The ISC Int'l Journal of Information Security

ISeCure

January 2013, Volume 5, Number 1 (pp. 5–36)

http://www.isecure-journal.org

Invited Paper

A Survey on Digital Data Hiding Schemes: Principals, Algorithms, and Applications

Mohammad Ali Akhaee [1,*], Farokh Marvasti [2,3]

[1] *Department of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran*
[2] *Electrical Engineering Department, Sharif University of Technology, Tehran, Iran*
[3] *Advanced Communications Research Institute, Sharif University of Technology, Tehran, Iran*

ARTICLE INFO.

Article history:
Received: 14 September 2013
Revised: 15 October 2013
Accepted: 15 October 2013
Published Online: 27 November 2013

Keywords:
Data hiding, Watermarking, Capacity, Robustness, Security, Steganalysis.

ABSTRACT

This paper investigates digital data hiding schemes. The concept of information hiding will be explained at first, and its traits, requirements, and applications will be described subsequently. In order to design a digital data hiding system, one should first become familiar with the concepts and criteria of information hiding. Having knowledge about the host signal, which may be audio, image, or video and the final receiver, which is Human Auditory System (HAS) or Human Visual System (HVS), is also beneficial. For the speech/audio case, HAS will be briefly reviewed to find out how to make the most of its weaknesses for embedding as much data as possible. The same discussion also holds for the image watermarking. Although several audio and image data hiding schemes have been proposed so far, they can be divided into a few categories. Hence, conventional schemes along with their recently published extensions are introduced. Besides, a general comparison is made among these methods leading researchers/designers to choose the appropriate schemes based on their applications. Regarding the old scenario of the prisoner-warden and the evil intention of the warden to eavesdrop and/or destroy the data that Alice sends to Bob, there are both intentional and unintentional attacks to digital information hiding systems, which have the same effect based on our definition. These attacks can also be considered for testing the performance or benchmarking, of the watermarking algorithm. They are also known as steganalysis methods which will be discussed at the end of the paper.

1 Introduction

The problem of embedding hidden messages has a history of thousands of years. By the development of the Internet and easy transmission of multimedia products, digital data hiding was reborn and became one of the hottest research topics all over the world during the late 90s and the first decade of the 21st century. Among various fields of this topic, information hiding in images gained more attention. However, the significance of audio packets and their transmission gradually increased the importance of audio watermarking

* Corresponding author.
Email addresses: akhaee@ut.ac.ir (M.A. Akhaee), marvasti@sharif.edu (F. Marvasti)
ISSN: 2008-2045

[1–3]. Owing to the interdisciplinary nature of the subject, research enthusiasts of signal processing, image, audio, and video processing, computer science, and even applied mathematics were attracted to this field. Early methods were proposed heuristically, but the theoretical aspects of the schemes enhanced with time and some data hiding articles were published in the journals of Information Theory. It is worth noting that data hiding systems improved very fast, which was mainly due to employing existing experiences in the field of data compression as well as using fundamentals of communications and coding theory [4, 5].

In fact, steganography or watermarking is embedding an amount of data, called secret message or watermark, into a cover medium, which may be audio, speech, image or video signal in an imperceptible way. In watermarking systems the focus is on the cover signal while in the steganography attention is paid to the secret message and the cover is just regarded as a career. The watermarking and steganography together form the information hiding. It can be easily inferred from the definition that imperceptibility is the first principle of any data hiding systems. Considering the requirements of an information hiding system, five significant features can be listed [2]. There are also other assortments, which may be found in some books [5]. These five principal traits are transparency, robustness, capacity, security, and implementation complexity that will be explained in detail in the next section. From these mentioned features, the first three ones are more important and their trade-offs have been examined from information theoretical point of view. Thus, it is not possible to improve all of them simultaneously. That is, to increase the transparency and robustness, the embedding rate should decrease and vice versa. In some papers, these three attributes have been considered as the three vertices of a triangle [4].

Most of the materials explained above are true for all signals. In order to design for instance an audio watermarking system, besides general information about data hiding, knowledge on audio and speech signals as well as theirbiological models are required. The same is true for the image watermarking and the knowledge about human visual system (HVS). In other words, a watermarking scheme will be efficient only if it is designed with human auditory system (HAS) or HVS in view. Since HAS is more sensitive than HVS, audio watermarking poses more challenges. In the next sections, we will briefly explain HAS and HVS. For more explanation about HVS, the readers may refer to [6–9].

