

Securing Technique Using Pattern-Based LSB Audio Steganography and Intensity-Based Visual Cryptography

Pranati Rakshit¹, Sreeparna Ganguly¹, Souvik Pal², Ayman A. Aly³ and Dac-Nhuong Le^{4,5,*}

¹JIS College of Engineering, Kalyani, West Bengal, 741235, India

²Global Institute of Management and Technology, Nadia, West Bengal, 741102, India

³Department of Mechanical Engineering, College of Engineering, Taif University, Taif, 21944, Saudi Arabia

⁴Institute of Research and Development, Duy Tan University, Danang, 550000, Vietnam

⁵Faculty of Information Technology, Duy Tan University, Danang, 550000, Vietnam

*Corresponding Author: Dac-Nhuong Le. Email: ledacnhuong@duytan.edu.vn

Received: 11 September 2020; Accepted: 10 November 2020

Abstract: With the increasing need of sensitive or secret data transmission through public network, security demands using cryptography and steganography are becoming a thirsty research area of last few years. These two techniques can be merged and provide better security which is nowadays extremely required. The proposed system provides a novel method of information security using the techniques of audio steganography combined with visual cryptography. In this system, we take a secret image and divide it into several subparts to make more than one incomprehensible sub-images using the method of visual cryptography. Each of the sub-images is then hidden within individual cover audio files using audio steganographic techniques. The cover audios are then sent to the required destinations where reverse steganography schemes are applied to them to get the incomprehensible component images back. At last, all the sub-images are superimposed to get the actual secret image. This method is very secure as it uses a two-step security mechanism to maintain secrecy. The possibility of interception is less in this technique because one must have each piece of correct sub-image to regenerate the actual secret image. Without superimposing every one of the sub-images meaningful secret images cannot be formed. Audio files are composed of densely packed bits. The high density of data in audio makes it hard for a listener to detect the manipulation due to the proposed time-domain audio steganographic method.

Keywords: Information security; visual cryptography; audio steganography; secret image; reverse steganography

1 Introduction

Secure transmission of confidential digital data via shared networks is a challenging task in itself. Shared channels like the internet and other local or wide area networks are often considerably fast and cost-effective ways of data transmission. The amount of data that changes hands each day through these media is huge as well. Instead of sending confidential data in



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

its original form, we can bring recoverable distortions to it. This way the data would seem meaningless to the eyes of interceptors while a proper decoding algorithm can bring back the original data. Cryptography and steganography are two such schemes of information security. *Cryptography* produces meaningless data out of the secret message and *steganography* hides it in some other digital data like images, audio files, etc.

Cryptography is a genre of information security where a manipulation, replacement or haphazard ordering scheme is applied on a particular message data to get a meaningless cryptic data or cipher. This cryptic data can be decoded only with the correct decryption technique or cryptanalysis which is logically opposite to the applied encryption scheme [1]. According to Rivest [2], cryptography defines the study of information exchange in a contending environment.

The use of cryptography as a security measure has a rich history [2–9]. It was a common wartime communication technique in the Caesarian era [3]. Claude Shanon's mathematical information coding theory or *the theory of theoretical secrecy* [4,5] is considered as the foundation stone for the modern cryptographic ciphering-deciphering schemes. This combinatorics-based theory mathematically showed the possibility of information coding with cryptography and its practical application in one to one secret communication. In private key cryptography [6,7], the sender and the intended receiver of the secret message agree upon a key before the communication of the actual message and then the secret message is encrypted with the key. The receiver holds a similar key to decipher the data. In this method, the key is similar at the sender and the receiver end so this method is also called as symmetric-key cryptography. On the other hand, in the public key cryptosystems, there are two types of keys; public keys and private keys [8,9]. Public keys are common to all the participants of the system. Using this key, the message is encrypted by anyone of the sending ends. Each of the participant ends of the system holds a distinct private key for decryption. This key can only decrypt the cipher if the holder of that particular private key is listed as an intended receiver by the sender. In this technique, the public and private keys are not at all similar so the system is also called asymmetric key cryptosystem. Cryptographic methods can produce incomprehensible cryptic messages or cipher messages out of secret information. The messages, though juxtaposed, are still present or visible in its original form. Not only the sender and intended receiver but also the third parties in the communication media are aware of the secret message. Steganography can be used to overcome this fundamental drawback of cryptography. In *steganography*, the secret message is hidden within a chosen media (text, image, audio, video, IP datagram, etc.) keeping the characteristics of the media unchanged [10–15]. The intended receiver of the message uses an appropriate decoding algorithm to extract the message out of the cover or carrier media. Since our system is based on audio steganography, our main theme of discussion would focus on this particular category.

The main aim of steganography schemes is to provide protection against detection and to protect against removal [16,17].

Protection against detection is immensely important in data hiding based secret communication schemes like steganography where the secret data may be hidden in its original form. Any steganographic algorithm tries to keep the difference between original cover media and the manipulated or secret data-filled cover media as small as possible. This way it becomes almost impossible to detect the presence of secret data and thus the secrecy is maintained in the communication. Simon's Prisoner's problem [18] is an efficient way-out to ensure detection security. This method revolves around the covert communication between two prisoners Bob and Alice who want to form a prison-break plan. All their messages need to pass through Warden Willie just like all secret messages need to be communicated in real-life schemes through shared networks. This scheme

works using an image as media for hiding secrets. Steganography techniques, these days are being used for various document marking strategies for protecting *removal*. Intellectual property theft detection is still a very challenging task due to the bulk of data that circulates through the World Wide Web.

A simple scheme of *Digital watermarking* or fingerprinting can ensure protection against the removal or misuse of copyrighted data [19,20]. In digital watermarking, an identity symbol or signature data is embedded in a copyrighted document using the methods of steganography. This invisible piece of author information authenticates the data. *Digital fingerprinting*, on the other hand, embeds a serial number or any other serial data on the document [21–33]. If the data is duplicated without consent, the fingerprint information does not get copied. It becomes easier to detect pirated documents in this manner.

In this work, these two methods are combined to produce a two-layer system that is more secure than only cryptographic or steganographic systems. The increased complexity reduces the chances of interception or detection of hidden data. Visual cryptography and audio steganography are two fundamental blocks of the proposed system. A brief idea about these two techniques is necessary for a better understanding of the system into consideration. *Visual cryptography* is a special type of cryptographic method which considers images as matrices of binary octets in terms of their pixel intensity.

2 Literature Survey

The following are some literatures which used mainly visual cryptography and audio steganography.

2.1 Visual Cryptography

Visual cryptography is a special type of cryptographic method that works only on *visually comprehensible secret* images. This method considers images as matrices of binary octets in terms of their pixel intensity. Using arithmetic means the pixel values are divided into some integer values and from these values, several meaningless component images are formed. When each of these images is combined by *mathematical superposition*, the original image is retrieved. In Fig. 1, we see a simple grayscale image that has only 8 bits of data per pixel. Using a visual cryptographic scheme this image is divided into two meaningless components. If we combine these two by superposition, the original image will be regenerated.

