

BLIND LINEAR CORRELATION TECHNIQUE FOR IMAGE WATERMARKING

SAJJAD BAGHERI BABA AHMADI

NOVEMBER 2014

BLIND LINEAR CORRELATION TECHNIQUE FOR IMAGE WATERMARKING

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES OF ÇANKAYA UNIVERSITY

BY SAJJAD BAGHERI BABA AHMADI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN THE DEPARTMENT OF COMPUTER ENGINEERING

NOVEMBER 2014

Title of the Thesis: Blind Linear Correlation Technique for Image Watermarking.

Submitted by Sajjad BAGHERI BABA AHMADI

Approval of the Graduate School of Natural and Applied Sciences, Çankaya University.

Prof. Dr. Taner ALTUNOK Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Murat SARAN Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

 114111

Assist. Prof. Dr. Reza ZARE HASSANPOUR Supervisor

Examination Date: 03.11.2014

Examining Committee Members

Assist. Prof. Dr. Reza ZARE HASSANPOUR (Çankaya Univ.)

Assist. Prof. Dr. Tansel ÖZYER

(TOBB Univ.)

Assist. Prof. Dr. Abdül Kadir GÖRÜR

(Çankaya Univ.)

STATEMENT OF NON-PLAGIARISM PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Sajjad BAGHERI BABA AHMADI

Signature 9 Fajad B. Date : 03.11.2014

iii

ABSTRACT

BLIND LINEAR CORRELATION TECHNIQUE FOR IMAGE WATERMARKING

BAGHERI BABA AHMADI, Sajjad M.Sc., Department of Computer Engineering Supervisor: Assist. Prof. Dr. Reza ZARE HASSANPOUR

November 2014, 81 pages

In digital environment, make, change, update, distribute and store digital data are convenient, therefore as much abuse of digital data is added. This calls for a method to prove the ownership right on digital contents and to avoid unauthorized users to tamper and distribute digital data. Thereby to achieve this security requirement, watermarking schemes are introduced that have applications in all three forms of media, i.e., video, music and image. This thesis aims to test blind linear correlation technique by Stirmark benchmark 4.0 that contains sixteen different tests, such as attacks and distortions which are well-known in image processing. Those tests are applied on the images which are watermarked by blind linear correlation technique. In this thesis, results of this experiment are discussed and analyzed.

Keywords: Digital Watermarking, Noise Reference Pattern, Linear Correlation.

KÖR DOĞRUSAL İLİŞKİ TEKNİĞİ İÇİN FOTOĞRAF WATERMARKING

BAGHERI BABA AHMADI, Sajjad Yüksek Lisans, Bilgisayar Mühendisliği Anabilim Dalı Tez Yöneticisi: Doç. Dr. Reza ZARE HASSANPOUR Kasım 2014, 81 sayfa

Dijital ortamda, değişim, güncelleme, dağıtım ve dijital veri depolaması çok kolaylıkla yapılmaktadır, aynı şekilde dijital verilerin kötüye kullanımı çoğalmıştır. Dijital veri Mülkiyet hakları kanıtlamak ve Manipülasyon ve dijital verilerin yetkisiz kullanıcılar tarafından önlemek için bir yöntem gerekmektedir. Bu güvenlik gereksinimleri gerçekleştirmek için Watermarking yöntemleri tanımlanmıştır ve tüm 3 form medyada örneğin: video, müzik ve fotoğrafta uygulanmaktadır. Bu tez, kör doğrusal ilişki tekniğini fotoğraf Watermarking üzerinde test etmek amacıyla yapılmıştır, Stirmark benchmark'ın 4 üncü versiyonu 16 tür deney içermektedir. Örneğin saldırı ve çarpıtma görüntü işlemede bilinmektedir. Deneme Watermark olmuş fotoğrafların üzerinde kör doğrusal ilişki tekniğini ile uygulandı. Bu tezde, elde edilen deneme sonuçları analiz ve tartışılmıştır.

Anahtar Kelimeler: Dijital Watermarking, Gürültü Referans Desen, Doğrusal İlişki.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis advisor Assist Prof. Dr. Reza Hassanpour, who has encouraged and guided me throughout this thesis.

It is a pleasure to express my special thanks to my friends for their valuable support.

I would like to express my deep gratitude to my family for their endless and continuous encourage and support throughout the years.

TABLE OF CONTENTS

CHAPTERS:

LIST OF FIGURES

FIGURES

FIGURES

LIST OF TABLES

TABLES

LIST OF ABBREVIATIONS

- LC Linear Correlation
- PN Pseudo Random Noise
- IP Intellectual Property
- MSE Mean Square Error
- HVS Human Visual System
- LSB Least Significant Bits
- DCT Discrete Cosine Transform
- DWT Discrete Wavelet Transform
- DFT Discrete Fourier Transform
- DVD Digital Video Disc
- ISO International Organization for Standardization
- MPE Multivariate Power Exponential
- GGD Generalized Gaussian Distribution
- RMSE Root Mean Square Error
- SDMI Secure Digital Music Initiative
- CDMA Code Division Multiple Access
- NTSC National Television Video
- PSNR Peak Signal-To-Noise Ratio
- JPEG Joint Photo Graphic Experts Group
- CPTWG Copy Protection Technical Working Group
- DSCDMA Direct Sequence Code Division Multiple Accesses

CHAPTER 1

CONCISE OVERVIEW ON WATERMARKING

The nature of digital information implies making, changing, distributing, and manipulating digital data are convenient, therefore as much abuse of digital data is present. this calls for a method that offers ways to authenticate users for providing their possession right on the digital contents and for avoiding tampering and illegitimate distribution by the not permitted users. To achieve this security requirement, watermarking has played a significant role in each portion of digital type, i.e., video, music and image. Community discovered the advantages of watermark and its useful role in the digital content. This thesis plans to test blind linear correlation technique by exposing this technique to diverse attacks and distortions. For this purpose, the Stirmark benchmark 4.0 is used which is a wellknown benchmark for testing watermarking schemes. In the first chapter, a concise overview on watermarking and explanations about its basic principles are presented. Secondly in the next chapter, the conceptual watermarking models and two main steps and attacks of watermark systems are discussed. Next, deeper details about watermarking algorithms and domains are given in the third chapter. Ultimately, the last chapter of the thesis is dedicated to the experiments, results and conclusion.

1.1 History of Watermarking

The word watermarking or the mark of water originated seven hundred years ago in traditional factories of producing paper in Italy. A damp fiber was pressed firmly on a piece of paper by a stamp in order to put a colorless mark of stamp on the paper permanently. It is unclear when the first time the digital watermarking was brought up. During 1979, a machine-detectable model, discussed by Szepanski was inserted on the documents used for anti-counterfeiting reasons. After nine years a new method of embedding a recognition code in the audio signals was described by Holtet al. As a matter of fact, in 1988 Komatsu and Tominaga were the first people who used the term digital watermarking. With the beginning of the communication age and the progression in computers, digital watermarking progressed dramatically. In 1995 there were thirteen essays about watermarking but in 1998 this number had increased to one hundred and seven essays.

Figure 1.1 shows a histogram about the quantity of essays published on this subject. In the 1996 the fist workshop of information hiding was held and digital watermarking was one of its chief subjects. After few years, several organizations started to use watermarking technology for different purposes and in various standards. For example, for protecting video on DVD, the Copy Protection Technical Working Group (CPTWG) had practiced watermarking systems.

Figure 1: The quantity of articles had published on Steganography and watermarking subjects by the IEEE

Watermarking was a central component of the system in the Secure Digital Music Initiative (SDMI) on behalf of protecting music. the European Union in order to test watermark in broadcast monitoring, supported two projects, VIVA and Talisman. The International Organization for Standardization (ISO) became interested in the knowledge about the content of designing highly developed MPEG standards. For marketing watermarking products, several companies were established in the late 1900s. More recently, many companies' applications consist of watermarking technology which is discussed in this chapter.

1.2 Steganography versus Watermarking

Steganography is a method to carry a secret message through a cover in a form of hidden. From this aspect that both steganography and watermarking are similar to each other; however, they have fundamental philosophical differences as shown below:

First of all, the aim of design is different in both of them. The stegonography has aims to carry a confidential message that is not related to the host cover and the robustness in steganography is not important. But; on other hand, watermarking has aims to protect ownership right by inserting a message that is related to the host cover and the robustness has high importance to watermarking. So, watermarking should provide strong security against attempts for removing or modification of the hidden message.

Second, in steganography the main purpose is to do one to one communications via hiding message in the cover but, in watermarking the goal is to do one to many communications to satisfy its applications purposes.

1.3 Watermarking Frame Work

Watermarking procedure has two main steps after the watermark signal is ready; first of all, by using an encoder the watermark signal will be inserted into the cover. Furthermore, in order to prove the ownership right or other purpose of watermarking, we use a decoder to detect the watermark signal and in figure 2 the generic watermarking procedure is given.

Figure 2: A general watermarking systems [2]

Let's assume the watermarking procedure is function E (), multimedia content is I, watermarking signal is W, and watermarked cover is I'. The watermarking procedure can be shown by following equation (1.1):

$$
E(I, W) = I'
$$
 (1.1)

1.4 Common Concepts And Terms In Watermarking

The process of watermarking has many concepts and terms that are important to watermarking procedure. Some important of them are listed below:

• Stegomedium/Host/Carrier

A cover or a type of the media, which can be used for embedding the watermark message inside it. That is called Stegomedium, Host, or Carrier.

• Steganogram

The watermarked cover is called a steganogram.

• Steganoanalysis

The analysis on a steganogram to detect and recover the watermark.

• Data rate capacity

The maximum amount of capacity which the host can provide for watermarking.

• Robust / Fragile

Robust: Changing the contents of a file will not damage the watermark.

Fragile: Changing the contents of a file will damage or may even remove the watermark.

Public / Private

Public: Users are permitted to identify and retrieve the watermark.

Private: Users are not permitted to identify and retrieve the watermark.

1.5 Evaluation Parameters

To evaluating a watermarking technique, it is necessary to consider the following properties of watermarking system, and based on these a watermarking algorithm can be judge that these properties are listed below [5]:

• Fidelity

Fidelity means the conceptual resemblance between the watermarked digital cover and original cover work. If this similarity is high and the difference is imperceptible, so we say the percent of fidelity is high and good. However, most applications use more powerful watermark signal for increasing the robustness that may result in loss of fidelity. In this case, it is necessary to balance fidelity and robustness by decreasing them to a required level. For visible watermarks, fidelity does not have meaning and the watermark may spread throughout or in imperative areas of the image for preventing to be deleted.

A video signal, transmitted over NTSC, is not very high quality. For this reason, it does not consider the watermark fidelity as a huge difficulty in the transmission using NTSC and could be small comparatively. Nevertheless, DVD and HDTV videos call for extremely higher fidelity watermarks because their signals contain very high quality.

• DataPayload

This term indicates the number of bits that can be embedded inside a time unit or work unit by a watermark scheme.

• Robustness

The capability for discovering the watermark signal after applying regular signal processing procedures is known as robustness. A watermark signal it considered as robust watermark if it can resist probable distortions and stays noticeable later than applied attacks. The robustness criteria are different and it is depending on the type of application.

Security

A watermark signal is secure if cannot be eliminated and stay detectable after being exposed to attacks which have full understanding about the specific used embedding and detector algorithms except the knowledge of the used secret key. In addition, at least such attacks have the awareness of one carrier with concealed message.

Computational Cost

Computational cost has sufficient role in applications which must embed and detect the watermark signals in real-time. For instance, in broadcast monitoring applications, the media production must not be effected by watermark embedding

operation and become slow down; furthermore, the watermarking detector have to operate in real-time at the same time as observing the broadcasts. For this reason, it requires practical watermarking schemes, which do not need many computational efforts. In contrast, it is not extremely significant for a detector to be used as a proof of ownership, since such detectors have usages for the period of ownership disputes [3].

False Positive Rate

During process time to detect the watermark signal, there may happen a mistake in discovering and the watermark detector may find a false watermark or even may not find the watermark signal. These are called as false positives. The quantity of false positives which are anticipated to occur in a specified quantity of detector runs is considered as the false positive rate [3].

1.6 Different Watermarking Aspects

Watermarking techniques can be categorized according to a few aspects which are shown in the figure 3. According to the domain, there are two parts, Spatial Domain and Frequency Domain.

