

# Steganographic Technique for Hiding Secret Audio in an Image

<sup>1</sup>Aiswarya T, <sup>2</sup>Mansi Shah, <sup>3</sup>Aishwarya Talekar, <sup>4</sup>Pallavi Raut

<sup>1,2,3</sup>UG Student, <sup>4</sup>Assistant Professor, <sup>1,2,3,4</sup>St John of Engineering & Management, Mumbai, India.

**Abstract** - Information security is one of the most important factor which is to be considered when secret information has to be communicated between two parties. Cryptography and steganography are the two techniques used for this purpose. Cryptography scrambles the information, but it reveals the existence of information. Steganography hides the actual existence of information so that anyone else other than the sender and recipient cannot recognize the transmission. In steganography the secret information to be communicated is hidden in some carrier such a way that the secret information is invisible. In our paper an image steganography technique is proposed. Audio signal is hidden in the transform domain of a cover image using wavelet transform. The audio signal in any format (MP3 or WAV) is encrypted and carried by the image without revealing the existence to anybody. When secret information is hidden in the carrier the result is stego signal. The quality of stego image is measured by Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE).

**Keywords** — DWT, MSE, PSNR, Steganography, Transform domain, WAV.

## 1. INTRODUCTION

The internet allows for easy spreading of information over large areas. This is both a blessing and curse, since friends all over the world can view the information but, so can everyone else. Encrypting data has been the most popular approach for protecting information, but this protection can be broken with enough computational power. An alternate approach to encrypting data would be to hide it by making these information look like something else. This way, only the sender and the receiver of the data would realize the true content. In particular if the important data is hidden inside an image then everyone would view it as a picture. At the same time the receiver could still retrieve the true information. This technique is often called data hiding or Steganography. The term steganography has been derived from two Greek words Steganos and graphia, where “steganos” means covered or secret and “graphic” means writing or drawing. Using steganography, information can be hidden in carriers such as images, audio files, text files, videos etc.

In this paper we proposed a new framework of an image steganography system to hide an audio sample. We primarily concentrated on the data security issues when sending the data over network using steganography techniques. The main objective of our paper is that the hidden message carried by stego media should not be sensible to human beings. The other goal is to avoid drawing uncertainty to the existence of

a hidden message or data. This approach of information hiding technique has recently become important in many application areas.

LSB (Least Significant Bit) [1] is one of the most common and easiest methods for message hiding. In this method, a message is hidden in the least significant bits of image pixels but due to its simplicity unauthorized user can easily guess and change the LSB's of the image pixels, thus original message gets destroyed. It also causes some distortion in the original image and Scaling, rotation, cropping, addition of noise, or lossy compression to stego image will destroy the secret message. Another method is masking and filtering [2] which are more robust than LSB in terms of some image processing algorithms like compression, cropping etc. which makes it suitable in lossy JPEG images. Masking images involves changing the luminance of the masked area. But the drawback in this method is that masking and filtering is mostly used for 24 bit and Grey scale. In Grey Level Modification [3] method it embeds information by modifying the grey level values of the grey scale image pixels. As it uses a concept of even and odd pixels, so the drawback associated with this scheme is that the noise, due to hardware imperfection or transmission medium, can make the odd pixels to even and vice versa resulting in annihilation of bit stream. In DWT [4] the transform decomposes the signal into mutually orthogonal set of wavelets. It decomposes the

image into four sub bands which ultimately creates more space for the secret information to be hidden and even improves the security. Encryption can also be performed to increase the level of security. The main constraint of our paper is if the audio file is of greater length then it becomes difficult to hide it in a cover image because the size of the cover image is less as compared to that of the large audio file. For resolving this problem we came up with a technique of initially compressing the audio using DCT (Discrete Cosine Transform), followed by encryption and then hiding it using DWT.

## II. METHODOLOGY

### 2.1 Audio compression using DCT

A discrete cosine transform (DCT) expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. Audio compression is data compression designed to reduce the transmission bandwidth requirement of digital audio streams and the storage size of audio file.

Compression is of two types:

**1. Lossless compression:** This algorithm mainly deals with representing the sender's data more precisely without an error. Lossless compression schemes are reversible so that the original data is reconstructed. This method is used for text compression. Compression ratio is low for this method.