Depending on the type of the image on which visual cryptography is being applied, visual encryption schemes are classified into three categories; *binary image encryption*, *grayscale image encryption*, and *color image encryption* [10]. The basic scheme of visual cryptography is a secret sharing method was first proposed by Naor et al. in [11]. Their model demonstrated a k out of $n(k, n)$ combinatoric system where k number of identical size share images were constructed from a binary secret image. If and only if at least n out of k shares are combined through visual superpositioning, the original image can be reconstructed. This particular scheme can hide only one secret image but there are multiple secret image hiding schemes for binary images as well [12]. Fig. 2 depicts how Naor and Shamir's algorithm produces share images from the secret image pixel by pixel.

Shyu et al. [13] first introduced visual cryptanalysis in RGB color images. Their method was based on pixel division method. In this method, each pixel of c color image is divided into several sub-pixels where each sub-pixel has c components. One of the components is used for encryption

and the other components were filled with black color's value. This method had the disadvantage of pixel expansion $c \times 3$ so the recovered image was distorted to a certain extent.

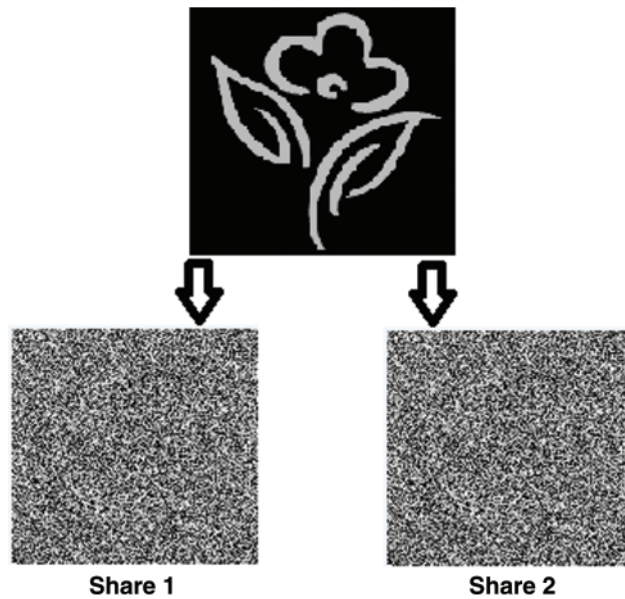


Figure 1: Visual cryptography on a grayscale image

Among the other genres of visual cryptography on color images *region-based visual cryptographic methods* [14] can produce distortion-free reconstructed image at the time of decryption. In this scheme, each pixel of the secret image is applied with the same arithmetic function. Our proposed system is based on this technique.















Pixel	Probability	Share ₁	Share ₂	Share ₁ ⊗ Share ₂
	50%			
	50%			
	50%			
	50%			

Figure 2: Block diagram of the Naor and Shamir's algorithm

2.2 Audio Steganography

In audio steganography, compressed or uncompressed audio files are used as cover media to hide the secret message. Audio files are a collection of sampled binary bits captured from

continuous audio signals. These bits streams can be manipulated in the time domain, transform domain, or codec domain.

2.2.1 Time Domain Audio Steganography

Time-domain methods are generally based on least significant bit schemes, silence removal schemes, or echo hiding methods. Least Significant Bit steganography schemes manipulate the lower order bits of audio samples as changing those bits do not cause any significant change in the overall audio quality. Fig. 3 explains the working principle of a simple LSB based audio stego scheme where a 16-bit message is hidden in the lower order 8 bits of audio's sampled data values. First, the LSBs are filtered out so that the spaces become blank, and then it is filled with message data. Generally, audio files have a huge number of bits per sample, so the manipulation process becomes lengthy still audio is considered as better cover media than images due to its data bulk. This work employs a pattern-based LSB steganography scheme in one of its layers.

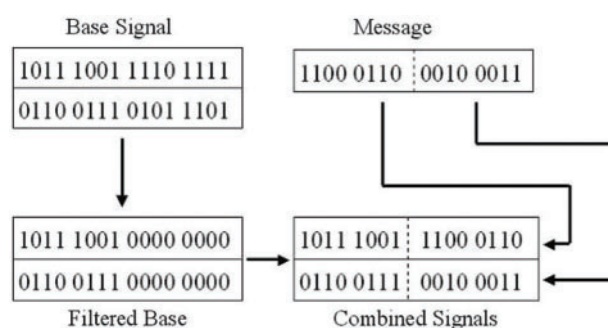


Figure 3: Working principle of LSB based steganography schemes

Silence removal is another kind of time-domain audio steganography scheme. It is based on the fact that audio signals like music or voice recordings generally have small pauses or silent zones in between. Reducing small time gaps out of those break time-lapses does not easily get detected by the human auditory system. This method applies to small secret data only plus uneven pauses increase the chance of detection.

The echo hiding method creates an inaudible echo from the original audio and hides data within it. This scheme loses its relevance if the echo becomes audible due to steganographic manipulation [15–17] so it is not a very secure technique.

2.2.2 Transform Domain Audio Steganography

Transform domain techniques consider an audio signal as a collection of frequencies and manipulate those frequency packets to hide the data. The human auditory system cannot detect the presence of a weak frequency in the neighbourhood of strong frequencies. If any insignificant changes are done in those weaker frequencies, it has a high chance to remain undetected. This type of technique is broadly categorized into 6 categories; spread spectrum, phase coding, discrete wavelet transform, amplitude coding, tone insertion, and cepstral domain steganography [18,19].

Spread spectrum methods distribute the parts of hidden data throughout the entire frequency spectrum of the cover audio using M-sequence codes and direct sequence spread spectrum (DSSS) method. Phase and Amplitude coding introduce phase and amplitude modification in the original audio. Small phase shifts are very hard to detect and provision of redundancy increases

fault tolerance. A discrete wavelet transform manipulates the LSBs of the wavelet coefficient of the cover audio. To ensure the inaudibility of the introduced noise or the hidden data a minimum hearing threshold has to be maintained. Among all the frequency domain audio steganography techniques, cepstral domain or log spectral domain methods are most efficient in terms of embedding, fault tolerance as well as protection against detection. In this method, data streams are concealed in a few selected cepstral coefficients.

2.2.3 Codec Domain Audio Steganography

Codec domain steganography schemes are employed at the time of data transmission by the sender. Here we manipulate the data rate at amplitude modulation and thus small differences are created in the sending and receiving rate of data. This technique shows high detection tolerance.