Spatial Domain: by changing the pixel values of original images, a watermark signal can be inserted in spatial domain of those images.

Frequency Domain: In natural conditions frequency domain means the image is fragmented into various frequency bands. For transferring the image to its transform representation, several reversible transforms can be used such as Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), or Discrete Wavelet Transform (DWT) [11]. In type of document view, we can categorize watermarking into four aspects. Based on our requirements, we can insert watermark into Text, Image, Audio, and Video. If we consider human perception, there are two categories, visible and invisible. Invisible part is divided into two subdivisions, Robust and Fragile. The robust algorithms aim to survive watermark signal after probable distortions such as filtering, noise additions, and compressions. However, the fragile watermarks have utility for detecting if there is any modification or manipulation on the digital content or not.

Figure 3: The category of different watermarking aspects [3].

These alterations modify or get rid of the watermark signals. In addition, this kind of watermarking has utility for content verification such as trustworthy camera. A watermark signal is inserted into the framework as soon as it is captured by the camera. By any modifying on image, the watermark signal will be missed, as a result confirming if the framework is the unique captured one or not [10]. Private and public watermarks are two classed of invisible robust watermarks as illustrated in the prior part. The original content is needed in private algorithms for detecting the watermarking signal; whereas, the original content is not required in public watermarks.

In accordance with the applications, the watermark is divided into destination and source based watermarks. The source based algorithms embed a unique watermark in all the copies and have utilities for possession recognition or verification. The watermark recognizes the possessor of the digital content. Nevertheless, the destination based watermarks or fingerprints are inserted individually to every copy and utilize to pursue the customers in a case of unauthorized actions.

Furthermore, Fingerprints have utilities in favor of broadcast monitoring as well. One sole watermark signal is placed into every audio or video clip prior to be broadcasted then programmed computers check out the broadcast and distinguish and report to advertisers that at what time and at which place their clips are broadcasted [11].

1.7 Practical Applications

The most usage of digital watermark is in fields of protecting copyright, identifying criminals and military purposes. The digital watermark can be used in military and police organizations or in medical centers with reasonable cost. Digital watermark also can help to implement the copyright laws in order to protect intellectual property. But the numbers of countries which are using digital watermark in an acceptable level are handful.

In some countries the usage of watermarking is more prosper. For instance in America, the military is used watermarking technology to protect itself radio communication. In some countries such as Switzerland where now this technology is used to perform the issuance of driving license, identity papers and control of entering and leaving the country. The active companies of this field are mostly American or European.

Digital Watermarking schemes have diversity of usages that can be listed as following: copy control, transaction tracking, device control, or proof of ownership, owner identification, broadcast monitoring, and authentication [2]. In following, the actual applications of them are listed below in form of brief.

Broadcast monitoring

Watermarking can be useful in broadcast monitoring, because there are many requests from advertises, musicians and actors that whether they received their purchased air time from broadcasting firms or not. An ignominy had broadened throughout Japan about television advertising during 1997. As a minimum 2 stations had been regularly overbooked air time. The advertisers paid thousands for commercials which never aired [9]. In order to achieve this requirement, first we

should to exert a sole watermark in every sound or video clip, after that the automatic checking stations by use those unique watermarks, can report to applicators (advertises, musicians and actors) that at what time and in which place your clips are broadcasted. There are several companies, such as Teletrax that provides watermark-base broadcast monitoring service from Philips [2, 10].

Owner recognition and Proof of ownership

The owner by embedding its own identification into its digital content that is an inseparable part of the content and nobody can remove it easily. By this identification watermark, the real owner will be identified. For example in order to address this request, The Digimarc Corporation promoted a watermarking system designed for this purpose. Their watermark embedder and detector are packaged with Adobe Photoshop which is a well-liked image processing program. As soon as the watermark detector discovers a watermark signal, it communicates with a fundamental database for identifying the watermarked content's owner who has to give a fee to maintain the information in the database.

Transaction tracking

Fingerprints or transactional watermarks permit an intellectual property or IP proprietor or content dispenser to recognize the starting place of an unauthorized copy by marking each legal document copy with a separate, sole watermark. If a document marked with a transaction watermark is abused (distributed illegitimately), the proprietor can discover who is accountable.

Authentication

Modifying digital contents is easier than proving the authentication of digital contents. It is difficult but by using the invisible fragile watermark, the problem is solved. As it is presented the properties of fragile watermark in the categories of watermark section, if somebody makes a slightly modification in the content, the fragile watermark will be destroyed and we can authenticate the content is valid or not.

• Copy control

For using watermarking in copy control of digital media, we need association of all the manufacturers of the recorder software; they should be able to implement watermark detection algorithms. If they have such capacity, so they don't let to recopy of digital contents which are watermarked [2,3]. There are commercial copy control softwares already in the market. In fact, what is offered by MarkAny is more than just preventing illegal copying. Their product relies on watermarks to control print, open, and download functions in relation to user right even later than the content is released by illegal user [8].

• Device control

Device control is a broad category of applications in which specially designed device respond to the watermarks they distinguish in the content [8]. Recently watermarking applications with purpose of device control, Digimarc's Mobile system is exclusive identifier for published and dispersed images such as magazine commercials, tickets, covering, and etc. After the image is captured through the camera of mobile, the watermark signal will be read by that software within the cell phone and the identifier code is employed to straight a web browser to an related web site [2].

CHAPTER 2

CURRENT BASED TECHNIQUES

Conceptual watermarking models are usually separated into two collections: First model collection considered watermarking signal as a communication method, and second collection is according to geometric visions on watermarking schemes. To be able to recognize the differences and similarities between traditional communications and watermarking, it is useful to have a concise review on the traditional communication systems and then discuss communication-based watermarking models. Furthermore in this chapter, we consider embedding and detecting procedures and their issues such as errors in detection. Ultimately, the different threats to watermarks and their attacks are discussed.

2.1 Communication System

There are some similarities and differences between traditional communications and watermarking. The procedure of the traditional communication form is illustrated in figure 4. In this scheme we want to convey a message, m, crossways a communication channel. In the first place, the channel encoder, that has duty to maps messages into a code word which can be transmitted over the channel, encodes the message. Then the word code is denoted as x. And it is transmitted over a noisy channel and received signal is denoted as y that is dissimilar from x. This alteration from x to y is as a result of the additive noise. In fact, the noise signal, n, is inserted to x. At the end of the channel, channel encoder encodes process and tries to fix communication errors. This purpose records conveyed message into m_n . The function of decoder is usually many-to-one. The possibility that deciphered message consists an error is negligibly miniature, under circumstance that the channel code is suitably equal to the particular channel [2].

Figure 4: The typical model of a communicating system.

2.2 Communication-Based Models of Watermarking

As a piece of fact, watermarking is a kind of communication. In this approach, we communicate a signal form watermark encoder to the watermark decoder as the receiver. Then we try to make watermark suitable for the traditional communication system. Here this sub-chapter, we observe three ways to live up to this necessity. These methods are different in the way of adding the cover work to the conventional communications model. The first model that is known as basic or primary model, considers the cover work like noise. The model number two is also considering the work cover like noise, but in this scheme the channel encoder is considered this noise as side information. Ultimately, the model number three has different policy and transmits a second message in company with the watermarking message in type of multiplexing.

2.2.1 Basic models

Figures 5 and 6 show a way to map watermarking into the frame of figure 4. Figure 5 demonstrates a way which makes use of an informed detector, and figure 6 demonstrates another way which employs a blind detector .In the mapping, watermarking considered as a communication channel from side to side that the watermarking message is transferred. The channel consists of the cover work. There are two basic steps in embedding process without considering that we use blind detector or an informed detector. At the first, we map the watermarking message into an additional pattern, w_a , that has the same dimensions as the original cover work, . For instance, in such case in watermarking images, the coder may create a 2 dimensional pattern that has the same dimension as the original cover image. We can use watermark key in this mapping. Next, we add w_a to the original cover work, c_0 , for producing the watermarked work c_w . In this type of embedding, the encoder ignores the cover work, so it is considered as blind embedder.

Figure 5: The watermarking system with an easy informed detector recorded into a communications model

Figure 6: Watermarking system with blind detector mapped into a communication model

After embedding the added pattern, we suppose the watermarked work c_w has been processed in somehow and then we form target processing as the noise addition. The processing kinds the Work may move through include broadcast, decompression, compression over a long channel, audio or image improvements, and so on. There are some malicious attempts that have plan for eliminating the watermark. All these processing are depend on the watermark work; thus, modeling their effects with additive noise is a simplification. In the case of using an informed watermark detector, there are two steps. In the first step, with the purpose of gain the noisy watermark pattern, n_w , the received work w_n is deducted by unwatermarked work. Then a watermarking decoder decodes that in the company of a watermarking key. There is only one difference between w_n and w_a that caused by the noise process. So, we can close the eyes to the additional cover work that denotes the watermark decoder, encoder, and the noise process shape a communication system.The whole unwatermarked cover work is not required in the more advance informed detection systems. In these systems, in detector, a usual data-reducing function is used to wipe up the noise effects that are presented by the additional work cover in the embedder.

In the figure 6, the blind watermark detector does not have unwatermark cover work, thus may not be removed prior to detecting. Some applications require robustness for special purposes, such as copy control or transaction tracking; in such case, we should increase the probability that spotted message will be matching for the embedded one. This scheme has the similar objective as the conventional communications. Nevertheless, we should consider that the goal in authentication application is not to communicate a message; in fact, they want to be aware that when and how a work has been changed from the time when a watermark signal has been embedded. Therefore, the models in figures 5 and 6 are not utilized in the verification systems.

We can create a easy model of an image watermarking scheme with a blind detector by using the model in figure 6.

2.2.2 Watermarking as communication with side information at the transmitter

Albeit in the figure 6, there are approaches in robust watermarking schemes in the company of blind detectors, but these models cannot satisfy all probable embedding algorithms and the encoded watermark must be autonomous of the cover work. Since, the embedder knows the unwatermark cover work c_0 , so it does not make sense to enforce this reaction. If the watermark encoder has permission to check up c_0 previous to programming the added pattern w_a , we can make much more effective embedding algorithms.

In the watermarking model in the figure 7, it has given permission to w_a to be dependent on c_0 . The only difference between watermark models in figure 6 and 7 is that the watermark encoder has an additional input c_0 . This modification provides facility to the embedder set c_w to any desired value by basically letting $w_a = c_w$ c_0 . This new watermark model is a type of communication system with side information at the spreader, if we keep on regard as the cover work as piece of the noise process $(c_0 + n)$ in the communication channel [16]. In fact, the embedder has ability to achieve much information regarding the channel noise, exclusively c_0 itself.

Figure 7: Watermarking like exchanges with side information at the spreader

2.2.3 Watermarking as multiplexed communication

In watermarking as multiplexed communication, instead of regarding the cover work like piece of communication channel, we transmit a subsequent message in company with watermark message in the same signal c_w . 2 different receivers will detect and decode the messages c_w and m. Those receivers are a watermark detector and human creature. Figure 8 is shown an alternative watermark model as communications. c_0 and m have been combined by watermark embedder into a signal c_w . There are some similarities between this combination and the multiple message transmission over a single line in the traditional communications. One of those differences is that the basic technology used, in traditional communication, for different messages is the same and a sole parameter such as code sequence, frequency or time separate the messages. But on other hand, in watermarking, watermark detection and human perception are used to separate messages. This is equivalent to via, state, spread spectrum coding for one message and frequency distribution for other message. Signal passes through the transmission channel and then arrives at either a conceptual system of humans or watermark detection.

Figure 8: The watermarking as simultaneous communication of two messages

The human observes c_{wn} and perceives an aspect near to the unique cover work without intervention from the watermark. In parallel, the detector detects a watermark in c_{wn} and gains the unique watermarking message without intervention from the cover work. In other word, that watermark detector can receive the unique cover work or a task of the cover work like a following input, if it is informed. This scheme of watermarking shows the balance among the watermark and the cover work. One of the techniques that this balance releases itself in the watermarking literature is the ability of using dissimilar utilities of the term signal-to-noise ratio (SNR).

2.3 Watermarking Embedding

In general watermarking system has three main components: watermark signal, embedding, and decoding. In between, the embedding procedure has importance because the watermarking properties extremely depend on the way that watermark is inserted within data. The watermark signal can be embedded into one of the following domains: Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), spatial and Fractals domains.

