Iraqi Journal of Physics, 2021 *Vol.19, No.49, PP. 53-61*

DOI: 10.30723/ijp.19.49.53-61

Studying Audio Capacity as Carrier of Secret Images in Steganographic System

Ahlam M. Kadhim, Huda M. Jawad

Department of Physics, College of Science, Mustansiriyah University, Baghdad, Iraq

E-mail: drhda222@uomustansiriyah.edu.iq

Corresponding author: ahlammajead@uomustansiriyah.edu.iq

Abstract

Steganography art is a technique for hiding information where the unsuspicious cover signal carries the secret information. Good steganography technique must include the important criterions e.g. robustness, security, imperceptibility and capacity. Improving each one of these criterions will effect the others, because these criterions are overlapped. In this work, a good high capacity audio steganography safely method has been proposed based on LSB random replacing of encrypted cover with encrypted message bits at random positions. The research also included a study for the audio file, speech or music, by safely manner to carrying secret images, so it is difficult for unauthorized persons to the suspect presence of the hidden image. Calculations of SNR, SNR segmental, SNR spectral, MSE and correlation show that, audio music cover file (2channales) is the safest as a carrier with replace the 9 number of LSB without noticeable noise. The capacity of the audio file that can be safely exploited is up to 28% of the total size of the music audio cover; this fact can be noticer from the values of measures of SNR, SNR_{Seg} and SNR_{Spec} (32, 28 and 31 dB). For speech cover audio the replacing LSB is safely uses LSB bits number 6, where the hiding capacity reach up to 37 % of size speech cover audio at 37, 36 and 39 dB for three type's measures of SNR. Correlation of cover samples was not affected as a result of hiding secret image, where its value is up to 0.99 for all hiding operations.

Key words

Audio Steganography, Secret Key Encryption, LSB.

Article info.

Received: Nov. 2020
Accepted: Dec. 2020
Published: Jun. 2021

دراسة سعة ملف الصوت كحامل للرسائل السرية في نظام الاخفاء

احلام مجيد كاظم، هدى محمد جواد

قسم الفيزياء، كلية العلوم، الجامعة المستنصرية، بغداد، العراق

الخلاصة

فن إخفاء المعلومات هو تقنية لإخفاء المعلومات حيث تحمل إشارة الغطاء غير المشكوك فيها المعلومات السرية. يجب أن تتضمن تقنية إخفاء المعلومات الجيدة المعايير المهمة، المتانة والأمان وعدم الإدراك والقدرة. إن تحسين كل معيار من هذه المعايير يؤثر على المعايير الأخرى ، بسبب تداخل هذه المعايير مع بعضها البعض.

في هذا العمل، تم اقتراح أسلوب إخفاء صوتي جيد عالي السعة بأمان استنادًا إلى استبدال LSB العشوائي للغلاف المشفر بتشفير بتات الرسائل في مواضع عشوائية. كما اشتمل البحث على قدرة على دراسة الملف الصوتي أو الكلام أو الموسيقى بطريقة آمنة لحمل الصور السرية، لذلك يصعب على غير المصرح لهم الشك في وجود صورة مخفية. تُظهر حسابات مقاييس SNR و SNR المقطعية و SNR الطيفية و MSE والارتباط أن ملف غلاف الموسيقى الصوتية (2 قناة) هو أكثر الاستخدامات أمانًا كحاجز مع استبدال الرقم 9 من LSB بدون ضوضاء ملحوظة. يمكن أن تخفي أجزاء الرسائل السرية ما يصل إلى 28٪ من إجمالي حجم صوت الغلاف الموسيقي ومقاييس الأنواع الثلاثة لـ SNR 32 و 28 و 31 ديسيبل بالنسبة لصوت غطاء الكلام، يتم استخدام LSBالبديل بأمان في LSB بت رقم 6، حيث تصل سعة إخفاء الصوت إلى 37 ٪ من حجم صوت غطاء الكلام عند 37 و 36 و 39 ديسيبل لثلاثة مقاييس من النوع .SNR لم يتأثر ارتباط عينات الغلاف بإخفاء الصورة السرية حيث تصل قيمتها إلى 0.99 لجميع عمليات الإخفاء.