3 Motivation

There are several schemes available to implement visual cryptography on images but ‘*k out of n*’ scheme of visual cryptography is perhaps the most explored method among them. Naor et al. [11] scheme and also its extension schemes fall into this category [12,13]. These schemes mostly work on binary black and white images or grey tone images and show low tolerance to pixel expansion. The purpose of this work was to propose a general visual cryptographic scheme that can be applied in any kind of uncompressed image be it binary, grayscale, or color image. To overcome the pixel expansion problem, we have used a ‘*region-based*’ visual cryptographic technique [14]. Region-based visual cryptography is a comparatively new strategy. This method works by dividing an image into various inherent subparts and applies similar or different encryption on each of the parts. By principle, region-based visual encryption can be applied to only those images where we can separate the object on focus and the background. Most of the real-life images it is not possible. This is one of the reasons why region-based schemes are not much explored although their efficacies are quite satisfactory for real-life usage. Our proposed system treats the entire secret image as a region and applies the same encryption scheme in each unit of the region. This way, the evenness is maintained throughout the image vicinity and the recovered image becomes free from pixel expansion. As for the steganography part of our work, we have used audio as our medium. Different audio steganography literatures are there [24–31]. Bit manipulation related steganography methods generally use the image as the media but due to the higher sensitivity of Human Visual System, chances of detection are quite high. Audio signals on the other side, have a much higher number of samples per unit runtime which ensures better scattering of a secret message within the vicinity. Here our secret is in the form of an uncompressed color image. Each pixel of the image consists of 24 bits of data. To hide such a big message, audio media is a far better choice than an image in both the time and frequency domain.

4 Proposed System

This security system generates 8 meaningless shares from a secret image using visual cryptography and hides each image-share in 8 separate audio files using a numeric pattern-based LSB audio steganography scheme in the sender side. These 8 shares can then be transmitted in 8 different ways. Each of the sub-images contains 1/8th share of the secret. The original secret image can be revealed if each share is extracted from their respective stego-audios using LSB stego-extraction process and then combined using a Visual decryption algorithm.

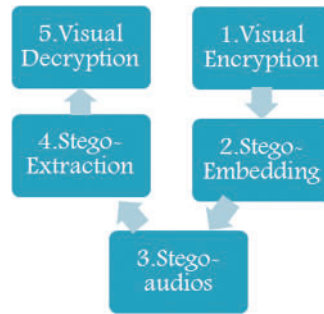


Figure 4: Block diagram of the proposed algorithm

4.1 Proposed Visual Cryptography Scheme

This particular visual cryptography scheme uses a simple pixel intensity division technique to produce the intended number of basis matrices whose size is the same as the image to be encrypted. Depending on the number of basis matrices, the original image intensity will be distributed. If the number of basis matrices is n , the algorithm will generate 2^n number of share-images. For this algorithm to work, value of n should be more than 2. Increment in the number of shares will increase the security as well as the computational complexity of the scheme. To design this particular system, we have considered n as 3 and that is why the total number of image shares is 8.

At the beginning of the algorithm, we have initialized the basis matrices B_1 , B_2 , and B_3 with (original image's pixel value/ n). So, for this case, it becomes, $B_1 = B_2 = B_3 = \text{original image's pixel value}/3$. Then B_1 , B_2 , B_3 are updated as $B_1 = 128 - B_1$; $B_2 = b^2$; $B_3 = 64 - B_3$. At this point, we have created a matrix K of the same size as the original image and filled it with random numbers generated by the uniform distribution. Finally, we constructed 8 share images using bitwise XOR operation (\oplus),

$$S_1 = K \oplus B_1, S_2 = K \oplus B_2, S_3 = B_2 \oplus B_3, S_4 = B_3 \oplus S_1 \\ S_5 = S_1 \oplus S_2, S_6 = S_2 \oplus S_3, S_7 = S_3 \oplus S_4, S_8 = S_4 \oplus S_5$$

where S_1, \dots, S_8 are matrices for the share images.

The same procedure is applied for red, green, and blue parts of the RGB image pixel. The decryption is symmetric but exactly opposite to the encryption scheme. We will have to perform simple bitwise XOR operation on all 8 sub-images and the random value-filled K matrix. Here bitwise XOR behaves like superposition operation. It combines all the shares and cancels out the randomness introduced in the encryption by the random matrix K .

4.2 Proposed Audio Steganography Scheme

Once the secret image is encrypted in component images using visual cryptography method, the system proceeds to the audio cryptography scheme. This layer encodes each of the 8 sub-images into 8 uncompressed audio files chosen by the user. Audio data are stored in groups of 16-bit pulse code modulated samples. Just like RGB image's r , g and b components, a stereo audio sample consists of 16 bits of left and 16 bits of right components. This algorithm starts information hiding from the left components of each sample and traverses through the samples sequentially. If the left samples are all covered, it moves to the right components. This algorithm introduces a novel audio steganography method using the least significant bit (LSB)

replacement strategy. Instead of replacing the same number of LSBs in each sample, we will vary the number of replaced bits using a predefined pattern. The pattern implemented here can be called 4-2-2-4. Here, the first and fourth of the audio samples will have 4 LSBs replaced. In the 2nd and the 3rd samples, the number of replaced bits will be 2 in each. The pattern will repeat itself as 4-2-2-4-4-2-2-4 and so on. The retrieval will require prior knowledge of the predefined pattern for the successful extraction of the secret images.

4.3 Proposed Algorithms

4.3.1 Sender Side Algorithms

Algorithm 1: Visual encryption

1. Select the RGB secret image as input.
 2. Prepare message image as 3 dimensions 8-bit binary vector whose dimensions are row, column, and color
 3. Select an arbitrary integer b (no. of basic matrices) ≥ 2 , considering the no. of shares to be generated (For this system $b = 3$, so no. of shares $= 2^3 = 8$).
 4. Initialize primary basis matrices B_1, B_2, B_3 using the formula $B_1 = B_2 = B_3 = \text{intensity value of each pixel in the secret image}/3$.
 5. Calculate actual basic matrices using, $B_1 = 128 - B_1; B_2 = b^2; B_3 = 64 - B_3$.
 6. Generate shares by the formula.
-

Algorithm 2: Audio steganography encoding

$$S_1 = K \oplus B_1, S_2 = K \oplus B_2, S_3 = B_2 \oplus B_3, S_4 = B_3 \oplus S_1$$

$$S_5 = S_1 \oplus S_2, S_6 = S_2 \oplus S_3, S_7 = S_3 \oplus S_4, S_8 = S_4 \oplus S_5$$

where, K is a uniformly distributed random number matrix of the size of secret image

1. Select a .wav file as carrier and a .png image as secret message;
 2. Open the carrier file;
 3. Prepare message image as 3D 8-bit binary vector whose dimensions are row, column and color;
 4. Prepare bytes of carrier as 16-bit binary column vector;
 5. Set counter = 1;
 6. **If** (counter == 1){
 7. Replace Least significant 4 bits of carrier with corresponding elements of message vector;
 8. Set counter = counter + 1; }
 - Else if** (counter == 4){
 9. Replace Least significant 4 bits of carrier with corresponding elements of message vector;
 10. Set counter = 1; }
 - Else**{
 11. Replace Least significant 2 bits of carrier with corresponding elements of message vector;
 12. Set counter = counter + 1 }
 13. Get the stego-file as output
-

