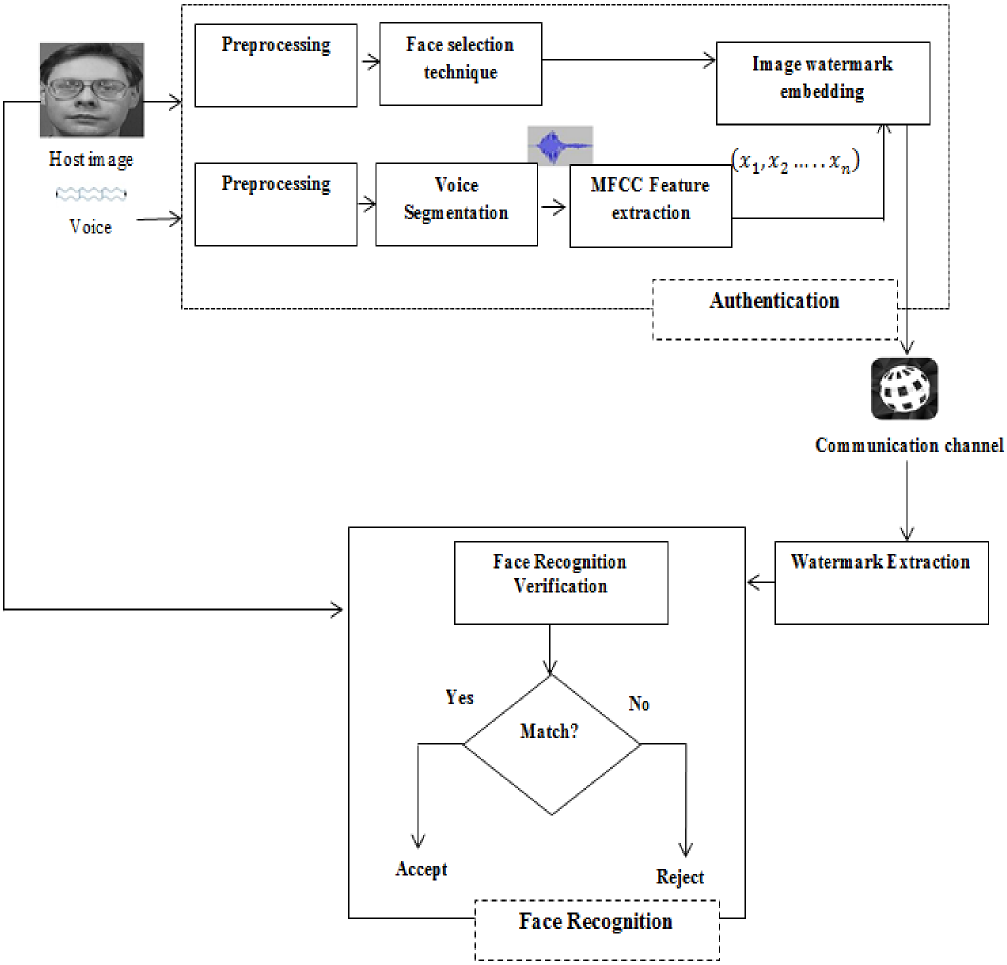

Figure 1: Overall Architecture

B. Voice Feature Extraction:

Voice samples are taken as the input for watermark embedding. Before embedding process, MFCC (Mel Frequency Cepstral Coefficients) features should extract from the voice samples. MFCC is based on cepstral representation and the speech signal spectrum and it is based on auditory processing. Most powerful feature of cepstrum is that repeated patterns or any periodicities. Generally, auditory processing system does not perceive any spectral components in linear scale, but it will perceive spectral components on a nonlinear scale.

The MFCC is defined as the short time power spectrum of a speech signal and the Mel scale wraps the frequency and allows better representation similar to the human auditory system. The Mel scale is defined as the mapping of frequency doubling to a human perception scale. The relation between Mel scale and Hertz scale is

$$f\,(\text{Mel}) = 2595\,\log_{10}\left(1+\frac{f(Hertz)}{7}\right) \qquad (1)$$

MFCC is reduced dimensional form of speech spectrum. In order to compute MFCC, the whole range of audio frequency is divided into frequency bands and also the energy of speech signal with in the band is computed. In MFCC, cepstral coefficients are obtained by computing the log of energy in each band.

C. Face Selection:

In face selection, person face is partitioned into number of patches: eyes, nose, mouth etc. The RGB face color space consists of three additive primary components: Red, Green and Blue. These color components are highly correlated and this will difficult to execute in some image processing algorithms. Many processing techniques are introduced and work on the face partition. In doing so, it produces poor results in all three color spaces. To mitigate this, we proposed Iterative Closest Point Selection algorithm. This algorithm performed with much ease on an image in the RGB Color space.

