هبة جبار عبد الواحد كريم العقابي .
جامعة النهرين/ كلية العلوم /قسم علوم الحاسبات
بكلوريوس علوم حاسبات ٢٠٠٢
ماجستير علوم حاسبات ٢٠٠٥
العنوان : محافظة ديالى/ بعقوبة/ التكية /خلف المحافظة –قرب مبنى كاتب العدل .
الهاتف:٠٢٥٥٢٦٧١٣
البريد الاليكتروني hiba_aleqabie@yahoo.com
تاريخ المناقشة ١٢-١٠-٢٠٠٥
المشرف:- د. ستار بدر سدخان
المناقشون :د.محمد زكي الفايز رئيسا"
د.قيس جميل الجميلي عضوا"
د.عبير متي يوسف عضوا"

Republic Of Iraq
Ministry of Higher Education
And Scientific Research
AL-Nahrain University
College of Science
Department Of Computer Science

Detector of Information Hiding

A THESIS
SUBMITTED TO THE
COLLEGE OF SCIENCE OF AL-NAHRAIN UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE IN
COMPUTER SCIENCE

By

Hiba Jebbar A. Al-Eqabie

(B.Sc. 2002)

Supervisor
Dr. Sattar B. Sadkhan

Jamad Al-akher 1426 July 2005

بسم الله الرحمن الرحيم

ن و القلم و ما يسطرون

صدق الله العلي العظيم

(سورة القلم -الاية ١)

Supervisor Certification

I certify that this thesis was prepared under my supervision at the department of computer science/ collage of science/Al-Nahrain University by ***Hiba Jebbar Al-Eqabie*** as a partial fulfillment of the requirements of the degree of Master of Science in computer science.

Signature:
Name: **Dr. Sattar B. Sadkhan**.
Title: **Assistant Professor.**
Date: / /2006

In view of the available recommendations, I forward thesis for debate by the examination committee.

Signature:
Name: **Dr. Sattar B. Sadkhan**
Title: Head of the department of computer science ,Al-Naharin University.

Date: / / 2006

Certification of the Examination Committee

We the examination committee certify that we have read this thesis titled "**Detector of Information Hiding** "and we have examined the student "***Hiba Jebbar Al-Eqabie*** " in its contents and what is related to it, and our option it meets standard of a thesis for the degree of **Master of science in computer science.**

Signature:
Name: **Dr. Mohammed Z. Al-Faiz.**
Title: **Assistant Professor**
Date: / /2006
(Chairman)

Signature:
Name: **Dr Kais J. Al-Jumaily**
Title: **Assistant Professor**
Date**: /** /2006
(Member)

Signature:
Name: Dr**. Abeer Mety**
Title: **lecturer**
Date: / /2006
(Member)

Signature:
Name**: Dr. Sattar B. Sadkhan**.
Title: **Assistant Professor.**
Date: / /2006
(Supervisor)

Approved by the dean of the collage of science, Al-Nahrain University

Signature:

Name: **Dr. Laith Abdul-aziz Al-Ani**

Title: Dean of The College of scince

Date: / /2006

الاهداء

إلى سر وجودي و الروح إلى كانت منبع الحب و الحنان **أمي**

إلى الذي تحمل مرارة الدنيا من اجلي **أبي**

إلى الذي ملء حياتي أملا" ونورا" **زوجي**

إلى زهور حياتي **أختي و إخوتي**

إلى برعم حياتي **ابنتي**

إلى كل من يسعده نجاحي

Acknowledgment

First of all I'd like to express my sincere gratitude and appreciation to my supervisor Dr. Sattar B. Sadkhan for his supervision and untiring efforts during the course of work.

Also thanks to Dr.Taha S. Bashaga.

Special thanks to my parents and husband Eng. Mohammed For thier kind ,patient support and encourages.

Abstract

This research aims to design and implement a steganalysis process through scanning and testing the tested images, each of 24-bit image, to find out if it contains a hidden information. The stego objects are embedded using S-Toll, and Developed Stegonography tool that modulate Least Significant Bit (LSB) of the pixel. Using of wavelet transformation of Haar wavelet type to produce feature vectors of coefficients ,these coefficients are mapped ,then using the ability of Probability Density Function(PDF) to minimize the features that will be use in the statistical tests:-The Standard tools Absolute Value Differences (AD), Mean Square Error(MSE), Signal- to- Noise Ratio(SNR), Peak Signal- to- Noise Ratio(PSNR), Normalized Cross –Correlation(NCC) , Correlation Quality (CQ).and in our research, we used the order statistics, such as:- (Mean, Variance, Skewness, Kurtosis) .

We tested 12 BMP images with different sizes, which had information hiding both steganography and watermarked. Though the system was tested 12 distinct images ,some were detected and some were not, 6 images had information hiding and 6 were clear, 3 of 12 (i.e. 33%)were pass as they were clear, while 2 images were not. The others were detected . The developed system is implemented using Visual Basic programming language version 6, provides by Windows environments (XP, Me), and the resulted obtained are encouraging.

Abbreviation	*Detailed*
AD	**A**bsolute **D**ifferences
BMP	**B**it-**M**a**p**
CQ	**C**orrelation **Q**uality
DCT	**D**iscrete **C**osine **T**ransformation
DWT	**D**iscrete **W**avelet **T**ransform
FFT	**F**ast **F**ourier **T**ransform
FLD	**F**isher **L**inear **D**iscriminate Analysis
HCF	**H**istogram **C**haracteristic **F**unction
HH	A Signal Or A N Image That Has Been **H**ighpass Filtered In Both Horizontal An Vertical Directions
HL	A Signal Or A N Image That Has Been **H**ighpass Filtered In Vertical Direction And Horizontally **L**owpass Filtered.
HPDM	**H**istogram **P**reserving **D**ata **M**apping
JPEG	**J**oint **P**hotographer **E**xperts **G**roup
LH	A Signal Or A N Image That Has Been Vertically **L**owpass Filtered and **H**ighpass Filtered In Horizontal Direction.
LL	A Signal Or A N Image That Has Been **L**owpass Filtered In Both Horizontal An Vertical Directions.
LSB	**L**east **S**ignificant **B**its
MSE	**M**ean **S**quare **E**rror
NCC	**N**ormalized **C**ross **C**orrelation
PDF	**P**robability **D**ensity **F**unction
PMF	**P**robability **M**ass **Function**
PSNR	**P**eak **S**ignal To **N**oise Ratio
RGB	**R**ed-**G**reen-**B**lue. An image color space where the image data is represented by red, green and blue bit planes of the image
SNR	**S**ignal To **N**oise **R**atio.
STFT	**S**hort-**T**ime **F**ourier **T**ransform
YCbCr	An image color space where the image data is represented by Luminance and red and blue color difference components. Most of the image information is in the Y component.
WT	Wavelet Transform

Table of content

Chapter One

Overview

1.1 Introduction.

Steganography literally means “Covered Message” and involves transmitting secret messages through seemingly innocuous files. The goal is that not only does the message remain hidden, but also that a hidden message was even sent .

There are many tools available (Steganography Software) that can hide messages in images, audio files and video, and steganography is now in common use. Steganography supports hiding messages amongst the huge volume of Internet traffic, in media files where the addition of a hidden message is difficult to detect with the human eye even if the file is viewed[3]

1.2 Literature Survey

Several researches in the Steganalysis field tried to detect the existing of an embedded information .The scope of this review is a general review of the researches that investigate in Steganalysis :

- Neil et al [18], presented an overview of Steganalysis and introduced some characteristics of Steganography software that point signs of information hiding. This work is a fraction of the steganalysis approach. Success in steganographic secrecy results from selecting the proper mechanisms. The more information is placed in the public's reach on the Internet, the more owners of such information need to protect themselves from theft and false representation. Systems to recover seemingly destroyed information and steganalysis techniques will be useful to law

enforcement authorities in computer forensics and digital traffic analysis.

- Hany Farid [7], presented the basic approach which was finding predictable Higher-Order Statistics of natural. images within a multi-scale decomposition, and then showing that embedded messages alter these statistics. This model includes basic coefficient statistics as well as error statistics from an optimal linear predictor of coefficient magnitude. These higher-order statistics appear to capture certain properties of natural images, and more importantly, these statistics are significantly altered when a message is embedded within an image. As such, it is possible to detect, with a reasonable degree of accuracy, the presence of Steganographic messages in digital images. To avoid detection, however, one need only embed a small enough message. In the examples shown here, the message was typically 5% the size of the cover image. As the message size decreases, detection will become increasingly more difficult and less reliable. There are several directions that should be explored in order to improve detection accuracy. The particular choice of statistics would be beneficial to optimize across a set of statistics that maximizes detection rates. The two-class *Fisher* Linear Discriminate analysis (FLD) should be replaced with a multi-class FLD that simultaneously distinguishes between no-steg images and steg images generated from multiple programs. However convenient FLD analysis is linear, and detection rates would almost certainly benefit format more flexible non-linear classification scheme. Lastly, the indiscriminant comparison of image statistics across all images could be replaced with a class-based analysis .One benefit of the higher-order models employed

here is that they are not as vulnerable to counter-attacks that match first-order statistical distributions of pixel intensity or transform coefficients**.** There is little doubt, however, that counter-measures will be developed that can foil the detection scheme outlined here. The development of such techniques will in turn lead to better detection schemes.

- Hany Farid and Siwei Lyu[9], presented and described an approach to detecting hidden messages in images that uses a wavelet-like decomposition to build higher-order statistical models of natural images. Support vector machines are then used to discriminate between clear and stego images. The techniques described here would almost certainly benefit from several extensions: (1) the higher-order statistical model should incorporate correlations within and between all three color channels; (2) the classifier should be trained separately on different classes of images; and (3) the classifier should be trained separately on images with varying compression rates. One benefit of the higher-order models employed here is that they are not as vulnerable to counter-attacks that match first-order statistical distributions of pixel intensity or transform coefficients.
- Jessica et al [3], they classified and reviewed current stego-detection algorithms that can be used to trace popular steganographic products. they recognized several qualitatively different approaches to practical Steganalysis-visual detection-detection based on first order statistics (histogram analysis), dual statistics methods that use spatial correlations in images and Higher-Order Statistics (RS Steganalysis), universal blind detection schemes, and special cases, such as JPEG compatibility Steganalysis. they also present some new results regarding their

previously proposed detection of LSB embedding using sensitive dual statistics.

- Niels Provos[20], presented an improved methods for information hiding. first method used probabilistic embedding to minimize modifications to the cover medium. Another method employed error-correcting codes, which allow the embedding process to choose which bits to modify in a way that decreased the likelihood of being detected. statistical analysis can reveal the presence of a hidden message. So they introduced two methods to improve the selection process. The first used a seeded pseudorandom number generator to determine the fewest modifications to the cover medium. The second used error-correcting codes to increase the flexibility in selecting bits without increasing the number of necessary changes. Together, these methods can be used to provide plausible deniability be embedding multiple hidden messages in the cover medium. Although the commonly used $\chi 2$-test is unable to detect modifications from the improved embedding process described in this paper, As a result, none of the presented statistical tests can detect the presence of steganography. they can use the estimate to quickly choose an image in which a specific message can be embedded safely. To evaluate the effectiveness of our approach, they presented statistical tests for the JPEG image format.
- Mohammed Alla' [16], he produced a system to attack and analyze hidden information by processes of three operations detection, extraction and distortion, also tried to extract the hidden information from the suspected images by concluding 1.the steganographic techniques that are used in the embedded process.2. the manner of choosing the pixels that are used in the

embedded process .3. the manner of choosing the bits that are substitute with the bits of the hidden information .

- Roman et al [23], improved Farid's [[7],[8],[9]] proposed algorithm such a detection algorithm based on higher-order statistics for separating original images from stego images. His method shows an astonishing performance on current Steganographic schemes. Starting from the statistical approach in Farid's algorithm they'll investigate the well known steganographic tool Jsteg as well as a newer approach , which relies on histogram-preserving data mapping. Both schemes show weaknesses leading to a certain detestability. Further analysis shows which statistic characteristics make both schemes vulnerable. Based on these results, the histogram preserving approach is enhanced such that it achieves perfect security with respect to Farid's algorithm. They have steganalysed the LSB embedding algorithm based on Histogram Preserving Data Mapping (HPDM) using Farid's universal blind detection method. Farid method using higher-order statistics gave us very interesting insights into both steganographic algorithms. Based on these insights they were able to modify HPDM to obtain a steganographic algorithm being perfectly secure with respect to Farid's detection method.
- Jeremiah Harmsen [14], presented a steganalysis of additive noise modelable information hiding. The process of information hiding is modeled in the context of additive noise. Under an independence assumption, the histogram of the stegotext is a convolution of the noise Probability Mass Function (PMF) and the original histogram. In the frequency domain this convolution is viewed as a multiplication of the Histogram Characteristic

Function (HCF) and the noise characteristic function. Least significant bit, spread spectrum, and Discrete Cosine Transform (DCT) hiding methods for images are analyzed in this framework. It was shown that these embedding methods are equivalent to a lowpass filtering of histograms that is quantified by a decrease in the HCF Center Of Mass (COM).

- Jessica et al [13], presented a steganalytic method that can reliably detect messages (and estimate their size) hidden in Joint Photographer Experts Group(JPEG) images using the steganographic algorithm F5. The key element of the method is estimation of the cover-image histogram from the stego-image. This is done by decompressing the stego-image, cropping it by four pixels in both directions to remove the quantization in the frequency domain, and recompressing it using the same quality factor as the stego-image. The number of relative changes introduced by F5 is determined using the least square fit by comparing the estimated histograms of selected DCT coefficients with those of the stegoimage. Experimental results indicate that relative modifications as small as 10% of the usable DCT coefficients can be reliably detected. The method is tested on a diverse set of test images that include both raw and processed images in the JPEG and BMP formats.
- Yanming et al [26], compared some steganalysis methods for attacking LSB-based steganographic techniques (logistic regression, the tree-based method {C4.5}, and a popular method Stegdetect). Experimental results show that the first two methods, especially the logistic regression method, are able to detect hidden information with high accuracy. They also studied the relationship between the number of attributes (the frequencies of quantized

DCT coefficients) and the performance of a classifier. Least Significant Bits(LSB)-based steganographic techniques (See appendix C) like Jsteg (See appendix B) change the statistical properties of the cover image when it embeds secret message in the image. Accordingly, such methods are vulnerable to statistical attack. Previous methods such as Stegdetect are based on Chi-square test. The accuracy of Stegdetect can be improved. When the size of the hidden message is small, it performs no better than random guess. In this paper they had proposed two steganalysis methods based on the logistic regression and the tree-based method {C4.5} for attacking LSB-based steganographic techniques. The experiments demonstrated that the performance of the logistic regression based technique is very impressive. When large amount of information is hidden, it can detect with very high accuracy. Even when the amount of hidden information is very small, it performs better than random guess. The tree-based method {C4.5} outperforms Stegdetect in the experiment where a relatively large amount of information is hidden. However, it does not perform well when the amount of hidden information is small. they suggest that one reason for {C4.5} not performing as well as the logistic regression is that it tends to produce boundaries that are parallel to the input variable axes, which in this case may not be appropriate. They also pointed out that the number of attributes used in classification can be related to a classifier's performance ,their experiments were carried out to break methods like Jsteg (see appendix B) that are employed to hide information. Their methods did not rely on the placement of the hidden information. Therefore they can be used without any modification on LSB based steganographic techniques that use random bit selection.