**2. Lossy compression:** This scheme accepts some loss of data in order to achieve higher compression ratio. This method is used for audio, image and video compression.

We proposed audio compression using DCT which is a lossy compression technique as it removes certain frequencies from audio data such that the size is reduced with reasonable quality. DCT is a first level approximation to mpeg audio compression. It takes a wave file as input, compress it to different levels or factors and assess the output of each compressed wave file. It gives an idea to hide audio signal in an image. Image is of .jpg, .bmp format whereas audio signal is .wav, .mp3 etc.

Steps for performing Audio compression using DCT:

- 1) Read the audio file.
- 2) Determine a value for the number of samples that will undergo a DCT at once that is the audio vector will be divided into pieces of this length.
- 3) Consider different compression rates and observe the compressed output.

### 2.2 Discrete Wavelet Transform

In numerical analysis and functional analysis, a discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discretely sampled. As with other wavelet transforms, a key advantage it has over Fourier transform is temporal resolution: it captures both frequency and location information (location in time).

In this method a steganography technique to hide an audio sample in a cover image is done. Image can be in any format like .jpg, .bmp etc. and audio also can be in any format like .wav, .mp3 etc. Since audio files contain large no. of samples even for small duration, the cover image has to be considerably large. Color images are suitable because of enough hiding space. Since YCbCr approach is more secure than RGB approach, YCbCr approach is used. The cover image is converted to YCbCr. Then Cb, Cr components and secret audio signal are transformed using LWT (Lifting Wavelet Transform). The first two bits of audio are swapped with last two bits of image component.

Steps for performing Audio hiding using DWT

- 1) Convert RGB image to YCBCR.
- 2) Take one band out of four for further processing.
- 3) Convert secret audio and cover image into 8 bit data.
- 4) Transfer the first two bits of audio to last two bits of image.
- 5) Combine all the bands, observe the stego image and calculate the performance parameters.

## III. RESULTS AND DISCUSSION

### 3.1 Characteristics of Steganography

**1) Robustness-** It refers to the ability of secret data to remain untouched if the stego image undergoes transformations such as linear and non-linear filtering, sharpening, blurring, addition of random noise or lossy compression.

**2) Imperceptibility:** It means invisibility of steganography algorithm.

**3) Payload Capacity:** It refers to the amount of secret information that can be hidden in the cover source.

**4) Security:** It is the measure of un-detectability. It is also the measure of the quality of a signal. For images it is measured in terms PSNR, MSE, SPCC, CQM etc.

### 3.2 MSE: Mean Square Error

It is defined as the cumulative squared error between the compressed and the original image.

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2 \quad \text{--- (1)}$$

### 3.3 PSNR: Peak signal-to-noise ratio

It is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation.

$$PSNR = 10 \times \log_{10} \left( \frac{MAX^2}{MSE} \right) \quad \dots (2)$$

### 3.4 CQM: Color Image Quality Measure

$$CQM = (PSNR_Y \times R_W) + \left( \frac{PSNR_U + PSNR_V}{2} \right) \times C_W \quad \dots (3)$$

Where  $PSNR_Y, PSNR_U, PSNR_V$  are the PSNR values of Y, U, V components of the color image respectively.  $C_W$  and  $R_W$  are the weights on the human perception of cone and rod sensors respectively. In HVS cones are responsible for chrominance perception and rods are responsible for luminance perception.  $C_W = 0.0551$  and  $R_W = 0.9449$  as specified by HVS. CQM greater value indicates greater image similarity. It is represented in dB.

### 4.5 SPCC: Squared Pearson Correlation Coefficient

It measures the similarity level between two signals. The higher the SPCC, the higher is the similarity level. Its range is between 0 and 1.

$$SPCC = \left[ \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \right]^2 \quad \dots (4)$$

Where  $x_i$  and  $y_i$  are the two signals,  $\bar{x}$  and  $\bar{y}$  are their averages.

- For Audio Compression

Parameter like PSNR and MSE were calculated and compared those with audio samples before and after compression. From Table 1 as shown below we can see that as the compression factor increases it results in reduction of mean square error and an increase in peak signal to noise ratio which means more the compression factor better is compression as shown below in Figure 2, providing audio sample without any loss in the data. For retrieving the original secret audio inverse of DCT is applied to compressed data. And it's observed that there is no much loss of audio sample and it shows a good similarity with the original secret audio. This way we require very less space to save that audio.