4.3.2 Receiver Side Algorithm

Algorithm 3: Audio steganography extraction

1. Select the .wav stego file as input;
 2. Open the stego-File;
 3. Get the size of secret image from stego-file;
 4. Prepare bytes of stego-file as 16-bit binary column vector;
 5. Set counter = 1;
 6. **If** (counter == 1){
 7. Extract Least significant 4 bits of the carrier into corresponding elements of message reconstruction vector
 8. Set counter = counter + 1;}
 - Else If** (counter == 4){
 9. Extract Least significant 4 bits of the carrier into corresponding elements of message reconstruction vector
 10. Set counter = 1}
 - Else**{
 11. Extract Least significant 2 bits of carrier into corresponding elements of message reconstruction vector
 12. counter = counter + 1}
 13. Regenerate the .png secret image
-

Algorithm 4: Visual decryption

1. Select all 8 image-shares S_1, S_2, \dots, S_8
 2. Construct a 3-dimension matrix K with the same size of secret images using random numbers in a uniform distribution
 3. Generate the secret image using,
- $$D = S_1 \oplus S_2 \oplus S_3 \oplus S_4 \oplus S_5 \oplus S_6 \oplus S_7 \oplus S_8 \oplus K$$
-

4.4 Implementation and Results

For the implementation of the proposed system, any .png format uncompressed color image can be used. For audio steganography, uncompressed stereo audios of .wav format can be used. Instances of the proposed system have been shown in the pictures below. In this demonstration, we have used an RGB image ‘onion.png’ of size 195×135 and for the cover audio, we have used ‘myfav.wav’. The proposed algorithm can retrieve the secret image without any loss in the intensity. If the cover media is exposed to noise, the distortion of the secret image will be proportional to the loss in the cover audio.

Figs. 5 and 6 depict the sender and receiver window of the proposed system. It is a standalone system so the sender and receiver side software can be deployed in any computer system that pertains to their basic requirements (Windows 7 and above, 2 GB RAM, etc.). The stego audios produced by the algorithm are also shown in the respective windows diagrammatically for the convenience of the users.

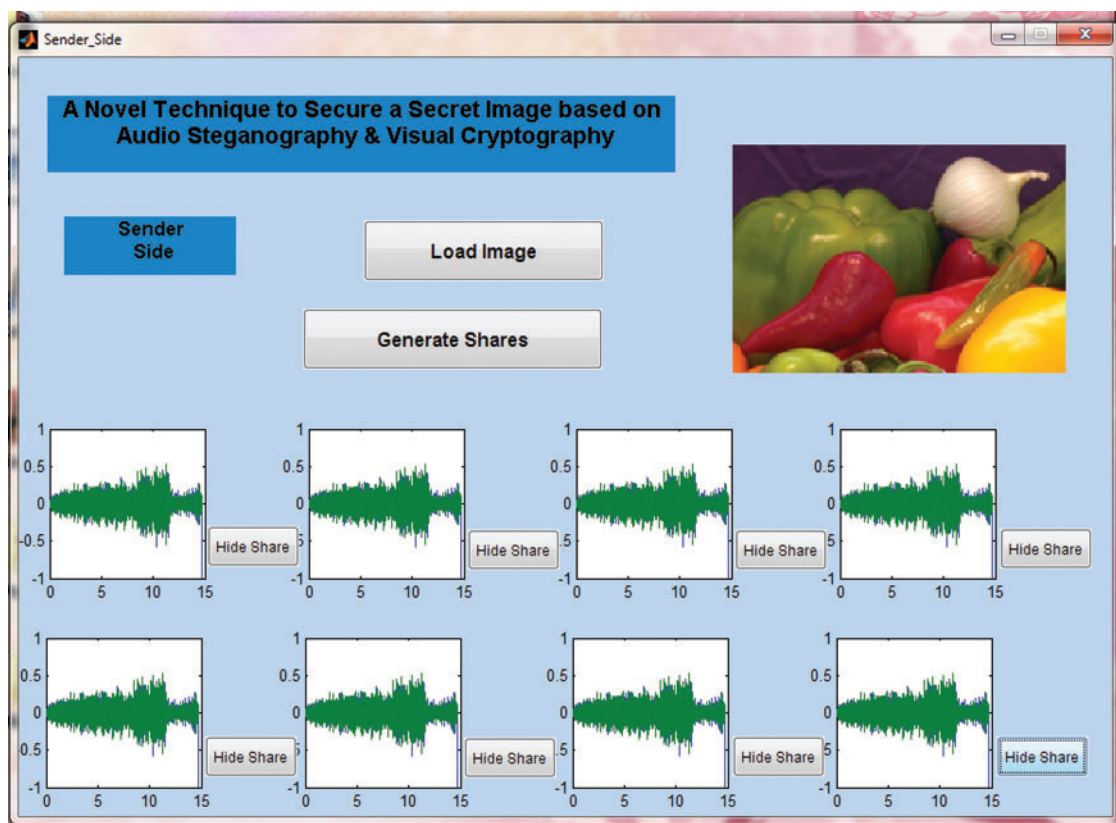


Figure 5: Implementation of the proposed algorithm on the sender side

4.5 Result Analysis

The spatial requirement and the efficiency of the proposed system majorly depend on the efficiency and size requirement of the two main components of the system namely the visual cryptic scheme and the audio steganography scheme.

4.5.1 Spatial Analysis of the Visual Cryptography Scheme

This system uses a region-based visual cryptography scheme. This particular scheme produces 8 component images from a secret image. Each sub-part pixel bears 1/8th share of the actual data of the original image pixels. This 1/8th part is not the 1/8th portion of the original images bit values; it is 1/8th portion of the data regarding that specific pixel's intensity and contrast. That means, a simple extraction and combination procedure on the visually encrypted images will not reproduce the original image. Active participation of every sub-image portions is an essential criterion for the extraction.

If the secret image is of size $m \times n$, the resultant cryptic images will also be of the same size. That means $m \times n$ number of pixels of the secret image will produce $8 \times m \times n$ number of cryptic pixels for a total of 8 sub-images. In the case of greyscale images, each pixel is composed of 8 bits of data so the resultant pixels will be of the size $8^2 \times m \times n$. For a RGB image, each pixel consists of 24 bits of data. So, for this case, the total number of resultant pixels will be $24 \times 8 \times m \times n$ or $3 \times 8^2 \times m \times n$. Since the resultant pixels are far greater in size, the chosen cover media should have a greater number of data bits to hide these secret data beyond suspicion. That

is the reason for which we have chosen audio as media where each sample consists of 16 bits of message data and the number of samples per second is in the order of 10^2 .