- T. Moerland [25], deled extensively with different methods of steganography applied to various (digital) media types like image (BMP, JPEG), audio (WAV, MP3) and plain text. Attacks, both visual and statistical will be considered as well. He showed how steganography has been used in the pre-digital age, and after that have constructed simple methods for hiding information in digital image, audio and text files. These methods however proved to be easily detectable, some even by obvious visual attacks like enhancing the least significant bit of the images. Several mathematical steganalytic methods have been reviewed, and he had shown how steganographic techniques could be improved to 16 evade these detections. Like in cryptography and cryptanalysis, this results in a cycle with both steganography and steganalysis trying to fool each other. Over the past one or two years sophisticated (mathematical) methods have been developed and steganography and steganalysis have grown into a mature science. Many interesting results are expected to be presented in the near future.
- George et al [6], proposed a machine learning (ML) approach to steganalysis. This work is based on a canvas representation of the media format that makes explicit all of the features that can be used for steganographic embedding. They have shown how this can be combined with a set of features selected from the canvas representation. In the current work, this includes value occurrence probabilities, and both unconditional and conditional entropies. These features were successfully used, for both GIF and JEPG formats, and by several different learning algorithms, to find hidden message bearing files. For JPEG format images, this approach outperforms one of the current state of the art

steganalysis techniques. Their system is certainly no panacea. They have shown examples of steganography algorithms for both GIF and JPEG formats that it cannot detect. While the current work uses straightforward features of the canvases and well-known learning methods, they indicate how it can be extended to more powerful representations and ML/DM methods. With these they anticipate being able to extend this work to use unsupervised anomaly detection approaches to steganalysis. These approaches should be able to detect the canvas features of clear media files, and hence should be able to distinguish those from stego-bearing files, regardless of the steganography method used. This should be able to detect steganography hidden using more powerful algorithms. In addition, by establishing general signatures of clean files, deviations from these signatures are a possible sign of steganographic embedding. While specific steganography algorithms would have specific deviations, any deviation raises the possibility of a hidden message. This holds the very exciting potential to transcend the current fragile nature of modern steganalysis – it may be possible to identify that a file has a hidden message, even if it is hidden using a new, previously unseen steganography algorithm. The results reported here show that ML algorithms work in both content- and compression-based image formats, outperforming at least one current hand crafted steganalysis technique in the latter. Their current work can detect previously seen (trained on) steganography techniques, and they discuss extensions that they believe will be able to detect steganography using more sophisticated algorithms, as well as the use of previously unseen steganography algorithms.

1.3 Aim of the Thesis

This thesis aims to design and implement a steganalysis process through scanning, testing suspected images to find out if it contained a hidden information ,depending on the use of Wavelet Transform(Haar wavelet transform) to extract features and use the ability of Probability density Function(PDF) to minimize the features to get the most important features that will be use in the statistical tests:-Standard tools(Average AD,MSE,SNR,PSNR,NCC,CQ)and proposed Statistics tools (Mean, Variance, Skewness, Kurtosis) these tests are useful in examine the suspected images if they contain a hidden information. The implementation this system is made on the BMP file. The embedded information is hided using LSB modification (Steganographic operation) and S-Tool (watermarked operation).

1.4 Chapters Overview.

- ***Chapter Two*** (Theoretical Background *of Steganography and steganalysis)* deals with image analysis, and information hiding ,Steganography, Steganalysis methods and wavelet transformation.
- ***Chapter Three*** (Proposed System For Steganalysis) the description of the proposed system is detailed describe.
- ***Chapter Four*** (System Interfaces of the Proposed system and Results) deals with the interfaces of the proposed system and the experimental results are detailed describe.
- ***Chapter Five*** (Conclusions And Future Works) consist of conclusion with recommended future works for the system.

Chapter Two
Theoretical Background *of Steganography and Steganalysis*

2.1 Introduction

Steganography is the art of hiding the presence of information by embedding secret messages into innocuous looking cover documents, such as digital images. Detection of Steganography, estimation of message length, and its extraction belong to the field of Steganalysis[12].

The first goal of Steganalysis is detection; there can be additional goals such as disabling, extraction, and confusion. Detection is more difficult than disabling in most cases, because disabling techniques can be applied to all files regardless of whether or not they are suspected of containing an embedded file [15].

Techniques and applications for information hiding have become increasingly more sophisticated and widespread. With high-resolution digital images as carriers, detecting the presence of hidden messages has also become considerably more difficult. It is sometimes possible, nevertheless, to detect (but not necessarily decipher) the presence of embedded messages [7].

The choice of cover images is important and influences the security in a major way. Images with a low number of colors، computer art, images with a unique semantic content, such as fonts, should be avoided. Some steganographic experts recommend grayscale images as the best cover-images .They recommend uncompressed scans of photographs or images obtained with a digital

camera containing a high number of colors and consider them safe for steganography.

2.2 Image Files

Image is an array of numbers that represent light intensities at various points (pixels). These pixels make up the image raster data. Such an image could contain about 300 kilobits of data .Digital images are typically stored in either 24-bit or 8-bit files [19].

All color variations for the pixels are derived from three primary colors: red, green, blue. One byte represents each primary color; 24-bit images use 3 bytes per pixel to represent a color value. These three bytes can be represented as hexadecimal number, actually three pairs represent red, green, and blue. A white background would have the value $(FF, FF, FF)_{16}$. Its decimal value is (255,255,255) .

This definition of weight background is analogous to the color definition of a signal pixel in an image. Pixel representation contributes to files size. *For example*, Suppose we have a 24-bit image 1,024 pixel wide bit 768 pixels high-a common graphics. Such an image has more than two million pixels, each having such a definition, which would produce a file exceeding 2 Mbytes [18].

2.2.1 Image analysis

Image analysis is a primarily a data reduction process. Images contain enormous amounts of data, typically on the order of hundreds of kilobytes or even megabytes. Often much of this information is not necessarily to solve a specific computer-imaging problem, so a primary part of the image analysis task is to determine

exactly what information is necessary. Image analysis is used in both computer vision and image processing application.

The image analysis process can be divided to three stages, as shown in Fig 2.1 [17].

- Preprocessing.
- Data Reduction, which includes two main sub-stages: Transformation, and Feature extraction.
- Features Analysis.

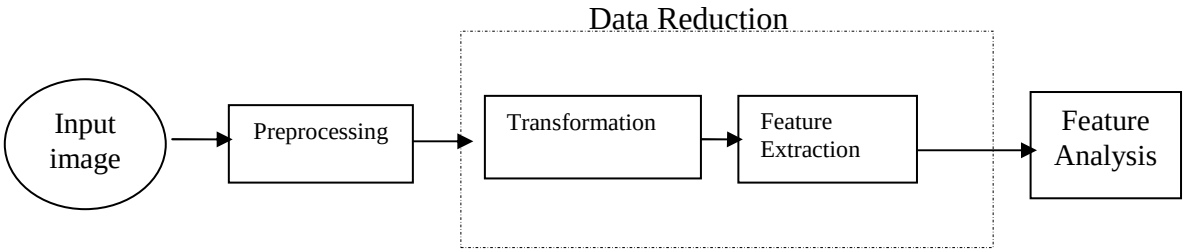


Fig 2.1 Image Analysis

a. Preprocessing

Techniques had operators that used to perform initial processing that makes the primary data reduction and analysis task easier. They include operations related regions of interest performing basic algebraic operations on images, enhancing specific image features, and reducing data in both resolution and brightness. Preprocessing is a stage where the requirements are typically obvious and simple, such as the elimination of image information that is not required for the application.

b. Transformation

Mathematical transformations are applied to signals to obtain further information (Features) from that signal that is not readily

available in the raw signal (time domain). Most of the signals in practice are *time-domain* signals in their raw format[17].

c. Feature Extraction and Analysis

The goal in image analysis is to extract information useful for solving application-based problems. This is done by intelligently reducing amount of image data including image transforms. After performing image transformation operation, image has been modified from the lowest level of pixel data into higher level representations. Features extraction can be considered as a useful operation for solving computer-imaging problems. The image transforms provide us with features based on spatial frequency information. [17].

2.3 Information Hiding

There are many approaches to hide the embedded file. The embedded file bits can be inserted in any order, concentrated in specific areas that might be less detectable, dispersed throughout the cover file, or repeated in many places. Careful selection of the cover file type and composition will contribute to successful embedding.

The number of bits in the cover file that get replaced will also affect the success of this method. In general, with each additional bit that is replaced the odd of detection increases, but in many cases more than one bit per cover file byte can be replaced successfully. Combining the correct selection of bits with analysis of the maximum number of bits to replace should result in the smallest possible impact to the statistical properties of the cover file. [19].

2.3.1 Steganography

Steganography is the art of invisible communication. Its purpose is to hide the very presence of communication by embedding messages into innocuous-looking cover objects. In today's digital world, much more versatile and practical covers for hiding messages digital documents, images, video, and audio files have replaced invisible ink and paper. As long as an electronic document contains perceptually irrelevant or redundant information, it can be used as a "cover" to hide secret messages [12].

Classical Steganographic systems depend on keeping the encoding system secret, modern Steganography tries to be undetectable unless secret information is known, namely, a secret key. [20].

Computer-based image Steganography is one way of data hiding which provides data security in digital images. It is considered as a technique inspired from ancient Steganography. The aim is to embed and deliver secret messages in digital images without any suspiciousness. The secret message might be a caption, a plain text, another image, a control signal, or anything that can be represented in bit stream form. The secret message may be compressed and encrypted before the embedding steps begin [5].

The goal of Steganography is to avoid drawing suspicion to the transmission of a hidden message. If suspicion is raised, then this goal is defeated [18].

Each Steganographic communication system consists of an embedding algorithm and an extraction algorithm. To accommodate a secret message, the original image, also called the cover-image, is slightly modified by the embedding algorithm. As a result, the stego-

image is obtained. The set of images from which cover-images are being drawn is part of the communication system [20].

2.3.2 Watermarking

It is a techniques used to imperceptibly convey information by embedding it into the cover-data. a popular application of watermarking is to give proof of ownership of digital data by embedding copyright statement. It is obvious that for this application the embedded information should be robust against manipulations that may attempt to remove it [18]

There has been a large amount of work in embedding data to maintain ownership of digital media. In this case the watermark should be as resistant to as many distortions as possible [10] [3].

2.4 Steganalysis

The broad goal of Steganalysis is to understand the effects of hiding data into a medium. This knowledge is typically used to either strengthen the hiding system or detect the use of data hiding [10][23].

Though the first goal of Steganalysis is detection, there can be additional goals such as disabling, extraction, and confusion. Detection of Stenography, estimation of message length, and its extraction belong to the field of Steganalysis. Detection is more difficult than disabling in most cases, because disabling techniques can be applied to all files regardless of whether or not they are suspected of containing an embedded file. However, if only a minute portion of all files is suspected to have embedded files then disabling in this manner is not very efficient.

Attacks and Analysis on hidden information may take several forms: detection, extracting, confusing (counterfeiting or overwriting

by an attacker, embedding counter information over the existing hidden messages), and disabling hidden information[24].

Cryptanalyst is one who applies cryptanalysis in an attempt to decipher encrypted messages, while Steganalyst is one who applies Steganalysis in an attempt to detect the existing of hidden information. In *Cryptanalysis,* Portion of the plain text and portions of the ciphertext are analyzed. In Steganalysis, Comparisons are made between the cover-object, the stego-object, and possible portion of the message. The end result in the cryptography is the cipher text, while the end result in Steganography is the stego-object.

In order to define attack techniques used for Steganalysis, corresponding techniques are considered in cryptanalysis. Attacks available ton the cryptanalyst are ciphertext only, known plaintext, chosen plaintext, and chosen ciphertext [18].

In *ciphertext only*, the cryptanalyst knows the ciphertext to be decoded. The cryptanalyst may have the encoded messages and part of the decoded messages which together may be used for a know plaintext attack .*the chosen plaintext*, it's the most favorable case for the cryptanalyst. In this case, the cryptanalyst has some ciphertext which corresponds to some plaintext chosen by the cryptanalyst has some ciphertext which corresponds to some *plaintext chosen* by the cryptanalyst. If the encryption algorithm and ciphertext are available, the cryptanalyst encrypts plaintext looking for matches in the ciphertext. This *chosen ciphertext* attack is used to deduce the sender's key [24].

2.4.1 Attacks available to the Steganalyst

Detection set signatures is based on the combination of carrier, stego-media, embedded messages, and steganography tools known by the analyst [19][24].***The associated attacks are***

1) **Stego-only Attack.** Only stego-object is available for analysis.
2) **Known cover attack.** The ”original” cover-object and stego-object are both available.
3) **Known message attack.** At some point, the Attacker may know the hidden messages. Analyzing the stego-object for pattern that correspond to the hidden message may be beneficial for future attacks against that system. Even with the message, this may be very difficult and may even be considered equivalent to the stego-only attack.
4) **Chosen stego attack**. The steganography tool (algorithm) and stego-object known.
5) **Chosen message attack.** The steganalyst generates stego-object from some steganography tool or algorithm from a chosen message. The goal in this attack is to determine corresponding pattern in the stego-object that may point to the use of specific steganography tool or algorithm.
6) **Known stego attack** .The steganography algorithm (tool is known and both the original and stego-object are available.

In order to develop a hiding scheme, which is difficult to detect, it is necessary to analyze the results of prospective methods. Comparing statistical changes introduced when embedding data typically does

this. If a method causes distinct predictable changes it will be fairly easy to detect and should be modified [10].

2.4.2 Steganalysis Methods

One of steganalysis technique is "the visible detection", which includes human observers detecting minute changes between a cover file and a stego file, or it can be done automatically. Additionally, since many Steganography tools take advantage of close colors or create their own close color groups, many similar colors in an image palette may make the image become suspect [15] [12].

Steganalysis can also involve the use of "statistical techniques". By analyzing changes in an image's close color pairs, the steganalyst can determine if LSB substitution was used. Close color pairs consist of two colors whose binary values differ only in the LSB. The sum of occurrences of each color in a close color pair does not change between the cover file and the stego file . This fact, along with the observation that LSB substitution merely flips some of the LSBs, causes the number of occurrences of each color in a close color pair in a stego file to approach the average number of occurrences for that pair [15]. These statistical techniques benefit from the fact that the embedding process alters the original statistics of the cover file and in many cases these first-order statistics will show trends that can raise suspicion of Steganography[15].

Fridrich[12] and others proposed a steganalytic technique called the RQP method. It is used on color images with 24-bit pixel depth where the embedded file is encoded in random LSBs. RQP involves inspecting the ratio between the number of close color pairs and all pairs of colors. This ratio is calculated on the suspect image, a test message is embedded, and the ratio is calculated again. If the initial and final ratios are vastly different then the suspect image was likely

clean. If the ratios are very close then the suspect image most likely had a secret message embedded in it [13].