2.3.1 Choice of host feature

Aside from that the watermark signals can be inserted into spatial or frequency domains that each has special feature of the host. In general, an efficient watermarking system should consider the suitability of host features for the embedding watermark information because the watermark modification should not diminish the perceptual quality of the host; in addition, it should be identical to other related perceptual principles such as visibility, audibility, or intelligibility. The second curtail feature set of host that plays an important role beside the embedding rules, is to provide a sufficient robustness against probable attempts to remove watermark signal. On other hand some application the host capacity has higher priority than robustness. That means the watermarking information amount that can be hidden and transferred with miner error probability [22].

As a piece of fact, there are culple major steps in the watermarking embedding process. In the first place, a set of host features is extracted and then modify them according to the embedding rules. These two steps have direct effect on the watermark robustness and imperceptibility, which are main factors in watermarking systems. We are used Human Visual System (HVS) due to its advantages for choosing a suitable way for hiding watermarking information with more energy and without degrading the visual image quality [8].

2.3.2 Algorithms for embedding watermark

There are many embedding algorithms that can be categorized into three main groups.

1) Linear additive algorithms: the host linear modification and the correlative processing in the finding stage characterized the additive embedding techniques. In such watermark algorithms, Wi is a number sequence of signature data that has length N which is inserted into a elected division of the host signal data coefficient f. this scheme has a basic and commonplace formula for embedding as equation 2.1 :

$$
F'(m,n) = f (m,n) (1 + a.Wi) \qquad (2.1)
$$

In this formula a stands for weighting factor and f' represents the watermarked host. In addition, there is an optional method of embedding which had suggested by Cox [21] such as equation 2.2:

$$
F'(m,n)=f(m,n)+a.Wi;
$$
 (2.2)

 In second alternative algorithm uses logarithms of the original coefficient, as equation 2.3 .

$$
F'(m,n)=f(m,n) e^{aw}i
$$
 (2.3)

- a) Gaussian Sequence Algorithm: these type algorithms provide watermarks in form of bitmaps, pretend random real number, or binary sequence. With regard to the author requirements and in order to embedding watermark , the image will be decomposed into two, three level, or wavelet transformation, and diverse coefficient which are selected for embedding the watermark signal. For instance, the Barni algorithm choices all coefficients in the maximum decree associate bands such as HL1, LH1, and HH1. The additive formula is simply done watermark embedding. Many algorithms utilized Gaussian sequence in order to develop the watermark embedding. Such as Barni algorithm, Dugad algorithm, or Corvi algorithm and J.R.Kim algorithm.
- 2) Nonlinear Quantization Embedding:

Quantization process is mapped a great probably infinite set of values to a so smaller set.

In fact, Nonlinear Quantization executes none linear modifications and quantizes the obtained examples to record them to the next-door rebuilding point with the purpose of detecting the embedded watermark signal. There are two mappings in quantizer, one for decode and other for encode. The source values range is divided into a quantity of gaps by the decoder. These gaps have unique codeword. The encoder stands for the complete source values plunge into a feature through assigning codeword to the gap. There could be many feasible separate samples which can go down in any given gap. The mapping is irreversible. For preventing this issue, the decoder puts a rebuilding value for each codeword produced by encoder [19] [21].

2.3.3 Spread spectrum coding

There is an idea of redundant embedding in the transform domain that leads to the celebrated spread spectrum model. Messages in the spread spectrum system are encoded into symbols and then transmitted as random sequences in form of 1s and 0s. In the next step, those random sequences of 1s and 0s are broadened crosswise of frequencies. As a result, if the signal is exposed to noise or filtering process that would damages only certain frequency bands and the message will still be discoverable.

There are two main characteristics in spread spectrum communication that have high importance to watermark. First character is that inserted signal energy of one frequency is so miner to create an observable artifact. Second special character is that the watermark has high robustness against various common signal distortions, because the watermark signal is scattered over numerous frequencies [8].

2.3.4 Multiple-Bit embedding technique

The elementary techniques of this scheme insert no more than 1-bite information into an image. In detecting process, if the watermark signal is detected, we will have logic-1 output. Otherwise we will have logic-0 output. For increasing the payload of this elementary technique, one simple way is insert a bits string such as $b_1b_2...b_L$

within a image that divides the image into L sub-images I_1I_2 ... I_L of volume mxn and to put in a same-size random watermarking pattern to each sub-image Ii, as shown in figure, later than doing the modulation of the pattern according to the matching bit value bi. These bits can be modulated the patterns in some ways. We may possibly put in the random pattern of volume mxn to the sub-image if the watermark bit is one. On other hand, if it is zero or -1, depart the sub-image unaffected.

Figure 9: Embedding multiple 25 bits to an image [23]

22 One another way for achieving multiple bits embedding is to use a structure of Direct Sequence Code Division Multiple Access (DSCDMA) spread spectrum communication. In the first place, according to this technique, it generates a separate random pattern of 1 and -1 for every message bit to be embedded. It is considered that our message is $b_1b_2...b_L$, so there are L independent random patterns which
have the size as host image and are known as $v_1v_2 \dots v_L$. Every v_i pattern is as well modulated by its relevant bit, b_i . Moreover, if b_i stands for a 0 then we use the pattern.

Otherwise, if b_i stands for a 1, we use the $-v_i$ pattern. For instance, the figure 9 gives you an idea about 1-dimantional example of the technique that generates a 7-bit watermark. During 2-dimentional the signal and watermark vectors are probably substituted by the m×n blocks of the host image and random -1s and 1s in that order. It is also possible to scale down this sum earlier than embedding to fit it within certain limits as shown in the equation 2.4.

$$
W = k. \{ \sum_{j=1}^{L} v_j \}
$$
 (2.4)

V_2 : 1 1-1-1 1-1-1 1 1-1 1 b2:0 V_3 : 1-1-1 1-1-1 1 1-1 1-1 b_3 :1 V_4 : -1-1 1-1-1 1 1-1 1-1-1 b4:1 $V_5:-1$ 1-1-1 1 1-1 1-1-1 1 bs:0 V_6 : 1-1-1 1 1-1 1-1-1 1 1 b_6 :1 $V_7: -1-1$ 1 1-1 1-1-1 1 1 1 b7:0	\rightarrow -Va: -1 1 1-1 1 1 -1-1 1-1 1 \blacktriangleright $-\mathbf{V}_4$: 1 1-1 1 1-1 -1 1-1 1 1 \blacktriangleright +Vs: -1 1-1-1 1 1 -1 1-1-1 1 \blacktriangleright -V6: -1 1-1-1-1 1 -1 1 1-1-1 \blacktriangleright + $\mathbf{v}_7:$ -1-1 1 1-1 1 -1-1 1 1 1
	w : -3 5 1 -3 1 3 -7 1 3 -1 3 I: 98 98 97 98 97 96 97 96 95 94 94
	I_w : 95 103 98 95 98 99 90 97 98 93 97

Figure 10: Generation of a 7-bit DS-CDMA watermark [23]

This technique of embedding multiple bits has flexibility to be expanded to any transform domain. In the case, we apply the algorithm on the transform coefficient blocks instead of image pixel blocks.

2.4 Watermarking Detection

The basic and main idea of watermarking systems is to inset some information into medium and then extract that information as reliably as feasible. If we assume the watermark embedder as a transmitter in a communication sequence, so the watermark detector can be considered as the receiver.

Figure 11: Diagram of generic watermark detection [8]

There is a detailed block diagram of detecting process in the figure 11. There are two tasks for detecting process: make decision whether the image under testing includes a watermark and extracting the message that may watermark signal carry.

2.4.1 Efficient watermarking detection

With regard to Gaussian noise, the optimal detector is linear correlation (LC) detector. In term of watermarking embedding, DCT or DWT domain coefficients are used and with the purpose of facilitating the shaping of embedding power according to HVS limitations and to allow choice of significant signal components. DWT and DCT natural image coefficients do not act upon Gaussian law in general. There are many proposed statistical models for frequency domain coefficients of video and image that establish the watermark detection statistic. Some of them are for Discrete Fourier Transform such as: Rayleigh and Weibull distribution model and some other are for DWT and DCT domain coefficients such as: Cauchy distribution and GGD model. Multivariate distributions, for instance Multivariate Power Exponential (MPE) distribution or multivariate Gaussian can model the correlated components of color image. One of the significant aspects of efficient watermark detection is the trade-off complexity between host model and detection performance with regard to computational effort [25].

G.Depovere, T.Kalker, J.Linnartz proposed a way to do filtering before correlation, in order to improve watermarking detection reliability. This prefiltering is going to gain optimal detection in the case of actual images wherein the power spectrum is not white. According to the experiment result of this new scheme, if we use filtering earlier than correlation, the detection reliability could be considerably developed. These developments were analyzed by theoretical model founded on detection theory and statistical communication [24].

2.4.2 Multiple-Bit watermark detection

First we dividing the image into a number of blocks or sub-images and then the multiple bits are embedded into each block; a signal bit b_i has control on I_i . In order to distinguish the watermark signal, the detector calculates the correlation between image blocks and related corresponding random pattern. In the second step, if the association goes over a definite threshold T, the detector allocates the value 1 to the constructed watermark bit, or else the watermark bit is allocated to zero.

It is possible to add two diverse random patterns P_0 and P_1 used for watermark bits 0 and 1, with the intention of avoiding of using a threshold. During this time, each sub-image is correlated with both different random patterns by the detector. In the next step, the bit value matching to the pattern that provides the maximum correlation with the watermarked image is considered as the received bit. Moreover, it is possible to use this method in a more reasonable way by choosing the patterns P_0 and P_1 in a way that they are different only in sign, which is $P_0 = -P_1$ [62]. So, in this case, the computation task on the detector is reduced and the detector must compute the correlation among the sub-image and one of the noise patterns, known as P_0 . If bit is determined as 0 that means the correlation is positive; otherwise, received watermark bit is gave to 1, if the correlation is negative.

2.5 Errors in Detection

Even the best-designed watermarking system may face errors and errors are inventible. In this study, we consider three error types: message errors, false negative errors, and false positive errors. False positive errors happen once the detector wrongly determines the presentence of watermark; on other hand, when detector mistakenly points out the watermark absence are false negative errors. In messages errors, the detector incorrectly decodes the message. For designing a watermarking system is crucial to conclude what error prices are acceptable throughout the requirement part of the design. So, it is obligatory to improve forms for the errors of interest. These kinds of models have aim to double. Firstly, in order to meet the specification, a model should allow choosing a detection threshold. Secondly, by experimental verification, we can be certain that the particular error rates won't exceed.

2.5.1 Message errors

In a case of using straight message coding, sometimes the detector wrongly deciphers one message once another main message was embedded. In the same way, in multi-symbol messages, the detector will mistakenly decipher one or extra symbols. These errors are known as bit errors, while a binary alphabet is using; furthermore, BER or bit error rate is considered as an evaluation of the frequency of bit errors. Once noise disfigures the embedded watermark signal and the signal is shifted to another detecting region and caused bit or message errors. For this reason, in order to increase the robustness against distortion, it requires to maximize the separation between codes.

To combat this issue, there are some forms of error detection and correction code for protecting multi-bit and multi-symbol messages. There are various codes for this purpose, and we choose one of them according to our error expectation in the watermarking application; also, according on the watermark design computational restrictions. For instance, cropping may perhaps be an ordinary distortion in the image watermarking applications. Assume, the message symbols are encoded via spatial multi-plexing in which every consecutive bit is inserted into the next spatial neighbor region; in this case, cropping attacks can bring about rupture errors. Which is a chain of consecutive bits will be damaged or removed. In such situation, it is wise to use an error correction code which is strong against rupture errors. Cycle codes are appropriate candidates, such as Reed-Solomon cod. Another solution to this problem is the randomizing of the spatial place of every enciphered bit. Although, in this case, cropping attack is not a serious threat anymore, but it may face random errors and in such situation, block codes are suitable solutions.

2.5.2 False positive errors

A false positive error happens once the detector incorrectly determines the presentence of watermark. A false positive likelihood of $10³$ shows one false positive occurs for each 1000 detecting efforts. The following figure 12 shows the reasons and how false positive error can happen. There are two curves in the figure, the first one from right side stands for the occurring frequency of every feasible value that can be outputted from the detector as soon as there is not watermark. Similarly, the second curve from left stands for the output values frequency of the detector when there is watermark in the content. The perpendicular line shows the decision border line with sample t. if the value of detector output is less than t, the watermark is confirmed not present; if not, the watermark is confirmed in attendance. A false positive error is feasible since there is a limited possibility that the detector will yield a value bigger than or alike to t once there is not watermark.