Next, we will have a general review on digital data hiding papers and divide the methods into seven categories based on their similarities. The pioneer scheme of each class together with its improved descendants will be examined, which includes the main idea, general advantages and disadvantages of each method. Apparently, several improved descendants have been obtained by applying the original scheme to various transform domains such as Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Cepstrum. Transform domain methods are usually more robust at the expense of more computational complexity. Robustness is the second principal, after imperceptibility, in data hiding systems. We consider any modification or intrusion to a watermarked/stego signal as an attack, which may be unintentionally made by the channel or intentionally applied to the signal by an attacker. It should be noted that in our literature, attack is referred to modifications which are not much perceptible; in fact, alterations which destroy the signal will also demolish the watermark. The latter is not the focus of our discussion. Anyways, a robust method should have sufficient resistance against attacks. On the other hand, attacks can be employed to assay the performance of data hiding systems. To this end, a combination of attacks can be designed to benchmark the proposed schemes in a standard way. Those intentional attacks that aim to examine the existence of the secret message in the stego signal are called steganalysis which will be discussed in a separate section.

The rest of the paper is organized as follows. Section 2 includes preliminaries that go over the concepts, definitions and applications of information hiding. The model used for human hearing and vision will be introduced in Section 3. In Section 4, conventional data hiding schemes along with their main characteristics have been proposed. By considering the strength and weak points of each technique, a general comparison of all the techniques are made. The extension of each pioneer method and its state of the art version are also introduced. The counter-measure of data steganography schemes called steganalysis has been introduced in Section 5. The generic block diagram of a stegnalyzer will be discussed and explained in this section. Finally, Section 6 finally concludes the paper.

2 Background

In this section, main characteristics of data hiding systems are explained and general applications for these systems, which determine the significance of each requirement, will be expressed. As previously stated, there are four properties of a digital information hiding system, i.e., transparency, robustness, capacity, and computational complexity. More explanation of these attributes are given below.

2.1 Data Hiding Attributes

2.1.1 Transparency

The primary requirement of a data hiding system, as inferred from its definition, is the imperceptibility or transparency of the watermark. In other words, there should not be any perceptible difference between the original and the watermarked/stego signals. Embedding the watermark definitely makes some distortions to the host signal; thus, a measure for evaluating the fidelity of the watermarked/stego signal to the original one is required. As men are the final users of the digital signal, human's ear and eye seem to be the best measure. Assessment of the original and the distorted signals has been considered in digital compression and codec problems. Data hiding and compression are indeed the two sides of the same coin. One tries to compress the signal through removing the redundancy while the other one attempts to embed the watermark by making alterations to the redundancy. In case of audio data hiding, subjective tests score the audio signal between one (extremely annoying) and five (imperceptible). In these tests, audio clips of sufficient duration from various genres are audited by several persons (usually with the ratio of two to one for respectively men and women) and the mean of their reported scores is obtained. This type of test is called Mean Opinion Score (MOS) [10]. The same can be done for image signals. After data embedding, the closer is the score to five; the better is the performance of the algorithm. To be more specific, any score other than five, in case that the host signal is absolutely clean, makes the embedding algorithm unacceptable. Another necessary test is distinguishing the original signal from the distorted one. When the two signals are very similar, which is often the case in data hiding, this test is more revealing. In this assay, clips or images are randomly played and the auditors or visitors should discern the watermarked ones. The closer is the test result to 50%, the more transparent is the data hiding scheme [7–10].

Most subjective tests are costly and time-consuming. Therefore, objective assays are usually employed. The Mean Square Error (MSE) and Signal to Noise Ratio (SNR) are the first objective measures; however, their results are not much consistent with that of subjective tests. As a simple illustration, an audio signal multiplied by -1 has a terrible MSE, while gaining the highest score in MOS test. For this reason, audio and image quality measurements which are highly consistent with HAS and HVS standards have been introduced. Spectral distance, Itakura distance [10], Perceptual Evaluation of Speech Quality (PESQ) for speech signals, and Perceptual Evaluation of Audio Quality (PEAQ) [11], and parameters from Noise to Mask Ratio (NMR) [12] for audio signals and Peak to Signal Ratio (PSNR) and Structural Similarity Index Measure (SSIM) [13, 14] for image signals are some instances of these measures. It is worth noting that in data hiding literature, PEAQ, NMR, PSNR and SSIM have been more frequently used. Below is a brief explanation of these standards.

PEAQ: In order to make the scoring procedure easy, an algorithm has been proposed which assigns a score between -4 and zero based on appropriate features extracted from the original and the distorted signals. The output is in fact the difference between the MOS's. Apparently, the closer the PAEQ value to zero is; the more transparent is the watermark. PEAQ is a freeware downloadable from [11]. More information about the feature selection and the scoring methods could be found in [15].

NMR$_{total}$: This quantity, presented in dB, is actually the average energy ratio of the difference signal related to the signal that is just masked. The less the value of NMR$_{total}$ is; the better is the quality of the distorted signal. For data hiding systems, values less than -8 dB are acceptable [12].

PSNR: This quantity is frequently used for evaluating the performance of image data hiding systems, and is the ratio between the power of an image with maximum allowable pixel intensity (255 for 8-bit images) to the power of the noise. The noise power is defined as the power the difference between original and watermarked images.