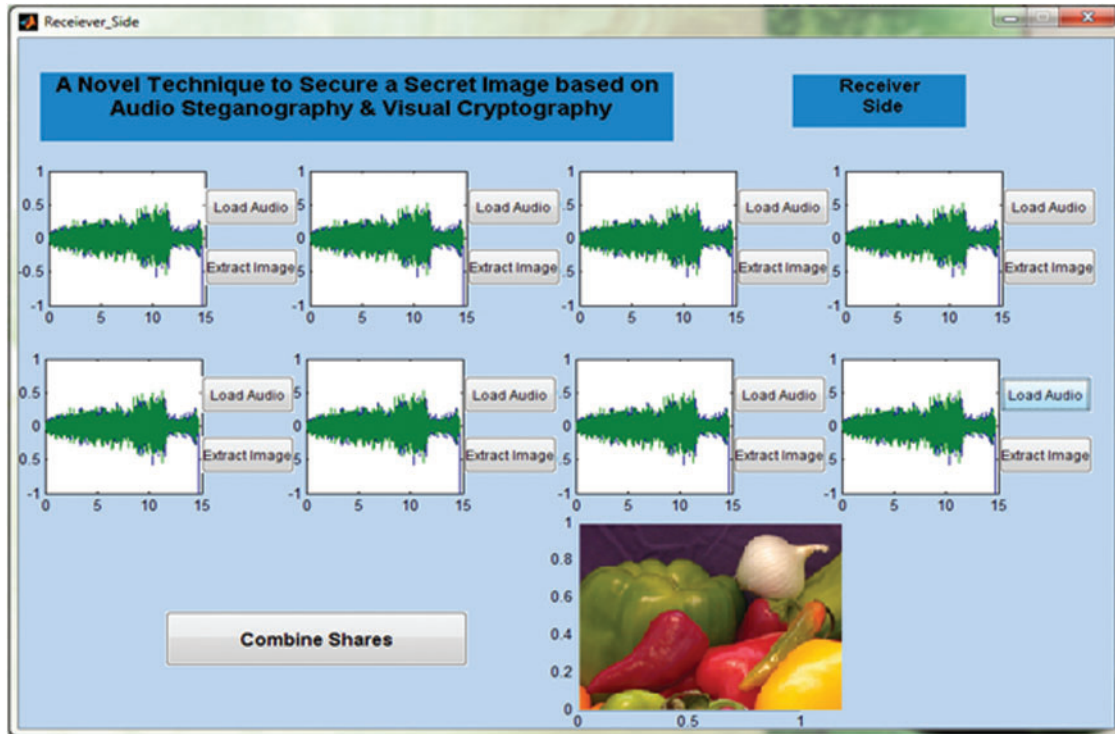


Figure 6: Implementation of the proposed algorithm on the receiver side

4.5.2 Spatial Analysis of the Audio Steganography Scheme

Here we have used an LSB based audio stego scheme that works in the time domain. In the time domain, the audio signal is sampled using 16 bits of binary data. This algorithm generates a simple repetitive number sequence $4 - 2 - 2 - 4 - 4 - 2 - 2 - 4 - 4 - 2 \dots$. According to the current value of this sequence data is being embedded in each sample of the cover audio. That means, if the sequence is currently at 4, the sample will be stuffed with 4 bits of secret message. To hide 8 bits of data we will require either 2 (for $\dots 4 - 4 \dots$ sequence values) or 3 (for $\dots 4 - 2 - 2 \dots$ or $\dots 2 - 2 - 4 \dots$) samples according to the sequence values.

If the secret image that is being hidden using this scheme, is a greyscale image with 8 bits of data per pixel data, it will require at least 2 or at the most 3 samples. If the entire image has $m \times n$ number of pixels, the lower bound of the required samples will be $(2 \times m \times n + 2)$ and the upper bound will be $(3 \times m \times n + 2)$. Here 2 is added as we will need two more samples to specify the width (m) and height (n) of the image.

If the image is a RGB based color image (as the case for our implementation), each pixel will have $3 \times 8 = 24$ bits of data values. To hide each pixel in the audio sample, we will need 3 samples for 8 bits of red component (sequence value $\dots 4 - 2 - 2$), 2 samples for 8 bits of green component (sequence value $\dots 4 - 4 -$) and again 3 samples for 8 bits of blue component (sequence value $\dots 2 - 2 - 4 \dots$). As the sequence will repeat itself as $\dots 4 - 2 - 2 - 4 \dots$ after that so for the next pixel

the sequence value for the red component will return to 4-2-2 as before. So, for this case, the lower bound = upper bound = 8 samples per pixel. For the entire image of size $m \times n$, the total no. of $(8m \times n + 2)$ samples are needed to completely hide the image within the audio.

4.5.3 Efficiency of the Proposed Scheme

The efficiency of any encryption algorithm is measured in terms of their chances of being detected. In this proposed system the secret is covered in the lowermost bits of the uncompressed audio signal in the time domain samples as noise. If the listener's ear can detect the slight distortions produced by hidden data, it will cause suspicion and the algorithm will fail.

Fig. 7 presents a comparative view of the original audio signal (before encoding with message data) and the encoded audio signal. Out of the left and right components of our stereo audio, here, only the left components are plotted as the left component bears the maximum portion of the hidden data. In the plotted graphs, the minor changes in stego-audio are almost invisible in such an amplified view. Our auditory system is far blunt than this fine-tuned view of the cover audio, so we can conclude that the steganography scheme will not be susceptible to detection in just by listener's perspective. Fig. 8 shows the fast Fourier transform of cover audio. Fig. 9 represents the cover audio in frequency domain. Stego audios are formed after LSB manipulation and that frequency domain representation is shown in Fig. 10.

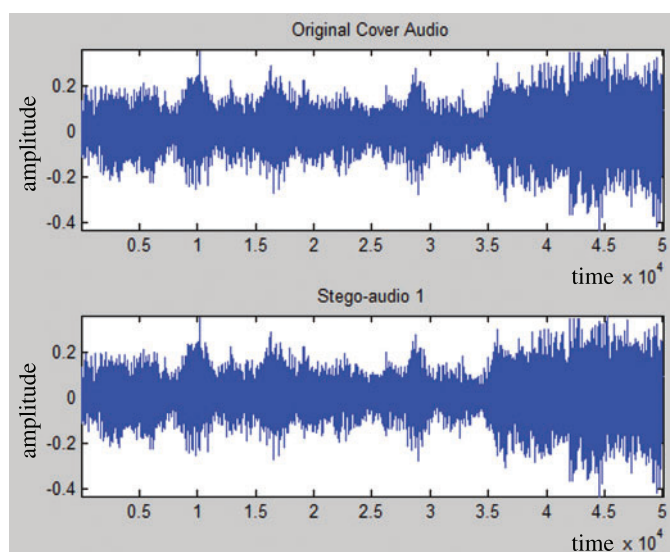


Figure 7: Original audio vs. Stego-audio in the time domain [only the left part of the stereo audio; right audio component is not shown]

Image of Fig. 7 depicts unmanipulated cover audio and stego-audio (a portion of image is hidden in it) in time domain. Here we have taken PCM sampling of an uncompressed .wav audio. As we can see, both the signals look almost identical in the time domain plotting. Our main motto was to minimize the difference between original audio and stego audio so that the noise remains low. Cryptanalytic attacks are time domain analysis based. Our algorithm keeps the noise margin low so that a common listener does not get suspicious by just hearing or doing standard cryptanalysis attacks on the stego audio.