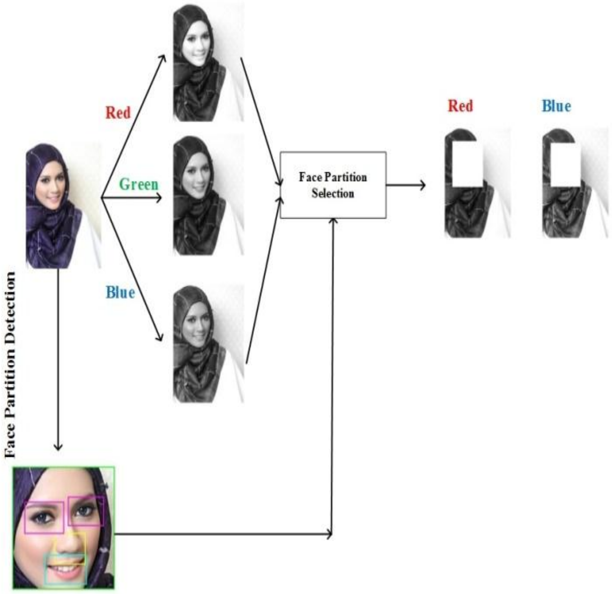

Figure 2: Face selection process

Consider the training samples including N faces (RGB color space). We divide each image into M Overlapping P*P patches. Each of them by a d-dimensional vector where d=P^2. Let i be the certain point and its corresponding closest point j is computed using distance formula. Learning weights $(w_{i1}, \dots \dots w_{ik})$ evaluate for the point (i, j). In this, the points which have the minimum weights have been eliminated in the system.

D. Watermarking

Digital watermarking in multimodal system is defined as the process of embeds biometric data (watermark) into another biometric image Further, watermark can be detected/extracted to make an assertion of the image. the host image denotes by H, the watermark W, the watermarked image by B and the extracted watermark w', N can be considered as attacks or noise and the E (H, W) and D(C) be the embedding and extracting function respectively.

the spatial location is computed by the following equation

$$\boldsymbol{I}\big(\boldsymbol{a}(\boldsymbol{r},\boldsymbol{\theta}),\boldsymbol{b}(\boldsymbol{r},\boldsymbol{\theta})\big) \rightarrow \boldsymbol{I}(\boldsymbol{r},\boldsymbol{\theta}) \qquad (\mathbf{2})$$

Where I represent the image, $a(r,\theta), b(r,\theta)$ are the coordinates of a, b of the original image and r is the radius lies on the interval [0, 1] and θ is angle between [0, 2π].

double watermarking is proposed using hybrid approaches. Here fragile watermarking for tampering detection and robust watermarking for transmitting the MFCC features to the system**.** The embedding and extraction of watermarking are briefly described in the following sections.

- **Nature Inspired Algorithms (DGA with SA):**

Searching best embedding co-efficient in image is a quite challenging and essential task. Nature inspired evolutionary searching algorithms introduced to select the best embedding point. In this scheme, we proposed Dynamic Genetic with Simulated Annealing for watermark embedding. As a result of this, the capacity is minimized by embedding the selected co-efficient. Dynamic Genetic Algorithm is an evolutionary based searching method approach that consists of natural genetics to solve the global optimization problems. Natural genetics are using operators such as mutation, crossover and selection for optimization. Chromosomes represent the candidate solutions with assigned the score to each of them. By applying genetic operators to chromosomes, we get the new offspring. After the generations of new offspring, chromosomes have better score values that considered being a sub-optimal solution. After applying DGA, SA is started. Simulated Annealing is another evolutionary based search method to find the best solutions. In mathematical side, SA initiated from the starting point and followed by next new candidate solutions is randomly generated. At this point, SA has taken long computational time. To avoid such issues, we combined Dynamic Genetic Algorithm and Simulated Annealing. At the starting point, parameters are initialized such as population size, number of variables, lower crossover and mutation rates, lower and upper bounds for each variable, annealing, selection method and temperature functions are defined. Then starts GA algorithm and stopping criteria are defined as a certain number of generations. At the end of this

Genetic algorithm, sub-optimal solution is generated and then SA is employed with the initial solution from the first part of the algorithm. Flowchart of the DG with Simulated Annealing is illustrated below.

Algorithm for Dynamic Genetic with Simulated Annealing

```
//Dynamic Genetic Algorithm
Step 1: Initialize the parameters N (population size)
Step 2: Evaluate the chromosomes
Step 3: Select the best chromosomes
Step 4: Repeat //Adaptive Determination of genetic parameters
1. Determine $P_c$ (Crossover Probability) and $P_m$ (Mutation probability), and
2. Perform crossover and mutation
3. Evaluate the chromosomes
Step 5: If average fitness of population is above the fitness average //user defined threshold (T)
{Return the solutions
} Else
Mutate top e solutions in the current population
4.        Iterate the step 2 through 5
Output: Sub-optimal solution
//Simulated Annealing
Step 1: Parameters initialization: Initial solution S, Iterations K=1…L of each T, and Initial Temperature T
Step 2: Update T for each iterations
Step 3: Calculate the increment $\Delta E = E(S') - E(S)$ where E(S) is the evaluation function
```

Step 4: Criterion for Metropolis: If ΔE is greater than 0 then accept S' as the new solution otherwise S' as the new solution with the acceptable probability P ($e^{\Delta E}/T$)

Step 5: If the termination condition is satisfied, the current optimal solution as the output. After that, terminate the program.