SSIM: If we denote the original and the manipulated watermarked images with o and m respectively, the structural similarity index metric (SSIM) Q is defined as: [13, 14]

$$Q = \frac{(2\hat{o}\hat{m} + C_1)(2\sigma_{om} + C_2)}{(\hat{o}^2 + \hat{m}^2 + C_1)(\sigma_o^2 + \sigma_m^2 + C_2)} \qquad (1)$$

where $\hat{o}$, $\hat{m}$, σ_o^2 and σ_m^2 are the mean values and variances of o and m respectively, and

$$Q = \frac{1}{N-1}\sum_{i=1}^{N}(o_i - \hat{o})(m_i - \hat{m}) \qquad (2)$$

where N is the number of samples. Moreover, parameters C_1 and C_2 are defined as:

$$C_1 = (K_1 l)^2, C_2 = (K_2 l)^2 \qquad (3)$$

where l is the dynamic range of the pixel values (255 for 8-bit grayscale images) and K1,K21 are small constants. The best value for Q is achieved if and only ifm_i=o_i for all i=1,2,,N. Since image signals are generally non-stationary and image quality is often spatially

varying, it is reasonable to measure statistical features locally and then combine them together. For instance, it is recommended to apply the quality measurement to non-overlapping BB block segments of the image, calculate a local index Q_j for each block, and find the overall quality index by arithmetic averaging over Q_j's from all blocks [13, 14].

2.1.2 Robustness

Robustness is the level of the watermark's resistance against modifications imposed on the watermarked/stego signal. It is important to retrieve the watermark or the secret message, intact as much as possible. A watermarked/stego signal may experience some intentional or unintentional modifications while transmitted over the channel. For example, a watermarked audio signal passing through a phone line (PCM channel) absorbs echo and some noise; or an audio CD given to the customer may undergo processing algorithms such as filtering, echo addition, warping, etc. Evaluating the robustness of a watermarking system is similar to the procedure we have in communication systems. That is, the bits of the message are embedded and the watermarked signal is sent over the channel; at the receiver end, the signal is examined to find out to what extent the message is retrievable. This reliability level is usually measured by Bit Error Rate (BER) or Message Error Rate (MER).

It should be pointed out that based on the application, e.g., copyright protection or for stgonagraphy, there exist two sorts of embedded information [16]. An embedded data belonging to the first category, which is used for authentication, serves as the authentication code (watermarking), while that of the second category serves as the transmission code (steganography). For the second type of the secret message, the performance is evaluated through BER. For the first type, the output of the decoder is the presence or absence of the watermark, and as in hypothesis test, there exist two kinds of error: the probability of lost watermark, i.e., not detecting the presence of the watermark while it is present, and the probability of false alarm, i.e., falsely detecting the presence of the watermark while it is absent. The former error is shown by Pmissed and Pfa presents the latter one.

A system may contain both kinds of watermarks. As a case in point, a broadcast monitoring system requires synchronization to discern watermarked frames; that is, the beginning and the end of the watermarked frames have to be specified. This is done using synchronization codes embedded in the signal. In this case, the decoder must detect the synchronization codes a priori in order to recognize the watermarked frames, and then, decode the frames. If the decoder misses a synchronization code, it will miss some bits of the secret message consequently. Similar situation is true for false alarm. In such applications, Pmissed and Pfa have to be very small (of the order of 10^{-5}) [17].

Another point is that in some usages, the purpose of data hiding is to distinguish any modification made to the watermarked signal. Such applications, usually belonging to signal authentication category, require another kind of watermarking which not only is not robust, but is also very sensitive to any little alteration. This sort of watermarking is called fragile watermarking [18]. In this case, the algorithm is designed in such a way that any intrusion breaks down the whole watermark and transforms it into some random bits, exactly like what happens to an encrypted signal. The fragility level can be adjusted in such a way that the watermark is robust against specific alterations (which are not intentional and take place throughout the channel) and is broken down by other kind of modifications. These algorithms are called semi-fragile [19, 20].

Sometimes both applications of watermarking, i.e., robustness against attacks and confidence that the watermarked signal has not been tampered with are required. In this situation, multi-purpose watermarking is employed [21], in which a robust and a fragile watermarking are perpendicularly and simultaneously embedded in the host signal. For further information about fragile and semi-fragile watermarking, refer to [22–27].