Table 1: Length, MSE and PSNR after different factors of Compression

Parameters	Audio samples before compression	Audio samples after compression		
		Factor 2	Factor 4	Factor 8
Length	264600 × 2	1×262144	1×262144	1×262144
MSE	0.184	0.0189	0.0187	0.0178
PSNR	65.4846	65.3699	65.4027	65.6217

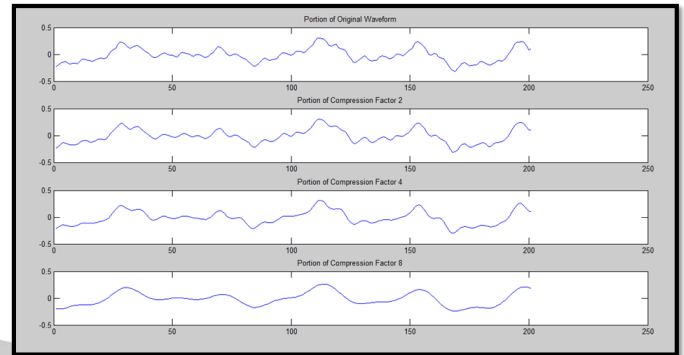


Figure 1: Audio Samples with compression factor 2, 4 and 8

- For Audio Hiding

This algorithm is tested by taking different color images of size 512 X 512 and one secret audio sample. While extracting, inverse wavelet transformation is performed. Since jpeg format is the most commonly used format, jpeg images are considered. There is no influence of the image format on the performance evaluation metrics because both cover and stego images will be in the same format and data hiding is done in the transform domain.



Figure 2: Original Image

In the above Figure 4 the given image is considered as the cover image (**wpeppers**) on which we are going to perform various steganography techniques. This RGB image consists of combination of hue, saturation, intensity along with luma and chroma components.

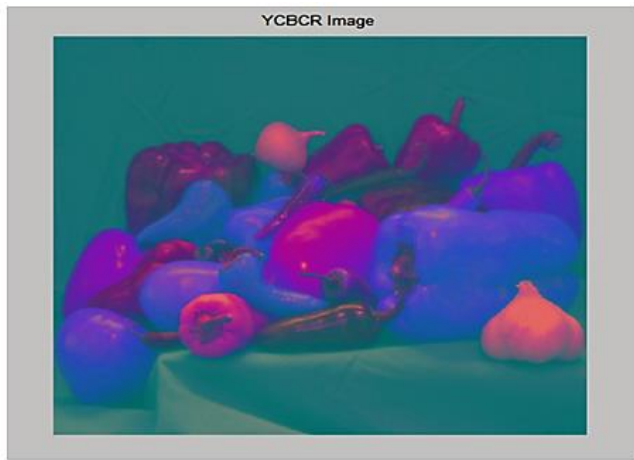


Figure 3: YCBCR Image

The image in Figure 4 is converted to YCBCR as shown in Figure 5 by using MATLAB inbuilt function. This conversion is done due to the fact that any operation done on RGB image will degrade its quality so in replacement of that we are converting it into YCBCR.