Image Domain tools encompass bit-wise methods that apply least significant bit (LSB) insertion and noise manipulation. The tools used in this group include StegoDos , S-Tools , Mandelsteg , EzStego , Hide and Seek (versions 4.1 through 1.0 for Windows 95) , Hide4PGP, Jpeg-Jsteg , White Noise Storm , see appendix B, and Steganos. The image formats typically used in such Steganography methods are lossless and the data can be directly manipulated and recovered.

The *Transform Domain* grouping of tools include those that involve manipulation of algorithms and image transforms such as Discrete Cosine Transformation (DCT) and wavelet transformation . These methods hide messages in more significant areas of the cover and may manipulate image properties such as luminance. Watermarking tools typically fit this categorization and the subset used are PictureMarc , JK-PGS , SysCop, and SureSign. These techniques are typically far more robust than bit-wise techniques[18].

However, if the cover-image, was initially stored in the JPEG format (as it is frequently the case), message embedding in the spatial domain will disturb but *not* erase the characteristic structure created by the JPEG compression and one can still easily determine whether or not a given image has been stored as JPEG lastly. Indeed, it is possible to recover the JPEG quantization table from the stego-image by carefully analyzing the values of DCT coefficients in all 8×8 blocks. After message embedding, however, the cover-image will become (with a high probability) incompatible with the JPEG format

in the sense that it may be possible to prove that a particular 8×8 block of pixels could not have been produced by JPEG decompression of any block of quantized coefficients. This finding provided a strong evidence that the block has been slightly modified [12].

With the exception of the "JPEG compatibility Steganalysis" which can be applied to any spatial Steganographic method, all previously proposed methods are tailored to a specific embedding algorithm or its slight variations. "Universal blind Steganalysis" is a detection method in the sense that it can be adjusted, after training on original and stego-images, to detect any Steganographic method regardless of the embedding domain. The trick is to find an appropriate set of sensitive statistical quantities (a feature vector) with “distinguishing” capabilities. Neural networks, clustering algorithms, Statistical methods, and other tools of soft computing can then be used to find the right thresholds and construct the detection model from the collected experimental data [12] [10].

2.5 Detection

The goal of detection is to get a decision for images if it contains a hidden data or not. Due to the adaptation of the following techniques in the developed system they will be discussed in the next sections.

1. Frequency Transform (Wavelet Transform)
2. Statistical Tests.

2.5.1. Frequency Transform.

2.5.1.1 Wavelet Transform

Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale [4].

Wavelet theory is based on analyzing signals to their components by using a set of basis functions. One important characteristic of the wavelet basis functions is that they relate to each other by simple scaling and translation. The original wavelet function, known as mother wavelet, which is generally designed based on some desired characteristics associated to that function, is used to generate all basis functions. For the purpose of multiresolution formulation, there is also a need for a second function, known as scaling function, to allow analysis of the function to finite number of components.

In most wavelet transform applications, it is required that the original signal be synthesized from the wavelet coefficients. This condition is referred to as perfect reconstruction. In some cases, however, like pattern recognition type of applications, this requirement can be relaxed. In the case of perfect reconstruction, in order to use same set of wavelets for both analysis and synthesis, and compactly represent the signal, the wavelets should also satisfy orthogonality condition. By choosing two different sets of wavelets, one for analysis and the other for synthesis, the two sets should satisfy the biorthogonality condition to achieve perfect reconstruction.

In general, the goal of most modern wavelet research is to create a mother wavelet function that will give an informative, efficient, and useful description of the signal of interest. It is not easy to design a uniform procedure for developing the best mother wavelet or wavelet

transform for a given class of signals. However, based on several general characteristics of the wavelet functions, it is possible to determine which wavelet is more suitable for a given application [1].

They have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. Wavelets were developed independently in the fields of mathematics, quantum physics, electrical engineering, and seismic geology. Interchanges between these fields during the last ten years have led to many new wavelet applications such as image compression, turbulence, human vision, radar, and earthquake prediction[1][4].

2.5.1.2. Wavelet Analysis of signal processing

In signal processing there are numerous examples of the benefits of working in the frequency domain. Fourier analysis remains a powerful technique for transforming signals from the time domain to the frequency domain. However, time information is hidden in the process. In other words, the time of a particular event can not be discerned from the frequency domain view without performing phase calculations, which is very difficult for practical applications[2].

The Fourier transform was modified to create the Short-Time Fourier Transform (STFT) in an attempt to capture both frequency and time information. The STFT repeatedly applies the Fourier transform to disjoint, discrete portions of the signal of constant size. As a result, wavelet analysis can better capture the interesting transitory characteristics of a signal.

Though the above discussion of Fourier analysis and wavelet analysis made reference to the time and frequency domains typically

associated with signal processing, the concepts also apply to the spatial and frequency domains associated with image processing.

2.5.1.3 Discrete Wavelet Transform (DWT)

There are different types of wavelet transforms, including the Continuous Wavelet Transform (CWT) and the Discrete Wavelet Transform (DWT).

The CWT is used for signals that are continuous in time and the DWT is used when a signal is sampled, such as during digital signal processing or digital image processing. The DWT has a scaling function **Φ** and a wavelet function **φ** associated with it. The scaling function **Φ** can be implemented using a low pass filter and is used to create the "scaling coefficients" that represent the signal approximation. The wavelet function **φ** can be implemented as a high pass filter and is used to create the "wavelet coefficients" that represent the signal details. If the DWT is used by scaling and shifting by powers of two (dyadic), the signal will be well represented and the decomposition will be efficient and easy to compute. In order to apply the DWT to images, combinations of the filters (combinations of the scaling function and the wavelet function) are used first along the rows and then along the columns to produce unique subbands [15][2].

The **L**ow pass filters are applied horizontally and vertically (LL) along the rows and columns and is commonly referred to as a course approximation of the image because the edges tend to smooth out. The **L**ow pass filter is applied horizontally and The **H**igh pass filter is applied vertically(LH) which is produced by low pass filtering along the rows and high pass filtering along the columns, thus capturing the horizontal edges.

The **H**igh pass filter is applied horizontally and The **L**ow pass filter is applied vertically(LH) which is produced by high pass

filtering along the rows and low pass filtering along the columns, thus capturing the vertical edges. The **H**igh pass filters are applied horizontally and vertically(HH) produced by high pass filtering along the rows and columns, thus capturing the diagonal edges. The LH and HL subbands are considered the band pass subbands and the LH, HL, and HH subbands together are called the detail subbands. These subbands are shown in Figure 2.2 By repeating the process on the LL subband, additional scales are produced.

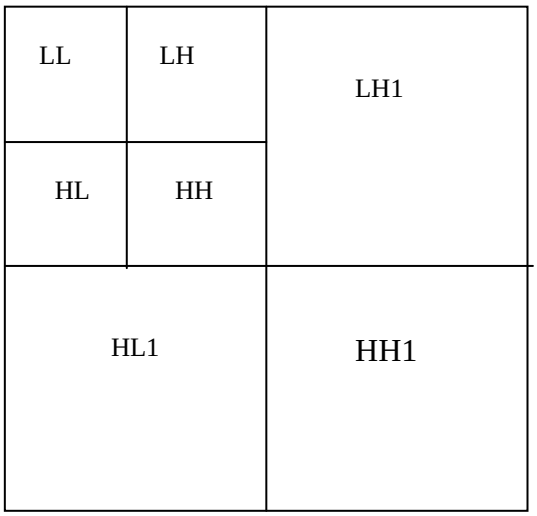


Fig2.2 Wavelet Subbands of an image

The statistics of the generated coefficients of the various subbands offer valuable results, According to Farid[[7],[8],[9]], A broad range of natural images tend to produce similar coefficient statistics. Additionally, alterations such as Steganography tend to change those coefficient statistics. The alteration was enough to provide a key for Steganography detection in Farid's research [7][8][9][23].

One set of statistics that Farid used consisted of the mean, variance, skewness, and kurtosis of the coefficients generated at the LH, HL, and HH subbands for all scales. He concluded that his method would be just as successful on other known methods. Also, the ratio of embedded file size to cover file size will typically affect

the accuracy of just about any steganalytic technique and this method is no exception. [15]

2.5.1.4 Discrete Wavelet Decomposition

A signal x of length N can be decomposed in any level to give a coarser approximation of the signal in the next level. The approximation coefficients at level j is given by:

$$c_j(k) = \sum_m h(m - 2k)\, c_{j+1}(m) \qquad (2.1)$$

$$d_j(k) = \sum_m g(m - 2k)\, c_{j+1}(m) \qquad (2.2)$$

Where c are called the scaling function or the approximation coefficients and d are called the wavelet or the details coefficients. h, g are both finite even length discrete values wavelet filters called the decomposition low-pass and high-pass wavelet filters respectively. The calculation of these filters will briefly be described in section 2.5.1.6.

Assuming that c with the highest resolution subscript is the original input signal. At each stage of the decompositions (2.1) and (2.2), the length of the resulting signals c_j and d_j is half the length of c_{j+1} because of the down-sampling process after each time in which the decomposition occurs.

The down-sampler (sometimes called a sampler or decimator) takes a signal x(n) as an input and produces an output y(n)=x(2n). The down-sampler is symbolically shown in figure (2.3).

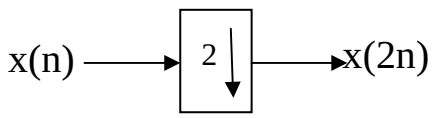


Fig 2.3 the Down-Sampler

In down-sampling, there is clearly the possibility of losing information since half of the data is discarded. The scale-j coefficients are filtered with the two low-pass h and high-pass g filters, after which down-sampling gives the next coarser scaling and wavelet coefficients. Both of them, having lengths equal to half the length of the input signal. In Figure 2.4 , the decomposition process is depicted for two decomposition stages[24][10][4][2].

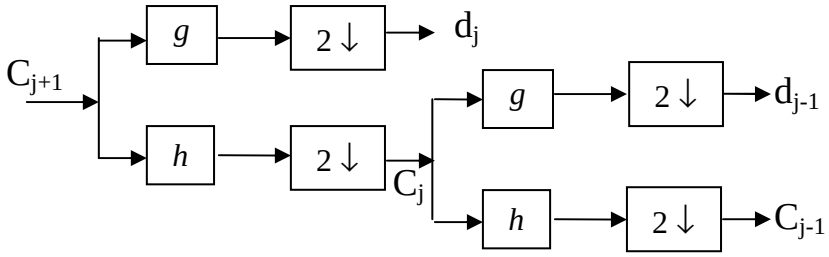


Fig 2.4 Two-Stage decomposition process

2.5.1.5 Discrete Wavelet Reconstruction

A signal considered at a resolution j+1, can be reconstructed from the combination of the scaling function and wavelet coefficients at a coarser resolution j. This can be written as:

$$c_{j+1}(k) = \sum_m c_j(m)\bar{h}(k-2m) + \sum_m d_j(m)\bar{g}(k-2m) \qquad (2.3)$$

Where $\bar{h}, \bar{g}$ are both finite even length discrete values wavelet filters called the reconstruction low-pass and high-pass wavelet filters respectively and are derived directly from the low-pass and high-pass decomposition filters.

At each stage of the reconstruction process given in equation (2.3), the length of the resulting signals c_{j+1} equals to the sum of the length of both c_j and d_j because of the up-sampling process after each time in which the reconstruction occurs [4].

The up-sampling means that the input to the filter has zeros inserted between each of the original terms. In other words y(2n)=x(n) and y(2n+1)=0.

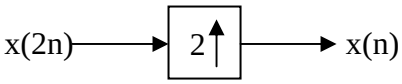


Fig 2.5 The Up-Sampler

The input signal is stretched to twice its original length and zeros are inserted. The up-sampler is symbolically shown in Figure (2.6).

In Figure 2.7 , the reconstruction process is depicted for two reconstruction stages.

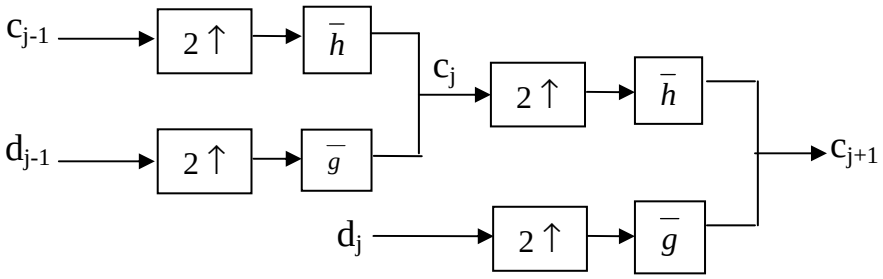


Fig 2.6 Two-Stage reconstruction process

2.5.1.6 Wavelet Filters Calculation

The wavelet filters are even finite-length real numbers coefficients that are convolved with the input signal to produce the next level approximation or details coefficients. The high-pass filter coefficients should be orthogonal to the low-pass filter coefficients. The decomposition high-pass, reconstruction low-pass, and reconstruction high-pass filter can be derived from the decomposition low-pass filter as shown below:

$$g(n) = (-1)^n h(N-n+1) \qquad (2.4)$$

$$\bar{h}(n) = h(N-n+1) \qquad (2.5)$$

$$\bar{g}(n) = (-1)^{n+1} h(n) \qquad (2.6)$$

Where n=1,2,…,N, and N is the finite even length of the wavelet filter.

2.5.1.7 Wavelet Transform Characteristics

Wavelet transform have proven to be very efficient and effective in analyzing a very wide class of signals and phenomena. The reasons for that are:

1. The size of the wavelet expansion coefficients drop off rapidly with its indices for a large class of signals. This property is being called an *unconditional basis* and it is why wavelets are very effective in signal and image compression, denoising, and detection.
2. The wavelet expansion allows a more accurate description and separation of signal characteristics. A wavelet expansion coefficient represents a component that is itself local and is easier to interpret. The wavelet expansion may allow a separation of components of a signal that overlap in both time and frequency.
3. Wavelets are adjustable and adaptable. Because there is not just one wavelet, they can be designed to fit individual applications. They are ideal for adaptive systems that adjust themselves to suit the signal.
4. The generation of wavelets and the calculation of the discrete wavelet transform is well matched to the digital computer. The defining equation for a wavelet uses no calculus. There are no derivatives or integrals, just multiplications and additions operations that are basic to a digital computer.[4]

2.5.2 Statistical Tests

2.5.2.1 Standard Tests

The statistical analysis (Test) is very useful for detection of embedded data in an image. These analysis or test can expose abnormality in an image that is not visible to human eyes.

These tests need the original and the suspected images for getting the correct results.

1. *Average Absolute Difference(AD)test*

The average of the difference between color of pixels in original image and suspected image can be calculated as:

$$AD = \frac{1}{MN} \sum_{x,y} \left| I_{x,y} - I`_{x,y} \right| \quad \text{..... (2.7)}$$

Where

I is the original Image, I' is the Suspected Image

M is the height, N is the width,

x, y are indexes

The absolute values used for difference to get accurate result of summation difference output.