Output: Optimal solution (F^{Best})

- **Blind Semi-fragile watermarking**

Blind semi-fragile watermarking can be performed using Integer Wavelet Transform and it can be invisible for users. It has a shift invariant behavior and the errors of reconstruction will decrease after embedding in cover image. As a results of the robustness of the watermarking method will increase.

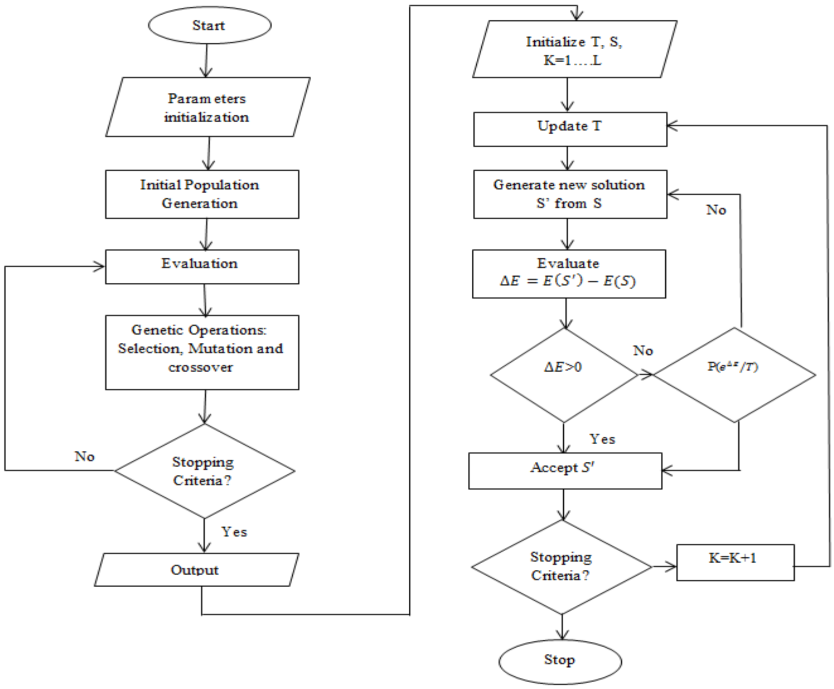
Start
Param eters initialization
Initial Population Generation
Evaluation
Genetic Operations: Selection, Mutation and crossover
Stopping Criteria?
No
Yes
Output
Initialize T, S, K=1….L
Update T
Generate new solution S' from S
No
Evaluate $\Delta E = E(S') - E(S)$
ΔE>0
No
P($e^{\Delta E}/T$)
Yes
Accept S'
Stopping Criteria?
K=K+1
Stop

Figure 3: Flowchart for DGA with SA

Algorithm: IWT based blind semi-fragile watermark embedding algorithm

```
Input: Host image and MFCC features $(x_1, x_2, \dots x_n)$
Output: Watermarked image
Begin
  Read the image and convert it to gray scale
  Decompose the image using IWT
  Using the 5th bit, choose the subbands
  Compress 5th bit data using arithmetic coding
  Compressed data insertion and the watermark features into the host image
  Compute IIWT (Inverse-IWT) to get the watermarked image
  End.
```

- **Robust Watermarking**

For robust watermarking, we select discrete contourlet transformation for watermark embedding because it captures the directional edges and smooth contours from the image and its better than the other conventional transforms namely DWT, DCT and DFT. Human Visual system is less sensitive to the image edges and watermarking is applied on the contourlet domain, which denotes image edges. It improved robustness due to selecting the embedding point using algorithm () and optimum control of its quantization scalar factor. However, the perceptibility of the watermarked image degrades. In our scheme, QIM (Quantized Index Modulation) is applied all sub bands when we reconstruct the watermarked image. Because the novel arrangements of the subbands preserve the robustness, so the proposed scheme is highly robust against various low-frequency attacks.

- **Discrete Contourlet Transform**:

Discrete Contourlet Transformation is a new enhanced image decomposition method to embed the watermark effectively. It outperforms than the DWT, RDWT and DCT. Contourlet transformation is categorized into two phases: LP (Laplacian Pyramid) decomposition and DFB (Directional Filter Banks). In DCT, the number of directional subbands at each level is set to 2^n where n is a positive integer number where n=1, 2,3,4,5 then we get the 1, 2, 4,8,16 subbands as shown in figure 4. The Energy of subband S (i,j) is computed by the following expressions:

$$E = \sum_i \sum_j |S(i,j)|^2 \qquad (3)$$

- **Quantization Index Modulation**:

QIM is a class of embedding methods, termed the quantization index modulation technique. It achieves efficient tradeoffs among information-embedding capacity and robustness of embedding. A particular quantizer is chosen from a set of possible quantizers by using watermark information as an index and then applied to the host information to embed the watermark. Assume that one bit s ∈ {0, 1} is to be embedded and m denotes the host signal. Two qunatizers will be considered and generated Q_i (m), where i=0, 1. Watermark bit identifies the selection of the quantizer Q_i (m) with a step size Δ, which can be computed as follows:

$$Q_i(m)=Q(m-d_i)+d_i,\ i=0,1$$

Where Q(m)= Δ × Round (m/Δ),

$d_i(d_0)= -\Delta/4$ and $d_i(d_1)= \Delta/4$ round (num) rounds num to the nearest integer.