2.1.3 Capacity

Another significant factor in digital data hiding is capacity. Capacity is usually expressed in bits per pixel in images or bits per second and sometimes bits per sample in audio watermarking. Data hiding algorithms with high capacity and low robustness are called steganography techniques, while the general term of watermarking usually refers to a low-capacity robust data hiding scheme. There exist many image steganography algorithms with adaptive varying capacities [6]. But due to the more sensitivity of HAS compared to HVS, and because of one-dimensionality of audio signals, embedding rate in these signals is less than that of images and embedding information with high bit rate is a challenging problem. For example, in [28], a method with high capacity and acceptable robustness against noise has been proposed. Moreover, in [29], a speech data hiding scheme with very high capacity but low robustness has been introduced. Data insertion in this method has been done on unvoiced part of the speech signal. Pixel Value Difference (PVD) methods are examples from high capacity image steganography algorithms. These schemes aim to embed most of the information, by hiding more bits

in the sharper edges of the original image to be compatible with the HVS [30–33]. In [34], an approach is suggested to reduce the capacity in exchange for increasing the robustness of the watermark which is suggested to be used for transmitting the flight information within aviator's conversation with a watchtower. In the field of image steganography, the Least Significant Bit (LSB) algorithms which offer moderate capacity of one bit per pixel via modifying the pixel value by one have been thoroughly discussed [35–37]. In most applications of audio watermarking, where the purpose is not covert communication, the common rate is about a few bits per second, taking high robustness against attacks into consideration [38]. In speech signals, because of successive periods of silence, this rate decreases substantially and sometimes becomes one bit per second. In image signals, several methods have been proposed to efficiently reduce the embedding capacity in exchange of robustness. The matrix embedding algorithms such as F5 [39] and nsF5 [40] are examples from this group which can offer an almost unchanged image where the embedding rate is very limited.

2.1.4 Security

The matter of security was at first a challenge among researchers entering data hiding from the area of cryptography or other fields of study. At the first glance, according to Kerckhoffs' principle, security should be independent of the algorithm and has to be based on the secret key. In this condition, embedding and extracting algorithms can be publicly known. From another point of view, data hiding was considered as an art, the importance of which was lying in its covertness and not catching attentions of others. For this criterion, only the sender and the receiver are aware of the algorithm. In general, it can be said that a data hiding scheme is secure if the content of the watermark cannot be detected within a short time if the embedding algorithm is known. To this end, watermarks are usually embedded and extracted based on a key and the malicious attacker will extract a set of random bits if he does not have the key. From another viewpoint, which is mostly discussed in covert communication and steganography, security means that steganalysis methods cannot detect the existence of the secret message in the host signal. In this case, the algorithm can be publicly known or unknown; in either case, a steganalysis scheme should not be able to distinguish a stego signal from a clean one. As the last point, note that to increase the security, the secret message is usually encrypted before being embedded into the host signal [6].

2.1.5 Computational Complexity

The last criterion in digital data hiding is the computational complexity of the algorithm and its implementation on different hardware. The importance of this factor is highly dependent upon the application. As an illustration, in watermarking applications such as copyright protection, the complexity of embedding and extracting algorithms does not matter much, whereas in steganography applications such as data transmission or broadcast monitoring, real time implementation is of high significance. In digital data hiding, it is important that on what hardware with how much processing power the algorithm is to be implemented. Furthermore, transformations should be carefully employed, since in some usages requiring synchronization (broadcast monitoring) the load of applying transformation has to be performed sample by sample, which results in a high processing volume [17].

2.2 Data Hiding Applications

After stating the requirements and attributes of digital data hiding, it is time to discuss its applications. As mentioned before, by the development of the internet and digital broadcast, data hiding has become popular. One of the most important topics in this field is copyright protection, supporting authors and, the producers of multimedia products. Before we enter the details of data hiding applications, we should remind again that the significance of the mentioned requirements is determined by the application. In other words, except for transparency, which has to be sufficiently valued in all applications, the weight of other criteria is dependent on the application. For instance, in watermarking applications such as copyright protection, capacity is not that important and embedding an identification number or a logo seems enough recognition. In addition, the level of complexity, in either embedding or detection stages, is insignificant and the procedure may take several hours. Furthermore, robustness against processing manipulations is very important. In contrast, for steganography applications such as covert data transmission, capacity is very significant. Low complexity is also vital in cases where online communication is necessary; however, robustness is unimportant, because the embedded data is not supposed to confront anything other than the channel.

2.2.1 Copyright Protection

One of the most substantial applications of data hiding is copyright protection of digital products. For this purpose, the name and the information of the company as well as the identification number or logo

are embedded in different parts of the digital signal. In this usage, the algorithm is required to be robust against all intentional and unintentional attacks and the watermark has to be detectable with acceptable accuracy. The logo and other information of the company should have been previously registered in a judicial bar so that the probable criminal can be officially accused when the copyright is violated. In copyright protection, the watermark must be embedded in such a way that there remains no space for extra embedding. In other words, all available capacity should be exploited so that embedding another watermark by the probable attacker leads into the destruction of the digital signal. Moreover, to prevent collusion attack, using a constant watermark, especially one which is independent of the host signal, must be avoided [6].

