The Republic of Iraq
Ministry of Education and Scientific Research
University of Babylon\\ College of Science for Girls
Department: Computer Science

اخفاء نص داخل صور ملونه

Graduation project for the Department of Computer Science at the College of Science for Girls, University of Babylon, as part of the requirements for obtaining a Bachelor’s degree in Computer Science

By the student: Hadeel Mahdi Gouda Ali

Supervised by: M. M. Rafif Mazhar

2023-2024م

بِسۡمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيمِ a

﴿الْحَمْدُ ّ هللِ ّ الهذّي أَنْزَلَ عَلَى عَبْدّهِ الْكّتَابَ وَلَمْ يَجْعَلْ لَهُ عّوَجًا ﴾
[الكهف:1]

صدقَ ٱللهُ العليُّ العظيم

اقرار المشرف

اشهد ان اعداد هذا المشروع (*Compression Text By Coding and Hiding In Color Image*)قد جرى في قسم علوم الحاسوب في كلية العلوم للبنات/جامعة بابل ,وهو جزء من متطلبات نيل شهادة البكالوريوس في علوم الحاسبات من قبل طالبة المرحلة الرابعة (هديل مهدي جوده).

توقيع المشرف

اسم المشرف :رفيف مظهر المرتبة العلمية

:

التاريخ: 2023-2024

الأهداء

وصلت رحلتي الجامعية إلى نهايتها بعد تعب ومشقةً.. وها أنا ذا

أختم بحث تخرجُّي بكل هَّمة ونشاط،

وأمتُّن لكل من كان له فضل في مسيرتي، وساعدني

ولو باليسير،

الأبوين، والأهل، والأصدقاء، والأساتذة المبُجلَّين.. أهُديكم

بحث تخرجُّي.

الشكر و التقدير

الشكر والثناء لله تعالى

على اتمام هذا البحث وعلى اتمام الدراسة وارجو ان تنال رضاه فالحمد لله على هذه النعم ، ومن ثم اتقدم بالشكر والتقدير الى:

اساتذتي الكرام في قسم علوم الحاسوب وبالاخص الاستاذه "رفيف مظهر" الذي تفضلت بالاشراف على هذا البحث ، شكر كبير لها على نصحها وارشادها ومساعدتها لي ... الى احباب قلبي أمي و اخوتي واخواتي

الى من كانوا خير صحبة ورفقة لي خلال مسيرتي الدراسية في الجامعة . الى كل من قدم لي الدعم والتوجيه

كل الشكر والامتنان لكم على كل شيء.

Abstract

Data security and protection is one of the most used technologies in the field of computer science, because of its great importance in all fields, whether political, economic, or military. The main goal of this paper is to develop and design a systematic, effective and secure method for transferring data, in addition to reducing the time and cost when sending data over the Internet.

Therefore, an intelligent system was proposed in this paper that combines firstly compression data by using coding method to reduce the amount of data sent and helps in fast transfer while using slow Internet or taking up small disk space. Secondly hiding bits sets in cover RGB image using the LSB method to implant bits of the compressed text into the cover image.

Contents

Chapter one:
General Introduction

1.1 Introduction

The process of securing the information transmitted between two parties and protecting it from intruders and tampering has become on top of the priorities that must be met and ensured especially [1]. as the world is witnessing great progress in the field of Communication in addition to the increase of methods that have been developed to obtain information in an unauthorized manner Legitimacy . Securing data by compressing it makes it interesting and suspicious of the existence of information a task that should not be viewed by unauthorized persons so concealing this information is important behind a cover that is not suspicious is better[2][3].

Steganography is a security method introduced in 1499 by Johannes Trithemiu [4]. it is the manner used to hide information and data in different mediums such as images, video and sound. There are many domains that include an application of steganography: secure transfer of confidential information between national and international governments, tampering, online banking assurance, voting methods and time stamps [5]. The methods of steganography are divided into two fields one of them is spatial steganography and its characterized by ability to absorb a lot of information to hide, LSB method belongs to this kind of steganography [6].

1.2 Previous works:

The topic of hiding information is one of the topics that was worked on early and developed rapidly and advanced with the introduction of the world of digital images, which was the focus of interest of many researchers. Below are a number of studies in this field:

- Researcher Eman (2008)[7]. proposed a method to hide data in binary images, whereby points are first identified Optical images that can be flipped without causing visible distortions in the embedded image by using A set of rules by which all points adjacent to the center point of each irregular sector are examined Then the center point is changed only if the sector matches these conditions and this property allows By discovering embedded data without referring to the original image, experiments have shown different results for images binary..
- The researcher Latef Abdul A.A (2011)[8]. presented a method for masking color images by dividing them into Four equal parts, each part consisting of three channels (Blue, Green, Red). Choose one of theseChannels for each part depending on the high color percentage in that part, then a wavelet transform is applied On the selected part, the message to be hidden is also divided into four parts and DCT is applied to it Then each part of it was included in the high-frequency wavelet transform of one of the parts of the cover image to obtain On the secret photo.
- Researcher Yong (2011)[9]. proposed a scheme to hide secret data inside the image using a transformation curvelet, where the image to be hidden is encoded using the Radon transform and using coefficients High-frequency curvelet conversion for data embedding.