2. *Mean Squared Error(MSE)Test*

To calculate the MSE between the original and the suspected images, we must et the difference of pixel color in the two images. The resulted will be the square amount of error depending on the size of these images, as shown:

$$MSE = \frac{1}{MN} \sum_{x,y} \left(I_{x,y} - I`_{x,y} \right)^2 \quad \text{.....(2.8)}$$

3. *Signal-to-Noise Ratio(SNR)Test*

$$SNR = \sum_{x,y} I^2 \; x,y \Big/ \sum_{x,y} \left(I_{x,y} - I`_{x,y}\right)^2 \quad \text{……..(2.9)}$$

4. *Peak Signal-to-Noise Ratio(PSNR)Test*

We use (PSNR) to calculate the maximum peak signal -to- noise ratio between the two images as follows

$$PSNR = MN * \max*(I^2 x,y \Big/ \sum_{x,y} \left(I_{x,y} - I`_{x,y}\right)^2 \quad \text{……(2.10)}$$

5. *Normalize Cross Correlation (NCC)*

This metrics is important for see the amount the correlation between the original and suspected images.

The (NCC) can be measured from the following equation:

$$NCC = \sum_{x,y} I_{x,y} * (I`_{x,y} \Big/ \sum_{x,y} I^2 x,y) \quad \text{…….(2.11)}$$

6. *Correlation Quality (CQ) Test*

Quality of correlation can be measured between the original images and the suspected images depending on the size of these images by using the following equation:

$$CQ = \sum_{x,y} I_{x,y} * (I`_{x,y} \Big/ \sum_{x,y} I_{x,y}) \quad \text{……..(2.12)}$$

2.5.2.2 The Developed Tests

1.*Mean*:

Mean is a simple, intuitive and easy to implement method of *smoothing* images, *i.e.* reducing the amount of intensity variation between one pixel and the next. It is often used to reduce noise in images.

The mean of input image I (m x n) is define as the total of its brightness over the product of the dimensions M an N.

$$\textbf{Mean= } \mu = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} I\ (i,\ j) \quad \text{……..2.13}$$

$$\textbf{Or=} \frac{1}{MN} \sum_{i=0}^{255} i * His\ (i) \quad \text{…….…..……2.14}$$

***2.Variance*:**

Variance measures the average of the squared summations of frequencies.

$$\textbf{Variance}=\sigma^2= \frac{1}{MN} \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} \left(I_{(i,j)} - \mu\right)^2 \quad \text{.... .....2.15}$$

$$\sigma^2= \frac{1}{MN} \sum_{i=0}^{255} (i - \mu)^2 His\ (i) \quad \text{……..(.2.16)}$$

3. Skewness:

Skewness is the deviation of the frequencies distribution curve over similarity

$$\varsigma= \frac{1}{MN} \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} \left(\frac{I_{(i,j)} - \mu}{\sigma_x} \right)^3 \quad \text{…….(2.17)}$$

There are two types of Skewness Positive Skewness and Negative skewness,

If $\mu > I(i, j)$ then the curve will be positive Skewness else the curve is negative

4. *Kurtosis:*

Kurtosis is the deviation of the top of the frequencies distribution curve over similarity.

$$\kappa = \frac{1}{MN} \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} \left(\frac{I_{(i,j)} - \mu_x}{\sigma_x} \right)^4 \quad \text{……..(2.18)}$$

There are several types of Kurtosis i.e. Lepto, Meso, and Platy Kurtosis.

if $\kappa=0$ then the top will be Meso Kurtosis.

$\kappa>0$ then the top will be Lepto Kurtosis.

$\kappa<0$ then the top will be platy Kurtosis.

2.6 Histogram and Probability Density Function

In statistics, a **histogram** is a graphical display of tabulated frequencies. That is, a histogram is the graphical version of a table which shows what proportion of cases fall into each of several or many specified categories. The categories are usually specified as nonoverlapping intervals of some variable.

There are different ways to display the same table, and two kinds of histograms are there. One shows the number of cases per unit interval, so that the area under the curve is the total number of cases. Other shows the number of cases per unit interval divided by the total number of cases, so that the area under the curve is exactly 1.

In mathematics, a **Probability Density Function (PDF)** serves to represent a probability distribution in terms of integrals. Any function that is everywhere non-negative and whose integral from $-\infty$ to $+\infty$ is equal to 1 is a probability density function.

If a probability distribution has density $f(x)$, then intuitively the infinitesimal interval $[x, x + dx]$ has probability $f(x)\,dx$. Informally, a probability density function can be seen as a "smoothed out" version of a histogram: if one empirically measures values of a continuous random variable repeatedly and produces a histogram depicting relative frequencies of output ranges, then this histogram will resemble the random variable's probability density (assuming that the variable is sampled sufficiently often and the output ranges are sufficiently narrow).[2].

Chapter Three

Proposed System for Steganalysis

3.1 Introduction

This Chapter concerns with the description of the design and implementation of the Proposed System. The system's layout will be presented. Next, the implantation of the main parts of the system will be discussed.

3.2 The Proposed System

The Proposed System as shown in Fig 3.1 can be divided up to four parts as:-

- Hiding process
- Image Preprocessing
- Features Extraction and Analysis Processes
- Discrimination Processes

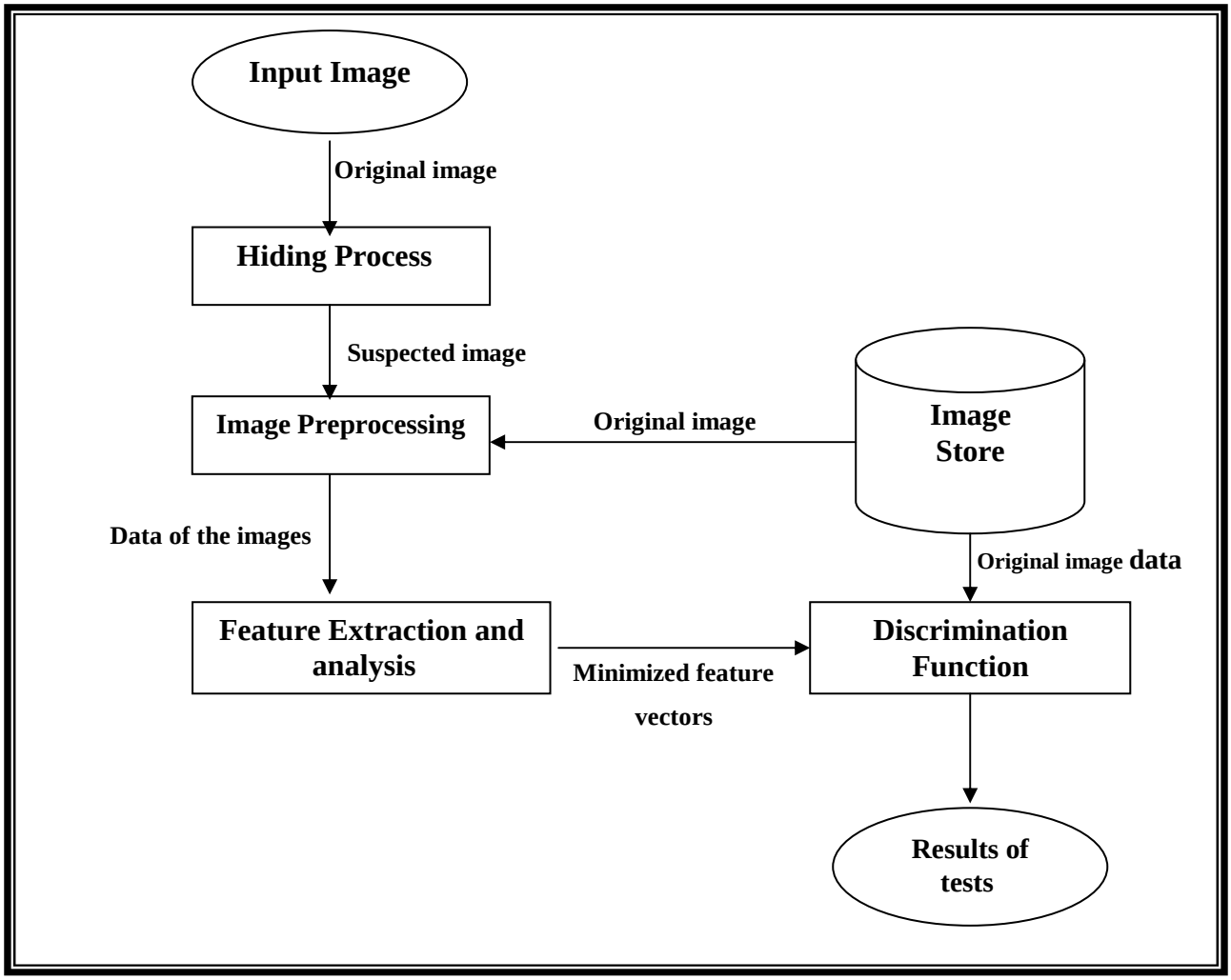


Fig 3.1 The overall System Model

3.2.1 Hiding process

In **LSB** modification, The embedded object is embedded in the spatial domain i.e. in two dimensional space of an image .At the beginning the color-plane (R, G, B, or all of them) is chosen, where the selected plane will carry the embedded object, by substitution the original bit with embedded bit.

The embedding process requires the original image (I), the embedded object(S) and the three parameters (color plane, pixels position, bit position) as input, the output of embedding process are stego image (I),as will be seen in the Algorithm 3.1 ,the block of LSB modification is illustrated in Fig 3.2,This operation is once performed .

Algorithm 3.1 LSB modification

```
INPUT
  I ( )   // image's data
  S ( )  //block of embedded object
  sz     //  size of block
OUTPUT
   I'( )  // data of stego image
Begin
    Loop s=1 to sz
        I'(s).blue= I(s).blue And 254
        I'(s).blue= I(s).blue Or S(s)
     End Loop
End
```

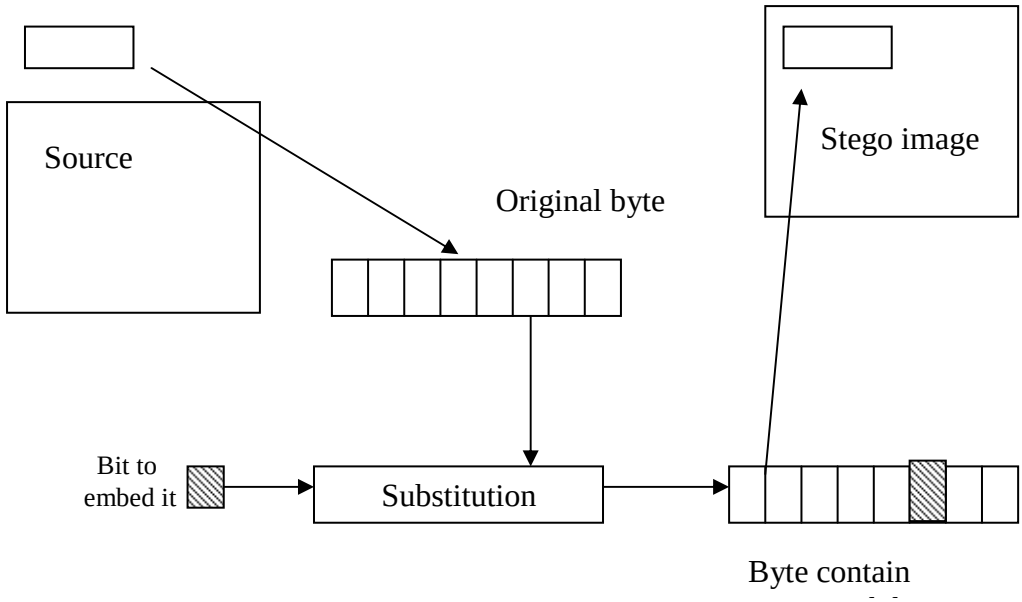


Fig 3.2 LSB insertion

3.2.21Image Preprocessing

Preprocessing algorithm, techniques and operators are used to perform initial processing that makes the primary data reduction and analysis task easier. In our system it represent opening the suspected image that will be tested in the system as the input image. The images are saved as BMP format that start out with header followed by sequence of byte .The size of BMP header is 54 bytes and the data of the image is beginning from 55 to the end of the image pixel. The implemented image is describe as 24 bit per pixel, each pixel have three band (red, Green, Blue)as shown in fig 3.3, the detailed description of BMP is given at appendix A, Algorithm 3.2 search for the original image through the database folder assuming it's already exist. And save it bitmap data for the next process.

Generally, Images came in different sizes and different Aspect ratio; therefore, we will scale the images into uniform size. Each image will be scaled to power of two (i.e. 48*48) so to put in a uniform size, and for both the suspected **I'** and the original image **I** the algorithm 3.3 will be applied.

The data at the BMP file is stored as reverse sequence; we can notice them when viewing the BMP image it will appear from the bottom of the image .So reordering the data to make it appropriate to processing. Reordering the bitmap data is stated in Algorithm 3.4.

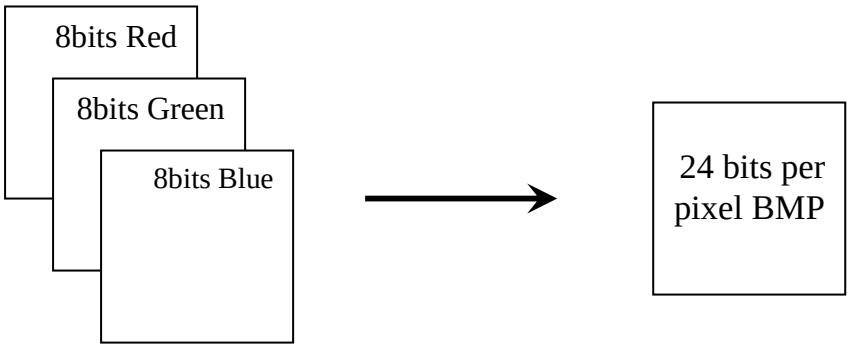


Fig 3.3 BMP Format

Algorithm 3.2 Open BMP File

INPUT

I is the original image, **I'** is the suspected image.

OUTPUT

D ( ) //Array of one dimension of the bitmap data of the Original image

D'( ) // Array of one dimension of the bitmap data of the Suspected image

Begin

Search For the original image in the database and open it as **I'**.