Introduction

Nowadays in the information era techniques, information penetration can be avoided by achieving steganography and /or cryptography. Steganography is the process of including confidential messages in the cover tag to reduced illegal detection [1]. The difference between Steganography and cryptography is in terms of message vision. The secret message in Steganography is totally hidden compared with cryptography where the secret message is still visional [2].

Steganography is frequently used in secret communication like government and military communications. The main requirements that must be met for good steganography algorithms include the perceptual transparency, durability and payload or capacity [3]. High capacity is an important aspect of Steganography compared with watermark. For watermarking, durability must be a dominant factor. Optimizing one of the mentioned requirements will weaken the other's performance due to their discrepancy according to Magic Triangle [4]. In recent years several techniques algorithms have been developed to hide information [5-7].

Most of these technologies use either video or image media but audio signal is scarcely used as a cover signal special at the high rate of data hiding, most likely due to the human auditory system (HAS) which is more sensitive compared to the human system (HVS) [8]. Nevertheless, the adoption of sound signals as a carriers signal may result in inferior inaudible execution; there are still appropriate features like unpredictability and transmission which make the audio signal as a suitable safe cover signal.

Generally, watermark and sound audio steganography can be distributed according to embedding domain either in transformation domain or time domain. Cryptography encrypts the message in a way that makes it incomprehensible, while cryptography hides a secret message signal in the cover signal without attracting the attention. Sending an encrypted message might lead to suspicion of eavesdropping, while message hidden in the cover signal does not. However, these two technologies can be combined to obtain a message protection in higher level [9]. Steganography differs from cryptography that it takes advantage of the perceptual limitations of the HAS or HVS, which fails to recognize the difference between the cover signals and the Stego signals [10]. The steganography method often uses media files such as video, images, or audio as host signals to embed the secret data. Generally, the use of the audio signal as cover signal is less common than the image, because of the HAS is more susceptible to hear the noise signal than the HVS [10].

In this paper we proposed a new embedding algorithm with high capacity and high output quality. The proposed algorithm is based on embedding an encrypted message within encrypted sound cover with using less significant bits (LSB). The sound stego-signal has Signal to Noise Ratio (SNR) above of 37 dB of speech sound for the input capacity 37 % with occupied 6 LSB in safely manner. For the music sound the safe

occupied LSB reaches to 9 bits for the same image which is hidden in speech sound with capacity 28 % and SNR above of 38 dB. The message extraction process does not need an original sound cover signal and this algorithm also has good security because of the randomization of bits blocks of secret messages in the hiding process instead of distributing the embedded messages as the known replacement positions in many stego algorithms.

Image encryption techniques

For the time being, there are many of encrypted algorithms like arnold map, tangram algorithm, baker's transformation, magic cube transformation, and affine transformation etc [11]. In some different available cryptography techniques algorithms, secret key of encryption technique are not safely and are ineffectively separated. This does not support the demined requirements for the cryptographic technique and are exposed to several attacks by any individual. It is in demand to develop an effective secret key for any new cryptographic system, particularly for real time domain secure communication of color image over an open internet networks.

To apply this, different kinds of image cryptography system have been proposed. One of these is the theory based on choice algorithm that has suggested a new and efficient method for heading these problems by encrypted color images at a very fast and secure level [12]. The main properties of chaotic dynamical systems like mixing property, sensitivity to initial conditions , periodicity, system parameters that can be considered analogous to some effective cryptographic features such as diffusion, confusion , avalanche properties and balance [13, 14]. One of the effective chaotic techniques is the Logistic map which has been recently at last for cryptographic systems. The equation of logistic map is the following [14-16]:

$$x_{n+1} = r\, x_n(1-r) \qquad (1)$$

Values of x_n are in the range (0, 1), parameter r is a positive number taking value up to 4. The value of r is determined and explores the behavior of the logistic map. The cryptographic system varied features with varies values of r which is called bifurcation parameter as shown in Fig.1 where the horizontal axis shows the values of the parameter r and the vertical axis shows the x_n values.