$$Q_i(m') \in (Q_0, Q_1)$$

The watermarked signal value $m^{,}$ computed using two quantizers (Q_0) and (Q_1) using the following expression:

$$m' = \begin{cases} Q_0(m), & s = 0 \\ Q_1(m), & s = 1 \end{cases}$$

In the watermark extraction process, the S’ can be extracted from the signal of m” by resolving the optimization problem.

$$S' = arg \min_{s\in\{0,1\}} \|m'' - Q_S(m'')\| \qquad (4)$$

- **Watermark Embedding:**

In the proposed scheme, the watermark embedding can be formulated in the following process:

Step 1: The host image H is transformed into the contourlet domain. The lowpass subbands of the coefficients of the host image are selected to embed the watermark. $H_{S,D}$(i,j) is the host image contourlet coefficients where S is the resolution scale and D is the frequency direction. We embed the watermark in the best embedding point because best embedding point is improving the performance of security.

Step 2: Apply QIM on each subbands of the host image.

Step 3: Use quantizers to produce watermark embed information for both bits 0 and 1.

Step 4: Before embed the watermark, strength of the watermark embedding(α) is derived from the visibility function. It is expressed by:

$$\text{VF} = \frac{1}{1+\sigma_1^2(a,b)} \qquad (5)$$

Where $\sigma_1^2(a,b)$ is embedding regions local variance and for every local embedding region, the watermark embedding strength is adaptively modulated. Then the adaptive embedding strength is derived from the following strength

$$\alpha = (1 - \text{NF}).\beta \qquad (6)$$

Where β is represents the predefined embedding strength for local watermark regions.

Step 5: After that, MFCC features sequence $(x_1, x_2, \dots x_n)$ to the number of best selected contourlet coefficients points using the following function:

$$H'_{S,D}(i,j) = \begin{cases} \frac{W.\alpha}{F^{Best}(i,j)}.\ H_{S,D}(i,j),\ F^{Best}(i,j) \neq 0 \\ H_{S,D}(i,j), \qquad F^{\,Best}(i,j) = 0 \end{cases} \qquad (7)$$

Where $H'_{S,D}(i,j)$ and $H_{S,D}(i,j)$ denotes original and watermarked coefficients respectively. F^{Best}(i,j) is the best embedding point of the host image H.

Step 6: Apply IDCT on the modified QIM contourlet coefficients to obtain the watermarked image.

Step 7: Obtain the watermarked image H'

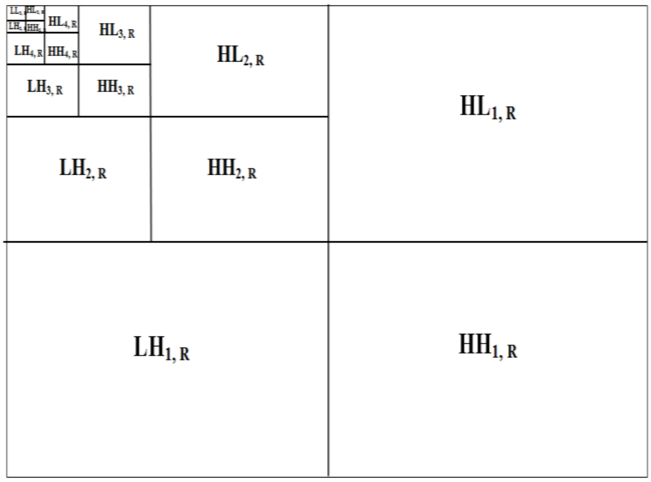

Figure 4: DCT Decomposition with 5-levels

Algorithm: DCT with QIM based Robust Watermark embedding algorithm

Input: Host image, MFCC features $(x_1, x_2, \dots x_n)$ and F^{Best}

Output: watermark image

Begin

Read the watermark image and apply DCT

Decompose the image into 5-levels $\{LL_{1,R}, \dots\dots LL_{5,R}\}$,$\{HL_{1,R}, \dots\dots HL_{5,R}\}$ $\{HH_{1,R}, \dots\dots HH_{5,R}\}$, $\{LH_{1,R}, \dots\dots LH_{5,R}\}$ using DCT

```
Choose lowpass subbands from the decompose image
$F^{Best}$ is selected using DGA-SA
Apply QIM on each subbands
Inserting $(x_1, x_2, ... x_n)$ into watermark image
Arrange Quantize levels
Compute IDCT (Inverse-DCT) to get the watermarked image
End
```

- **Watermark Extraction:**

The steps of watermark extraction process are the same steps of the process of watermark embedding. Watermark extraction are performed as follows: First, detect the voice segments in watermarked signal and then extract the watermark information from the embedding locations depending on the best embedding locations that generated using DG with SA as in the watermark embedding.