For both **I** and **I'** do

Open **I, I'** file for read

Read **I, I'** headers

Read the bitmap data of **I** and **I'** as **D** and **D'** respectively

End

Algorithm 3.3 Resize the images

INPUT

Let image be $\mathbf{I_1}$ with size $h_1 * w_1$

OUTPUT

The scaled image $\mathbf{I_2}$ with size h2*w2

Begin

For each pixel (x_2,y_2)in $\mathbf{I_2}$ do

Find the corresponding pixel (x1,y1) in $\mathbf{I_1}$ to have the value of $\mathbf{I_2}(x_1,y_1)$ pixel

$$X_1 = x_2 * ((w_1 - 1) / (w_2 - 1))$$

$$Y_1 = y_2 * ((h_1 - 1) / (h_2 - 1))$$

$\mathbf{I_2}(x_2,y_2) = \mathbf{I_1}(x_1,y_1)$

END

Algorithm 3.4 image Reordering

```
INPUT
  readimage with size N*M
OUTPUT
reordered image  I
 // the pixel consist of three colors(Blue, Green ,Red) (R,G,B)respectively,
one byte is allocated to each color plane
Begin
For y = 1 to M
 For x = 1 to N
      I (h - y + 1, x).b = readimage (3 * x - 2, y)
      I (h - y + 1, x).g = readimage (3 * x - 1, y)
      I (h - y + 1, x).r = readimage (3 * x, y)
 End
```

3.2.3 Feature Extraction and analysis Processes

Feature extraction refers to the process of forming a new set of features from the original and suspected features set, and find a mapping that reduces the dimensionality of pattern by extraction some numerical measurements from raw input pattern. There is no well-develop theory feature extraction; most is application oriented.

The extraction of feature vector are derived from the wavelet transformation process, For each suspected and original image, theses features are coefficients produced in this transformation and assumed to be fixed, do not changed after several image processing operation.

There are different types of wavelet transformations depending on the bases functions used in the transformation, in our work, Haar transformation is used to implement wavelet transformation. The Haar bases vectors are :-Lowpass : $\frac{1}{\sqrt{2}}[1,1]$,Highpass : $\frac{1}{\sqrt{2}}[1,-1]$

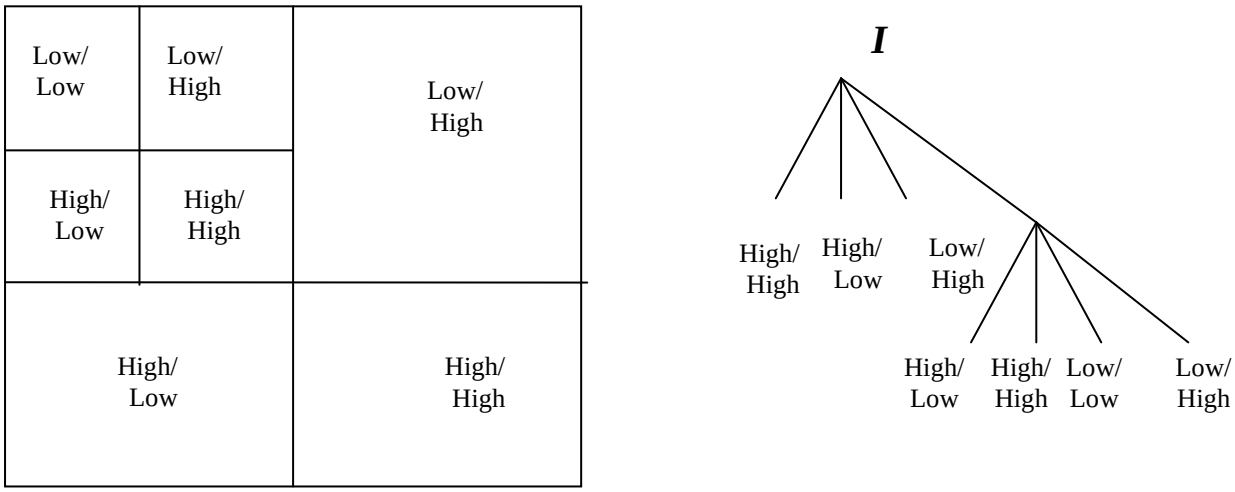


Fig 3.4 a. 2-D wavelet Transform, b. Tree presentation of 2-D wavelet Transform

Decompose a given image with 2-D wavelet transform into 4 images, as indicated in Fig 3.4 the image is divided into four subbands after wavelet transform: horizontal, vertical, diagonal subimages, and low resolution subimages, which can be viewed as tree in Fig3.3.b

The algorithm used for extracted feature vector is stated bellow:-

Algorithm 3.6 Wavelet Transformation

INPUT

image

OUTPUT

Feature Vector.

Begin

1. Convolve the lowpass filter with the rows and save the results
2. Convolve the lowpass filter with the columns of the results from step1 and sub sample this result by taking every other values; this give us the lowpass filter version of image.
3. Convolve the result from step1, the lowpass filtered rows, with the highpass filter on the columns, subsample by taking every other value to produce the low-pass highpass image.
4. Convolve the original image with the highpass filter on the rows and save the result.
5. Convolve the result in step 4 with the lowpass filter on the columns; subsample to yield the highpass-lowpass version of the image.
6. Convolve the columns from step 4 with highpass filter to obtain highpass version of the image.

End.

Now we find the Histogram, It is a simple graph that displays where all of the brightness levels contained in the scene are found, from the darkest to the brightest. These values are arrayed across the bottom of the graph from left (darkest) to right (brightest). The vertical axis (the height of points on the graph) shows how much of the image is found at any particular brightness level.

Then Apply "Probability Density function ", that Distribute the data using Probability density function as a way to minimize the feature vector in order to choose the best features that required and important in the work

Algorithm 3.7Histogram and PDF

INPUT

Subband image **I**

OUTPUT

PDF ( ) // minimized vector.

$$\sum_{y=0}^{h-1}\sum_{x=0}^{w-1} \quad \text{His (I(x, y)) =His (I (x, y)) +1}$$

$$\text{PDF (j)} = \left(His\ (\ j)\right) \Big/ \sum_{i=0}^{255} His\ (i)$$

3.2.4 Discrimination Processes

After specifying the features vector by performing feature extraction and analysis, methods for comparing two features vectors need to be determined. These methods are either to measure the differences between the two or measure the similarity**.** These are the Statistical Tests as mentioned previously in chapter two.

Chapter Four
System Interfaces of the Proposed System And Results

4.1 Introduction

This chapter will concern System Interfaces of the Proposed System With discussion of the results produced by the proposed system to perform the operation of Steganalysis . Calculations of the results mentioned in this chapter based on samples taken from a group of images testing it with all options . The proposed system is checking the images (deals with BMP image file format) if it contains a Stego object or not.

4.2System Interfaces

The steganalysis system was designed to detect the existing of hidden images in images (BMP image file format). Main diagram is shown in Fig 4.1

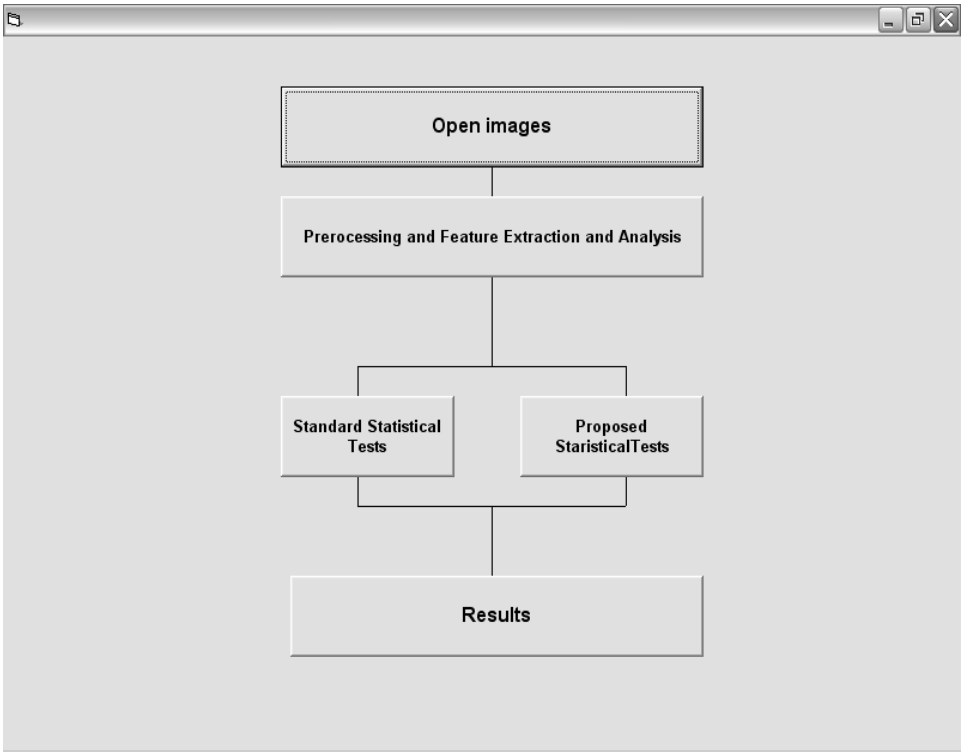


Fig 4.1 Main Diagram

It represents The proposed system , the first part is “Open images ”, when fig (4.2) is appear to start processing on the suspected image and searching for the original image , then the stages of preprocessing and feature extraction and analysis, then choosing Discrimination Function which is either standard test(4.8) , or proposed statistics (4.11 and 4.12)

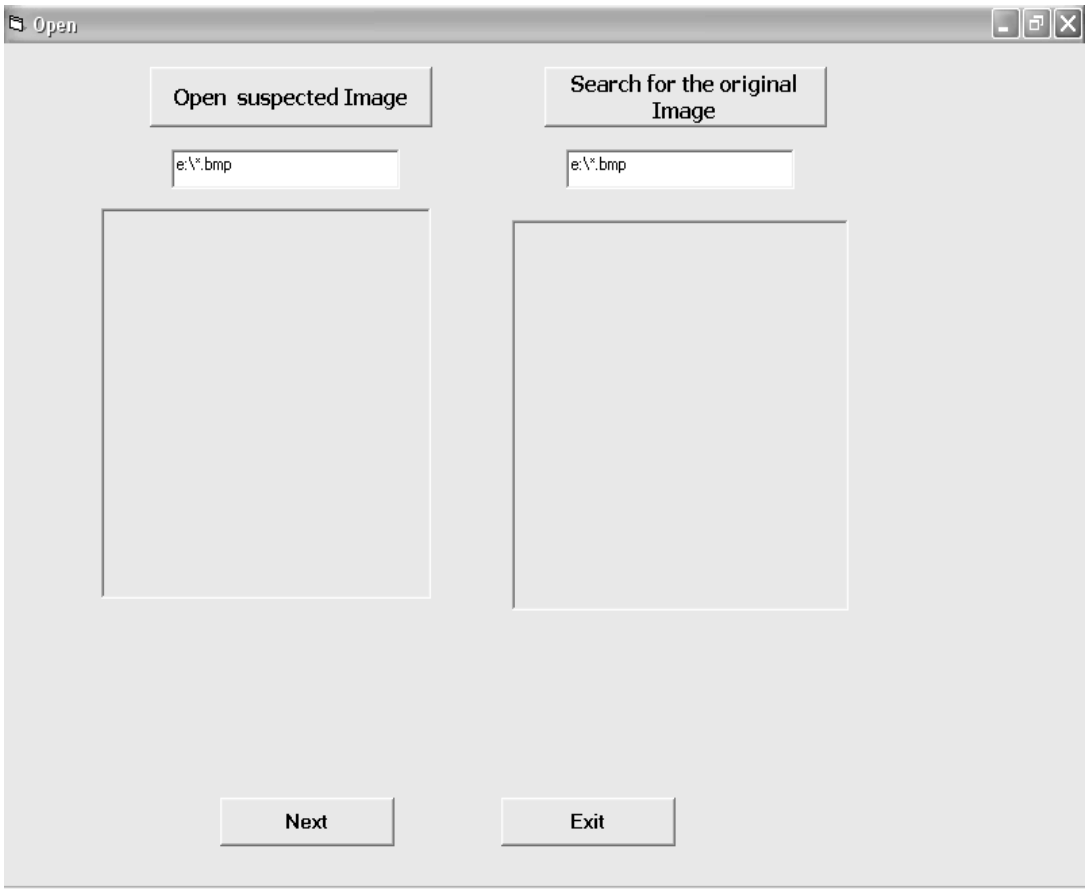


Fig 4.2 Open Images

"Open suspected image" icon is referring to get the suspected image to start testing. Fig 4.3 shows the opening process, now choose the suspected image The image will appear in the dialogs shown later in (Fig 4.5).

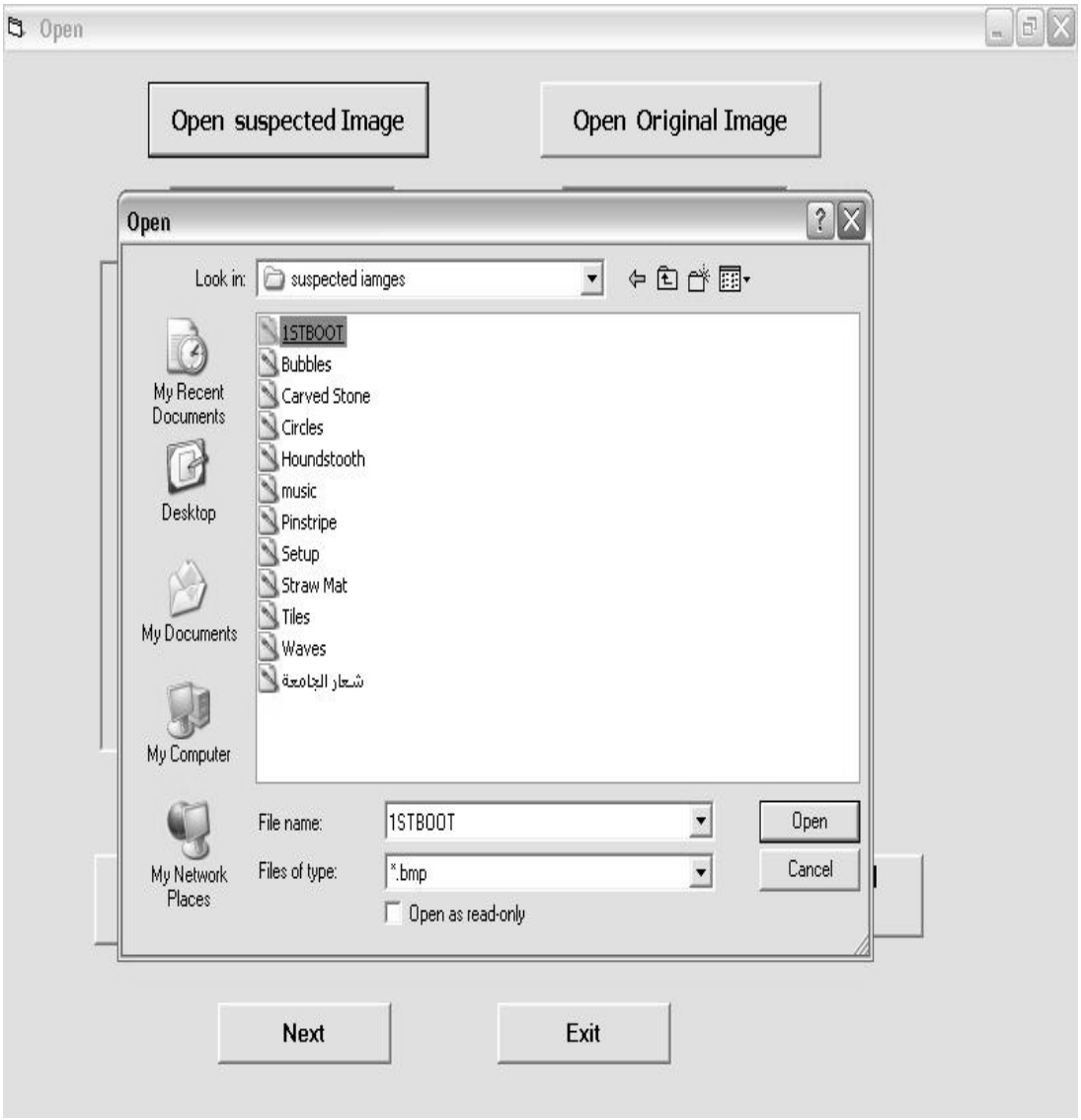


Fig 4.3 Get Suspected Image

Search for the Original image " icon is refer to searching for the original image assuming that it's within the image folder Fig 4.4 shows the opening process, Here Searching for the original image in the database, the image will appear in the Fig(4.5).