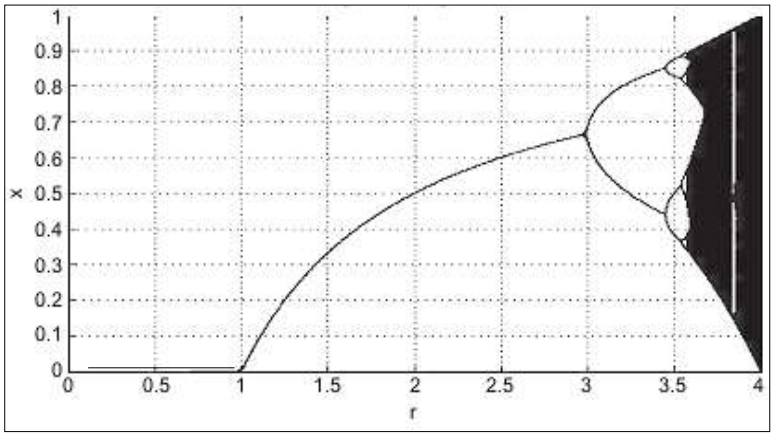

Fig.1: Bifurcation for the logistic map.

Capacity calculations

The main problem in most steganographic systems is the designing system which achieves the balance between the requirements that are necessary to get the successful hiding technique. If there is a small size message embedded within audio signal, the message may not be exposed to attack by unauthorized persons. But when the size of image is large, embedding process may generates noise which raises suspicion about the carrier cover, which leads to easy extraction of the confidential image. So, there is an urgent need to a steganographic system that achieves balance between the three requirements which are undetectability, robustness and capacity.

If times of sound audio cover, frequency, and recording channels are T sec, F Hz and Ch respectively, then the size of cover (Siz_C) is expressed mathematically as follows:

$$\boldsymbol{Siz_C = T * F * Ch} \qquad (\mathbf{2})$$

Total segments number of audio cover is calculated as follows, where segment length is (Seg):

$$\boldsymbol{NSeg} = \frac{\boldsymbol{Siz\_C}}{\boldsymbol{Seg}} \qquad (\mathbf{3})$$

The total size cover in bits ($\boldsymbol{Siz\_CB}$) is expressed mathematically as follows, where B (bits /sample) is bits resolution:

$$\boldsymbol{Siz\_CB = Siz\_C * B} \qquad (\mathbf{4})$$

In this study the Seg is equal to 216 samples by replacing selected number of LSB the size of embedded secret message in bits (Siz_MB) is calculated as follows:

$$\boldsymbol{Siz\_MB = Seg * LSB} \qquad (\mathbf{5})$$

Assume that the secret color image have equal number of rows (R) and columns (C) the size of secret color image in pixels is calculated as follow:

$$\boldsymbol{Siz\_M = R * C * 3} \qquad (\mathbf{6})$$

If the bits resolution of image is 8 bits / pixel, the dimensions of embedded secret image can be fined as follows:

$$\boldsymbol{R = C} = \sqrt{\frac{\boldsymbol{Siz\_MB}}{\mathbf{3 * 8}}} \qquad (\mathbf{7})$$

The bits rat for the audio cover can be calculated as follow:

$$\boldsymbol{Bit\ Rat}\ \% = \frac{\mathbf{100} * \boldsymbol{Siz\_MB}}{\boldsymbol{Siz\_CB}} \qquad (\mathbf{8})$$

The relation between cover size and message size identifying the capacity of steganographic system as in the following equation:

$$\boldsymbol{CP}\ \% = \frac{\mathbf{100} * \boldsymbol{Siz\_M}}{\boldsymbol{Siz\_C}} \qquad (\mathbf{9})$$

In this paper, capacity of audio cover of 4 minutes time period has been calculated using Eq.(9). The relation between the message size and audio cover size was calculated by assuming that all segments of audio cover where used.