Figure 12: A models of detector output distribution and a detection threshold. The sheltered part stands for the possibility of a false positive error [2]

2.5.3 False negative errors

The statistical chance that false negative error could happen is the false negative probability; and the frequency of occurrences is measured by the false negative rate. In the figure 13, false negative error happen due to the detector output distribution, as shown by first curve from right side, intersects the threshold t. Therefore there would be a limited possibility that detecting out will get less value than threshold t, even though at a situation that watermark exists in the computing work.

Figure 13: A model of detector output distribution and a detection threshold. The sheltered part stands for the possibility of a false negative error [2]

There is a same criterion for analyzing false positive and false negative probability. But contrasting the case of false positive possibilities, before analyzing the false negative probability, we must consider much more variables. This is because such probabilities are extremely dependent on the both watermark detector and embedder; in addition, they are dependent on what occurs in a work during embedding and detecting processes.

2.6 Attacks on Watermarks

In the watermarking literacy an attack is described as every processing which have aim to impair watermark detection or the information communication that convey via the watermark signal. So processing watermarked information is known as attacked information. In watermarking schemes, the robustness against attacks has a high importance. Perceptual quality measures the effectiveness of an attacked data and some other criteria such as channel capacity, miss probability, or bit error probability that are used for measuring the amount of watermark impairment. An attack can be successful in overcoming a watermarking system if it damages the watermark signal further than acceptable restrictions at the same time as keeping the perceptual quality. There are many types of watermark attacks that can be categorized into four groups: cryptographic, geometric, protocol and removal attacks as shown in Figure 14.

Figure 14: Watermark attack categorization [13]

2.6.1 Removal attacks

In such attacks, the main is to deteriorate or eliminate the watermark signal from its related contented, at the same time as preventing the content from being damage or become ineffective after the attack is ended. This type of attacks contains remodulation, denoising, collusion, and quantization attacks. In quantization and denoising, the procedure is to damage the watermark quality maximally, while preserving the attacked data quality high enough. The effect of Lossy compression is the same as denoising. The strategy in re-modulation is to anticipate or to predict the watermark. This conducted through a subtraction of the watermarked image median filtered version from the watermark image itself. Next, the expected watermark is eliminated from the original watermarked image. Collusion can be implemented if there are many given reproductions of a data set, every marked with a distinctive watermark, could be achieved by an attacker. This scheme of attack could be successful via averaging the entire reproductions or taking just minute portions from every diverse reproduction.

2.6.2 Geometric attacks

Geometric attacks are particular usage for images and videos comprising operations as cropping, rotation, scaling, translation, and etc. This type of attacks is different and has not aim to get rid of the inserted watermark signal, but plan to remove the watermark detector harmonization that is associated along with the inserted information. In fact, if a perfect synchronization is re-obtained, the detector could regain the embedded watermark signal. For combating this issue, new watermark methods are used invariant domains, templates, image characteristic reliant techniques or self harmonizing watermarks.

2.6.3 Cryptographic attacks

The strategy in cryptographic attacks is to break the security watermarking schemes in order to get rid of the embedded watermark information or to add deceptive watermarks. There is another attack in this type of attack that is known as the supposed Oracle attack, which has an ability to produce a non-watermarked signal while a watermark detector machine is obtainable. As piece of fact, these type attack applications are limited because of their high computational complexity.

2.6.4 Protocol attacks

The watermark inversion attack was introduced by Craver et al that generate a fake watermark scheme that is implanted on the watermarked image to resulting in uncertainty about which one of watermark signals was embedded firstly. Copy attack is a further protocol attack that the watermark is evaluated by means of a watermarked data and this evaluated watermark is inserted into new information by adjusting the regional attributes to convince its invisibility.

CHAPTER 3

WATERMARKING DOMAINS

Watermark signals could be embedded in different domains and by different algorithms which are specialized for their domains. Chapter three has categorized watermarking algorithms in term of their domains as following.

3.1 Spatial Domain

By changing the pixel values of original images, a watermark signal can be inserted in spatial domain of those images. There are chiefly three categories:

3.1.1 LSB modifications

The simplest watermark embedding method inserts the watermark signals directly into the least significant bits of the cover pixels. Since every pixel is available, smaller objects are embedded in order to combat the cropping problem. This method because of its simplicity has a number of drawbacks in the watermarked images. The watermark signal can be completely removed if the entire the LSBs of the watermarked image are place to 1; furthermore, it is susceptible to various intermediate attacks; for instance, lossy compression or every totaling of noise. There is another trustworthy method that uses a pseudo-random noise sequence producer in order to choose the pixels to be embedded that is depends a certain and secret key or seed. In other word, the watermark security would be developed as the watermark may not be exposed to the intermediate parties [14].

3.1.2 Correlation-based techniques

There is another watermarking embedding method which exploits the correlation features of extra pseudo-random noise patterns at the same time as inserted into an original image.

A PN pattern $W(x, y)$ is attached to the face image $I(x,y)$, in accordance with the following equation 3.1 as be seen below.

$$
Iw(x, y) = I(x, y) + k * W(x, y)
$$
 (3.1)

In this equation, k represents an increase feature, and IW the caused a watermarked image. Increasing k, raises the watermark robustness at the cost of the quality of a watermarked image.

An equal PN noise producer algorithm in order to retain the watermark uses the identical key and the correlation among the noise pattern and the watermarked image. The watermark will be spotted, and a sole bit will put if the correlation goes above a definite threshold T. This method has a capability to expand to a multiple bit watermark, if we segregate the image pixels into cantons and implementing this process separately on every canton.

3.1.3 CDMA- based techniques

CDMA techniques in spatial domain, has aim to disperse every bits randomly all through the cover image that caused to rising in ability and developing confrontation to cropping. In the first place, the watermark is designed as an extended string more willingly than 2D images. For every watermark value, a PN sequence is produced by an autonomous key seed. Such keys can be hoarded or themselves produced in the course of PN methods.

The total of these pseudo-random sequences stand for the watermark signal that in next step is balanced and inserted to the original cover image. In the detecting procedure, we use each seed to produce its pseudo-random sequence and then are connected with the full image.

That bit in the watermark is place to 1, if the correlation is excellent, differently a 0. This process is reiterated for all other watermark values. Although, the CDMA-based technique requires more calculation, but it increases the watermark robustness substantially [12].

3.2 Frequency Domains

In a straightforward definition, frequency domain means the image is fragmented into various frequency bands. For transferring the image to its transform representation, several reversible transforms can be used such as Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), or Discrete Wavelet Transform (DWT). In fact, every transform has its own features and illustrates the watermarked image in distinctive modes [13].

3.2.1 Discrete cosine transform

Discrete cosine transform watermarking is categorized into blocks based DCT and Global DCT watermarking.

$$
X_{k} = \sum_{n=0}^{N-1} x_{n} \cos \left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right]
$$

K= 0, 1, 2 ..., N-1 (3.2)

In the above quotation 3.2, X_k stands for discrete cosine transform of X. DCT is particularly employed for lossy information compression, due to the its tough energy compaction possessions. There was one of the primarymethods exhibited employed global DCT advance to insert a strong watermark in the conceptually important segment of HVS. Inserting the watermark signal in the conceptually significant segment of the image that has several benefits, since a large amount of compression schemes eliminate the perceptually not important part of the image.

3.2.2 Discrete wavelet transform

The DWT breaks down the image into 3 spatial trends as following: vertical, straight, and slanting. In term of computation, DWT is efficient and can be performed by means of ordinary filter complication. In the first place, the signal is analyzed at diverse frequency bands with unlike decrees by DWT through breaking down the signal into a coarse estimations and detail information. DWT uses scaling functions that employ low pass filters and wavelet functions that associated with high pass sifts. The signal disintegration into diverse frequency bands is basically gained by consecutive low pass and high pass filtering of the time domain signal. In this procedure, the main signal $x[n]$ is firstly crossed through a high-pass filter g[n] and then through a low-pass filter h[n].

$$
y_{\text{high}}[k] = \sum_{n} x [n].g[2k - n]
$$
(3.3)

$$
y_{\text{low}}[k] = \sum_{n} x [n].h[2k - n]
$$
(3.4)

In the equations 3.3 and 3.4, $y_{high}[k]$ is the output of high-pass filtering and $y_{low}[k]$ is the output of low-pass filtering.

Coefficients of DWT have bigger magnitude in the lowest bands LL at every disintegration level, but it is lesser in other bands HH, LH, HL. If scale if wavelet coefficient is large that shows it is more significant. In term of computation, detecting at lower decrees is much more efficient, for the reason that there are not many frequency bands involved in every consecutive resolution levels. The images in wavelet are multi-resolution, so can be represented in different resolution levels and can be processed from low to high decrees. Discrete wavelet transform is much more calculation use than DCT [13].

CHAPTER 4

BLIND LINEAR CORRELATION TECHNIQUE AND EXPERIMENTS

Here more specific detailed information about correlation and blind linear correlation techniques are given that is useful to have deeper understanding before starting the experiment. Next in the experiment for testing blind linear correlation technique, the Stirmark benchmark version 4 is used by applying 16 different tests on the image sets with different values, intensities, or degrees. Stirmark is a fair benchmark for testing watermarking schemes and exerts a diversity of alterations to a watermarked content to estimate the security and robustness of the watermark [26]. The used metrics are certainty of extraction, PSNR, and the visual quality of attacked images. More details about attacks, results, and discussions are given in the second part of this chapter. Finally, with respect to the experiment results and discussions, a conclusion is drawn at the end of this chapter.

4.1 Blind Linear Correlation Technique

Although in previous chapter, it has mentioned about correlation based techniques in brief, but it would be useful to have a deeper understanding of correlation based techniques, before we start to explain blind linear correlation technique. There are different kinds of correlations which have been used in watermarking. They can be categorized into three groups as flowing: linear correlation, normalized correlation, and correlation coefficient. The most straightforward is linear correlation. Other correlation types are obtained by a number of normalization applied to vectors previous to the calculation of their internal product. We gain the normalized correlation; if two vectors are normalize to unit magnitude. Subtracting their means pervious to computing normalized correlation provides the correlation coefficient between them.

Linear Correlation is the average product of elements between two vectors w_r and c, as shown in the equation (4.1).

$$
z_{1c}(v, w_r) = \frac{1}{N} \sum_{i} v[i] w_r[i].
$$
 (4.1)

It is widespread practice in watermarking communications to check for the presence of a transmitted signal w_r in a received signal cover v by computing z_{1c} (v, w_r) and contrasting it to a threshold. In fact, this perform is referred to be as matched filtering and is well-known to be an most advantageous method of detecting signals in the presence of additive, Gaussian noise.

Normalized Correlation: there are some problems with linear correlation that one of them is that detection values are extremely dependent on the magnitudes of vectors took out from works. This shows that even when reference marks are drawn from white Gaussian distribution; it would be difficult to predict linear detector's false positive probability.

Such problems can be solved through the normalization of the picked out mark and reference mark in order to make a united magnitude pervious to calculating the internal product between them. That is shown in equation 4.2.

$$
\check{\mathbf{v}} = \frac{\mathbf{v}}{|\mathbf{v}|}
$$

$$
\overline{\mathbf{w}}_{\mathbf{r}} = \frac{\mathbf{w}_{\mathbf{r}}}{|\mathbf{w}_{\mathbf{r}}|}
$$

$$
\mathbf{z}_{\text{nc}}(\mathbf{v}, \mathbf{w}_{\mathbf{r}}) = \sum_{i} \check{\mathbf{v}}[i] \overline{\mathbf{w}}_{\mathbf{r}}[i]. \tag{4.2}
$$

Correlation Coefficient: the final type of correlation we consider in this study is correlation coefficient which is obtained by deducting out the means of two vectors pervious to computing the normalized correlation between them as shown in the following equation 4.3.

$$
\check{\mathbf{v}} = \mathbf{v} - \check{\mathbf{v}}
$$

$$
\overline{\mathbf{w}}_{\mathbf{r}} = \mathbf{w}_{\mathbf{r}} - \overline{\mathbf{w}_{\mathbf{r}}}
$$

$$
\mathbf{z}_{\text{nc}}(\mathbf{v}, \mathbf{w}_{\mathbf{r}}) = \mathbf{z}_{\text{nc}}(\check{\mathbf{v}}, \overline{\mathbf{w}_{\mathbf{r}}}).
$$
 (4.3)

One advantage of this scheme is to providing robustness against changes such as the addition of a continuous intensity to the all image pixels [2].