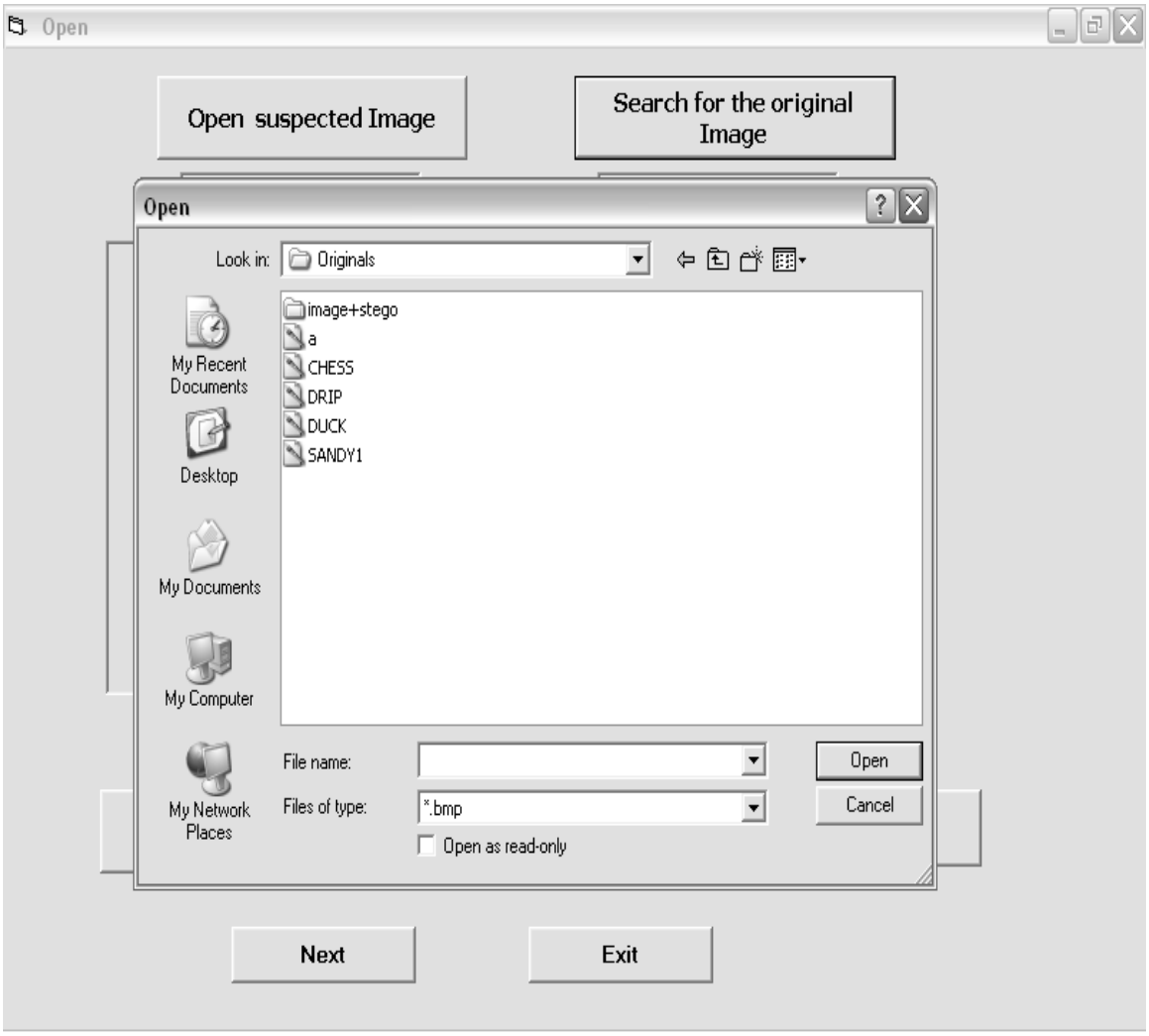


Fig 4.4 Search for the original image

Fig 4.5 both of the suspected and the original images are opened

The icon "Preprocessing and Feature extraction and analysis" mentioned at (Fig4.1)is executing the steps of both preprocessing and Feature extraction and analysis, Which are :

Preprocessing includes" read bitmap of the images, Scaling, reordering. Feature extraction and analysis includes "Wavelet Transformation" as extraction method, that apply on both the original and suspected images, which produce Feature vectors ,these vectors are minimized in order to choose the best features that required and important in the work in analysis step using Histogram, Probability density Function as mentioned at chapter three.

The Icon "Standard Statistical Tests" is a Discrimination step, the statistical test that were mentioned in chapter two.thenthis window will appear (Fig 4.6)

After specifying the features vector by performing feature extraction and analysis, methods for comparing two features vectors need to be determined.

There are six statistical tests build in our developed system. These test if the images are normal or not as shown in (Fig 4.8).

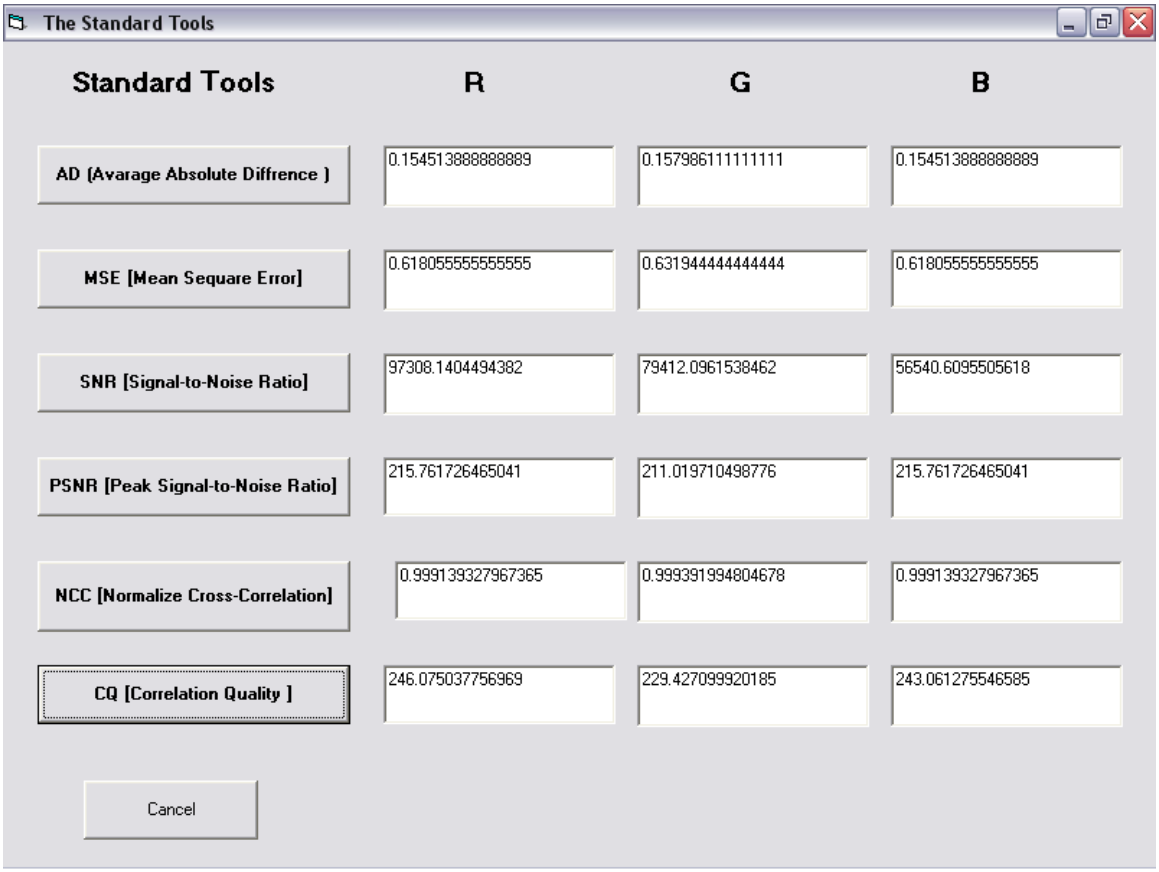


Fig 4.6 Standard Statistical Tests

Now Choosing "Proposed Statistical Test,", We develop here a statistical test a higher order statistics Mean, Variance, Kurtosis, Skewness that will be applied instead of the standards tools.

Now in Fig 4.7 a sketch of decomposed image using 2-Dwavelet decomposition, where we will check and test each level a side.

If we show in (fig 4.7) press level one then it will execute the feature vectors extracted in the first level only as show in (fig 4.8), else pressing level 2 then it will execute the feature vectors extracted in the second level show in (fig 4.12).

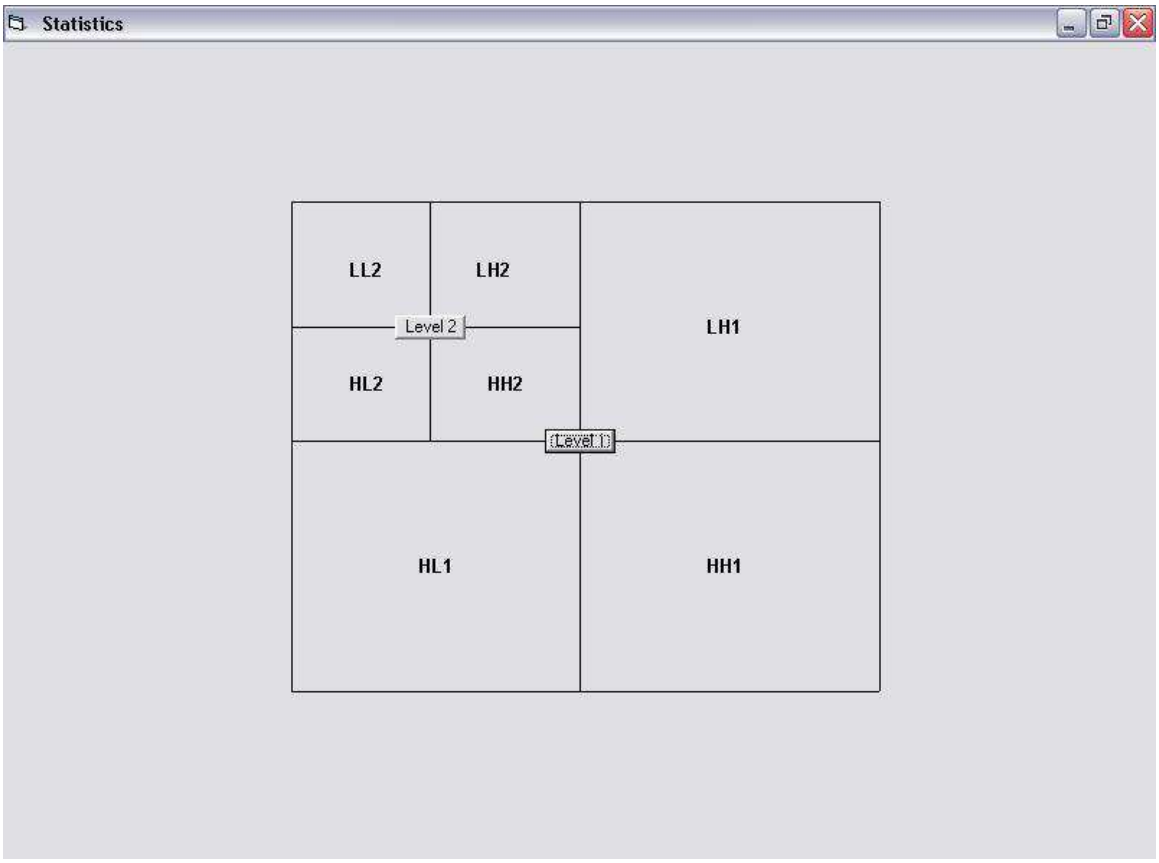


Fig 4.7 Choosing the Proposed Statistical Test

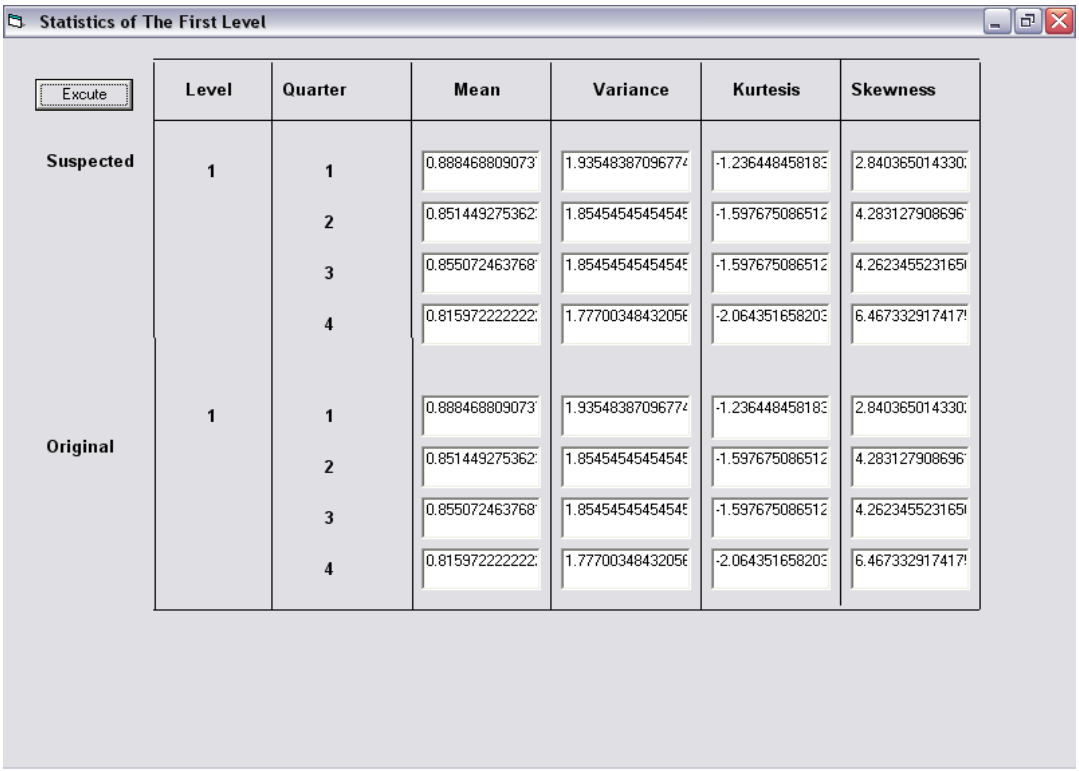


Fig 4.8 The statistics on level 2

This window show us the Results tests between the suspected and the original images that lies at level 2 of all the quarters .