Proposed hiding method

The main structure for the proposed hiding method is shown in Fig.2. The figure illustrates the three main stages after preprocessing the input audio cover and secret images message. Cover audio signal was segmented to (NSeg) segments; each segment has a length of 216 samples. For secret message preprocessing, the secret image was also segmented to (Seg) bits segments of equal length; each bits segment

is reshaped as a 1-D binary array of 216 bits. The third stage was repeats a number of times to embed each secret message segment within cover segment at random samples positions of audio cover. The audio frequency of the audio cover was 8000 Hz/sec for the speech audio and music audio. Tow speech audios were recorded using 1 channel for female and male speakers of 4 minutes period time. Single recording channel and bit resolution for digital representation was 16-bit per audio sample. The hiding method was tested for 8 color images.

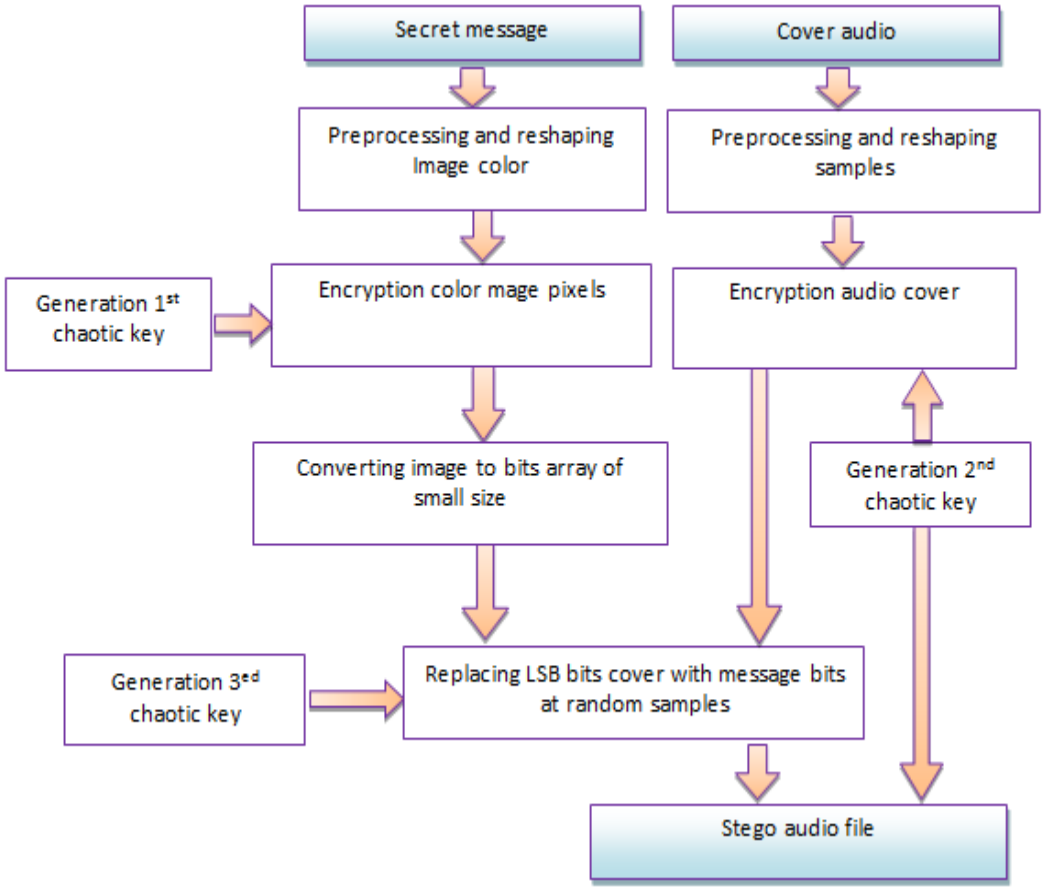

Fig.2: General block diagram of the proposed method.