Blind linear correlation technique embeds a pseudo-random noise pattern which has same size and dimensions as the original image; furthermore, this pattern comes into view to be random, but in point of fact it is absolutely deterministic and is based on specific algorithm and watermarking key to be generated. Then this technique exploits linear correlation features between a watermarked image and regenerated PN pattern for detecting watermark signal and does not require original image during detecting process as shown in figure 4.1.

In another word, by using the same pseudo-random noise producer algorithm and watermark key, the PN pattern can be regenerated in order to restore the watermark signal by doing linear correlation between the noise pattern and the watermarked image. If the correlation goes above a definite threshold T the watermark signal will be detected, and a single bit will set. This method has a capability to be expanded for a multiple-bit watermark, if we dividing the image pixels into blocks and implementing this procedure separately on every blocks. The author of this scheme is Nazim A. Fates [26, 27] and the last update of this scheme had been released in February 2002.

4.1.2 Pseudo random number

In blind linear correlation technique, the watermark is inserted in a form of Pseudo random noise (PN). A PN sequence contains of binary numbers of $+1$ and -1 which appears to be random, but in fact it is deterministic. This deterministic sequence of binaries repeats itself subsequent to its period. Here this pattern is determined by a seed key and its repetition period can be extremely extended, even million of binary numbers. It is impractical to predict this pseudo random number, without having knowledge of its algorithm or PN generator and the key. Pseudo random number algorithm starts from a key which is considered as initial seed. It constantly produces the same PN sequences when it is initialized with that the key or initial seed; in fact, different keys generate singular sequences. The most popular PN sequences maximal length sequences, or Gold, Barker, and Kasami codes. For example, if the key we want to use has K registers. If it has a length of 2 k-1, it is considered as maximal length sequence. Such sequence over one period has 2k-1 zeros and 2k-1 ones. A period, N, has a sequence p1 p2 p3...pN and its autocorrelation function is $Rxx(k)$ which can be seen in equation 4.4.

$$
R_{xx}(K) = \frac{1}{N} \sum_{n=1}^{N} P'_i P'_{i+k} \qquad (4.4)
$$

In above equation, P'_i equals to 1-2 pi and k stands for k-th moved version of that sequence. If $R_{xx}(K)$ is equivalent to 1, k has a values of 0 and $-\frac{1}{N}$ $\frac{1}{N}$ otherwise. In a more clear concept, if k is not equal to 0 ($k\neq 0$), the produced sequence by PN generator is uncorrelated to every part of its round shifts. Let's assume, we have a watermark W which is a binary message and has L bites of b1, b2, b3…,and bL. Every bi element is enciphered to one zero mean PN vector of span, N. since every symbol bi has two statuses, consequently two pseudo random sequences of span, N, are employed; in another word, the first sequence is corresponded to status 0 and its complementary to status 1. The matching of every symbol bi and its pseudo random sequences produces the enciphered watermark Ws. Ls=NL (4.5).

Figure 15: Blind linear correlation technique procedures

4.1.3 Embedding

Here, it codes only one bit of information in order to keep things uncomplicated; therefore, m can be 1 or 0 as shown in equation 4.6. With regard to m, we select a random reference pattern which has same dimensions or size as the original image cover and it can be seen as W_r in equation 4.6. This pattern's components are extracted from a random Gaussian distribution in the interval of -1 and 1. We use the key as a seed to start the pseudo random number producer which produces the random reference pattern.

$$
W_{\rm m} = \begin{cases} W_{\rm r} & \text{if } m = 1\\ -W_{\rm r} & \text{if } m = 0 \end{cases}
$$

$$
W_{\rm a} = \alpha W_{\rm m}
$$

 $C_w = C_o + W_a$

Message pattern analyzing is depending on what we are embedding; in fact, embedding 0 results in tacking negative to obtain the message pattern that can be seen as $-W_r$ in equation 4.1. On other hand, embedding 1 causes the pattern to be leave as it is. α in above equation has been used as a controller of embedding strength and higher values for α mean more robust watermark signal, but it has its

consequence which we experiment it in PSNR test. According to the last row of equation 4.1 and in order to obtain the watermarked image, we add original image C_0 to the balanced message pattern W_a . Figure 16 shows an example of the blind linear correlation embedding process with embedding strength $\alpha=1$. As it can be observed from figure 15 the watermarked image did not get distorted due to the embedding process and there is no perceptual difference between watermarked and original image.

Figure 16: An example of blind linear correlation embedding process

4.1.4 Detecting

In detecting procedure we proceed as equation 4.2, at the first, the linear correlation between the received watermarked image C and the initial reference pattern w_r should be calculated. By using the watermarking key which is known as initial seed, the initial reference pattern can be restored.

$$
z_{Ic}(v, w_r) = \frac{1}{N}C.w_r = \frac{1}{N} \sum_{x,y} C[x, y]w_r[x, y].
$$
 (4.2)

$$
m_n = \begin{cases} 1 & \text{if } z_{\text{Ic}}(c, w_r) > t_{\text{Ic}} \\ \text{no watermark} & \text{if } -t_{\text{Ic}} \le z_{\text{Ic}}(c, w_r) \le t_{\text{Ic}} \\ 0 & \text{if } z_{\text{Ic}}(c, w_r) < -t_{\text{Ic}} \end{cases}
$$
(4.3)

With regard to the result of the linear correlation computation, we can determine what the watermark message is. According to equation 4.3, if the linear correlation value m_n is higher than the threshold, it is declared that the message is a 1. On other hand, if the linear correlation value m_n is less than the negative of the threshold, it is declared that the message is a 0. However, it is declared that there is no embedded message, if the linear correlation value is between positive and negative threshold.

4.2 Experiments and Results

In the experiment part, the Stirmark benchmark version 4 is used for testing blind linear correlation technique sets by applying 16 different tests on the image sets. For doing a fair comparison, it is imperative to test an images watermarking scheme on different images; in addition, the same sample image must always be used. It is unfeasible to obtain a comprehensive list of image classes and it is so difficult to achieve an agreement of satisfactory index for using stock photo companies. However, at least we can use images which have been used in watermarking comparisons for years and which are interested from point view of image processing. Five images in this experiment with formats of bmp and ppm are used which can be categorized into classic (which have been used for years), photo with edge and lines, landscape and grid or patterns image. Each single test or attack is applied on all image sets with different values, intensities, or degrees. In fact, 109 tests in 16 different groups have been applied on 5 watermarked images and as results of this experiment there are 438 attacked outputted images. It might come to view that the image sets do not cover a broad range of different types of images, but because many different attacks have been applied on images and the nature of images have changed, so they have converted into many different types of images; for example, the tested image sets does not include bright color images, but at output of convolution filtering test Lena image becomes bright color image, and still retains its watermark signal. Those sixteen tests include PSNR, Embedding and extracting time, Additive noise ,JPEG, Convolution Filtering, Self-Similarities, Remove Lines, Cropping While, Rescale, Rotation, Rotation and Cropping, Rotation and Rescaling, Affine Transformations, Small Random Distortions, and Latest Small Random Distortions. Stirmark is a fair benchmark for watermarking schemes and applies a variety of distortions to a watermarked content to evaluate the robustness and security of the watermark [26]. Here, the certainty of extraction is used as a measure for robustness and PSNR as a measure for distortions on images; in addition, the visual quality of attacked images is considered as a measure for the effectiveness of attacks.

4.2.1 PSNR test

Images get distorted when they are transmitted from one place to another, due to the noise present in the channel which is called PSNR. This engineering term is an abbreviation of Peak signal-to-noise ratio; in another word, it is the maximum possible power of the signal to the power of corrupting noise ratio. A high PSNR value shows a good reconstruction.

Figure 17: The outputted watermarked images of PSNR test

Decibel or dB is PSNR unit and our image sets are 24 bits, and their square of the peak value is shown as Max in equation 4.4. For calculating PSNR, we divide MAX by mean square error or root mean square error which is shown in equation 4.4 as MSE and RMSE respectively.

$$
PSNR = 10 \log_{10} \left(\frac{\text{Max}^2}{\text{MSE}} \right) \text{ or } PSNR = 20 \log_{10} \left(\frac{\text{Max}}{\text{RMSE}} \right) \quad (4.4)
$$

Assume the damaged image is C_{noisy} and the original image is C_0 , so we calculate MSE or mean square error and RMSE or root square error as equation 4.5.

$$
MSE = \frac{\Sigma (C_0 - C_{noisy})^2}{\Sigma 1} \quad \text{and} \quad RMSE = \sqrt{MSE} \quad (4.5)
$$

If the RMSE value equals to Max, we obtain a PSNR value of zero and for the RMSE value more than Max, we obtain a negative PSNR value. The simplicity of this metric (PSNR) has caused to be considered as the most popular distortion measure in the research area of image, compression and video coding. In this test, different embedding strengths (from 10 to 100) are inserted into image sets and the amounts of PSNR and certainty have been measured. The results are shown in table 1.

	Set3/Sample.bmp	67.8197	27.402
	Set1/Sample.bmp	69.4687	VV
60	Set1/Lena.bmp	159.481	24.5416
	Set1/Sample.ppm	69.4634	25.5724
	Set2/skyline_arch.bmp	94.5055	25.3791
	Set3/Sample.bmp	69.4665	25.5722
	Set1/Sample.bmp	71.9354	23.9907
	Set1/Lena.bmp	162.362	23.0302
70	Set1/Sample.ppm	71.93	23.991
	Set2/skyline_arch.bmp	97.0651	23.8187
	Set3/Sample.bmp	71.9331	23.9908
	Set1/Sample.bmp	73.5788	23.0661
	Set1/Lena.bmp	164.254	22.1442
80	Set1/Sample.ppm	73.5734	23.0664
	Set2/skyline_arch.bmp	98.765	22.9151
	Set3/Sample.bmp	73.5765	23.0662
	Set1/Sample.bmp	76.0407	21.8533
90	Set1/Lena.bmp	167.033	21.0196
	Set1/Sample.ppm	76.0388	21.8536
	Set2/skyline_arch.bmp	101.3	21.7187
	Set3/Sample.bmp	76.043	21.8533
100	Set1/Sample.bmp	77.6825	21.1354
	Set1/Lena.bmp	168.847	20.3708
	Set1/Sample.ppm	77.6805	21.1357
	Set2/skyline_arch.bmp	102.982	21.0071
	Set3/Sample.bmp	77.6847	21.1354

Table 1: The result of PSNR test

The result table of PSNR test shows that when we use weak embedding strength 10, the PSNR value of image sets are approximately 39 that shows a good reconstruction. But as the embedding strength is increased to 20, the PSRN amounts are decreased to approximate 35 in all tested images. This procedure is continuing by increasing embedding strength and decreasing PSNR. That means there is an inverse direct relation between watermark embedding strength and PSNR. As mentioned in equation 4.3, increasing a (watermark embedding strength) result in more distortions in visual quality and subsequently decreasing in PSNR value. In addition, as shown in the figure 17 and as we expected the images with lower watermark embedding strength have better quality. It can be observed that by increasing the watermark embedding strength, images' visual qualities are decreased. That shows there is tradeoff between quality of images and watermark embedding strength. But on other

hand, as shown in the figure 16 and table 4.1, when the values of PSNR are changed, it does not decrease or increase the visual quality of images; in fact, we can observe that PSNR values do not have direct effect on the visual quality of images. In another word, watermarked images with bad PSNR value could have high visual quality and vice versa. All in all, as the PSNR test results illustrated, blind linear correlation technique has almost reasonable PSNR values even at the high watermark embedding strength.