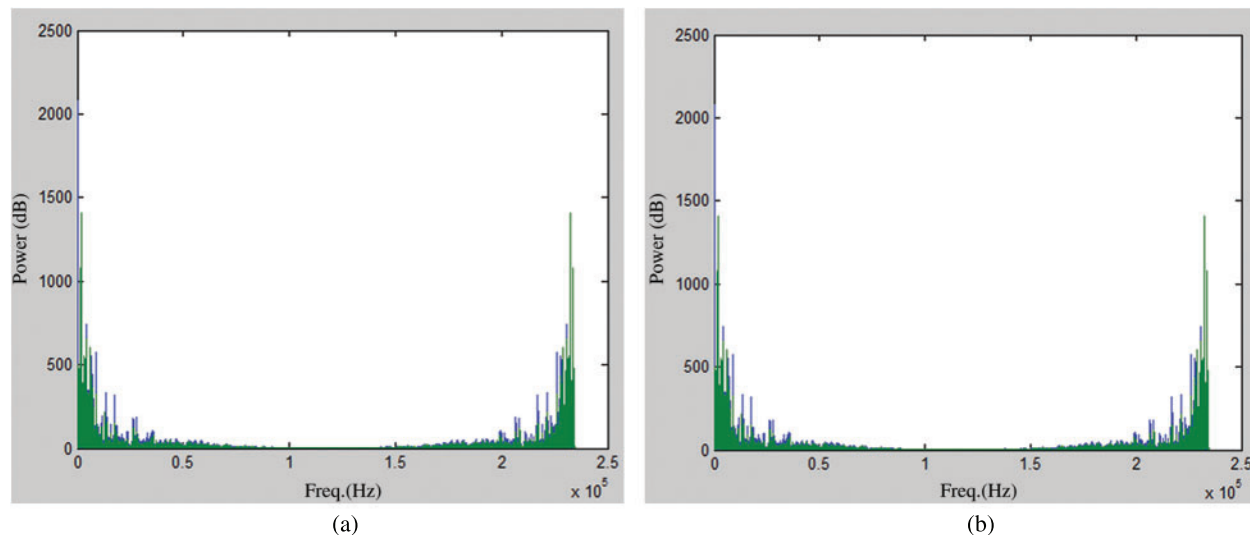


Figure 8: Fast Fourier transform of original cover audio (a), audio after LSB manipulation (b) [left and right component of the stereo signal]

In the image of Fig. 8, we have presented the fast Fourier transform of time domain audio signals shown in Fig. 7. Fig. 8a is the frequency domain plotting of original audio while Fig. 8b is for stego audio. Hiding the presence of noise or secret data is easier in the time domain. However, a frequency analysis of the same data easily reveals the presence of foreign data in the media. In Fig. 8, both original signal and stego signal look identical; we are not seeing any significant difference even after stegano-manipulation. This increases the security of our algorithm even more.

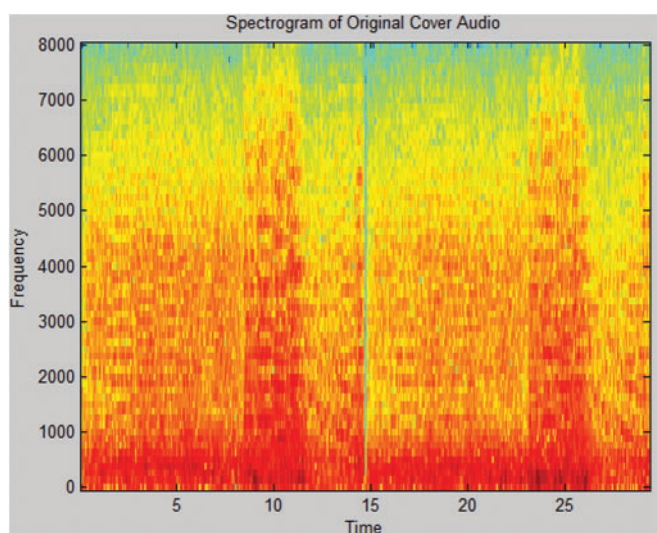


Figure 9: Cover audio in the frequency domain

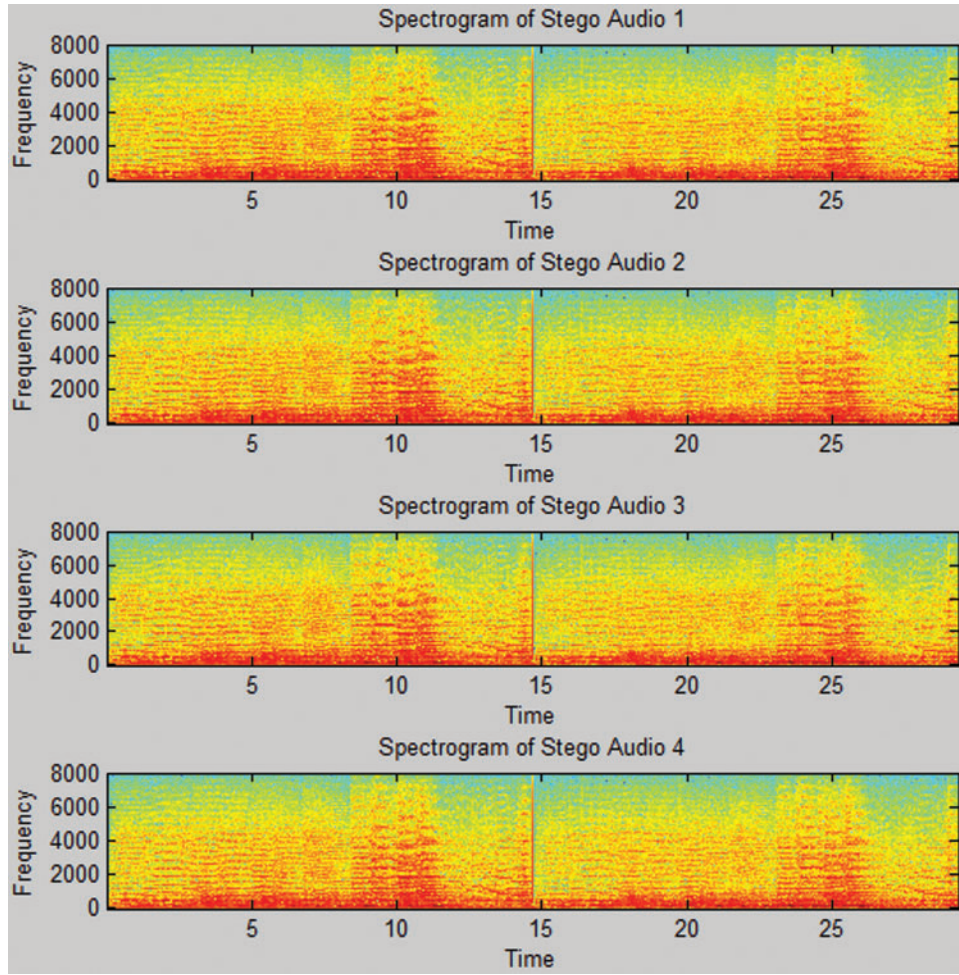


Figure 10: Frequency domain representations of stego-audios after LSB manipulation