The embedding proposed method

In this paper, encrypted color image has been hidden and processing using chaotic key produced an encrypted cover which is using another secret key to select sample position of replaced bits. The generation technique of secret keys number was used to mix pixels of image respectively. By changing x secret images and samples of sound cover, this process increases the security levels of the proposed technique.

Fig.3 shows the histogram of original, encrypted and stego audio that has been encrypted using 2^{nd} secret key number, where initial value of secret key was (x=0.6716) and control parameter (r=3.7567). N is the key length which is equal to the length of audio cover or total size (0.00000001 to 1.0) and r (0.3000000 to 4.0), the number of keys can be increase to the value $(10^7)^6$.

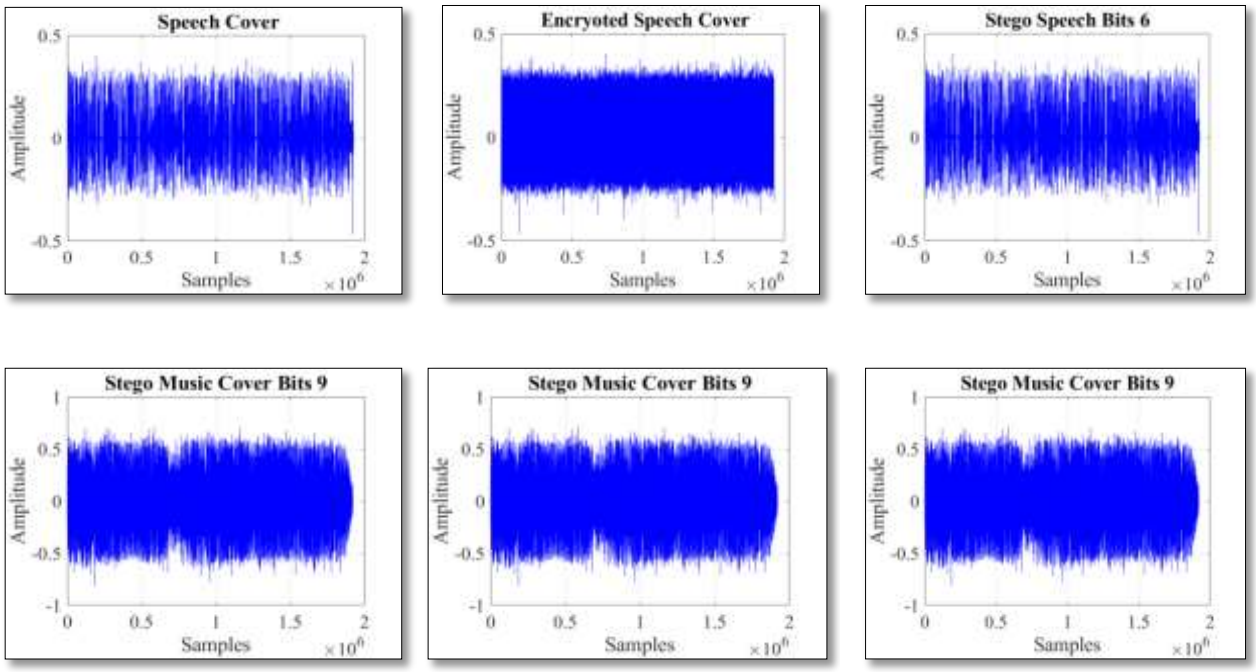

Fig.3: The original, encrypted and stego audio cover signals.

The embedding and extracting are the two algorithms for steganography technique. Fig.3 shows histogram of original, encrypted and stego image. It is noticed that from encrypted image histogram, that the image cannot be identified, if it was viewed by the naked eye. The extracted image after using extraction algorithm is very similar to the original image before embedding within sound-carrying file. The encrypted image will not be decrypted if the key matrix is not available. To reconstruct the original image, the users should be provided with security NEK which help them to obtain the correct color images. The proposed technique was performed using the following algorithm:

Procedure

Step 1: Reading audio cover.

Step 2: Input and reading secret color image that will be hidden within audio cover.