2.2.2 Authentication and Tamper Protection

In some applications, we need to be sure that the received audio or image signal is from a specific person. That is, we want to be certain of the senders authenticity. For this criterion, a pre-determined watermark (watchword) can assure the receiver from the senders authenticity. For example, suppose that some radios wish to talk and another radio tries to speak on the same frequency while disguising itself as a member of the friends group. Allocating a set of watermarks to the groups members, the intruder will easily be recognized. In this usage, robustness against the channel, including compressor, channel fading, noise, andanalog to digital convertor. along with instantaneous implementation of the algorithm on the radios processor are highly important. In another application, the purpose of data hiding is to prevent any tamper to the transmitted audio signal. In this case, not only the watermark should not be robust, but also it has to be fragile so that any alteration becomes completely evident. In some usages, it is necessary that modifications caused by the channel are distinguished from intentional intrusions and forgeries. In this situation, the watermark is embedded in a way that is not fragile against channel attacks. This type of data hiding requires high capacity while security does not matter [6].

2.2.3 Image Tampering Protection and Self-Recovery

Although several primary data hiding algorithms were designed to detect and locate the image tampering, yet recovering the original data in the tampered area is a very recent trend of image data hiding. These self-recovery techniques aim to accomplish the task of integrity verification, tampering localization and data recovery using only a single watermark. In order to fulfill this purpose, the self-recovery watermarking algorithms usually conceal a representation of the original image into itself. This watermark must be fragile to enable the receiver to detect any malicious modification when its integrity is violated. Moreover, the extracted watermark helps the receiver to retrieve the original image data in the lost area to an extent depending on the amount of manipulation. The watermark bits used for the integrity verification are called check bits, while the reference bits are exploited to recover the lost data. In [41], DCT coefficient or reduced color-depth version of the host image are embedded in the LSB of the original image. This representation of the original image can also be the first few DCT coefficients of each block [42], a binary image generated from the difference between the host image and its chaotic pattern [43], hash function of the original image [44], watermark derived from approximation subband coefficients of wavelet transform [45], a vector quantized [46] or halftone [47] version of the original image.

The problem of image self-recovery is about finding appropriate trade-offs among these three parameters: watermarked image quality, content recovery quality, and tolerable tampering rate (TTR). The size of watermark determines the amount of imposed distortion and quality of watermarked image. On the other hand, more watermark bits are required to achieve higher TTR or better quality in the recovered area. As examples for this tradeoff, some methods provide almost error free restoration at the expense of very limited TTR [48–50] or very low quality of watermarked image [49]. On the other side, some techniques sacrifice the restoration quality to deal with high tampering rate [51, 52]. Content adaptive methods have been recently proposed to compromise between TTR and restoration quality based on required application [53, 54].

Very recently in this field, the compressive sensing techniques are used to retrieve a representation of the original image from the embedded watermark which is lost to some extent due to the image tampering [55]. In [56], fountain codes [57] are applied to protect the watermark against the image tampering. Finally in [58], a set partitioning in hierarchical trees (SPIHT) [59] is used to generate a compressed version of the original image. After being protected by Reed-Solomon (RS) channel codes [60] of long blocks and over the large fields, this compressed version is embedded in the original image itself.

2.2.4 Finger Printing and Traitor Tracing

In some applications, we need to trace the digital product. In this condition, a particular watermark (for instance, the customers identifier) is embedded in the

signal. As a result, whenever an unauthorized copy of the product is detected, it can be easily realized who has done it. Here, the bit rate is not of much importance, but rather the robustness against attacks, especially collusion attack is substantial. In this kind of attacks, a group of,K, colluders, who have digital products of the same kind but with different watermarks, try to obtain an average of the K products. In this way, the watermark is almost removed and the person or persons who start the unauthorized copy procedure cannot be detected. Finding algorithms robust against collusion attack (at least for up to a determined number of colluders) is a hot and interesting topic in data hiding. The general idea of such algorithms is to prepare different positions for the watermark and to embed the watermark in only some of them (similar to PPM modulation) [6].

2.2.5 Broadcast Monitoring

An important application of digital data hiding is broadcast monitoring. It happens many times that we need to implement a new service on an old infrastructure (for instance an analog one) without modifying its basis. In this case, data hiding becomes helpful. Since it does not take any particular bandwidth, there exists no need to change the facilities, and most importantly, the quality of the signal is not affected. One of such cases is broadcast monitoring, which is employed by people who order radio/TV commercials, and makes it possible to automatically monitor and control the broadcasting of commercials. In this application, a specific identification code is embedded in the first (sometimes the last) second of the commercial; broadcasted radio programs are automatically scanned and by detecting the watermark, the number of times that particular commercial is broadcasted, as well as the time and the duration of those broadcasts can be determined. In another usage, the lyrics of the music or the subtitle of the movie can be broadcasted along with the music/video itself so that the audience can watch the lyrics or read the subtitle, some information about the composer, singer, director, and even some commercials on the monitor of their radios/TVs. Using this technology, we can also obtain some statistics on the listening rate of specific programs [6].