- Researchers Abdelwahab and Hassan (2008)[10]. suggested using the first level of wavelet transforms In hiding and embedding the data, but the extracted data was not completely identical to the embedded version.

1.3 The aim of the research:

The data protection process acquires special importance that increases with the degree of sensitivity of the message and the sending parties. This project aims to secure the text message exchanged between the two parties by compressing the message before hiding it in another text file.

1.4 Research Layout

This research is organized as follows

Chapter Two: The overall objective of this chapter is to present fundamentals, details, and characteristics of all approaches which have been used in proposed work, where this chapter starts with short introduction to steganography methods that have been used and then it gives an explanation it.

Chapter Three: This chapter presents the designed steps of the entire system's stages, displays the implementation results of experiments, and a discussion on the obtained results.

Chapter Four: The derived conclusions from the proposed work and some suggestions to enhance the proposed work have been presented in this chapter.

Chapter Two

2.1 **INTRODUCTION**

In the current scenario, secret messages can be sent by hiding in an image or a text so nobody other than sender and receiver can read or see the message. Steganography is used to conceal the existence of a message using cover text [11]. Data hiding is a significant side of communication and data security technology. Information hiding is an important method of steganography communication. In other words, the important data to be sent should be embedded in the carrier, so that it cannot be easily found secret data. There are many carriers applied information hiding, such as text, image, audio and video

Data compression is important to information security because compressed data is more secure and easier to annotate. Effective data compression techniques create efficient, secure, and easy-to-communicate data. We also compress data before sending it over the Internet in order to reduce time and cost. The simultaneous use of a number of technologies in the same system makes it difficult for attackers to penetrate the system. Combining data compression, encryption, and information hiding in one integrated system enhances security and efficiency in data transfer and storage, as it reduces the size of the data, storage space, and transmission time[12].

2.2 Compression

Data compression, source coding,[13]. or bit-rate reduction is the process of encoding information using fewer bits than the original representation[14].Any given pressure is lossy or lossless. Reduces lossless compression by identifying and removing statistical redundancy. No information is lost in lossless compression. Lossy compression reduces bits by removing unnecessary or less important information[15] .Typically, the device that compresses the data is referred to as the encoder, and the device that reverses the process (decompression) is referred to as the decoder.

The process of reducing the size of a data file is often referred to as data compression. In the context of data transmission, it is called source encryption; Coding that takes place at the data source before it is stored or transmitted[16]. Source coding should not be confused with channel coding, for error detection and correction, or line coding, means of correlating data onto a signal.

Compression is useful because it reduces the resources required to store and transmit data. Computational resources are consumed in compression and decompression operations. Data compression is subject to a space-time complexity trade-off. For example, a compression scheme for video may require expensive hardware in order for the video to be decompressed quickly enough to view while it is decompressed, and the option to decompress the entire video before viewing it may be inconvenient or require additional storage space. Designing data compression schemes involves trade-offs between various factors, including the degree of compression, the amount of distortion introduced (when using lossy data compression), and the computational resources required to compress and decompress the data[17][18].

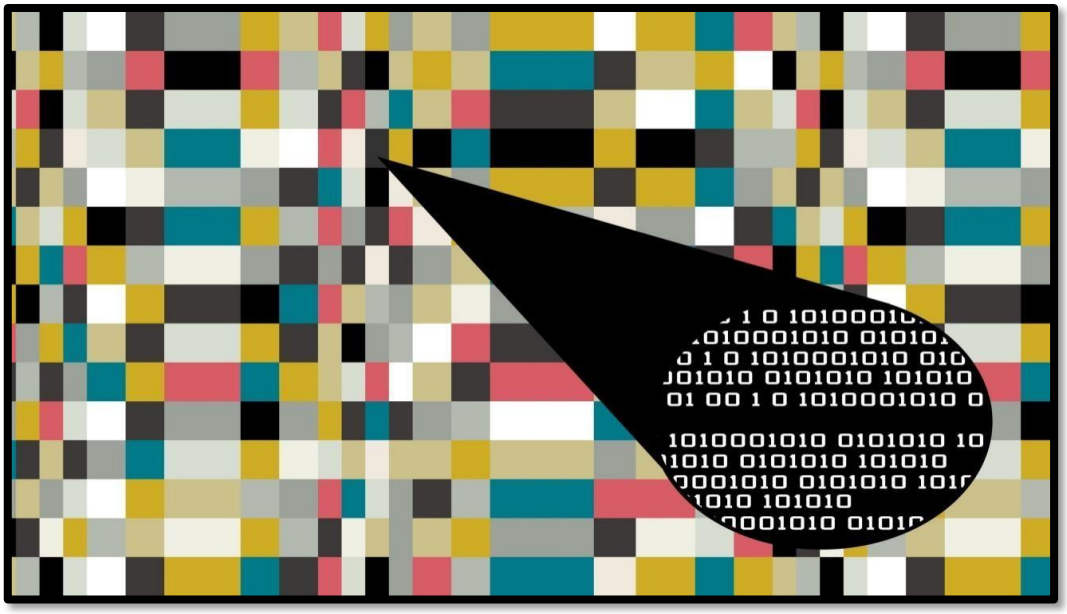
2.3 Hiding Information:

There is art that aims to completely hide data for communication between two parties in a way that is not apparent to a third party, and this is what Known as steganography, it is a method or technique for blocking and hiding data within a digital medium, so that it is hidden that there is a communication or exchange of information that takes place in secret, and only the people concerned are aware of this communication.

The word steganography is originally derived from a Greek word meaning “hidden writing.”

The issue of sending a hidden message by concealing that something was sent in the first place is an old method (and idea). It has a historical story that began in the days of the ancient Greek Empire, when letters were written on the heads of slaves at that time

after shaving the slave's hair. A certain secret message was written on his head, and when his hair grew back again, the secret message was hidden under his thick hair, and then it was sent to the other party, who would do his part. By shaving the Eid's head again so that he could read the message to him, and so were the beginnings of using this method to hide a message or information under a cover or something so that there would be no knowledge that any secret communication was taking place between two or more people.



2.4 least significant bit (LSb)

There are a number of methods and algorithms for hiding data, where you can hide one type of data (such as texts and files) inside another type of data. For example, you can hide an image inside another image, or an image inside a video file, or text inside an audio file, and so on without anyone watching. There is a change in the image. Data hiding algorithms manipulate the contents of the file, modifying the bits of the file You manipulate it so that it does not affect the contents of the file, and at the same time you inject into it the bits of the other file. When you see an image that contains text inside it,

for example, you will not see the text, as the text is represented inside the image.You will also not notice a change in the quality of the images in bits.All data stored inside the computer is basically represented by the binary counting system. The binary counting system was given this name because it contains only two numbers (symbols) to represent the data, which are 0 and 1. To display data that the user understands (such as texts, images, and numbers), the system is transformed Binary to other counting systems to facilitate data representation.

The LSB algorithm, or Least Significant BIT, is one of the most famous algorithms for embedding data over an image. The least significant bit in the image is changed to a bit that represents the secret message that must be Include it in the picture.

Chapter Three

3.1 PREPOSED METHOD

Explain the processes of embedding the secret message text in a color image in this section. The sender's side contain several operations as follows to protect the hidden text from theft:

Stage1:- Sender side "The stage of including text in a cover image"

Step 1: Read the text of the message to be protected.

Step 2: Compress the message text: characters are represented using the ASCII code each character is represented by 8 bits (In the normal case). In this work, each letter is represented according to Table 1 by taking the letter of the message and the number corresponding to it , thus converting a message from a string of letters to a string of decimal digits between(0-25).The series of decimal numbers is converted to the binary system each number in a string into a 5-bit binary equivalent because the maximum representation of the decimal number 25 in binary equivalent is 5 bits.

TABLE 1. Decimal number for each character

Letter	Code no.	Letter	Code no.
A	0	N	13
B	1	O	14
C	2	P	15
D	3	Q	16
E	4	R	17
F	5	S	18
G	6	T	19
H	7	U	20
I	8	V	21
J	9	W	22
K	10	X	23
L	11	Y	24
M	12	Z	25

Step 3: Read the RGB cover image

Step 4: Embedding data process

4-1 Divide the RGB cover image at three bands: R band, B band, and G band.

4-2 Each time take 3 bits from the binary string and embedded them into the three bands of the cover image sequentially using LSB. Where the first bit is in the R range, the second bit is in the B range, and the third bit is in the G range .Then, the RGB image is sent over the communication channels, and the random key is also sent. Figure (1) appears the particular block diagram for the sender and recipient.