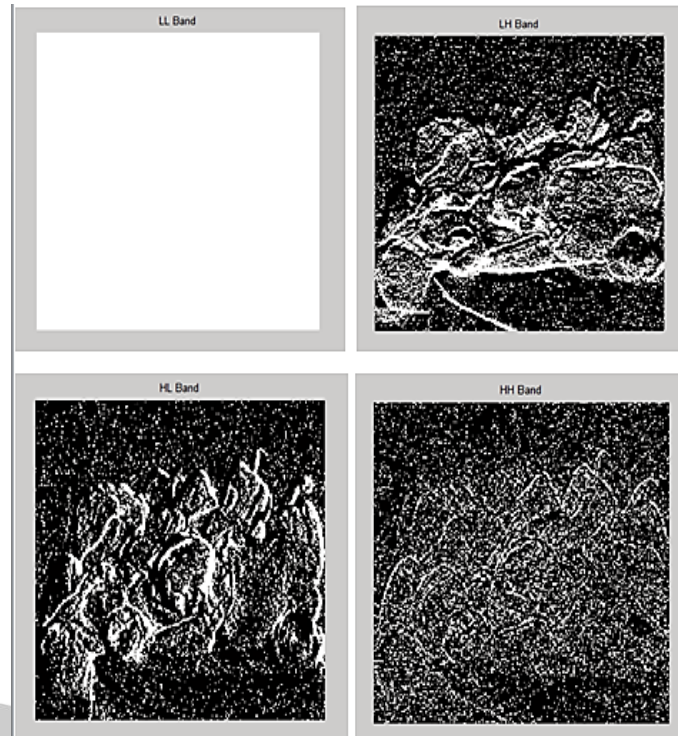


Figure 5: [LL, LH, HL, and HH] bands of CB component

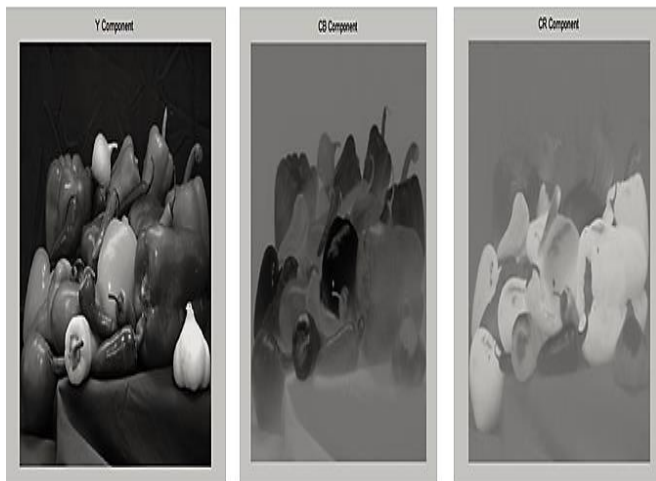


Figure 4: Separation of Y, CB, CR components

From Figure 5 we got three separate components as shown in Figure 6. The Y component indicates the luminance part of an image. It represents the achromatic image. CB and CR represent the colour information i.e the chrominance part of an image.

From Figure 6 we have obtained the 4 bands of CB component named [LL, LH, HL, and HH] as shown in Figure 7.

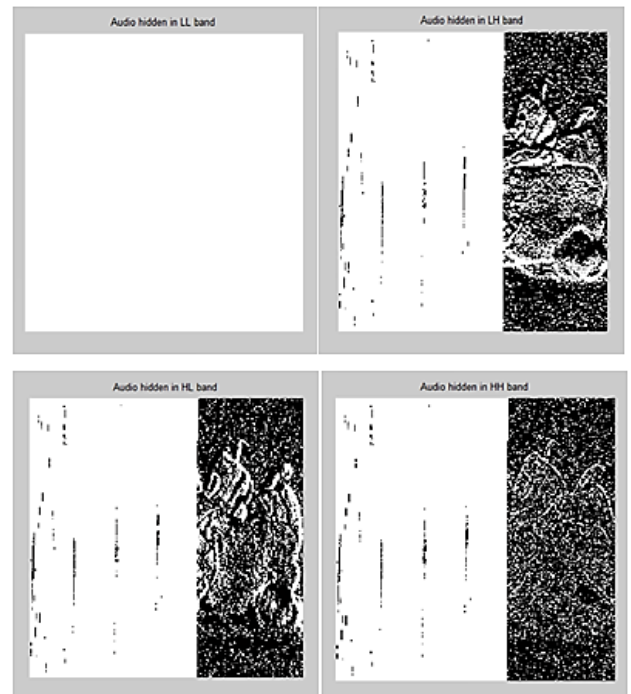


Figure 6: Audio encrypted in [LL, LH, HL, and HH] band

Figure 8 shows final sub bands of CB components with audio encrypted in its LSB (Least Significant Bit). The left hand side portion of [LH,HL,HH] shows the distorted portion which is due to encryption of audio into it.