2.2.6 Other Applications

Generally, it is possible to embed some side information in a digital signal, which may be later used. For example, watermarks containing the production number, time, etc., are useful in constructing and sorting a database [61]. In another interesting application, which has been recently proposed, the music is recorded in stereo format and the information of one channel is compressed and embedded in the audio of the other channel. Wherever stereo playing is possible, the information of the second channel is extracted from the existing audio and after decoding, it plays simultaneously with the first audio (in stereo format). Note that this usage is still at research stage [62]. Another interesting application of the data hiding has been found in the flight control. In this case, the watchtower conceals the flight information into the seemingly-normal conversation to the aviator. In this way, the secret flight information is unnoticed or inaccessible to the eavesdropper who wants to illegally receive this information [34].

3 Human Perception Modeling

Digital data hiding is in fact manipulating a cover signal such as adding pseudo random noise for data embedding. The major principle in data hiding is imperceptibility of the watermark. In order to satisfy the transparency condition and also to have the most achievable power of watermark (for the sake of robustness), it is necessary to know human hearing and visual structures.

3.1 Human Visual System

The HVS is too complex to be understood from all aspects. The main interesting aspect of HVS in data hiding is to find out whether or not a certain stimulus is perceivable by HVS. As a generalization of the regular observations, three rules of thumb are found for HVS. First of all, it turns out that the highly textured regions are the most suitable area of the images for the watermarking purposes. The second rule states that data embedding is much less perceivable on edges than on flat areas. The third rule suggests that the very bright or very dark areas of the image as the proper candidates for data embedding.

Since every stimulus can be decomposed to a sum of sinusoidal stimuli, we are interested to investigate how sinusoidal stimuli are perceived. This quality is measured by just noticeable difference (contrast), or simply JND. Suppose that a uniform background of luminance L_0 is superimposed by a sinusoidal stimulus of spatial frequency v, orientation θ and amplitude ΔL. The spatial luminance of the image can be formulated as:

$$L(x, y) = L_0 + \Delta L \cos(2\pi v(x \cos\theta + y \sin\theta)) \quad (4)$$

Then, the luminance of the sinusoidal is increased until the observer perceives it. The value of the luminance of the sinusoidal stimulus which is just noticeable is named as "just noticeable visibility threshold" (ΔL_{jn}). To achieve a measurement independent of the

viewing distance, the angular frequency $f = \frac{\pi d}{180} v$ is defined which is expressed in cycles/degrees, where d is the distance between the observer and the image. The just noticeable difference or contrast is obtained from just noticeable threshold according to the following equation:

$$LC_{jn} = \frac{\Delta L_{jn}}{L_0} \tag{5}$$

The application of the inverse of JND is also common and is called Contrast Sensitivity Function (CSF) denoted by S_c. There are many analytical expressions in the literature for the CSF; the one obtained by Barten [6] is one of the most widely used:

$$S_c(f, L_0, W, \theta) = \\ a(f, L_0, W) f e^{-r(\theta) b(L_0) f} \sqrt{1 + 0.06 e^{b(L_0) f}} \tag{6}$$

where

$$a(f, L_0, W) = \frac{540(1 + \frac{0.7}{L_0})^{-0.2}}{1 + \frac{12}{W(1+\frac{f}{3})^2}} \tag{7}$$

$$b(L_0) = 0.3(1 + \frac{100}{L_0})^{0.15}$$

$$\Gamma(\theta) = 1.08 - 0.08 \cos 4\theta$$

where f measured in cycles/degree, is the angular frequency of the stimulus, W is the observer viewing angle in degree, L_0 is the mean local background luminance in $\frac{cd}{m^2}$, θ in radian is the orientation of the stimulus.

In Figure 1, the plots of the CSF with respect the angular frequency are presented for a horizontal stimulus and for an observer viewing angle $W = \frac{180}{(\pi\sqrt{12})^5}$. The eye exhibits the maximum sensitivity in the middle range of angular frequencies. It is also perceived that the stimuli of higher frequencies are less perceivable by HVS [6].

HVS modeling and JND information can be exploited for the sake of image data hiding. As an example, discrete cosine transform decomposes the image into a summation of sinusoidal signals. Therefore, considering the JND information for all spatial frequencies, the amount of modification tolerable by each DCT coefficient can be determined. One can guarantee an imperceptible data embedding with the highest possible rate; as long as the modification is below the JND for each DCT coefficient [63].