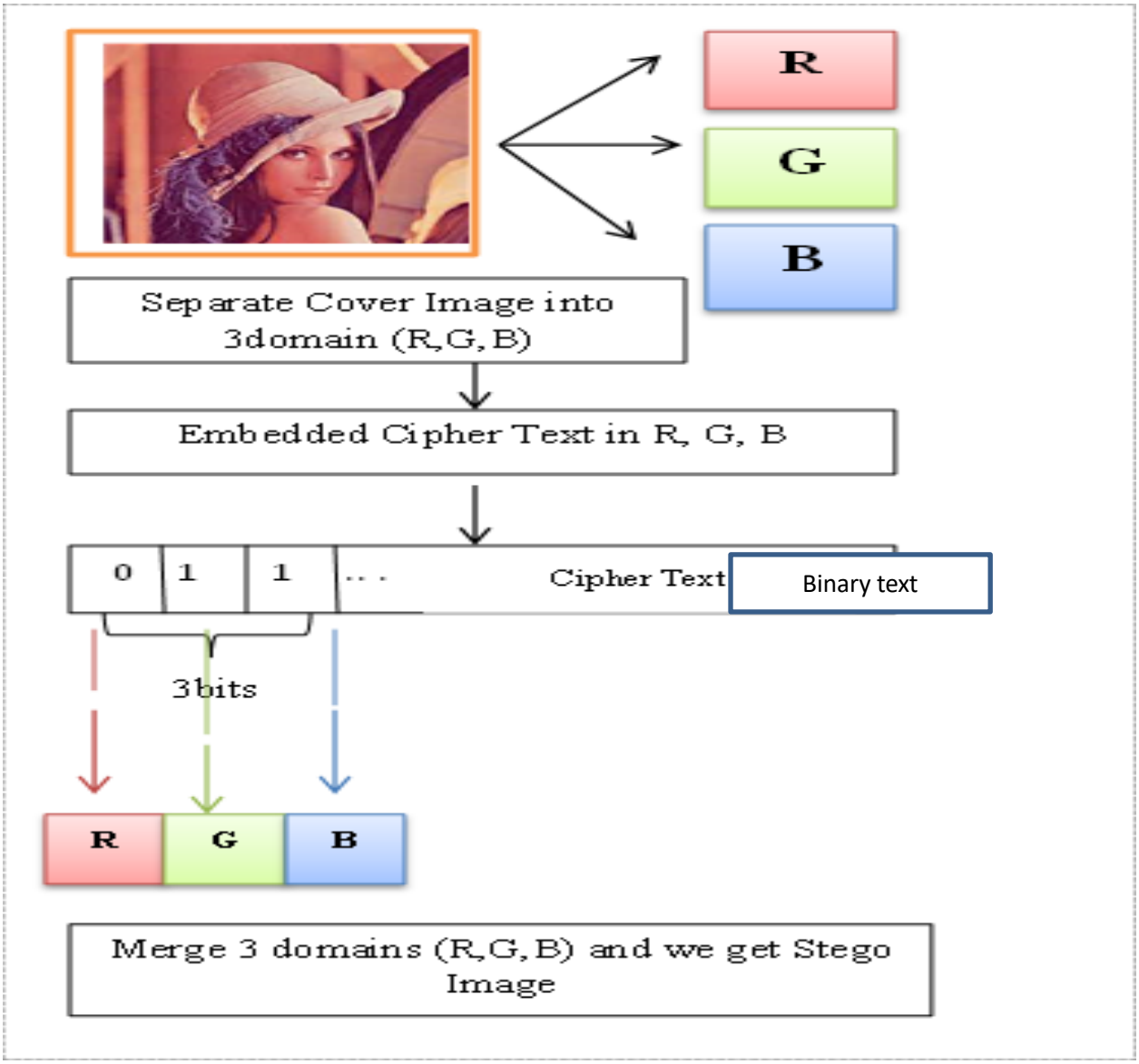


Figure (1)Embedded cipher text.