Figure 7: Stego images consisting of audio in different sub bands

In Figure 9 it is seen that distortion of image occurs when audio is encrypted and hidden in LL band. This is due to the fact that LL band of image component comprises of important information so any kind of changes made in LL band will result in deformation of final image. This way an intruder can easily detect the presence of secret information. Meanwhile when audio is hidden in LH band, the level of distortion decreases. Finally when audio was checked with HH band, it showed very less deformity as HH band contains less important data so any kind of changes made did not affect the final image much. This way we concluded for using HH band to hide our secret audio samples in it.

Table 2 shown below shows different set of images that have undergone wavelet decomposition. It's observed that for a same audio sample we get different values of MSE and PSNR. The mean square error ranges from 1 to maximum up to 5 and it's observed that values more closely to zero are considered to be better. PSNR is the ratio between maximum power of a signal and the power of corrupting noise that affects fidelity of its representation. It ranges from 40,000 to 50,000 and it's observed that higher PSNR value generally indicates that reconstruction is of higher quality. So we can conclude that Image 6 is better for transmission.

Table 2: MSE and PSNR of stego image.

Cover Image	Audio Sample	MSE	PSNR
Wpeppers.jpg	40,231	2.5784	44.0173
Image1.jpg	40,231	5.6139	40.6381
Image2.jpg	40,231	4.3421	41.7538
Image3.jpg	40,231	4.9675	41.1694
Image4.jpg	40,231	3.2775	42.9754
Image5.jpg	40,231	5.5435	40.6929
Image6.jpg	40,231	1.2485	47.1669
Image7.jpg	40,231	5.3887	40.8159
Image8.jpg	40,231	5.8899	40.4297



Figure 8: Different images on which algorithm was tested

Figure 11 shown below is the final output of **Wpeppers** image. It has got CB component from which HH band was extracted using wavelet decomposition. The audio sample was initially encrypted using swapping method. The audio bits and image bits were swapped and combined to get final stego image. We then extracted the audio file by doing inverse procedure and it was observed that extracted audio and image showed a good similarity with original audio and image.

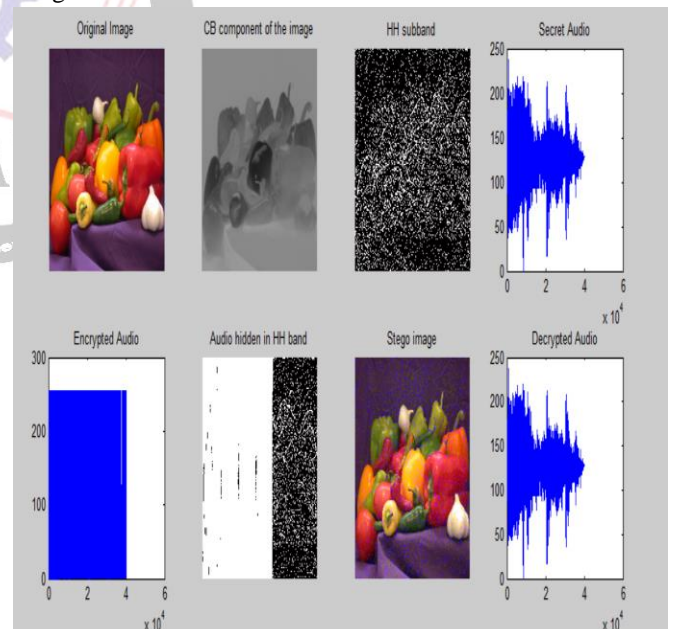
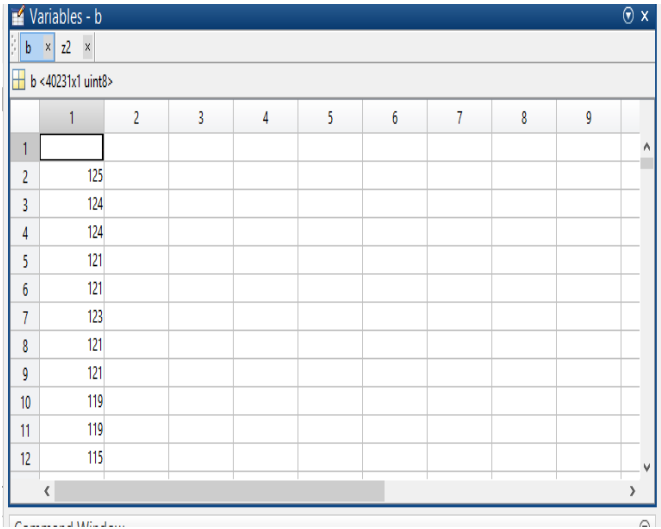