4.2.2 Embedding and extracting time test

This test is done by embedding 5 random keys for per media and computing the average embedding and extracting times for each image. In addition, this test provides the average PSNR value for each image as it can be seen in table 2. The embedding average times for the color images, such as Lena and Skyline Arch, are 146 and 123 ms respectively which are much higher than the embedding average time of grid images which have average embedding time of 40 ms. This is ordinary for color images to have more embedding time, because they have much more colors per pixel. Moreover, the PSNR values of those colorful images are approximately 25 which are less than PSNR values of grid images that are approximately 27. The color images have more complex colors per pixel and get more channel noise that is why in such image reconstruction is lower than those images which are grayscale or pattern images.

In extracting time test, the average extracting times of colorful images Lena and Skyline Arch are 44 and 36.2 ms which are again higher than the average extracting time of graphic images. Furthermore, PSNR values of colorful images Lena and Skyline Arch are 25.3726 and 26.6726 respectively that are a bit less than PSNR values of graphic images which have PSNR values 27.402. This test has revealed that blind correlation technique in both extracting and embedding times or in processing time is reasonable and has low computing fee, because one of the main advantage of linear correlation based techniques is their low computing cost due to its the straightforwardness.

	Tested images	Average time (ms)	PSNR(dB)
Embedding Time Test	Set1/Sample.bmp	39.8	27.402
	Set1/Lena.bmp	146.2	25.3726
	Set1/Sample.ppm	40.6	27.402
	Set2/skyline_arch.bmp	123.4	26.6726
	Set3/Sample.bmp	40.8	27.402
	Set1/Sample.bmp	11.6	27.402
Extracting Time Test	Set1/Lena.bmp	44	25.3726
	Set1/Sample.ppm	10.4	27.402
	Set2/skyline_arch.bmp	36.2	26.6726
	Set3/Sample.bmp	13	27.402

Table 2: The result table of embedding and extracting time test on image sets

4.2.3 Add noise test

This test is fairly streamlined and applies additional noise on the image sets, in order to stop watermark detecting process. In view of the fact that every pixel of image has tolerance for noise amount that can be given and still stay invisible. This test by adding extra noise, tries to use that tolerance value to give the greatest amount of uncertainty that a watermark detector will must deal with. Different noise levels have been applied on images in this test; furthermore, the Certainty and PSNR are computed on each image with the noise level from 0 to 100. The results can be seen in the table.3 and figure 18. With no level of noise, the watermarked images have high Certainty. But as the noise level has increased to 20, the watermarked images' certainty amounts are decreased substantially. Because this attack is directly targeted the tolerance to be changed. For instance, as can be seen the certainty amount of Skyline Arch image is decreased from 92.7884 to 55.1775; in addition, the images with noise level of 20 have gotten distorted perceptually. Moreover, PSNR values of those images have decreased too.

Noise Level	Tested Images	Certainty	PSNR(dB)
	Set1/Sample.bmp	67.8197	$1.$ #INF
	Set1/Lena.bmp	157.543	$1.$ #INF
$\overline{0}$	Set1/Sample.ppm	67.8169	$1.$ #IN F
	Set2/skyline_arch.bmp	92.7884	$1.$ #INF
	Set3/Sample.bmp	67.8194	$1.$ #IN F
	Set1/Sample.bmp	44.1471	12.3051
	Set1/Lena.bmp	84.0784	9.07065
20	Set1/Sample.ppm	44.146	12.2171
	Set2/skyline_arch.bmp	55.1775	11.3399
	Set3/Sample.bmp	44.239	12.2284
	Set1/Sample.bmp	45.9442	9.66423
40	Set1/Lena.bmp	73.0864	7.61133
	Set1/Sample.ppm	45.8787	9.60244
	Set2/skyline_arch.bmp	53.3152	9.37902
	Set3/Sample.bmp	45.4834	9.60033
	Set1/Sample.bmp	51.3379	8.48619
	Set1/Lena.bmp	70.0575	7.19091
60	Set1/Sample.ppm	51.2691	8.48135
	Set2/skyline_arch.bmp	56.6485	8.61636
	Set3/Sample.bmp	51.2847	8.50547
80	Set1/Sample.bmp	54.0519	7.97841
	Set1/Lena.bmp	68.4483	6.97718
	Set1/Sample.ppm	54.463	7.9755
	Set2/skyline_arch.bmp	58.3521	8.24036
	Set3/Sample.bmp	54.5528	7.97433
100	Set1/Sample.bmp	56.3439	7.67836
	Set1/Lena.bmp	67.4586	6.86204
	Set1/Sample.ppm	56.3717	7.6639
	Set2/skyline_arch.bmp	59.4079	8.06952
	Set3/Sample.bmp	56.4485	7.71125

Table 3: The result of additive noise test on the image sets.

By applying noise level of 40, the certainty amounts are decreased again, but this time slightly. For example, Lena's Certainty decreases from 84.0784 to 73.0864; also its PSNR value is reduced from 9.07065 to 7.61133. This procedure happened on other watermarked images too with increasing noise level.

Figure 18: Some outputted watermarked images of additive noise test

In another word, increasing noise level has direct negative effect on certainty and PSNR; in addition, the image qualities got distorted significantly as shown in the figure 18. This is typical to see such consequences of adding large amount of additional noise to images, because this large amount of noise causes severe variations in brightness or color of images that result in reducing of both certainty and PSNR values. It can be concluded that watermarked images based on blind linear correlation technique are sufficient robust against additive noise attacks, because by adding even noise level of 100 the watermarked images show good certainties more than 50 which is enough for watermark signal to be detected.

4.2.4 JPEG test

In the test, JPEG compression is applied on watermarked image sets with different range quality level of 15 to 100. Currently, one of the most popular lossy compression algorithms for images is IPEG compression. Nowadays, in order to lessen file size of images, lossy compressed has been used popularly. Therefore, any reasonable image watermarking system should be flexible to some level of lossy compression.

Table 4: The result of JPEG test on the image sets.

As shown in table 4, the quality of JPEG compression on watermarked image sets is started from 15. At this level of compression quality, the tested images are resilient and have good amount of certainties, normal PSNR values; for instance, Skyline Arch image has very high certainty of 92.78.27 that shows a good robustness, and with a PSNR value of 31.6774. This shows a good reconstruction and retaining of watermark signal. When JPEG quality level has come up to 30, the certainty amounts are almost unchanged. In another word, the certainty amounts are decreased very slightly or vice versa at different JPEG quality level. For example, according to the results of table 4, the certainty amount of Set/Sample image increased from 67.0776 to 67.976; furthermore, the certainty amount of Lena image has reduced fractionally from 157.479 to 157.534.

Figure 19: Two outputted watermarked images of JPEG test

By continuing this trend and increasing JPEG quality degree to 100, it can be observed that the certainty amounts are decreased very slightly or vise versa. In addition, PSNR values are increased significantly that shows a good reconstruction and also shows that root mean square error (RMSE) has lower value then Max. that is why PSNR values at JPEG quality degree of 100 are high. For instance, Lena image, at JPEG quality level 100, has a PSNR value 49.9876 that is less than its PSNR value at JPEG quality level 15 by a difference of 16.1887. These results prove this point that blind linear correlation technique has high robustness against JPEG attack. Furthermore, with regard to HVS and as shown in the figure 18, Lena image at JPEG quality level of 100 lost its visual quality slightly and became a little

brighter in a way that is so difficult for human perception to recognize it. That shows this technique can conceal watermark more efficiently to HVS and it is so resilient watermarking system for image compression.

4.2.5 Median cut filtering test

Median cut filter is a kind of nonlinear filtering methods that has aim to get rid of noise. This technique is used popularly in digital image processing, so digital image watermarking techniques should be compatible with such filtering which is very widely used. In current test, median cut filter with different filter size from 3 to 9 are applied on the watermarked image sets and the results are shown in table 5 and figure 20.

Table 5: The result of median cut filtering test on the image sets

As table 5 shows us, the PSNR values with filter size of 3 are reduced substantially for grid images, this happened since this filtering on such images removed mistakenly many points of patterns because it considered them as noises ; for instance, the PSNR value of Set1/Sample.bmp image at filter size of 3 has decreased to 9.27999. This low PSNR value of Sample image shows a bad reconstruction that may have bad effects on its visual quality. PSNR values of colorful images have decreased partially such as Set1/Lena.bmp which its PSNR values decreased to 36.8774. But the main outstanding result at this level of filter size is that the certainty amounts of all images are almost remained in high level; such as certainty amounts of Set1/Lena.bmp image that is 157.697 at filter size of 9.

Figure 20: Some outputted watermarked images of median cut filtering test

With regard to HVS and as shown in figure 20 at this filter size, the images get distorted very slightly for all images except grid images which got distorted dramatically. According to the table, as filter size is increased to 5, 7, and 9, both certainty and PSNR values are decreased very slightly; albeit, as shown in figure 4.6 images at those filters sizes are got distorted and lost their quality, but except grid images all other images have still high certainty values such as Lena and Skyline arch images. It can be seen that blind linear correlation technique is not fully robust against median cut filtering.

4.2.6 Convolution filtering test

Gaussian and sharpening filtering are two convolution filtering types which are used in this test. The values for those two specific filters are as following in this test of Stirmark benchmark:

First filtering, Gaussian values: 3 3 9, 1 2 1, 2 4 2, 1 2 1.

Second filtering, Sharpening values: 3 3 9, 0 -1 0, -1 5 -1, 0 -1 0.

The first two digits stand for width and height, and last digit stands for division factor. The results of this test are shown in table 6 and two outputted image of those filters are shown in figure 21. With respect to the result table 6, the images at first filtering or Gaussian filtering have low PSNR values that caused bad reconstructions on images. In fact this filtering changes every pixel of original image into bight pixel in order to get rid of noise by its filtering values that caused RMSE to get very lower values than Max and results in low PSNR values. But they have very high certainty values which show images' resistance against Gaussian filtering. In addition, the visual quality of images got disturbed and became slightly as can be observe in figure 21.

Figure 21: Some outputted watermarked images of convolution filtering test

The certainty values of outputted images at second filter or sharpening filtering are decreased sharply and images become so dark; furthermore, PSNR values are reduced significantly. For instance, PSNR value of Set1/Sample.bmp image is - 7.33256, and its certainty amount is 17.6324. This image gets a negative PSNR value because its root mean square error (RMSE) is more than its peak value (Max), such RMSE value results in a negative PSNR value.

Filter No.	Tested Images	Certainty	PSNR(dB)
1	Set1/Sample.bmp	109.323	11.4535
	Set1/Lena.bmp	224.45	11.6843
	Set1/Sample.ppm	109.327	11.4534
	Set2/skyline_arch.bmp	120.705	16.3893
	Set3/Sample.bmp	109.326	11.4535
2	Set1/Sample.bmp	17.6324	-7.33256
	Set1/Lena.bmp	84.0784	9.07065
	Set1/Sample.ppm	14.75	1.2835
	Set2/skyline_arch.bmp	10.3038	-5.59144
	Set3/Sample.bmp	14.7493	1.36991

Table 6: The result of convolution filtering test on the image sets

According to the result, it can be seen that Blind linear-correlation based watermarked images are more robust against Gaussian filtering rather than against sharpening filters. Albeit, at second filtering tested images lost their certainty amounts sharply, but because of losing their visual qualities significantly, we cannot say this technique was defeated against this attack. In another word, an attack can be successful if it can reduce certainty values without disturbing visual quality of the attacked image.

4.2.7 Self-Similarities test

This test is based on the correlation between image pixels that is called selfsimilarities. The aim of this attack is to substitute some image pixels with other similar neighbor pixels of image, in order to stir the watermark signal. This test is applied on image sets with different percentage of swap from 1 to 3. The results are shown in the table7 and more details are in following.

Swap	Tested Images	Certainty	PSNR(dB)
Percentage			
	Set1/Sample.bmp	68.6078	16.1343
	Set1/Lena.bmp	157.161	41.374
1	Set1/Sample.ppm	68.6062	16.1347
	Set2/skyline_arch.bmp	92.8362	31.4008
	Set3/Sample.bmp	68.6094	16.1343
$\overline{2}$	Set1/Sample.bmp	67.7518	34.9654
	Set1/Lena.bmp	157.384	51.5908
	Set1/Sample.ppm	67.7493	34.8987
	Set2/skyline_arch.bmp	92.7652	51.1258
	Set3/Sample.bmp	67.7548	34.9138
3	Set1/Sample.bmp	67.6918	17.1357
	Set1/Lena.bmp	157.224	35.5704
	Set1/Sample.ppm	67.6891	17.2056
	Set2/skyline_arch.bmp	92.4259	31.9016
	Set3/Sample.bmp	67.694	17.1386

Table 7: The result of self-similarities test on the image sets

Figure 22: Some outputted watermarked images of self-similarities test.