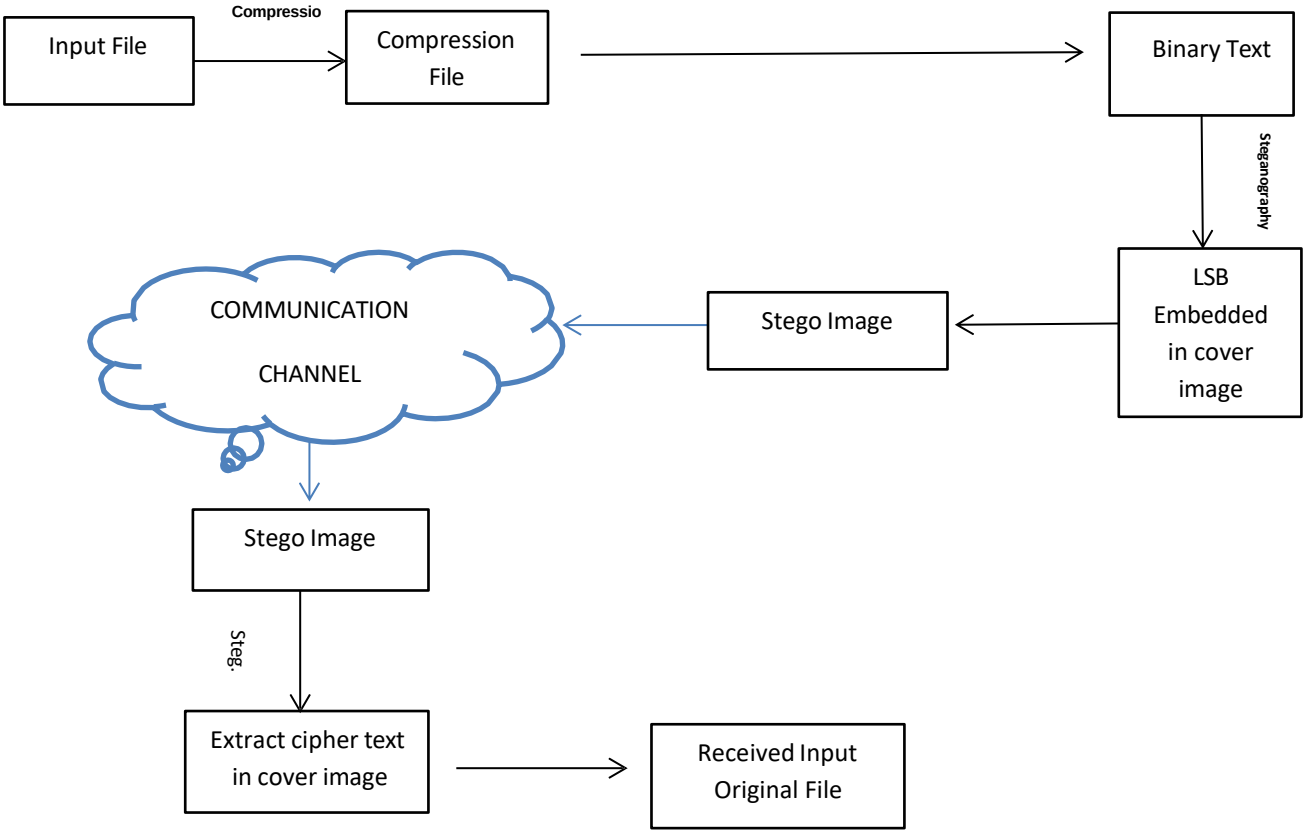


FIGURE 2. Block diagram of the sender side and receiver side.

Stage 2: Recipient: Take the message text out of the stego image.

At this stage, the recipient receives a stego image and divides the image into three bands to retrieve secret data. The first bit comes from the bits of the secret data string from range 1 (R), the second bit comes from range 2 (b), and the third bit comes from range 3 (G). In this manner, we obtain the bits of the secret data string and keep going until the data string is recovered. Next, we get the binary data string . Then, we take 5 bits from the data string and convert it into its decimal representation. Ultimately, the original text is recovered by taking the letter corresponding to each decimal value in the data string.

3.2 Experiment results:

The proposed procedures were implemented in Matlab 2012. The masking method was successfully implemented with the introduction of some plain text of a specified size. A screenshot of the text and image is represented before the masking process is performed:

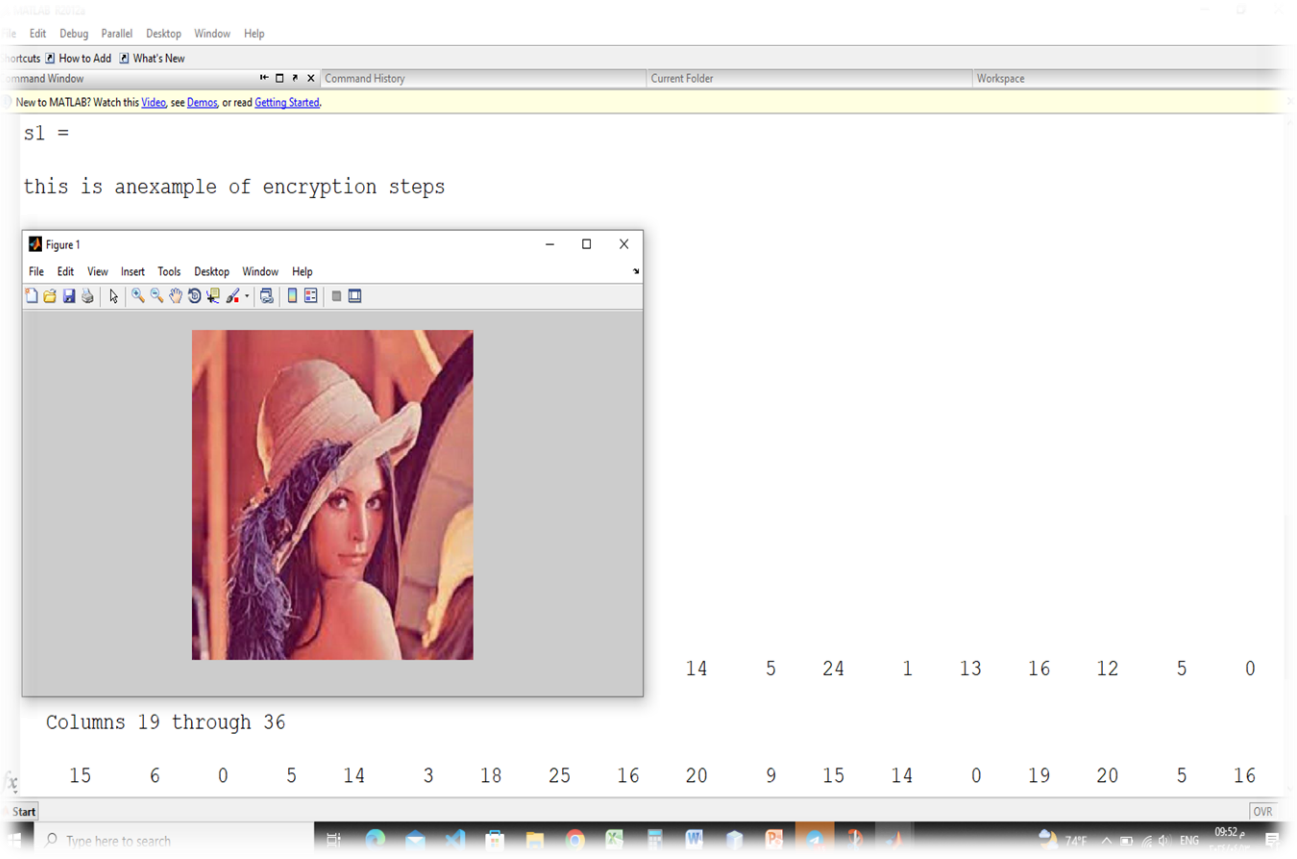
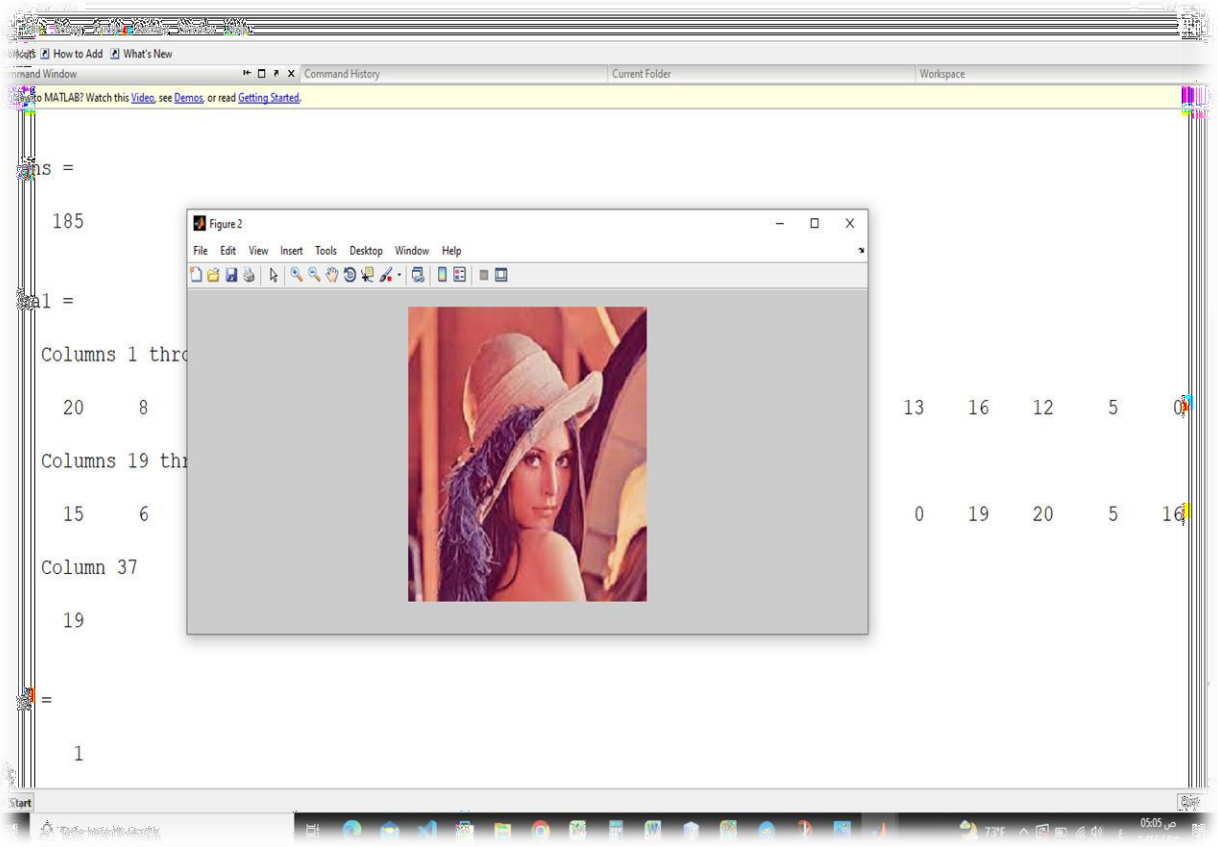


Image No. (1): The original text and Original color image



Picture No. 2 cover image after compression text.