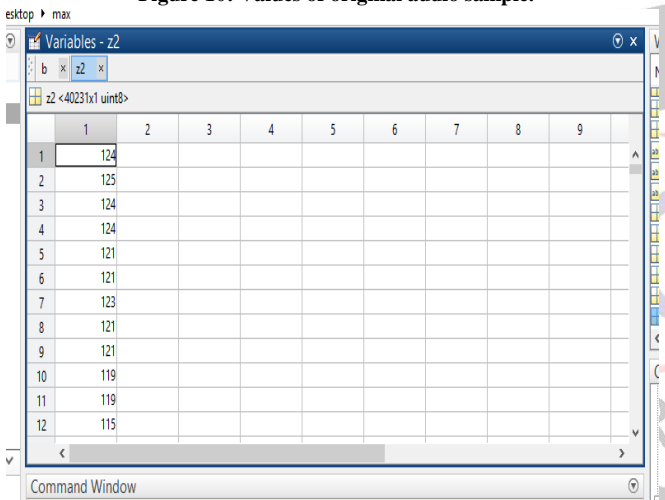
Figure 9: Final output consisting of Original Image, CB component, HH band, Secret Audio, Encrypted Audio, HH band after encrypting audio in it, Stego Image and Decrypted Audio

When the audio sample was decrypted, values of original audio sample and decrypted one remained same indicating that there is no loss of data while transmission. It is shown in Figure 12 and 13 below.



	1	2	3	4	5	6	7	8	9
1									
2		125							
3		124							
4		124							
5		121							
6		121							
7		123							
8		121							
9		121							
10		119							
11		119							
12		115							

Figure 10: Values of original audio sample.



	1	2	3	4	5	6	7	8	9
1	124								
2		125							
3		124							
4		124							
5		121							
6		121							
7		123							
8		121							
9		121							
10		119							
11		119							
12		115							

Figure 11: Values of decrypted audio sample.

#### IV. CONCLUSION AND FUTURE WORK

As experimental result shows, it has been observed that steganography technique plays a dominant role. The wavelet decomposition proved to be an efficient, well organized method in image steganography. Initially we converted RGB image to YCBCR and took one component of that image and performed wavelet decomposition. We achieved greater similarity and uniformity between original and stego image and unauthenticated users will fail to identify between zero content image and important information content image as our aim was to hide the secret audio in the least information band (i.e. HL) of image component. Before hiding we encrypted the audio for better security. And there is a good resemblance between encrypted and decrypted audio. We

even calculated performance parameters like PSNR and MSE which helped us in choosing the efficient method.

The future scope of our project is to have secure communication which can be even done by using different hiding algorithms like Diffie-Hellman, Elliptical Curve Steganography, RSA, Asymmetric key algorithm. The quality of the stego image can also be measured by Structural Similarity Index Metric (SSIM), Universal Image Quality Index (UIQI), Color Image Quality Measure (CQM) etc.

#### REFERENCES

- [1] LSB method by Neil F Johnson and Sushil Jajodia on “EXPLORING STEGANOGRAPHY: Seeing the Unseen”.
- [2] Masking & Filtering method by Amandeep Kaur, Rupinder Kaur and Navdeep Kumar on “A Review on Image Steganography Techniques”.
- [3] Gray level modulation method by Ahmad T. Al-Taani and Abdullah M. AL-Issa on “A Novel Steganographic Method for GrayLevel Images”.
- [4] DWT method by Hemalatha S on “Wavelet transform based steganography technique to hide audio signals in image”,
- [5] Introduction on steganography by Rashi Singh and Gaurav Chawla on “A Review on Image Steganography”
- [6] Image Steganography method by Vikranth.B.M on “A SURVEY OF IMAGE STEGANOGRAPHY”
- [7] DWT method by Shikha Choudhary and Chaten Panwar on “Key Based Image Steganography using Dwt and Chaotic Map”
- [8] Parameter analysis by Adel Almohammad on “Steganography-Based Secret and Reliable Communications: Improving Steganographic Capacity and Imperceptibility” .