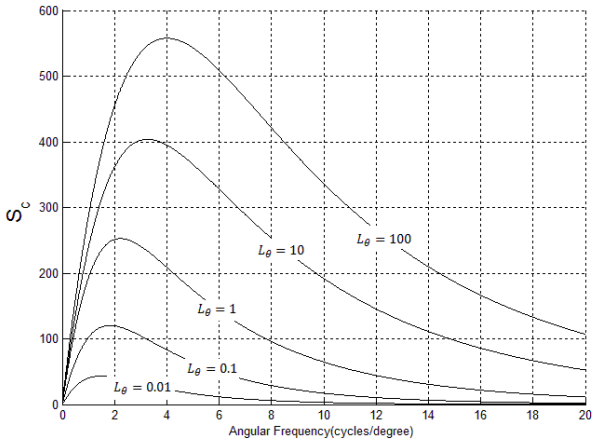

Figure 1. Plots of the CSF with respect to the angular frequency for values of background luminance [6]

3.2 Human Auditory System

It is worth mentioning that audio data hiding is much more challenging than image data hiding due to the vast dynamic range of the HAS in terms of power and frequency. A human being is able to hear signals of power from 1 to 109 and of frequency ranges from 50 Hz to 20 KHz. A little amount of white Gaussian noise (SNR=60 dB) is quite audible which makes data embedding so difficult. On the other hand, the HAS is not much sensitive to the phase and if the phase continuity is maintained, its alteration will not be noticed. An illustration is an audio signal multiplied by -1 (phase change of 180?) which has no effect on hearing. Moreover, the human auditory system has a narrow differential range that masks low volume sounds by the loud ones. Furthermore, some kind of distortions sound natural to the HAS, owing to its internal structure [2].

Investigations around the HAS indicates that human does not hear all frequencies in the same way. As a matter of fact, it operates like a filter bank. A group of bands towards which the HAS has the same sensitivity are called critical bands. The bandwidth of these critical bands is about 100 Hz for low frequencies and up to 5 KHz for high frequencies. Consequently, the human auditory system can be considered as a non-uniform filter bank. Generally, the whole frequency range is divided into 26 bands, named bark. The transformation from the frequency domain to the bark domain is nonlinear and is modeled as follows [10]:

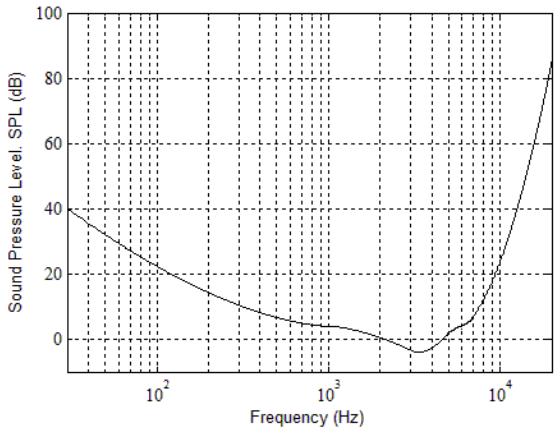

Figure 2. The sound pressure level in silence

$$Z(f) = 13arctan(7.6 * 10^{-4} f) + 3.5arctan((\frac{f}{7.5 * 10^3})^2) \quad (Bark) \tag{8}$$

Another substantial parameter is the Sound Pressure Level (SPL), i.e., the level which the HAS can hear in silence. This threshold depends on the frequency. Figure 2 shows this threshold. As shown, the hearing sensitivity in the middle frequencies (2-4 KHz) is more. Besides SPL, there is another concept, called masking, which is used in speech coding systems [64]. This phenomenon occurs when a low level sound is not heard in the presence of a stronger sound. There exists a threshold for each audio signal up to which it is inaudible in the presence of another signal. There are two sorts of masking, the temporal masking and the frequency ones, which are frequently exploited in data hiding systems. Below is a brief explanation for these types of masking.

3.2.1 Temporal Masking

Typically, in the HAS, a more powerful signal (masker) can make a weaker signal (maskee) not to be heard. In one case, the weaker signal is masked before the powerful signal takes place (pre-masking) and in the other case, the weaker signal is masked after the occurrence of the stronger signal (post-masking). This phenomenon is often exploited in MP3 compression. Whenever the Sound Pressure Level (SPL) of the maskee is below the threshold depicted in Figure 2, no sound is heard. According to this Figure, post-masking has a longer duration than the pre-masking (almost 10 times); thus, a longer period is available for embedding data after a loud sound rather than before it. In audio data hiding, for the sake of transparency, both of these periods should be used.

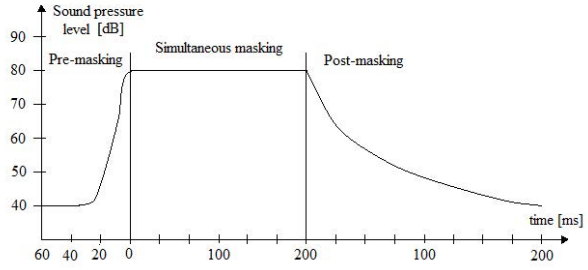

Figure 3. The temporal masking phenomenon (both pre-masking and post-masking)[2]