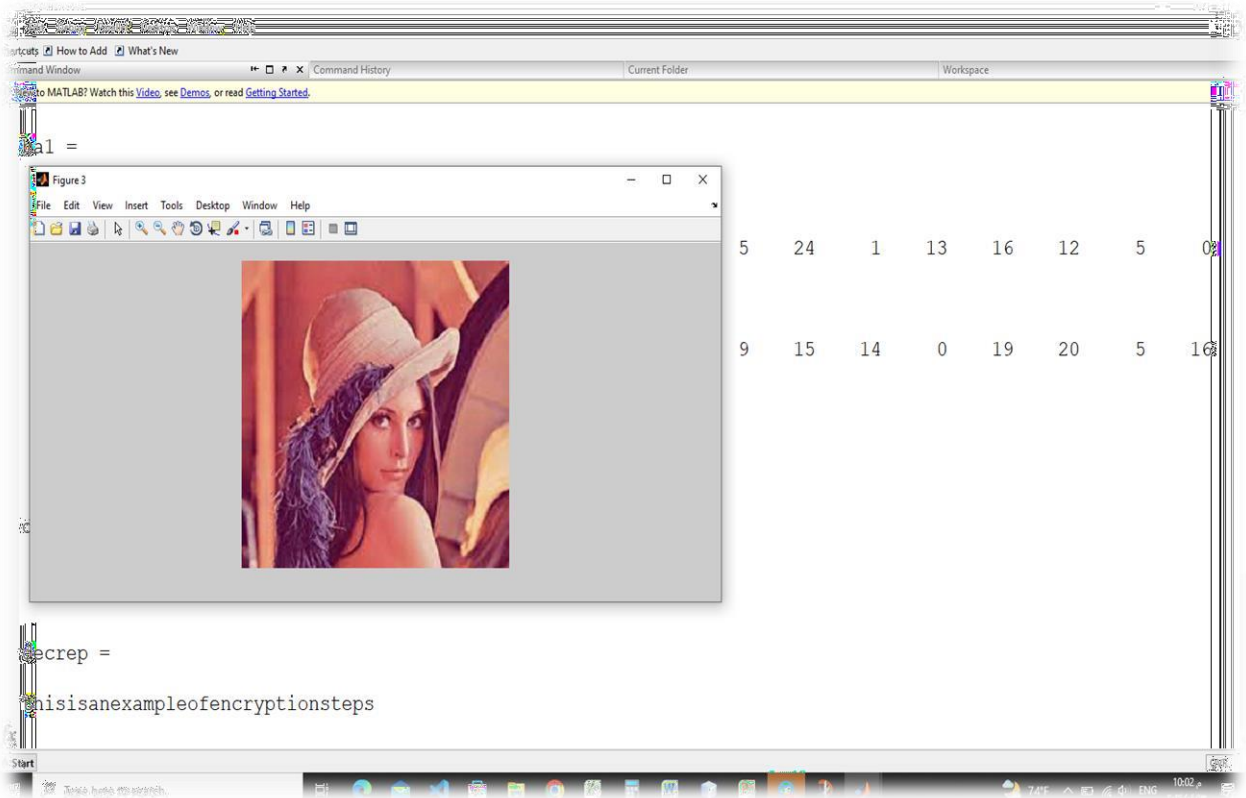


Image No. (3) retrieving the original text and image

3.3 Quality standards

Peak signal-to-noise ratio (PSNR) is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed as a logarithmic quantity using the decibel scale.

PSNR is commonly used to quantify reconstruction quality for images and video subject to compression.

Through the following equation:

PSNR= 10 log10 ((MAX ^2)/MSE)

TABLE 2.The PSNR value

Text File size in bits	**Dimension of Image (512*512)**	**PSNR**
200 bit		89.4531
1800 bit		76.7012
5048 bit		75.2320

TABLE 3.The PSNR value

Text File size in bits	**Dimension of Image (512*512)**	**PSNR**
200 bit		89.1635
1800 bit		76.5875
5048 bit		75.0707

Chapter Four

4.1 Conclusions

Steganography and compression are two ways to ensure the confidentiality of information. In compression, the number of bits used is reduced, while in steganography, the secret text remains as it is, but it is embedded in another format of data. Today, in light of the presence of strong communication systems, protecting confidential information from hackers is a difficult task. This research relies on a text hiding method to protect text messages sent between two parties by using hiding techniques. The standard image Lena (256*256) was used to hide The text in the image cover using the least significant bit method. This is the proposed method that can be successful in this direction of increasing the security of data transmission.

This research relies on the method of compressing the text message and then hiding it to protect the text message sent between the two parties, by using the least significant bit method to hide the bits of the compressed message.