According to the table 7 and figure 22, images at swap percentage 1 have high certainty amounts and the changes are imperceptible; in addition, PSNR values are low for grid images because such images have very exact edges and lines that even small substitution of pixels with other similar neighbor pixels causes to lose their visual quality and increases RMSE value which results to low PSNR values. But PSNR values are higher for other colorful images and that may causes better reconstruction. For example, Set1/Lena.bmp image is colorful and has PSNR value of 41.374 which is higher than PSNR values of Set1/Sample.bmp image and as a result, we have better visual quality at Lena image.As can be observed in figure 4.8, by increasing swap percentage to 2, images lost their HVS quality slightly; in other word, the images get a little distortion, but as table 4.7 show us that images are still robust and have a high certainty amount. Furthermore, PSNR values are increased for all tested images; for instance, PSNR value of Set1/Sample.bmp image has increased from 16.1343 to 34.9654, but its certainty value decreased slightly from

68.6078 to 67.7518.By swap percentage 3, all tested images become opaque but their certainty amounts are remained high and their PSNR values are increased slightly. It can be concluded that blind linear correlation technique has high robustness against self-similarities attacks and has high certainty values even at high percentage of swap.

4.2.8 Remove lines test

The given images in this test will be involved in a procedure of removing columns and rows at different frequencies from 10 to 100. For instance, at frequency 10, one line will be removed in every 10 lines. The results of this test are shown in the table 8 and figure 23.

Figure 23: Some outputted watermarked images of remove line test

With regard to results in figure 23 and table 8, at frequency 10, images get distorted especially grid images in a perceptible way. Even though, this high frequency removed many lines, the attacked images illustrated a high resistance and all images have high certainty values. For example, Set2/skyline_arch image has certainty value 92.7956 and its quality got distorted slightly. Also at other frequencies, attacked images show almost the same resistance and have high certainty values. It can be observed that blind linear technique is robust again remove line attack at different frequencies.

70	Set1/Sample.ppm	65.4656
	Set2/skyline_arch.bmp	92.7473
	Set3/Sample.bmp	65.4709
	Set1/Sample.bmp	65.9143
	Set1/Lena.bmp	157.585
80	Set1/Sample.ppm	65.9125
	Set2/skyline_arch.bmp	92.8767
	Set3/Sample.bmp	65.9134
	Set1/Sample.bmp	66.1448
	Set1/Lena.bmp	157.602
90	Set1/Sample.ppm	66.1402
	Set2/skyline_arch.bmp	92.7335
	Set3/Sample.bmp	66.1402
	Set1/Sample.bmp	66.4858
	Set1/Lena.bmp	157.626
100	Set1/Sample.ppm	66.4856
	Set2/skyline_arch.bmp	92.726
	Set3/Sample.bmp	66.4835

Table 8: The result of remove lines test on the image sets.

4.2.9 Cropping while test

This test has aim while target and maintain the center part of image, crops images based on cropping ratio. With regard to figure 24 and table 9 at cropping ratio 1, even though only central minor part of image has been kept and the rest has removed, the attacked images have high certainty values such as Set2/skyline_arch which its certainty value at cropping ratio one is 129.517.

Figure 24: Some outputted images of cropping while test

By increasing cropping ratio, certainty values remained high with only small swing; in another word, by increasing cropping ratio, sometimes certainty values increased and vice versa. For instance, Set1/Sample.bmp image's certainty value at cropping ratio two is 53.3333 and at cropping ratio five is increased to 66.7778; but, at ratio 10 is decreased to 55.3538. There is not any direct effect of cropping ratio on images' certainty values, because the reference pattern was spread across a long of the images, not just a special part of images, that is why even by small part of image we can find the watermark signal through correlation between regenerated pattern and attacked images. The main point is that blind linear correlation watermarked images are robust against Cropping While attack.

	Set2/skyline_arch.bmp	
	Set3/Sample.bmp	59.8023
	Set1/Sample.bmp	57.3161
	Set1/Lena.bmp	163.963
25	Set1/Sample.ppm	57.3151
	Set2/skyline_arch.bmp	146.749
	Set3/Sample.bmp	57.3151
	Set1/Sample.bmp	80.2692
50	Set1/Lena.bmp	158.622
	Set1/Sample.ppm	80.2771
	Set2/skyline_arch.bmp	134.688
	Set3/Sample.bmp	80.2771
75	Set1/Sample.bmp	77.0847
	Set1/Lena.bmp	156.252
	Set1/Sample.ppm	77.0847
	Set2/skyline_arch.bmp	77.0718
	Set3/Sample.bmp	108.707

Table 9: The result of cropping while test on the image sets

4.2.10 Rescale test

This test or attack rescales watermarked images with different percentages of 50, 75, 90, 110, 150, and 200, in order to destroy or damage watermark signal. The results of the test are shown in table 10 and figure 25.

Figure 25: Some outputted images of rescale test

Tested images lost their visual quality at high rescale percentage as the figure 25 shows us because of high rescaling enlarges pixels that causes to reduction of visual quality of images. But all images have very high certainty values at different rescaling percentage. For instance, Set1/Sample.bmp image at different rescaling percentages has certainty values about 67 or 68. These high certainty values show us that those blind linear correlation watermarked images have high resistance and are robust significantly against rescaling attack.

75	Set1/Sample.ppm	67.974
	Set2/skyline_arch.bmp	92.9295
	Set3/Sample.bmp	67.9684
	Set1/Sample.bmp	68.681
	Set1/Lena.bmp	157.538
90	Set1/Sample.ppm	68.6862
	Set2/skyline_arch.bmp	92.8665
	Set3/Sample.bmp	68.6799
	Set1/Sample.bmp	67.8502
	Set1/Lena.bmp	157.552
110	Set1/Sample.ppm	67.8546
	Set2/skyline_arch.bmp	93.0017
	Set3/Sample.bmp	67.8479
	Set1/Sample.bmp	67.7156
	Set1/Lena.bmp	157.502
150	Set1/Sample.ppm	67.7181
	Set2/skyline_arch.bmp	92.9832
	Set3/Sample.bmp	67.7193
	Set1/Sample.bmp	67.9761
	Set1/Lena.bmp	157.514
200	Set1/Sample.ppm	67.9777
	Set2/skyline_arch.bmp	92.9182
	Set3/Sample.bmp	67.9758

Table 10: The result of rescale test on the image sets

4.2.11 Rotation test

This test rotates watermarked images by angles of -2, -1, -0.75, -0.5, -0.25, 0.25, 0.5, 0.75, 1, 2, 5, 10, 15, 30, 45, and 90 in order to damage watermark signal detection. The results of this test are shown in table 11 and figure 26. With regard to table 4.11, attacked images by low or negative angle degree have very high certainty values; for instance, Lena image has certainty values 152.732 and 147.589 at angle degree -1, and 5 respectively. But by increasing angle degree to 30 and 45, certainty values have been reduced sharply, because our technique is based on linear correlation and effectiveness of computing correlation at those sharp angle degree are reduced and when images are rotating at angles which are closer to straight line the effectiveness computing correlation is increased.

Figure 26: Some outputted images of rotation test

For example, Skyline Arch's certainty value at angle degree of 0.25 is 92.156 but at angle degree of 45 it is reduced to 44.7367. On another hand, when angle degree increased to 90, images restore high certainty values again because at such angle degree, computing linear correlation is more effective. It can be observed that blind linear correlation watermarked images are robust against rotation attack. But due to their sharp reduction of certainty values at angle degrees of 30 and 45, it can say the watermarked images by this technique are effectible at those angle degrees but still can retain the watermark signal.

5	Set1/Sample.ppm	57.8902
	Set2/skyline_arch.bmp	78.2574
	Set3/Sample.bmp	57.8893
Set1/Sample.bmp		50.576
	Set1/Lena.bmp	117.852
10	Set1/Sample.ppm	50.5719
	Set2/skyline_arch.bmp	67.9614
	Set3/Sample.bmp	50.5757
	Set1/Sample.bmp	45.255
	Set1/Lena.bmp	105.401
15	Set1/Sample.ppm	45.2535
	Set2/skyline_arch.bmp	60.4148
	Set3/Sample.bmp	45.253
	Set1/Sample.bmp	36.252
	Set1/Lena.bmp	84.5352
30	Set1/Sample.ppm	36.2505
	Set2/skyline_arch.bmp	48.0188
	Set3/Sample.bmp	36.252
	Set1/Sample.bmp	33.8227
	Set1/Lena.bmp	79.0291
45	Set1/Sample.ppm	33.8241
	Set2/skyline_arch.bmp	44.7367
	Set3/Sample.bmp	33.821
	Set1/Sample.bmp	67.5096
	Set1/Lena.bmp	157.221
90	Set1/Sample.ppm	67.5107
	Set2/skyline_arch.bmp	92.7607
	Set3/Sample.bmp	67.5087

Table 11: The result of rotation test on the image sets.

4.2.12 Rotation and cropping test

This test rotates images by minor angle degrees from -2 to 2 and at the same time crops the corners of images. Table 12 and figures 27 show the result of this test.

Figure 27: A number of outputted images of rotation and cropping test

Certainty values of all images at different rotation and cropping angle degree are very high, because of two reasons. First of all, the reference pattern is spread all over the images and cropping cannot remove watermark signal. Secondly, rotating at such low angle degrees (which are close to straight line) does not affect the computing linear correlation. For instance, Lena image has approximate certainty value of 157 at all angle degrees. As can be seen in figure 27, the damage of the rotation and cropping on images are obvious; but in fact, this technique is completely robust against rotation and cropping attack.

	Set1/Sample.bmp	69.4056
	Set1/Lena.bmp	157.667
	Set1/Sample.ppm	69.4077
	Set2/skyline_arch.bmp	95.6749
	Set3/Sample.bmp	69.4067

Table 12: The result of rotation and cropping test on the image sets

4.2.13 Rotation and rescaling test

This test has the same procedure as pervious test but with an addition of rescaling to the original input image size. The outcome and results of this test are shown in table 13 and in figure 28.

Figure 28: Some outputted images of rotation and rescaling test

At the lowest angle degree of -2, all attacked images have high certainty amounts and low PSNR values such as: Skyline Arch image which has certainty amount of 95.351, but low PSNR value of 17.7584. Because this attack by rotating and rescaling decreased the RMSE value that resulted in low PSNR value and weak reconstruction. According to table 13, at other rotation and rescaling angle degrees all attacked images have still good certainty, but very low PSNR values; for example, Lena image has an approximate certainty and PSNR values of 157 and 23 respectively. With regard to the result, it can be determined that blind linear correlation watermarked images has very high resistance against rotation and rescaling attack, but it has low PSNR values.

	Set3/Sample.bmp	68.3909	7.54742
	Set1/Sample.bmp	69.5903	7.02619
	Set1/Lena.bmp	157.697	19.1355
	Set1/Sample.ppm	69.5903	7.02623
	Set2/skyline_arch.bmp	95.764	17.6603
	Set3/Sample.bmp	69.5911	7.02619

Table 13: The result of rotation and cropping test on the image sets

4.2.14 Affine transformation test

This general test is used for subjective affine image transformation. the parameters should be specified as a, b, c, d, and e for the inverse transformation matrix of the equation form 4.6:

$$
\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d \\ e \end{pmatrix} \quad (4.6)
$$

In this test, 8 different matrixes with above equation form at three shearing-angles have been used for testing the images. Those 8 matrixes with shearing-angles have been shown in table 14. The results of this experiment are displayed in table 15 and figure 29.

More general affine transformation	Mat6=1.010 0.013 0 0.009 1.011 0 Mat7= $1.0070.010000.0101.0120$
	Mat8=1.013 0.008 0 0.011 1.008 0

Table 14: Matrixes have been used in the affine transformation test

Figure 29: Several outputted images of affine transformation test

According to table 15 and figure 29, images get distorted perceptually at different matrixes, but they preserved the watermark signals. As it can be observed, at all affine transformation matrixes, all tested images have high certainty values such as Lena image which has an approximate certainty value of 150 at all matrixes because the modifications caused by those matrixes on attacked images do not damage the linear correlation between watermarked image and regenerated pattern. In accordance with the result, it can be concluded that blind linear correlation watermarked images are robust against affine transformation attack and have shown high resistance.