4.3 Experimental Result

This part concerned with experimental result for samples of images being tested (as discussed previously, several images are taken as shown fig 4.9),where these images have different sizes and type BMP with 24-BPP.The test is done on BMP cover image and BMP image as stego object in the LSB modification And a text file in the S-Tool . These BMP are 24bits true color . Types of hiding were used LSB and S-Tool and adding a simple noise.

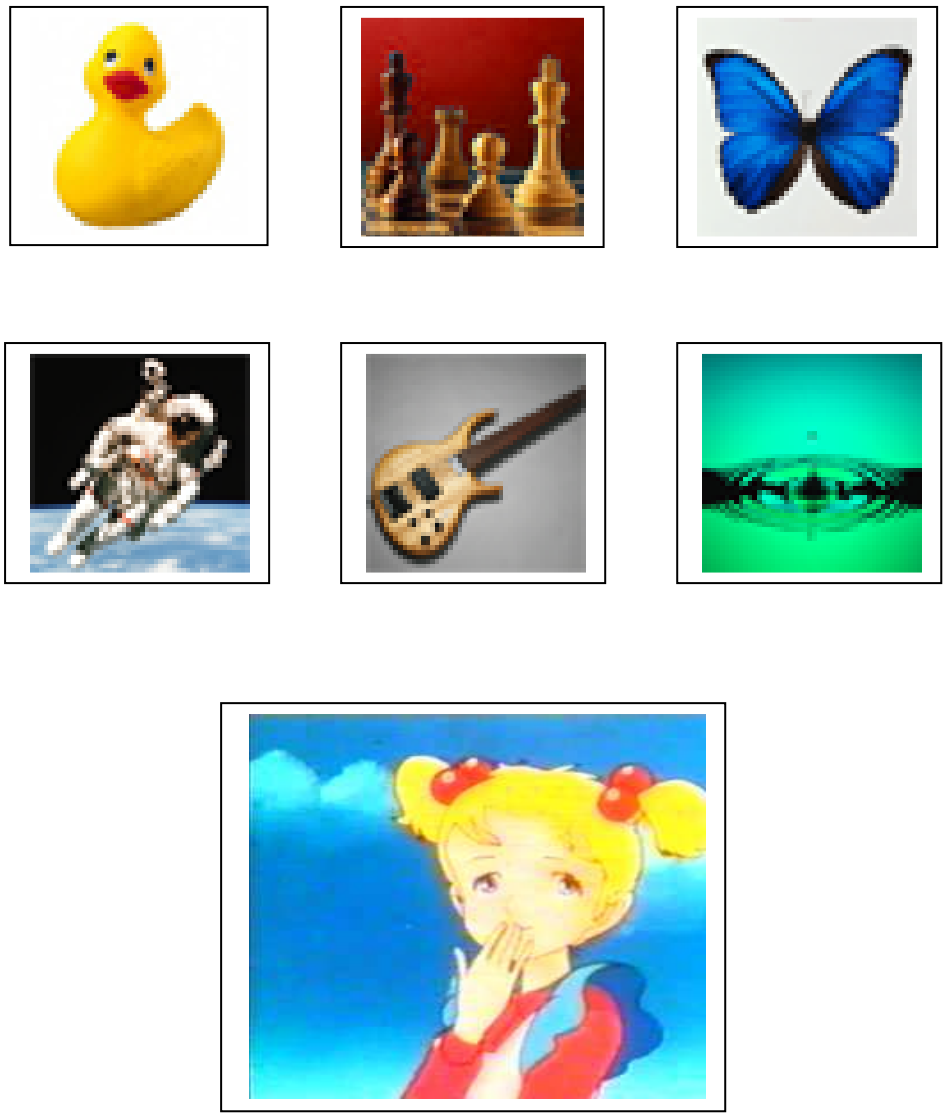
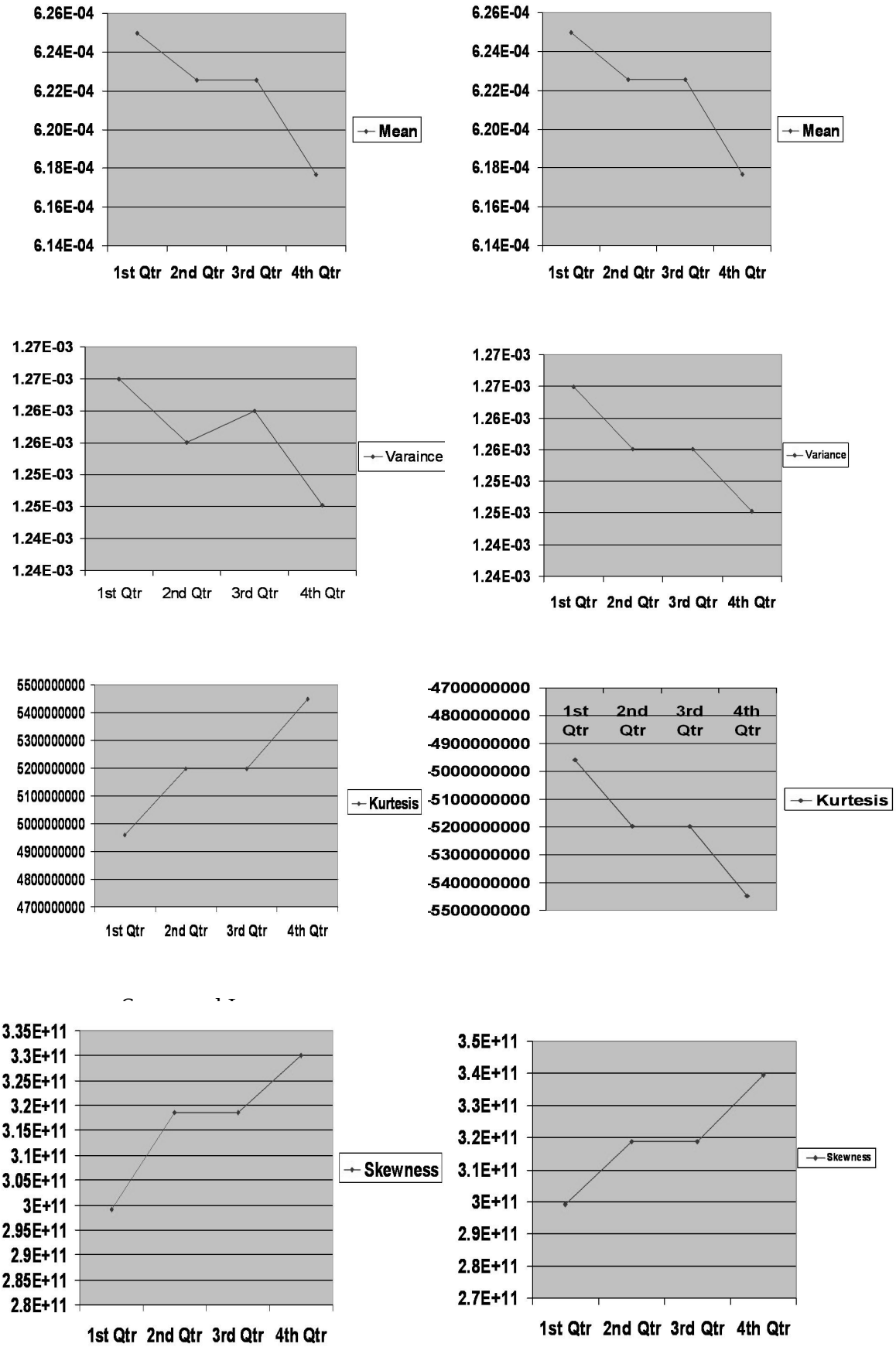


Fig 4. 9 Set of images

Table 4.1 Standard Statistical Tests

No.	*Image*	*Tests*	*R*	*G*	*B*
1.	Duck +Duch1	**AD**	0.1545	0.1579	0.15451
2.	=	**MSE**	0.6180	0.6319	0.6180
3.	=	**SNR**	97308.14045	79412.0962	56540.6096
4.	=	**PSNR**	215.7617	211.01971	215.7617
5.	=	**NCC**	0.9991	0.9994	0.99914
6.	=	**CQ**	246.07509	229.4271	243.0613
7.	Duck+Duch2	**AD**	9.2014	0.09375	8.11632
8.	=	**MSE**	9.20138888888889E-02	0.09375	8.11631944444444E-02
9.	=	**SNR**	653616.9434	535296.3519	430555.2300
10.	=	**PSNR**	1449.2671	1422.4292	1643.0198
11.	=	**NCC**	1.0000	1.0000	1.0000
12.	=	**CQ**	246.2261	229.5789	243.2935
13.	Sandy bell+ Sandy bell1	**AD**	0.00515	0.0055	0.0044
14.	=	**MSE**	0.0205	0.02221	0.0178
15.	=	**SNR**	1210115.2760	1576468.8434	2281166.91780
16.	=	**PSNR**	6502.5333	6002.33843	7482.3670
17.	=	**NCC**	0.999	0.9999	0.9999
18.	=	**CQ**	212.8974	195.2398	213.5197
19.	Sandy bell+ Sandy bell2	**AD**	2.88391	0.00299	0.00311

20.	=	***MSE***	2.88391	0.00299	0.00311
21.	=	***SNR***	8605264.1851	11710911.4081	13060798.8235
22.	=	***PSNR***	46240.2368	44588.7997	42840.2193
23.	=	***NCC***	0.99999	0.999999	0.9996467125979
24.	=	***CQ***	212.9011	195.24367	213.5230
25.	Drip+Drip1	**AD**	0	0	0
26.	=	***MSE***	0	0	0
27.	=	***SNR***	0	0	0
28.	=	***PSNR***	Division by Zero	Division by Zero	Division by Zero
29.	=	***NCC***	1	1	1
30.	=	***CQ***	5.486039	214.2329	152.75908
31.	Guitar+ Guitar1	**AD**	0	10.2478	
32.	=	***MSE***	0	513.30425	0
33.	=	***SNR***	overflow	overflow	overflow
34.	=	***PSNR***	Division by zero	Division by zero	Division by zero
35.	=	***NCC***	1	0.930604	1
36.	=	***CQ***	163.6459	144.9521	148.2131



Suspected Image Original Image

Fig 4.10 the order statistics

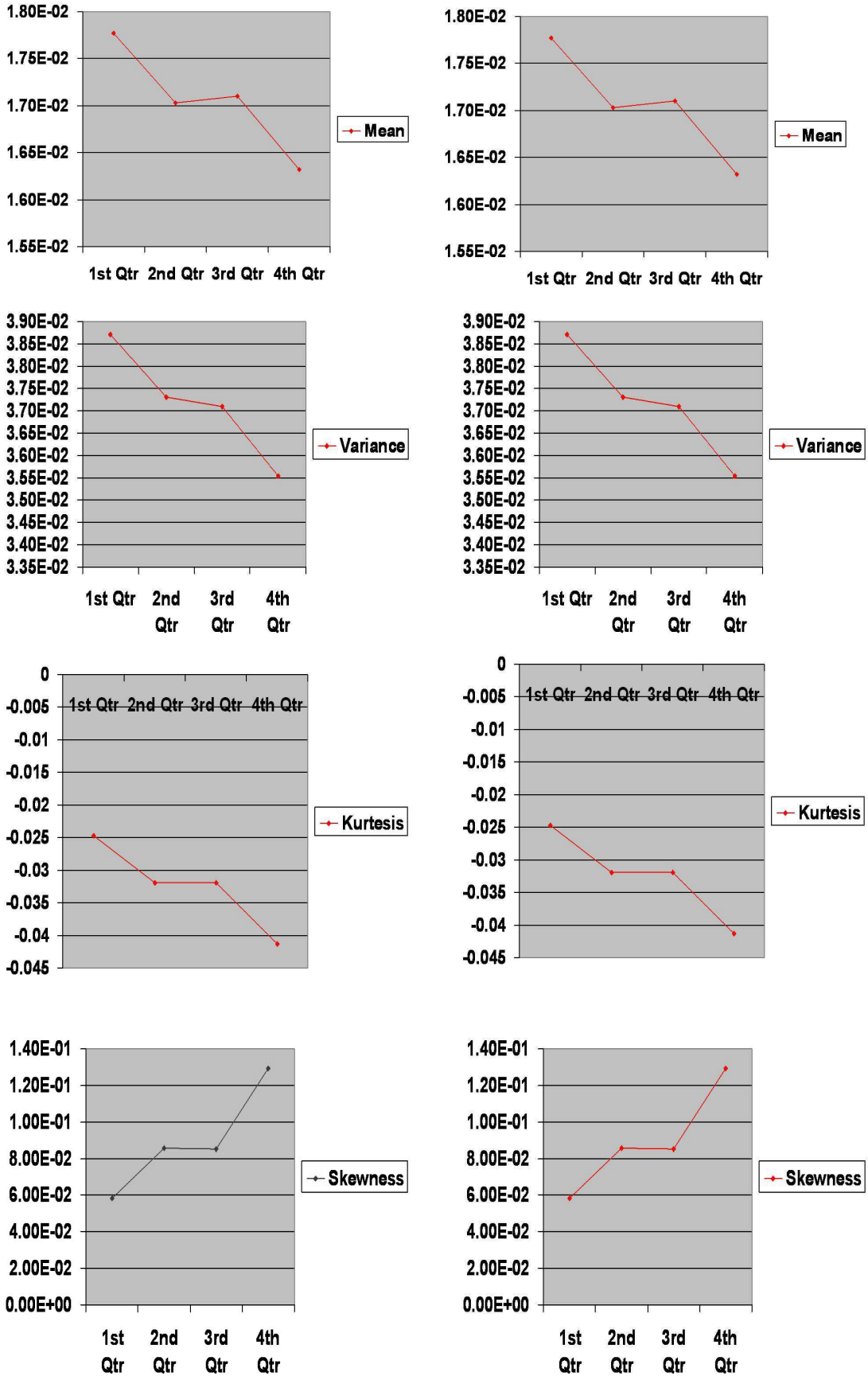


Fig 4.11 Clean Image

4.4 Discussion

In the first the system were tested with several samples, the images are Sandy bell that has a stego object using (LSB),Sandy bell2 has a stego object using (Stool)),Duck with no hidden information, drip with noise and some images with LSB, S-Tool ,clear …etc. Table 4.1 shows the result of the comparisons of several images

In the first step table 4.1 represent the statistical tests (AD, MSE, SNR, PSNR, NCC, CQ), these test are applied on different stego objects and the result are differ from image to image depending on the Steganography algorithm that used on insertion , we can notice the differences

If we take the image Sandy bell1 at table 4.2 we'll find

Sandy bell and Sandy bell 1 ➔	R	G	B
AD	0.00515	0.0055	0.0044
MSE	0.0205	0.02221	0.0178
SNR	1210115.2760	1576468.8434	2281166.91780
PSNR	6502.5333	6002.33843	7482.3670
NCC	0.999	0.9999	0.9999
CQ	212.8974	195.2398	213.5197

if we take the Guitar and Guitar1 the resulted of

Guitar and Guitar1 ➔	R	G	B
AD	0	10.2478	
MSE	0	513.30425	0
SNR	overflow	overflow	overflow
PSNR	Division by zero	Division by zero	Division by zero
NCC	1	0.930604	1
CQ	163.6459	144.9521	148.2131

Then the Figure 4.10 shows the order statistics (Mean, Variance, Kurtosis, Skewness) for the Image Sandy that has hidden information using LSB modification .You can notice the differences between those two images. where the statistics is little different and seems to be the same but when you focus you can notice it.