Step 3: Separate secret image bands RGB. Reshaping each band to 1-D array and combining the three RGB in 1-D array.

Step 4: Encryption the output 1-D array of step 3 using 1st secret key, by mixing the pixels at random position.

Step 5: Encryption cover audio using 2nd secret key.

Step 6: Converting encrypted image pixels of image to their values of digital binary form, and reshaping the result to 1-D array and then to small segments of length 216 bits.

Step 7: Chose the number of LSB audio cover that will be replace in each samples segment with secret image bits.

Step 8: Replace image segments bits with LSB bits of segment samples cover (length 216) at random samples positions using 3^{ed} secret key.

Step 9 End: Decrypted carrier audio cover using the same 2nd secret key, to get stego audio.

The extraction process is performed to reconstruct the color image in reverse manner to that of the embedding algorithm. The stego audio is encrypted by samples mixing of the carrier audio using the 2nd secret key, and then the embedded bits are extracted from random samples positions using the 3ed secret key. When the bits extraction is complete, the secret image can be obtained, but it is an encrypted image. The correct arranging of image pixels is restored using the 1st secret key.The sender and receiver must agree on the following in order to properly recover the hidden image:

1. The correct numbers for the three secret keys.
2. The replacement door number with the hidden image bits
3. The size of the secret image hidden inside the audio cover.

Fig.4 shows the original secret image and encrypted image resulted after using secret key (NEK). The encrypted image was embedded within audio cover, and then extraction process was performed to obtain encrypted image. Fig.4 shows the image after preforming decryption operation for the extracted image using the same key (NEK).

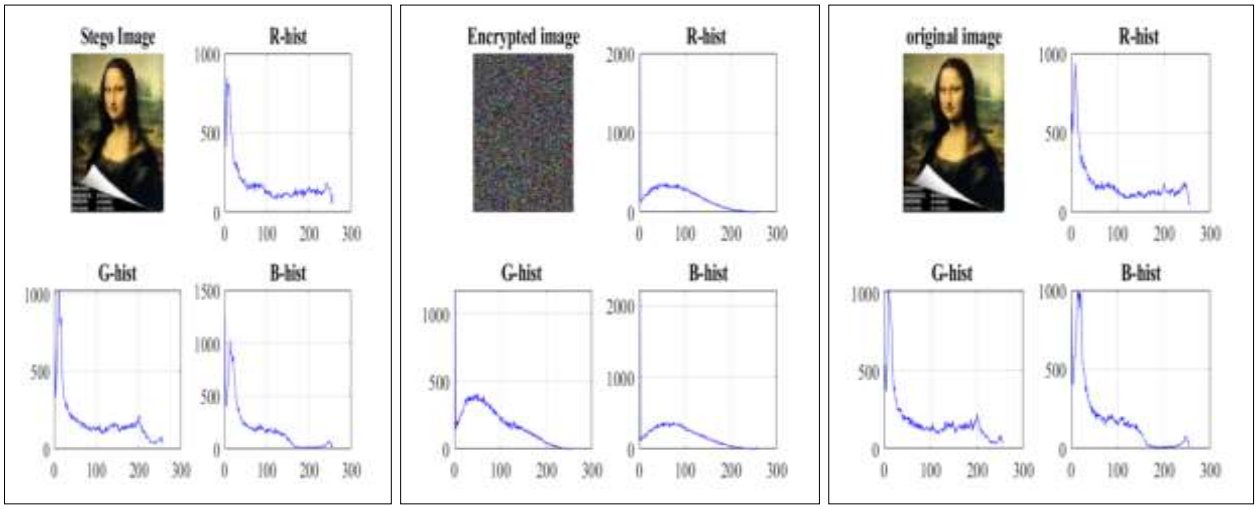

Fig.4: The original, encrypted and stego secret image to be hidden within cover audio.