Table 15: The result of affine transformation test on the image sets

4.2.15 Small random distortion

Small random distortion test has aim to do a simulation of resampling process; in another word, this test produces perturbations; for instance, some errors usually happen when an image is printing and subsequently it should be scanned again.

Figure 30: A number of outputted image of small random distortion

The entries of 0.95, 1, 1.05 and 1.1 have been used in this test; albeit, this entry is not used yet but it has to be present in order to perform the test. The experiment results are shown in table 16; in addition, some output images of this test are shown in figure 30. At entry of 0.95, all tested watermarked images have low PSNR values, but high certainty values except the Set3/Sample.bmp image which has a certainty value of 0.197031 that is so low to be detected as watermarked image. Since, this attack causes on RMSE value of Set3/Sample.bmp image to be increased to almost equal value with its Max value that results in a very low PSNR value near to zero. Furthermore, as can be observed in figure 28, this grid image lost its visual quality

completely. On other hand and as shown in figure 30, colorful images such as Lena lost its visual quality slightly, but also it has high certainty amount but low PSNR value because this attack increased the RMSE values and subsequently we got low PSNR value, but this attack could not damage linear correlation in this images and that is why we have high certainty value in this color image.At other entry degrees, the certainty amounts of images except the Set3/Sample.bmp, are decreased very slightly; for example, the certainty amount of Lena image has decreased from 147.529 to 146.261 at entry of 1.1. Moreover, the Set3/Sample.bmp has a certainty amount less than one and a negative PSNR value because its RMSE value increased more than its Max value that results in a negative value of SPNR. Regarding the results, some blind linear-correlation watermarked images are robust against small random distortion attack; on other hand, some watermarked image may not be robust such as Set3/Sample.bmp image which its certainty amount is not sufficient to be detected that can be considered as a disadvantage of this technique against such attacks.

Table 16: The result of small random distortion test on the image sets

4.2.16 Latest small random distortion

This test had been introduced in version 4.0 of stirmark benchmark and is an evolved modification of small random distortions test. This test, through a distortion procedure, substitutes a couple of processing paces of higher frequency displacement and global bending.These two main paces are depended in sine-functions to decide the individual distortion of every point of the outputted image. In another word, in this test, every point of outputted image will be exposed to the distortions of x and ycoordinates. Here, the same entry as pervious has been used and results are shown in table 17 and some outputted images of this experiment are shown in figure 31.

Entry	Tested Images	Certainty	PSNR (dB)
	Set1/Sample.bmp	65.3091	6.85594
	Set1/Lena.bmp	151.774	16.4369
0.95	Set1/Sample.ppm	55.463	6.26 178
	Set2/skyline_arch.bmp	49.4878	7.93895
	Set3/Sample.bmp	3.35954	6.81434
	Set1/Sample.bmp	65.249	6.8202
	Set1/Lena.bmp	151.566	16.1946
1	Set1/Sample.ppm	54.944	6.2599
	Set2/skyline_arch.bmp	47.3276	7.93895
	Set3/Sample.bmp	2.44988	6.82833
	Set1/Sample.bmp	65.1765	6.7987
	Set1/Lena.bmp	151.361	15.9675
1.05	Set1/Sample.ppm	54.4753	6.25393
	Set2/skyline_arch.bmp	45.1287	7.93895
	Set3/Sample.bmp	1.74882	6.84214
1.1	Set1/Sample.bmp	65.131	6.77074
	Set1/Lena.bmp	151.154	15.8205
	Set1/Sample.ppm	53.6723	6.30116
	Set2/skyline_arch.bmp	42.924	7.6826
	Set3/Sample.bmp	1.20621	6.85398

Table 17: The result of latest small random distortion test on the image sets

With regard to the result table 17, all tested images except Set3/Sample.bmp image, at entry of 0.95, have high certainty amounts but low PSNR values as in previous test. But here, images are a little bit more robust and have a bit higher certainty

amounts. At other entry degrees, the similar results as in previous test can be observed, but according to figure 31, colorful and grid images get less distorted than previous test. Although, the certainty amount of Set3/Sample.bmp image at most intensive entry is 1.20621 and is so low to be considered as a watermarked image, but the image got distorted sharply in term of visual quality. So, that is not a failure for blind linear-correlation technique, but it is a disadvantage for this technique against latest small distortions attack as well.

Figure 31: Few outputted images of small distortion test

CHAPTER 5

CONCLUSION

This thesis aims to test blind linear correlation technique by exerting a diversity of alterations on the image sets with help of Strimark benchmark 4.0. This technique embeds watermark signal in a form of PN pattern which is based specific key and PN generator. This technique does not require original image during detecting process and use the key to regenerate the reference pattern and then computes linear correlation between watermarked image and the reference pattern for detecting the watermark signal. For doing comparison, the used metrics in tests are PSNR, Certainty, and visual quality of attacked images. As we experimented, blind linear correlation technique shows a high potential of robustness against most attacks. For example, in PSNR test, tested images even at high watermarking strength have good PSNR values and at low watermarking strength have very high certainty values. Furthermore, this technique has an outstanding processing time and low computational cost that can be considered as one of the most important advantages of this scheme. Another significant feature of this scheme is its high resistance against some attacks such as cropping and affine transformations. It is seen that watermarked images still have high certainty value after these manipulations. Although in literature, an attack is successful if can remove watermark signal or disable the detector to find the watermark signal without any perceptual damage on the image, and none of tested attacks was successful from this point. But as a matter of fact, the robustness of some watermarked images based on this scheme is affected significantly by some geometrical attacks. For instance, at small random distortion attack, although the image get distorted dramatically, but it has very low certainty value. This issue can be considered as a weakness besides its advantages for this scheme. Finally, with regard to other related experiments, none of the proposed watermarking techniques are fully robust against geometrical distortions.

REFERENCES

- **1. Nidhi K., Anıl K.,** "*Digital Watermarking – A Solution for Copyright Protection of Multimedia Data*", IJCSMS International Journal of Computer Science and Management Studies, vol. 12, no.02, pp.151-154.
- **2. Ingemar J., Matthew L., Jeffrey A., Jessica F., Ton K., (2008),** *"Digital Watermarking and Steganography"*, Second Edition, Elsevier Publisher, USA, pp.33-256.
- **3. Tolga G., (2005),** *"Template Based Image Watermarking in The Fractional Fourier Domain"*, A Thesis Submitted to the Graduate School of Natural and Applied Sciences Of Middle East Technical University, in Electrical and Electronics Engineering ,pp.151-154.
- **4. Cayre F., Fontaine C., Furon T., (2005),** *"Watermarking Security: Theory and Practice,"* IEEE Transmission on Signal Processing. vol.53, no.10, pp.3976-3987.
- **5. Anand B., Nikhil D., Aditya B.,** *"Competitive Analysis of Digital Image Watermarking Techniques"*, International Journal of Recent Technology and Engineering (IJRTE), vol.1, no.2, pp. 198-200.
- **6. Chaw-Seng W., (2007),** "Digital Image Watermarking Methods for Copyright Protection and Authentication", A PhD Thesis Submitted to the Information Security Institute Faculty of Information Technology, Queensland University of Technology, pp.11-27.
- **7. Osman E., (2006),** *" Quantization Index Modulation Based Watermarking Using Digital Holography"*, A Master Thesis Submitted to the Graduate School Of Natural and Applied Sciences of Middle East Technical University, in Electrical and Electronics Engineering, pp. 2-16.
- **8. Sal_H erfen Balci, (2003),** *"Robust Watermarking of Images"*, A Master Thesis Submitted to the Graduate School of Natural and Applied Sciences of Middle East Technical University, in Electrical and Electronics Engineering, pp.17-31.
- **9. Kilburn D., Linen D., (1997),** "*Dark Secrets*". Adweek Publisher, USA, pp. 10-15.
- **10. Ingemar J. C., Matt L. M. and Jeffrey A. B., (2000),** *"The Properties and Applications of Watermarkin"* Published in the International Conference on Information Technology, Las Vegas, pp. 7-9.
- **11. Anand B., Nikhil D., Aditya B., (2012),** *"Competitive Analysis of Digital Image Watermarking Techniques"*, International Journal of Recent Technology and Engineering (IJRTE), vol.1, pp.3-5.
- **12. Ramashri T., Narayana R. S., (2009),** *"Robust Image Watermarking Algorithm Using Decimal Sequences"*, International Journal of Wireless Networks and Communications, vol. 1, pp. 1-8.
- **13. Vidyasagar M., Song H., Elizabeth C., (2005),** *"A Survey of Digital Image Watermarking Techniques"*, Proceeding IEEE International Conference on Industrial Informatics (INDIN), pp.2-6.
- **14. Sutaone M. S., Khandare M.V., (2008),** *"Image Based Steganography Using LSB Insertion Technique"*, IEEE WMMN, pp. 146-151.
- **15. François C., Caroline F., Teddy F., (2005),** "*Watermarking Security: Theory and Practice"*, IEEE transactions on signal processing, vol.53, no.10, pp. 5-7.
- **16. Shannon C. E., (1958),** *"Channels With Side Information at the Transmitter"*, IBM journal of research and development, pp. 289-293.
- **17. Shelby P., Sviatoslav V., Maribel M., Stephane M., Thierry P.,(2001),** *"Second Generation Benchmarking and Application Oriented Evaluation"*, In Information Hiding Workshop 3, Pittsburgh, PA, USA, pp. 1-16.
- **18. Jidong Z., (2006),** *"Watermark Embedding and Detection"*, Doctoral Thesis, the Department of Computer Science and Engineering, Shanghai Jiaotong University, pp. 15-32.
- **19. Peter M., (2001),** *"Digital Image Watermarking in the Wavelet Transform Domain"*, A Master Thesis, The Department of Computer Science and Engineering, Salzburg University, pp. 8-22.
- **20. Navnidhi C., Basha S. J., (2012),** *"Comparison of Digital Image Watermarking Methods DWT & DWT-DCT on the Basis of PSNR"*, International Journal of Innovative Research in Science, Engineering and Technology, vol. 1, no. 2, pp. 1-2.
- **21. Ingemer J., Joe K., Tom L., Talal G., (1997),** *"Secure Spread Spectrum Watermarking for Multimedia"*, IEEE Proceeding International Conference on Image Processing, vol.6, pp 1673-1687.
- **22. Barni M., Podilchuk C., Bartolini F., Delp E., (2001),** *"Watermark Embedding: Hiding a Signal Within a Cover Image",* IEEE Communications Magazine, pp.102-108.
- **23. Ng K. S., Cheng L. M., (1999),** *"Selective Block Assignment Approach for Robust Digital Image Watermarking"*, in Proceeding SPIE/IS&T International Conference Security and Watermarking of Digital Multimedia Contents, vol. 3657, pp. 14-17.
- **24. Depovere G., Kalker T., Linnartz J. P., (1988),** *"Improved Watermark Detection Using Filtering Before Correlation",* Proceeding 5th IEEE International Conference on Image Processing, vol. 1, pp. 5-8.
- **25. Peter M., (2010),** *"Digital Watermark Detection in Visual Multimedia Content"*, A Doctoral Thesis, the Department of Computer Science and Engineering, Salzburg University, pp.26-31.
- **26. Fabien A., (2000),** *"Watermarking Schemes Evaluation"*, IEEE Signal Processing, vol. 17, no. 5, pp. 58–64.
- **27. Fabien A., Ross J., Markus G., (1988),** *"Attacks on Copyright Marking Systems"*, in David Aucsmith (Ed), Information Hiding, Proceedings, LNCS 1525, Springer-Verlag Publisher, pp. 219-239.

APPENDICES A

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Bagheri Baba Ahmadi, Sajjad **Date and Place of Birth:** 18 September 1988, Masjed Solaiman **Marital Status:** Single **Phone:** +90 507 574 7901 **Email:** Sajad_bagheri67@yahoo.com

EDUCATION

FOREIN LANGUAGES

Fluent English, Beginner Turkish

PUBLICATIONS

1. "*Digital Image Watermarking For Intellectual Property Protection"*, International Scientific Conference of Iranian Academics in Turkey, Hacettepe University, vol.4, pp.85-86, (2014).

HOBBIES

Traveling and Swimming.