E. Face Recognition:

The process of face recognition is carried out by using Double-Deep-Q-Network. Consider the classification in multimodal biometric decision. Suppose we have N learning samples $\{a\}_{i=1}^{N} \in R^{l}$ and corresponding class labels $b_i \in \{0, 1\}$ where 0 represents an impostor and 1 represents genuine person. Double Deep-Q-Network is the reinforcement algorithm that newly created in neural networks. The major advantage of this algorithm is to generalize learning across actions without changing of reinforcement algorithm. Value functions are considered here. First we can describe about Deep-Q Network. There are two key aspects of networks are considered that are model free in sense by the environment and experience replay. The network parameters are Q (m, x, θ) with parameters of θ. To evaluate this network, loss functions must be optimized for n iterations.

$$L_n(\theta_n) = E_{m,x,r,m'}\left[\left(y_n^{DQN} - Q(m,x,\theta_n)\right)^2\right] \quad (8)$$

Where y_n^{DQN}= r + γ $\max_{x'} Q(m',x',\bar{\theta})$

Where $\bar{\theta}$ denotes the parameters of a separate and fixed target network. In online, we use the Q-Learning for the network parameters$Q(m,x,\theta_n)$. Specific gradient is used to update the parameters of network.

$$\nabla_{\theta_n} L_n(\theta_n) = E_{m,x,r,m'}\left[\left(y_n^{DQN} - Q(m,x,\theta_n)\right) \nabla_{\theta_n} Q(m,x,\theta_n)\right] \quad (9)$$

In learning, database D of experiences (m_t, x_t, r_t, m_{t+1}) from many occurrences. During training, the current experience is prescribed from D uniformly at random manner.

$$L_n(\theta_n) = E_{m,x,r,m'\sim u(D)}\left[\left(y_n^{DQN} - Q(m,x,\theta_n)\right)^2\right] \quad (10)$$

In DQN Q-learning algorithm, select and evaluate action gives the same values while max operator uses these same values for solving optimization problem. This will lead to the overoptimistic problem. To overcome this, Double Deep-Q-Network is proposed

$$y_n^{DQN} = r + \gamma\ Q(m', \underset{x'}{\arg max}\ Q(m',x',\theta_n);\ \bar{\theta}) \quad (11)$$

F. Euclidean Distance Measure*:*

To evaluate the similarity or the dissimilarity between the two images, Euclidean Distance is used.

Similarity matching for Face($S_{v(i,j)}$):

$$S_{f(i,j)} = 1\text{-d}\left(u_i, v_j\right) \quad (12)$$

Where u_i training samples and v_j is the testing sample and d$\left(u_i, v_j\right)$ is the Euclidean Distance between the training and testing sample and it is evaluated by the following expression:

$$\mathrm{d}\big(u_i, v_j\big)=\sqrt{\textstyle\sum_{Z}^{M=1}|u_z - v_z|} \qquad (13)$$

Where M is the total number of samples in database and u_z, v_z is the training and testing samples.

Similarity Matching for Voice($S_{v(i,j)}$):

$$S_{v(i,j)}= 1\text{-d}\big(u_i, v_j\big) \qquad (14)$$

Where u_i training samples and v_j is the testing sample and $\mathrm{d}\big(u_i, v_j\big)$ is the Euclidean Distance between the training and testing sample and it is evaluated by the following expression:

$$\mathrm{d}\big(u_i, v_j\big)=\sqrt{\textstyle\sum_{Z}^{R=1}|u_z - v_z|} \qquad (15)$$

Where R is the total number of samples in database and u_z, v_z is the training and testing samples.

G. **Score Fusion Method***:*

In our proposed system, score fusion method is carried out to combine each biometric score since this method is easy to access and combine the scores of different biometric modalities. Fusion at match score level uses sum method. Let S: R^1→ R is the hypothesis function mapping these pattern features onto a scalar measure for decision inference. $\text{Fusion}_{\text{score}}$ produces a continuous output then the output must be threshold in order to label each sample as $genuine_{user}$or Impostor-User. Given a decision threshold τ, and τ *between the* [0,1,…9]. Now the fusion score is expressed as,

$$F_{S\,(i,j)}= \; S_{f(i,j)} + \; S_{v(i,j)} \qquad (16)$$

$$Fusion_{Score\,=}\begin{cases}1(= genuine_{user}) \;\; if\ Fs\ (i,j) \geq \tau \\ 0(= impostor_{user})\ if\ Fs\ (i,j) < \tau\end{cases}$$

iv. EXPERIMENTAL RESULTS

This section describes the experimental results and the performance of the proposed system.

A. Experimental Setup

For our implementation, we used Matlab. The performance of the proposed multimodal biometric watermarking system has been tested upon ORL and TIMIT databases. Whenever an ownership claim is to be resolved, the face features and voice features are extracted from the suspected watermarked image and compared with the other training samples of the user stored in the two databases. If a match is found from the database, it is categorized as genuine otherwise it taken as impostor attempt. The most important performance factor for the success of any biometric system (uni/multimodal) is its recognition accuracy. In order to validate and verify the proposed system, various factors are considered and tested. The multimodal biometric traits have been described and tested in detail as follows:

B. Databases Description:

For the experimental studies, the multimodal biometric data of ORL and TIMIT databases are taken. ORL contains 40 distinct subjects with the size of 92 × 112 pixels and 256 grey levels per pixel. Some of the images are taken at different facial expressions like smiling, not smiling, closed eyes, open eyes, anger, etc. The entire facial images were taken at a dark homogenous background with the subjects in some side movements (upright, downright and frontal position). TIMIT corpus consists of speech data for the 630 subjects of 8 dialects, 6300 utterances and 10 sentences in American English. The corpus contains totally 5 hours of speech. All 630 speakers are native speakers of the United States. In addition, auxiliary subjects were recorded but are not considered in the CD-ROM.