4.6 Comparative Analysis

This section deals with a comparison analysis with different methods. Initially, we have made an attribute-wise comparison analysis with different techniques mentioned in [Tab. 1](#). Using this table, we have shown comparative analysis of different characteristics of our work with another state-of-the-art existing works.

- *Quantitative Analysis*

Peak Signal to Noise Ratio (PSNR) measures the maximum noise, the signal tolerates is given as

$$\text{PSNR} = 10 \log_{10} \frac{C_{\max}^2}{\text{MSE}} \quad (1)$$

$$\text{MSE} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N (S_{xy} - C_{xy})^2 \quad (2)$$

where C acts as a host image, S represents the stego image, C_{max} shows the maximum value of a pixel in both original and stego image, x , and y are subscripted variables, M and N indicate image resolution in pixels.

Table 1: Comparative analysis of different characteristics of our work with other states of the art existing works

Characteristics/attributes	SCC method [2]	PIT [5]	ST-FMM [8]	Karim's method [10]	CISSKA-LSB [12]	Proposed method
Two layers of security	No	No	No	No	Yes	Yes
Audio is chosen as cover file	No	No	No	No	No	Yes
Both the time domain and frequency domain analysis are done	No	No	No	No	No	Yes
High PSNR with high payload	No	No	No	No	Yes	Yes
Region based visual cryptography scheme is used	No	No	No	No	No	Yes

Tab. 2 deals with Quantitative results using PSNR for different images with different methods in different kinds of literature. We have compared the SCC Method, PIT, ST-FMM, Karim's Method, and CISSKA-LSB using PSNR and represent in a tabular format.

Table 2: Quantitative results using PSNR for different images with different methods in different literatures

	Image name	SCC method [2]	PIT [5]	ST-FMM [8]	Karim's method [10]	CISSKA-LSB [12]
1	F16jet	47.4852	45.6879	40.2347	47.4902	53.1665
2	House	51.1776	47.6956	40.2518	51.1564	52.7303
3	Trees	38.5418	38.2702	39.5397	38.5421	49.7496
4	Scene-2	40.0355	39.6545	40.0388	40.0353	46.8066
5	Flowers	28.5347	28.5169	29.6394	28.5347	42.0526
6	Baboon-2	33.932	33.8367	34.4471	33.9322	42.3607
7	Building-1	28.8451	28.8213	40.2552	28.8451	43.4071
8	Parrot	28.0434	28.0249	27.7969	28.0434	49.2153
9	Baboon	48.9531	46.5568	39.9997	48.9536	47.8747
10	Masjid	28.5361	28.5173	39.6331	28.5363	44.7425
	Avg. of 50 images	36.3208	34.7621	33.9232	36.3187	45.0309

We have presented better efficiency in Tab. 3, in which our proposed method has been compared with SCC Method, PIT, ST-FMM, Karim's Method, and CISSKA-LSB using PSNR.

In Fig. 11, we have shown the graphical representation, and we can say that the value of PSNR is greater than other schemes. This result analysis justifies the motivation of our work.

Table 3: Quantitative results using PSNR for the comparison between the proposed scheme and avg. PSNR of other schemes from [Tab. 2](#)

SCC [2] method	PIT [5]	ST-FMM [8]	Karim's method [10]	CISSKA-LSB [12]	Proposed method
36.3208	34.7621	33.9232	36.3187	45.0309	48.0843

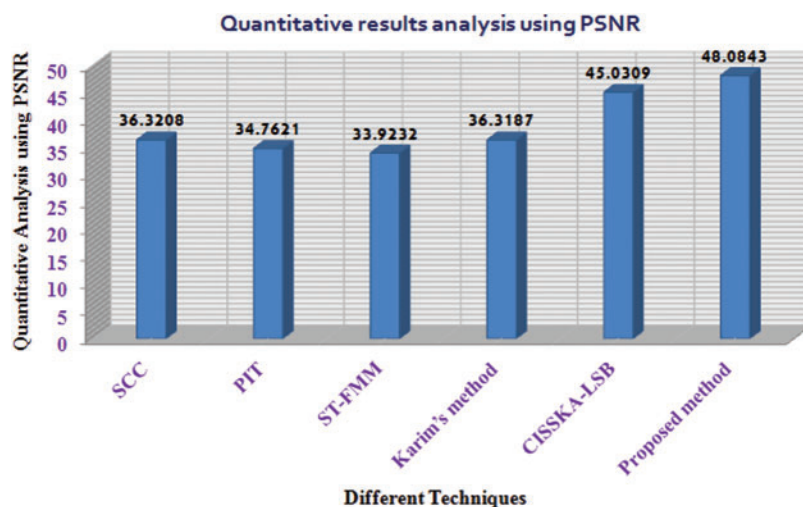


Figure 11: Quantitative comparison analysis using PSNR

5 Conclusion

In this work, an intensity-based visual cryptography scheme is used along with an additional layer of audio steganography in the time domain. Here, the intensity of secret image pixel is distributed among several basis matrices. Superimposition of various combinations of those basis matrices is used to generate 8 distinct share images. The shares are then hidden in audio samples of uncompressed stereo files. The method is quite secure as the data image is completely hidden within the audio. The change in sound quality due to manipulation is quite insignificant and thus listening to the audio would not spring up suspicion. Visual cryptography schemes often cause data loss due to pixel expansion [32–36] but this algorithm can reconstruct the secret image without data loss. Though the process of encryption and decryption takes more time due to the presence of a large number of data bits in both secret image and hiding medium, this scheme works for both colour and greyscale images in uncompressed form. The PSNR value 48.0843 is also quite high as compared to some other current works. The cover media can also be of mono or stereo form. This work can be improvised using transform domain techniques in place of LSB based time-domain methods.