Chapter Five
Conclusions and Future works

5.1 Conclusions

To produce a Steganographic attack system presents a challenge, because there is much Steganography software. So, producing this humble system to analyze hidden information in the suspected image when the original image is available. We tried to analyze it without the original image but I find it very difficult to implement because we did not have any information about the original image to compare it with the suspected.

The following are some concluded remarks of the system:-

1. Using statistical test is a good idea to detect the changing in image when original image is available.

2. In this System, It needs to save only the features vectors that need only several kilobytes per image.

3. Using wavelet features give better results than any other transformation.

4. Some hiding technique hide the secret message without any change in color pixel. So, the histogram of similarity and differences is not effective for these techniques.

5.2 Future Works

1. Try to use different image format (i.e. GIF, JPEG) instead of using BMP file format.
2. Increase the ability of the system using Neural Network as a classification method instead of statistical tests.
3. Extracting features using Independent component analysis (ICA).
4. Try to detect hidden information without the need of the original image.

References

1) Ali M. Reza," **Wavelet Characteristics, What Wavelet Should I use?",** an internet paper, Spire Lab, UWM,1999.privet communication through Internet .

2) Amars page www.wavelet.org. An Inernet site, http://en.wikipedia.org/wiki/Probability_density_function And, internt Website "http://en.wikipedia.org/wiki/Histogram

3) Ammar T. Albaka'a," **Image Copyright Protection Using Digital Watermarking**", a M.SC. Thesis submitted to Al-Nahrain university ,computer science department ,2002 .

4) Burrus C. S. , Gopinath R. A., and H. Guo, **" Introduction to Wavelets and Wavelet Transforms: *A Primer*"**, Prentice Hall, Inc., 1998.

5) Da-Chun W. a, Wen-Hsiang T., "**A Steganographic Method For Images By Pixel-Value Differencing**", privet communication through Internet .

6) George B., Ian D., Ming-Yuan D. and Goutam P. , **"Searching For Hidden Messages: Automatic Detection Of Steganography",** Computer Science Department , University at Albany, SUNY,1400 Washington Ave, Albany, NY 12222,2003.

7) Hany F., **"Detecting Hidden Messages Using Higher-Order Statistical Models"** Department of Computer Science, Dartmouth College, Hanover NH 03755,2001.

8) Hany F. ,**"Detecting Steganographic Messages in Digital Images",** ,Department of Computer Science, Dartmouth College, Hanover, an internet paper, privet communication through Internet . NH 03755,2001.

9) Hany F. and Siwei L. **"Detecting Hidden Messages Using Higher-Order Statistics and Support Vector Machines"** , Dartmouth College, Hanover NH 03755, In: 5th International Workshop on Information Hiding, Noordwijkerhout, The Netherlands, 2002.

10) Haider J.**"Wavelet-Base Steganography "**a M.SC. Thesis submitted to Al-Nahrain university ,computer science department ,2004 .

11) Hubbard B. B. W., A K Peters. **"The World According to Wavelets**", 1996. Privet communication through Internet.

12) Jessica F. and Miroslav G. "**Practical Steganalysis of Digital Images – State of the Art",**SUNY Binghamton, Department of Electrical Engineering, Binghamton, NY 13902-6000 ,2000.

13) Jessica F., Miroslav G., Dorin H.,"**Steganalysis Of JPEG Images: Breaking The F5Algorithm**", Binghamton, NY.2000

14) Jeremiah J. "**Steganalysis Of Additive Noise Modelable Information Hiding"**, A M.Sc. Thesis Submitted to the Graduate, Faculty of Rensselaer Polytechnic Institute, Rensselaer Polytechnic Institute,Troy, New York,April 2003.

15) Jacob T. ,Gregg G., Roger C.,Gary B, **"Blind Steganography Detection Using a Computational Immune System",** International Journal of Digital Evidence Winter 2003, Issue 1, Volume 4

16) Mohammed Ala'a Hammamie, **"Information Hiding Attack in image "**, a M.Sc.thesis submitted to Iraqi commotion for computer and informatics , 2002.

17) Mustafa D. T. ," **Design of A Fingerprint Recognition System Using Wavelet Transformation**", a M.SC. Thesis submitted to Al-Nahrain university ,Computer science department Computer science, Al-Nahrain university, 2001.

18) Neil J., Zoran D., Sushil J.," I**nformation Hiding: Steganography And Watermarking –Attacks And Countermeasures"**, ISBN 0-7923-7204-2 , a book published in 2000.

19) Neil J. and Sushil J., "**Steganalysis: The Investigation of Hidden Information**", Proceedings of the 1998 IEEE Information Technology Conference, Syracuse, New York, USA, September 1st - 3rd, 1998.

20) Niels P.**, "Defending Against Statistical Steganalysis**" Center for Information Technology Integration, University of Michigan, published in 10th NSENIX security Symposium, Washington DC, pp. 323-335,2001.

21) Robert Krenn ,"**Steganography Implementation & Detection"**, an internet paper , Privet communication through Internet,January 21, 2004.

22) Rioul O. and Martin V.. "**Wavelets and Signal Processing,"**,IEEE SP Magazine, 14–38,October 1991.

23) Roman T., Robert B., Johannes B. , "**Steganographic System Based on Higher-Order Statistics"** University of Erlangen-Nuremberg, Cauerstr., D-91058 Erlangen, Germany. published in

Proceedings of SPIE Vol. 5020, Security and Watermarking of Multimedia Contents V, Santa Clara, California, USA, 2003.

24) Stefan K., Fabien. P., "**Information Hiding For Steganography And Digital Watermarking** ", ISBN 1-58053-035-4 , published in 2000.

25) Moerland T.," **Steganography and Steganalysis "** ,an internet paper, Privet communication through Internet, May 15, 2003.

26) Yanming D., Huan L., Avinash R., and Arunabha S. **"Detecting Hidden Information in Images:A Comparative Study**", Department of Computer Science and Engineering , Arizona State University, Tempe, AZ 85287,2003.

The BMP file format

The BMP file format divides a graphics file into four major parts, these are:

- Bitmap file Header the bitmap file header is 14-bytes long and it formatted as :

UNIT	bfType	(hold the signature value 0x4D24, which identifies the file as BMP)
DWORD	bfSize	(holds the file size)
UNIT	bfReserved	(not used, set to zero)
UNIT	bfReserved	(not used, set to zero)
WORD	bfOffcet	(specifies the offset, relative to the beginning of the file, where the data representing the bitmap itself)

- Bitmap Information header the bitmap information header contains important information about the image .the windows format for this header is:

DWORD	biSize	(hold the header length in bytes)
LONG	biWidth	(identify the image width)
LONG	biHight	(identify the image Height)
WORD	biPlanes	
WORD	biBitCount	(identify number of bit/pixel in the image and thus the maximum number color that the bitmap can contain)
DWORD	biCompression	(identify the compression scheme that the bitmap employs. It will contain zero if the bitmap uncompressed)
DWORD	biSizeImage	(set to zero for uncompressed image, else it holds size(in bytes)of the bits representing the bitmap image for compressed image)
LONG	biXpelsperMeter	
LONG	biYpelsperMeter	

DWROD biClrUsed

DWORD biClrImportant

- Palette: The color table specifies the color used in bitmap. The BMP file comes in four color formats

  1. 2-color one bit per pixel
  2. 16-color four bit per pixel
  3. 256-color eight bit per pixel
  4. 16.7million-color 24bits per pixel

  The number of bits per pixel can be determined from the biBitCount shown above. In the 2-color, 16-color, 256-color BMP format, the color table contains one entry for each color. Each entry specifies the intensities of color's red, green, and blue components and it is of 24-bytes long as shown below:

  BYTE regBlue

  BYTE regGreen

  BYTE regRed

  BYTE regReserved

  Each color-table entry can specify range or red, green, and blue values from 0 to 255. True color BMP files do not contain color table, because a single color table with 16.7million entries of 4-bytes each would require 64MB of storage space.

- Bitmap Bits: the bitmap bits are set of bits defining the image –the bitmap itself. In the 2-color, 16-color, and 256-color BMP formats, each entry in the bitmap is an index to the color table. In 16.7milion-color bitmap, where there is no color table, each bitmap entry directly specifies a color. These

first 3-bytes in each 24-bit entry specify the pixel color and component, the second specifies green component, and the third specifies blue.

The bitmap bits representing a single line stored in left-to-right order, the same way that the pixel they represent line up on the screen. The first row pixel data in bitmap responds to the bottom row of pixel on the screen, the second row corresponds to the row of pixel second from the bottom, and on.

The size of one bitmap entry is determined by the number of bits per pixel as shown in the following table:

Number Of Colors	Number Of bits Per Pixel Required
2	1
16	4 (1/2 byte)
256	8 (1 byte)
16.7million	24 (6 bytes)

Popular Steganographic Tools

a. S-Tool

Is one of the most versatile steganography tools of the application tested. Version 3.00 includes programs that process GIF and BMP images. Version 4.00 incorporates images and sounds file processing into a signal program. In addition to support 24-bit images, S-Tool also includes encryption routines with options. 24-images provide good covers and processed quickly in S-Tool 4.00.

b. HIP(Hide In Picture)

HIP is a programs that allows users to hide files inside bitmaps. The pictures look like normal , but in fact there is secrete data stored in them .HIP hides file inside the picture by placing the bits of the file in the LSB of each byte in the picture. The difference between the stego image and the original images is at most 1 , so it is very difficult , if not impossible, for the human eye to identify any different from the original picture.

With 256 color pictures, the process is a little more complicated, because the bytes do not represent color intensities, but entries in the palette (a sequence of at most 256 different colors). HIP chooses the nearest color in the palette containing the appropriate LSB.

c. Hide and Seek

Hide and Seek creates stego-images with different properties depending upon the version applied.

This program takes data, usually text, and hides it in a GIF file. The data is hidden away where it will be hard to detect. The way it works is it uses the LSB of each pixel to hide characters, 8 pixel per characters, it uses dispersion to spread the data(and the picture quality digression) out a bit through out the GIF in a pseudo random fashion.

d. Hide4PGP

Hide4PGP uses 8-bit and 24-bit BMP images as cover images and provides a number of options for selecting how the 8-bit are handled or what bit leaves the data is hidden. The default storage area for hidden information is in the LSB of 8-bit images in the forth LSB(forth bit from the right) in 24 bit images.

e. EzStego, Stego on-line

the approach is taken by EzStego to hide information may introduce enough visible noise suspicion. However , if a carrier is carefully selected, then the embedded information may take some prying to reveal itself.

f. Jsteg-Jpeg

Jsteg is a publicly available steganographic tool written by Derek Upham. The tool has to be patched onto the fourth public release of the Independent JPEG Group's free JPEG software . Jsteg performs the wide-spread least significant bit (LSB) embedding technique. Installed in the JPEG encoder Jsteg overwrites the LSBs of (quantized) discrete cosine transform (DCT) coefficients with message bits. Jsteg does not only omit DCT coefficients equal to 0 but also omits DCT coefficients equal to 1[RT] .

g. Mendelsteg

This tool does not manipulate any preexisting cover images but generates Mandelbrot fractal graphics as cover images for hidden messages. Depending upon the parameter , the images may vary in color and size.

الخلاصة

يهدف هذا البحث للتطوير و تنفيذ طريقة تحليل الاغمار من خلال الفحص الدقيق للصور المشكوك باحتوائها على معلومات مخفية (كصور ال BMP) ، لتحديد فيما لو كانت تحتوي على اخفاء ولا تحتوي.

تم اغمار المعلومات المخفية باستخدام نوعين من الاغمار . النوع الاول نسخدم (S-Tool)و هو احد طرائق الاغمار المعروفة ،و طريقة مطورة تتعلق باغمار المعلومات باستخدام طريقة LSB.

و بالاعتماد على طريقة تحليل المويجة نوع هار (Harr Wavelet) كطريقة للتحويل الى المجال الترددي ثم استخلاص متجه المعالم (Feature Vector) الذي يحتوي على معلومات خاصة تتم معالجتها لاحقا،ثم استخدام امكانية دالة كثافة التوزيع(Probability Density Function) لاختزال المعلومات و تقليص عدد المعاملات اعتمادا على عمل الدالة حيث ستتم معالجة المعاملات في الاختبارات الاحصائية و التي هي على نوعين:طرائق تقليدية و طريقة مطورةو كالاتي:

الطريقة التقليدية هي عبارة عن عدة اختبارات هي: AD فرق القمة المطلقة ،MSE معدل مربع الخطا، SNRنسبة الاشار الى الضوضاء ، PSNR نسبة الاشار الى الضوضاء القمية،NCCالرتباط المتقاطع الموزون،CQ نوعية الارتباط.

و في الطرقة المطورة تم استخدام الاختبارات الاتية : Mean المتوسط ، Variance التباين، Skewness معامل الالتواء ، Kurtosis مقياس التفلطح.

تم اخبار عدة نماذج من الصور (كصور ال BMP) و باحجام مختلفة تحتوي على معلومات مخفية و اغمار مائي،لقد تم اختبار ١٢ صورة مختلفة بعضها تم تمييزه و اكتشاف احتواءه على اخفاء.٤ من ١٢ (٣٣%) دخلت النظام و لم تكتشف وكانها خالية من اي اضافة ، بينما الاصل هو ١ خالية و الاخر حاوية على اخفاء .تم بناء النموذج باستخدام (Visual Basic 6) المجهز بواسطة بيئة النوافذ . و تم اختبار النموذج في بيئة نوافذ XP و نوافذ ME و قد اثبت النظام كفاءة مشجعة.

جمهورية العراق

وزارة التعليم العالي و البحث العلمي

جامعة النهرين

كلية العلوم

قسم علوم الحاسبات

طريقة اكتشاف المعلومات المخفية

رسالة

مقدمة الى كلية العلوم في جامعة النهرين كجزء من متطلبات نيل شهادة الماجستير

في علوم الحاسوب

مقدمة من قبل

هبة جبار عبد الواحد العقابي

المشرف

د.ستار بدر سدخان

تموز ٢٠٠٥ جمادى الاخر ١٢٤٦

Chapter One

Chapter Two

Chapter Four

Chapter Three

Chapter Five

Appendices

Refrences