C. Parameters Description:

The performance of the system is generally based on the evaluating imperceptibility, robustness and security measures like PSNR, SSIM, EER and other metrics. But when it comes to a multimodal biometric watermarking system, we need to ensure the performance in terms of accuracy. A watermarked system should further enhance the security aspects of the biometric traits such as face and voice used without compromising in its quality and features. For the proposed algorithm, we have verified the performance based on with watermarking and without watermarking using face, voice and multimodal in terms of accuracy, equal error rate (ERR), PSNR and SSIM. The details are described as follows:

Peak Signal to Noise Ratio (PSNR): PSNR is widely used and accepted measure of the fidelity of the watermarking method and allows visual inspection between original images and watermarked or reconstructed images. It is evaluated by the following expression.

$$PSNR = 10log_{10}\frac{255^2}{MSE} \qquad (17)$$

Where $MSE=\frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left(f(i,j)-f'(i.j)\right)^2}{m\times n}$ where f and f` are the two images being compared.

Structural Similarity Index (SSIM): It is a similarity measure between two images of which one image is considered as of perfect quality where SSIM (x,y) is given as follows;

$$SSIM\,(x,y) = [f(x,y)]^{\alpha} * [c(x,y)]^{\beta} * [s(x,y)]^{\gamma} \qquad (18)$$

Equal Error Rate (EER): is also known as crossover rate/crossover error rate. The EER is used to evaluate the performance of our proposed multimodal biometric system. It is used find the value of threshold for FAR (False Accept Rate) and FRR (False Recognition Rate). The lower EER value indicates the better performance and also it improves the accuracy.

$$Equal\ Error\ Rate = \frac{False\ Accept\ Rate}{False\ Recognition\ Rate} \qquad (19)$$

Recognition Rate (RR): is also referred as accuracy. RR is the percentage of detected images from the total number of images.

$$Recognition\ Rate = \frac{No.of\ corrected\ identified\ images}{Total\ no.of\ images} * 100 \qquad (20)$$

D. Comparative analysis:

The comparative analysis of the proposed method with existing mechanisms is carried out using the EER, PSNR, accuracy and SSIM. The results of the existing mechanisms are taken from the existing techniques [15, and 11]. Here multimodal biometric techniques have been compared using various performance metrics. From Table 1, we understand that our proposed techniques showed the higher performance in terms of accuracy, PSNR, EER and SSIM.

Table.1 presents various multimodal methods based on digital watermarking techniques. As seen almost of the multimodal approaches have been applied face and voice metrics for biometric system as well all of the systems have some pros and cons. From the table, we have analyzed there were not strong and robust watermarking technique which can provide security, robustness, imperceptibility, capacity and memory criteria. On the other hand, watermark embedding effect is degrading the performance of the face recognition system

TABLE1. Comparison of Different Watermarking Techniques

Techniques	Merits	Limitations
LSB	1. Easy to use and understand 2. High perceptual transparency 3. Image quality low degradation	1. lack of basic robustness 2. Cropping and scaling vulnerabilities 3. Noise vulnerability

DCT	1.In digital watermarking, the watermark is embedded into the coefficients of the middle frequency. So the image visibility will not affected 2. Does not affects pixels themselves with each other 3. High robustness	1. Lack of invariance (block wise DCT destroys invariance) 2. Higher frequency
DWT	1. Good in localization 2. Higher compression ratio	1. Computation cost is high 2. High compression time 3. Blur/noise edges of images
RDWT	1. Invariance property achieved translation 2. Sufficient embedding capacity	1. Lack of spatial location 2. Best watermark embedding is required in L-Level decomposition.

Two image quality assessment metrics such as Peak Signal to Noise Ratio (PSNR) and Structural Similarity Matrix have been measured to test the performance of imperceptibility of the double watermarking mechanism. Both IQA metrics are well-known conventional metrics for pixel based image operation. From the table 2, we proved that our proposed system achieves good performance in both IQA metrics and also it improves the imperceptibility. Generally, SSIM lies between 0-1. Note that the higher IQA value always implies higher visual quality of the original watermarked image.