Embedding capacity tests

Perceptual quality is defined as the inability of human hearing to perceptible variation in the cover sound signal before and after hiding information. Perceptual quality can be measured either by listening to both the cover sound file and stego sound file (cover after message embedding) by several persons who have excellent hearing, or with using mathematical measures. The good mathematical method to measure a quality signal is to calculate the Signal- to - Noise Power Ratio (SNR) by regard to the difference between the cover signal file and the stego signal file as a noise as in the following equation [17]:

$$\mathbf{SNR} = \mathbf{10log_{10}} \frac{\sum_{i=1}^{x} \mathbf{C_i}^{\mathbf{2}}(\mathbf{j})}{\sum_{i=1}^{x} [\mathbf{C_i}\,(\mathbf{j}) - \grave{\mathbf{C}}_{\mathbf{i}}\,(\mathbf{j})]^{\mathbf{2}}} \qquad (\mathbf{10})$$

where C and C' are the cover and the stego signals respectively, x is he number of samples in each segment, and Seg is the number of segments.

In this research, the audio cover file was divided to a number of segments, SNR_{Seg} was calculated to each segment as in Eq. (11).

$$\mathbf{SNR_{Seg} = 10log_{10}\frac{\sum_{i=1}^{x}\sum_{j=1}^{seg} C_i{}^2(j)}{\sum_{i=1}^{x}\sum_{j=1}^{seg}[C_i(j) - \grave{C}_i(j)]^2}} \quad (11)$$

where: Seg the total number of segments that were used to embed message. SNR_{spec} was calculated in frequency domain after transforming the audio signal file using the Fourier transform, and then the audio file was divided to segments and the SNR_{Spec} was calculated using Eq.(11). The proposed algorithm is implemented by using matlab program 2018a. Two audio cover signals were used for testing using the proposed algorithm: music and female voice. The measures quality are illustrated within Tables 1 and 2. The quality of stego audio signal in each test was computed using three type measures SNR, SNR_{Seg}, andSNR_{Spec}.

Table 1: Capacity calculation of the speech (T=4 min, F=8000 Hz, ch=1, B=16).

LSB	Bit Rat	CP %	SNR	SNR_{Seg}	SNR_{Spec}	MSE	Corr
1	3.6450	7.2900	47.3753	67.1048	70.0234	e-084.0780	0.9999
2	7.2900	14.5800	47.2615	60.4970	63.4869	e-084.1862	0.9994
3	10.9350	21.8700	46.8274	54.2379	57.2310	e-084.6262	0.9974
4	14.5800	29.1600	45.4321	48.4184	51.4602	e-08 6.3790	0.9933
5	18.2250	36.4500	42.1043	42.4541	45.5065	e-071.3726	0.9837
6	21.8700	43.7400	37.1000	36.3609	39.3944	e-07 4.3448	0.9635
7	25.5150	51.0300	31.3190	30.5889	33.6151	e-061.6446	0.9341
8	29.1600	58.3200	25.1763	24.8257	27.7310	e-06 6.7661	0.9163
9	32.8050	65.6100	20.9945	19.8710	22.9691	1.7722 e-05	0.9029
10	36.4500	72.9000	20.9926	19.8745	22.9128	1.7730 e-05	0.9082
11	40.0950	80.1900	8.9516	7.5384	10.1215	2.8367 e-04	0.8029
12	43.7400	87.4800	2.8931	1.6170	3.8440	0.0011	0.7400
13	47.3850	94.7700	4.6429-	3.157-	3.0734-	0.0046	0.6836

Table 2: Capacity calculation of the music audio (T=4 min, F=8000 Hz / Sec, ch=2, B=16).