Funding Statement: Taif University Researchers Supporting Project No. (TURSP-2020/77), 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] M. Bellare, D. Pointcheval and P. Rogaway, "Authenticated key exchange secure against dictionary attacks," in *Int. Conf. on the Theory and Application of Cryptographic Techniques*, Bruges, Belgium, Berlin, Heidelberg: Springer, pp. 139–155, 2000.
- [2] K. Bailey and K. Curran, "An evaluation of image-based steganography methods," *Multimedia Tools and Applications*, vol. 30, no. 1, pp. 55–88, 2006.
- [3] R. L. Rivest, "Cryptography," in *Handbook of Theoretical Computer Science*, J van Leeuwen, (Ed.) Amsterdam, The Netherlands: Elsevier, pp. 717–755, 1990.
- [4] D. Kahn, *The Codebreakers: The Story of Secret Writing*. New York: Macmillan, 1967.
- [5] A. A. A. Gutub, "Pixel indicator technique for RGB image steganography," *Journal of Emerging Technologies in Web Intelligence*, vol. 2, no. 1, pp. 56–64, 2010.
- [6] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949.
- [7] D. Boneh and M. Franklin, "Identity-based encryption from the Weil pairing," in *Annual Int. Cryptology Conf., LNCS*, Springer, vol. 2139, pp. 213–229, 2001.
- [8] F. A. Jassim, "A novel steganography algorithm for hiding text in image using five modulus method," *International Journal of Computer Applications*, vol. 72, pp. 8887, 2013.
- [9] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Transaction on Information Theory IT-22*, vol. 6, pp. 644–654, 1995.
- [10] K. Muhammad, Z. Jan, J. Ahmad and Z. Khan, "An adaptive secret key-directed cryptographic scheme for secure transmission in wireless sensor networks," *Technical Journal, University of Engineering and Technology*, vol. 20, pp. 48–53, 2015.
- [11] M. Naor and A. Shamir, "Visual cryptography," *Advances in Cryptography: EUROCRYPT'94, LNCS*, vol. 950, pp. 1–12, 1995.
- [12] K. Muhammad, J. Ahmad, N. U. Rehman, Z. Jan and M. Sajjad, "CISSKA-LSB: Color image steganography using stego key-directed adaptive LSB substitution method," *Multimedia Tools and Applications*, vol. 76, no. 6, pp. 8597–8626, 2017.
- [13] S. J. Shyu, S. Y. Huanga, Y. K. Lee, R. Z. Chen and K. Wangand, "Sharing multiple secrets in visual cryptography," *Pattern Recognition*, vol. 40, no. 12, pp. 3633–3651, 2007.
- [14] E. Verheul and H. V. Tilborg, "Constructions and properties of k out of n visual secret sharing schemes," *Designs Codes and Cryptography*, vol. 11, no. 2, pp. 179–196, 1997.
- [15] D. R. D. Brabin, D. Venkatesan, D. Singaravelan and L. Rajendran, "Region based visual cryptography scheme for color images, International," *Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 3, pp. 1473–1477, 2013.
- [16] R. Popa, "An analysis of steganographic techniques," Master thesis, The Politehnica University of Timisoara, Romania, pp. 18-50, 1998.
- [17] D. Kahn, "The history of steganography," in *Proc. of the First Int. Workshop*, Cambridge, UK: Springer, pp. 1–5, 1996.
- [18] C. Cachin, "An information-theoretic model for steganography," in *Int. Workshop on Information Hiding*, Berlin, Heidelberg: Springer, pp. 306–318, 1998.
- [19] R. J. Andersen and F. A. Petitcolas, "On the limits of steganography," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 474–481, 1998.
- [20] C. Podilchuk and W. Zeng, "Image-adaptive watermarking using visual models," *Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 525–539, 1998.
- [21] R. B. Wolfgang, C. I. Podilchuk and E. J. Delp, "Perceptual watermarks for digital images," in *Proc. of the IEEE*, vol. 87, no. 7, pp. 1108–1126, 1999.
- [22] B. Chor, A. Fiat, M. Naor and B. Pinkas, "Tracing traitors," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 893–910, 2000.
- [23] J. Brassil, S. Low, N. Maxemchuk and L. O'Gorman, "Electronic marking and identification techniques to discourage document copying," in *Proc. Infocom'94*, Toronto, Ontario, Canada, pp. 1278–1287, 1994.

- [24] N. Cvejic and T. Seppben, "Increasing the capacity of LSB-based audio steganography," in *2002 IEEE Workshop on Multimedia Signal Processing*, USA, pp. 336–338, 2002.
- [25] S. Shirali-Shahreza and M. T. Manzuri-Shalmani, "High capacity error-free wavelet domain speech steganography," in *2008 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Las Vegas, NV, USA, pp. 1729–1732, 2008.
- [26] A. Westfeld and A. Pitzmann, "Attacks on steganographic systems," *Lecture Notes in Computer Science*, Berlin: Springer-Verlag, vol. 1768, pp. 61–75, 2000.
- [27] W. Bender, D. Gruhl, N. Morimoto and A. Lu, "Techniques for data hiding," *IBM Systems Journal*, vol. 35, no. 3–4, pp. 313–336, 1996.
- [28] K. Gopalan, "A unified audio and image steganography by spectrum modification," in *IEEE Int. Conf. on Industrial Technology*, Churchill, Victoria, Australia, pp. 1–5, 2009.
- [29] K. Gopalan, "Audio steganography by cepstrum modification," in *Proc. of the IEEE, 2005 Int. Conf. on Acoustics Speech and Signal Processing*, vol. 5, pp. 481, 2005.
- [30] X. Li and H. H. Yu, "Transparent and robust audio data hiding in cepstrum domain," in *Proc. IEEE Int. Conf. on Multimedia and Expo*, New York, NA, USA, pp. 397–400, 2000.
- [31] A. Nishimura, "Data hiding for audio signals that are robust concerning air transmission and a speech codec," in *IIH-MSP'08*, pp. 601–604, 2008.
- [32] D. N. Le, B. Seth and S. Dalal, "A hybrid approach of secret sharing with fragmentation and encryption in cloud environment for securing outsourced medical database: A revolutionary approach," *Journal of Cyber Security and Mobility*, vol. 7, no. 4, pp. 379–408, 2018.
- [33] B. Seth, S. Dalal, V. Jaglan, D. N. Le, S. Mohan *et al.*, "Integrating encryption techniques for secure data storage in the cloud," *Transactions on Emerging Telecommunications Technologies*, e4108, 2020, (Pre-online).
- [34] M. Sajjad, I. Mehmood, N. Abbas and S. W. Baik, "Basis pursuit denoising-based image super resolution using a redundant set of atoms," *Signal, Image and Video Processing*, vol. 10, no. 1, pp. 181–188, 2016.
- [35] M. Hussain, A. W. A. Wahab, Y. I. B. Idris, A. T. Ho and K. H. Jung, "Image steganography in spatial domain: A survey," *Signal Processing: Image Communication*, vol. 65, pp. 46–66, 2018.
- [36] P. Agarwal, D. Moudgil and S. Priya, "Encrypted transfer of confidential information using steganography and identity verification using face data," in *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, Singapore: Springer, pp. 155–166, 2020.