TABLE 2. Comparisons of the functionalities of our double watermarking method and related double watermarking mechanism

	Existing	**Proposed Approach**
Dual watermarks	Fragile + Robust	Fragile + Robust
Embedding domain	Spatial domain + DWT	DCT+QIM
Visibility	Invisible + Invisible	Invisible + Invisible
Blind Extraction	Yes + Yes	Yes +Yes
Target image	Color (RGB)	Color (RGB) and voice
PSNR dB	~ 40dB (40.32)	~ 48dB (48.9835)
Copyright Protection	Yes	Yes
Image Authentication	Yes	Yes
Average Recognition Rate (%)	99.97%	99.98
Robustness	No	Yes
Security	Yes	Yes
SSIM	0.9830	0.9879

TABLE3. Results for our Multimodal Biometric System

	EER (%)			**Recognition Rate (%)**
	Voice	Face	Multimodal	
Without watermarking	7.5714	8.25	4.25	97.90
with watermarking	7.5714	7.2574	3.333	99.98

TABLE4. Results for Existing Multimodal Biometric System

Mechanism	**Recognition Rate (%)**		
	Voice	Face	Multimodal
Without watermarking	86.8	89.4	94.0
with watermarking	86.8	89.4	94.0

To further demonstrate the system performance, we compared the proposed system recognition rate with existing system [29]. The comparison results of proposed and existing systems are shown in table 3 and 4. In table 3, the recognition rate of our proposed system is 99.98% which is better than the existing system.

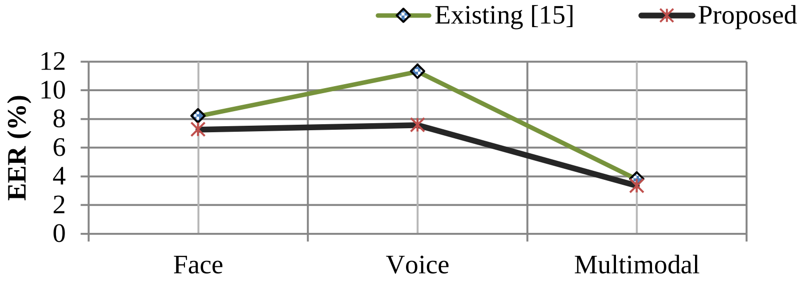

Figure 5: Results of EER

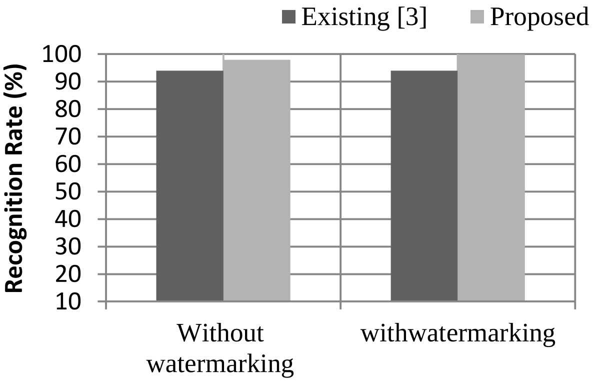

Figure 6: Results of Recognition Rate

v. CONCLUSION

In this paper, we have proposed a blind double digital watermarking method with multimodal biometric system to improve the previous works in area of spatial and temporal domain in terms of robustness, security, authentication and imperceptibility. An important aspect of the proposed digital image watermarking system has the development of watermarking for authentication and recognition. The developed dual watermarking is comprised with blind semi-fragile and robust watermarks into the facial image. Face selection process is proposed to preserve most of the discriminative features of face image and to secure face image by applying image watermarking process. This face selection process is not only enhancing the performance of face recognition system, but also improving the process of feature extraction. Double Deep-Q-Network is performed for face recognition process which reduces the complexity and improves the accuracy of recognition process. After completion of recognition process, score fusion method is expressed to identify the genuine or imposter user. our proposed double watermarking scheme exhibits better performance in terms of accuracy, EER, PSNR, SSIM,

REFERENCES

[1] Wadood Abdul, “Securing Biometric Authentication through Multimodal Watermarking”, 3rd International Conference on Artificial Intelligence, Modeling, and Simulation, 2015

[2] Bairagi Nath Behera, V.K.Govindan, “Improved Multimodal Biometric Watermarking in Authentication Systems Based on DCT and Phase Congruency Model”, International Journal of Computer Science and Network, Vol.2, Issue. 3 and 2013

[3] Mayank Vatsa, Richa Singh, Afzel Noore, “Feature based RDWT watermarking for multimodal biometric system”, Image and vision computing, 2007.

[4] A. Lathika, D. Devaraj, “Artificial Neural Network based Multimodal Biometrics Recognition System”, International Conference on Control, Instrumentation, Communication and Computational Technologies, 2014

[5] Piotr Stefan Nowak, Wojciench Sankowski, Pawel Krotewicz, “3D Face and Hand Scans Acquisition System”, 23rd International Conference on Mixed Design of Integrated Circuits and Systems”, 2016

[6] Tudor Barbu, Adrian Ciobanu, Mihaela Luca, “Multimodal Biometric Authentication based on Voice, Face and Iris”, The 5th IEEE International Conference on E-Health and Bioengineering, 2015

[7] Rohit M. Thanki, Ved Vyas Dwivedi, and Komal R.Borisagar, “Robust and Secure Watermarking using Sparse Information of watermark for Biometric data protection”, NIRMA University Journal of Engineering and Technology, 2016.