LSB	Bit Rat%	CP %	SNR	SNR_{Seg}	SNR_{Spec}	MSE	Corr
1	1.8225	3.6450	82.5809	77.5080	80.4248	e-101.3580	1.0000
2	3.6450	7.2900	75.5205	70.3458	73.2149	e-106.9013	1.0000
3	5.4675	10.9350	69.3220	64.1573	67.0726	e 092.8760	1.0000
4	7.2900	14.5800	63.2434	58.0379	61.0121	e-081.1658	1.0000
5	9.1125	18.2250	57.2033	52.1719	55.1265	e-084.6843	1.0000
6	10.9350	21.8700	51.0455	46.3294	49.2681	e-071.9339	1.0000
7	12.7575	25.5150	44.9870	40.2903	43.2993	e-077.8033	0.9999
8	14.5800	29.1600	38.9553	34.3243	37.0170	e-063.129	0.9998
9	16.4025	32.8050	32.8710	28.2896	31.2245	6 e-1.270	0.9990
10	18.2250	36.4500	26.9305	22.2402	25.2756	4.9881e-05	0.9958
11	20.0475	40.0950	21.0561	16.1890	19.0970	1.9292e-04	0.9835
12	21.8700	43.7400	15.0405	10.1654	12.6446	e-07.7078	0.9416
13	23.6925	47.3850	8.9718	4.3152	6.3171	0.0031	0.8283

Conclusions

The proposed embedding technique is an efficient method in completely recovering secret color image. Also, the sound that carries the message does not arouse suspicion until replacing the LSB number 6 for speech and the LSB number 9

for musical audio. Notice that, not all audio samples have been replaced; the replacing of bits is done as required to the size of secret color image. The strength of steganography system was done by three secret chaotic keys, and the user can use a very large number of them by changing the initial value and the control parameter.

Acknowledgment

The author's would like to express special thanks to the image processing laboratory staff in department of physics, College of Science, of Mustansiriyah University.

References

[1] N. Provos and P. Honeyman, IEEE Security and Privacy Magazine, 1, 3, June (2003) 32-44.

[2] H. Wang and S. Wang, Communications of the ACM magazine, 47, 10, October (2004) 76-82.

[3] A. Cheddad, J. Condell, K. Curran, P. M. Kevitt, Journal of Signal Processing, 90, 3, March (2010) 727-752.

[4] Y. Wang; P.Moulin, Information Theory IEEE Transactions, 54, 6, Jun (2008) 2706-2722.

[5] N. Cvejic, "Algorithms for Audio Watermarking and Steganography", MSc. Thesis, Department of Electrical and Information Engineering, Information Processing Laboratory, University of Oulu, Finland Oulu, Finland, (2004).

[6] K. Bailey and K. Curran, Journal of multimedia Tools and Applications, 30, 1, July (2006) 55-88.

[7] N. Meghanathan and L. Nayak, International Journal of Network Security & Its Application (IJNSA), 2, 1, January (2010) 43-55.

[8] S. Shahreza and M. Shalmani, IEEE International Conference on Acoustics, Speech, and Signal Processing, March 31 -April 4 (2008) 1729-1732.

[9] A. Cheddad, J. Condell, K. Curran, M.C. Kevitt, Signal Process, 90, 3 (2010) 727-752.

[10] X. Huang, Y. Abe, I. Echizen, J. Inform. Hiding Multimedia Signal Process., 1, 2 (2010) 72 -90.

[11] L. Singh and R. K. Bharti, International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), 3, 11 (2013) 1-11.

[12] Anjula Gupta and Navpreet Kaur Walia, International Journal of Engineering Development and Research (IJEDR), 2, 2 (2014) 1667-1672.

[13] Ahlam Majead Kadhim, Shaimaa H. Abd Muslim, Huda M. Jawad, IOP Conference Series: Materials Science and Engineering, IOP Publishing, 871, 1 (2020) 1-8.

[14] R. Tenny, L.S. Tsimring, H.D.I. Abarbanel, L.E. Larson, (Institute for Nonlinear Science), Springer, (2006) 191-229.

[15] Sayed Ahmad Salehi, Rasoul Amirfattahi, Discrete Wavelet Transforms-Algorithms and Applications, (2011) 41-56.

[16] Ahlam M.K, "Audio Steganography for Multimedia Files Using Wavelet Transform" Thesis, University of Mustansiriyah, (2016).

[17] H.I. Shahadi, R. Jidin, W.H. Way, Research Journal of Applied Sciences, Engineering and Technology, 7, 11 (2014) 2311-2323.