4.2 Advantages of the legislator

1-The project provided high protection for image data from illegal hacking

2-The project's ability to perform its functions as required

3- The speed of the method used to hide text and retrieve text

4- We notice that when retrieving the text, there is no loss of the secret text

4.3 Project disadvantages:

The language used does not support indicators, as indicators facilitate dealing with images and do not exhaust the memory

4.4 Future works:

- ✓ Modifying the negatives present in the project.
- ✓ Other algorithms will be adopted to hide the text in order to increase the security of the secret text by adopting different compression methods and adding an encryption method.

The References

1. NavdeepandMsNehaGoyal"Hide Text in Images Using Steganography and a Review of Methods and Approach for Secure Steganography", International Journal of Research in IT & Management Vol. 6, Issue 5, 2016.

2. Suhaila Mohammed, Shaymaa Ahmed, Ghusoon Mohammed, and Dhuha Abduljabbar, " Block-based Image Steganography for Text Hiding Using YUV Color Model and Secret Key Cryptography Methods ", Australian Journal of Basic and Applied Sciences, 11(7) , 2017.

3. Gourav Tiwari, Rameshwar Nath Pathak," Secret Information Transmission within Color Image using Wavelet Transformation", International Journal of Computer Science and Information Technologies, Vol. 8 (3) , 2017.

4. Dr. Rajkumar L Biradar and AmbikaUmashetty, "A Survey Paper on Steganography Techniques", International Journal of Innovative Research in Computer and Communication Engineering (A High Impact Factor, Monthly, Peer Reviewed Journal) Vol. 4, Issue 1, 2016.

5. Manivasagam Srinivasan and SrideviAnnadurai"Data Hiding And Image Compression Using SMVQ And DCT ", International Journal of Advanced Research in Basic Engineering Sciences and Technology, Vol.3, No.24 2017.

6. SujaraniRajendran and ManivannanDoraipandian" Chaotic Map Based Random Image Steganography Using LSB Technique" International Journal of Network Security, Vol.19, No.4, 2017.

7. Eman Th. Sedeek Al-obaidy,(2008), "An Algorithm for Data Hiding in Binary Images", Raf. J. of Comp. & Math's., Vol. 5, No. 2, 8.

8. A.A. Abdul Latef, (2011), "Color Image Steganography Based on Discrete Wavelet and Discrete Cosine Transforms", IBN AL- HAITHAM J. FOR PURE & APPL. SCI. VOL.24 (3).

9. Yong Hong Zhang, (2011), "Digital Image hiding using curvelet transform", IEEE Conference.

10. A.A. Abdelwahab, L.A. Hassan, (2008), “A discrete wavelet transform based technique for image data hiding”, in: Proceedings of 25th National Radio Science Conference, Egypt.

11. Suhaila Mohammed, Shaymaa Ahmed, Ghusoon Mohammed, and Dhuha Abduljabbar, " Block-based Image Steganography for Text Hiding Using YUV Color Model and Secret Key Cryptography Methods ", Australian Journal of Basic and Applied Sciences, 11(7) , 2017.

12.I. U. W. Mulyono, et al., “Encryption of Text Message on Audio Steganography Using Combination Vigenere Cipher and LSB (Least Significant Bit),” Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control, vol. 4, no. 1, pp. 63-74, 2018.

13. Wade, Graham (1994). Signal coding and processing (2 ed.). Cambridge University Press. p. 34. ISBN 978-0-521-42336-6. Retrieved 2011-12-22. The broad objective of source coding is to exploit or remove 'inefficient' redundancy in the PCM source and thereby achieve a reduction in the overall source rate R.

14.Mahdi, O.A.; Mohammed, M.A.; Mohamed, A.J. (November 2012). "Implementing a Novel Approach an Convert Audio Compression to Text Coding via Hybrid Technique" (PDF). International Journal of Computer Science Issues. 9 (6, No. 3): 53–59. Retrieved 6 March 2013.

15. ^ Pujar, J.H.; Kadlaskar, L.M. (May 2010). "A New Lossless Method of Image Compression and Decompression Using Huffman Coding Techniques" (PDF). Journal of Theoretical and Applied Information Technology. 15 (1): 18–23.

16. Salomon, David (2008). A Concise Introduction to Data Compression. Berlin: Springer. ISBN 9781848000728.

17. S. Mittal; J. Vetter (2015), "A Survey Of Architectural Approaches for Data Compression in Cache and Main Memory Systems", IEEE Transactions on Parallel and Distributed Systems (IEEE) 27 (5): 1524–1536, doi:10.1109/TPDS.2015.2435788.

18. Tank, M.K. (2011). "Implementation of Lempel-ZIV algorithm for lossless compression using VHDL". Implementation of Limpel-Ziv algorithm for lossless compression using VHDL. Berlin: Springer. pp. 275–283. doi:10.1007/978-81-8489-989-4_51. ISBN 978-81-8489-988-7. {{cite book}}: |work= ignored (help(