[8] Wioletta Wojtowicz and Marek R.Ogiela, “Biometric watermarks based on face recognition methods for authentication of digital images”, Security and Communication Networks, 2014

[9] Abdulmawla Najih, S.A.R. Al-Haddad, Abd Rahman Ramli, S.J. Hashim, Mohammad Ali nematollahi, "Digital image watermarking based on angle quantization in Discrete Contourlet Transform", Journal of Kind Saud University-Computer and Information Sciences, 2016.

[10] Rohit M. Thanki, Komal R. Borisagar, "Novel Approach for Multimodal Biometric System using Compressive Sensing Theory based Watermarking", International Journal of Computer Science Engineering, Vol. 3, Issue. 4, 81-90, 2013

[11] Priyanka Singh, Balasubramanian Raman, Partha Pratim Roy, "A multimodal biometric watermarking system for digital images in Redundant Discrete Wavelet Transform", Multimedia Tools and Applications, Springer, 2016

[12] Amit Kumar Singh, Nomit Sharma, Mayank Dave, Anand Mohan, "A Novel Technique for Digital Image Watermarking in Spatial Domain", 2nd IEEE International Conference on Parallel, Distributed and Grid Computing, 2012

[13] Mohd Rizal Mohd Isa, Salem Aljaresh, Zaharin Yusoff, "A Watermarking technique

[14] to improve the security level in Face Recognition Systems", 2016

[15] [14]. Su wang, Roland Hu, Huimin Yu, Xia Zheng, R.I. Damper, "Augmenting Remote Multimodal Person Verification by Embedding Voice Characteristics into face images", IEEE international conference on multimedia and expo workshops, 2013.

[16] Methaq talib Gaata and Refan Aamer Jaafer, "Adaptive Watermarking Technique for Speech Signal Authentication", International Journal of Computer Science and Information Technology, Vol.8, No.4, 2016

[17] Hashim, R. R. S. J., & Albannai, N. Matching Fingerprint Images for Biometric Authentication using Convolutional Neural Networks. [17]. Ala,E,O (2017). Joint MFCC-and-Vector Quantization based Text-Independent Speaker Recognition System. IEEE International Conference on Communication, Control, Computing and Electronics Engineering.

[18] Shashi Choudary and Naveen Choudhary, "Intensifying the Security of Multimodal Biometric Authentication System using Watermarking", Global Journal of Computer Science and Technology, Vol. 15, Issue. 4 and 2015

[19] Ibrahim A. El rube, Mohammed Abou El Nasr, Mostafa M. Naim, Mahoud Farouk, "Contourlet versus Wavelet Transform for a Robust Digital Watermarking technique", International journal of electrical, computer energetic, electronic and communication engineering, Vol. 3, No.12, 2009.

[20] Bali,M (2013).Face Recognition using Eigen Faces and Transmission of Hidden Data using Watermarking Authentication. International Journal of Engineering and Computer Science.2(6), 1827-1833.

[21] Sirvan Khalighi, Parisa Tirdad, and Hamid R. Rabiee, "A Contourlet based Image Watermarking Scheme with High Resistence to removal and geometrical attacks", EURASIP Journal on Advances in Signal Processing, 2010

[22] Amit Mehto, Neelesh Mehra, "Adaptive Lossless Medical Image Watermarking Algorithm based on DCT and DWT", International conference on Information Security and Privacy, 2015

[23] Yuqiang Cao, Weiguo Gong, Mingwu Cao, SenBai, "Robust Biometric Watermarking Based on Controulet Transform for Fingerprint and Face Protection", International Symposium on Intelligent Signal Processing and Communications systems, 2010.

[24] Yahya AL-Nabhani, Hamid A. Jalab, Ainuddin Rafidah Md Noor, "Robust watermarking algorithm for digital images using discrete wavelet and probabilisticneural network", Journal of Kind Saud University-Computer and Information Sciences, 27, 393-401, 2015

[25] BinMa., Yunhong, W., Chunlei, L., Zhaoxiang, Z., & Di, H. (2014). Secure multimodal biometric authentication with wavelet quantization based fingerprint watermarking. Springer Multimedia Tools and Applications, 72(1), 637-666.

[26] Mohd,R,M,I., Salem,A,Z,Y (2016). A Watermarking technique to improve the security level in Face Recognition Systems.

[27] Gil-Je Lee-Kee-Young Yoo, "An improved double image digital watermarking scheme using the position property", Multimedia Tools and Applications, Springer, 2014

[28] Vandana S Inamdar, and Priti P Rege, "Dual watermarking technique with multiple biometric watermarks", Sadhana, Springer, vol. 39, issue. 1, pp. 3-26, 2014

[29] Xiao-Long Liu, and Chia-Chen Lin, "Blind dual watermarking for color image authentication and copyright protection", IEEE transactions on circuits and systems for video technology, 2016.

[30] Rohit M. Thanki, Ved Vyas Dwivedi, Komal R. Borisagar, " A watermarking technique using Discrete Curvelet Transform for Security of Biometric Features", International Journal of Information Processing, Volume 10, issue 1, pp. 103-114, 2016.

[31] Ohud Nafea, Sanaa Ghouzali, Wadood Abdul, and Emad-Ul-Haq qazi, “Hybrid Multi- Biometric Template Protection using Watermarking”, The Computer Journal, 2015