3.2.2 Frequency Masking

Similar to temporal domain, a stronger signal in the frequency domain (such as a narrow-band noise or a single sinusoidal) can mask a weaker signal within a particular frequency band. Identical to what explained in Figure 2, in the absence of any masker signals, a signal with SPL below the absolute silence threshold (AST) is inaudible. This threshold is variable by the frequency. In the presence of a masker, the threshold increases whose value depends upon the type of the masker (narrow-band noise or single sinusoidal), the SPL of the masker, and the frequency. Figure 3 shows masking threshold for a single tone with SPL of 70 dB. It can be seen that masking threshold in low frequencies decreases more rapidly to reach AST,which implies that HAS is more sensitive in low frequencies. Another point is that the masking power of a single tone is more than a narrow-band noise; consequently, more data (pseudo-noise) can be embedded after a powerful sinusoidal signal. Generally, masking threshold in the frequency domain, called just noticeable distortion, is created by a set of tones and narrow-band noises of different powers. To calculate this threshold, first, the sum of the masking thresholds of all single tones and pseudo-noises is obtained for different frequencies. The total masking threshold is the sum of the effects of all the maskers and the AST. For more information about the HAS, the reader may refer to [64].

4 Data Hiding Schemes

In this section, well-known data hiding schemes are explained, including the LSB coding, Quantization Index Modulation (QIM), patchwork, phase coding, echo hiding, and spread spectrum watermarking. Among these schemes, phase coding and echo hiding are used only in audio data hiding, whereas the others are ex-

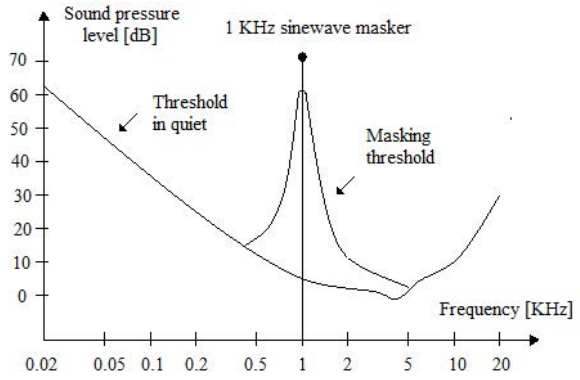

Figure 4. The frequency masking phenomenon [2]

ploited for both audio and image signals. It is worth noting that for all the above-mentioned methods, if the HAS or HVS is appropriately considered in data hiding and the most possible amount of data is embedded, the watermarked signal will be robust against common attacks, while remaining transparent. Another point is the difference between the watermark decoding and the watermark detection. Based on the application, sometimes it is needed to decode the watermark bit by bit to obtain the ultimate data; while in other cases, it is enough to verify the existence of the watermark. In this situation, observing the existence of the logo or the synchronization code suffices and there is no need to decode binary stream.

Here, we only describe the main idea of embedding procedures and detection methods. To learn more about the schemes or to investigate the performance of the algorithms, readers may refer to the appropriate references. After reviewing different techniques in this section, there exists a basic comparison of the introduced methods. At the end, multiplicative watermarking algorithms in particular are discussed as an example of very efficient data hiding techniques.

4.1 LSB Coding

This method [65] is one of the first and simplest information hiding algorithms, which is employed for various types of signals including audio, image, video signals. In this scheme, the LSB of each sample of the host signal is replaced by one bit of the secret massage. For instance, when the host signal is audio, the LSB of each sample of the audio, which has been digitized with for example 16 bits, is replaced with one bit of the secret binary sequence. This algorithm does not make use of HAS or HVS to form the stego signal. The detection is blind and is performed by reading the LSB or LSBs of the received signal samples. It requires precise synchronization of the watermarked signal. Image watermarking by LSB coding is performed in a similar fashion.

Using this method, it is possible to embed a large amount of data; however, it is not robust against various attacks. By repeating the watermark and consequently, reducing the embedding rate, robustness can be increased to some extent. Owing to its simplicity and high embedding rate, this type of data hiding algorithms has become popular. Some versions of LSB coding have been applied to transform domains such as DFT, DCT, and DWT [66]. The most significant one is JSteg, which embeds the bits in the JPEG coefficients of the image.

By the extension of LSB-based schemes, the methods analyzing them have been improved, as well. Most of them have the LSB drawback which makes each pair of adjacent values have close occurrence frequencies in the stego signal [67, 68]. In order to tackle this problem, several versions of JSteg algorithm, including JPHide, F5, and nsF5, were proposed [69–71]. Another improvement to LSB method are approaches to reduce the embedding distortion [35, 36, 72]. In [35], instead of embedding one bit in the LSB of one sample, two bits are embedded in the LSBs of two samples, where the replacement of the second bit depends on whether the first bit is 0 or 1. In this way, the probability of changing the LSB of each sample decreases from 50% to 37.5%, which results in less distortion and makes the analysis harder. In another algorithm, the replacement is performed on a set of samples with module four. This also leads to less watermarking distortion [72]. The probability of change per sample further decreased to 33% in one-third probability embedding [36]. In 2009, Li *et al.*, proposed a generalized version of the LSB algorithm [37]. It was shown that working on the groups of two or three pixels can decrease the percent of changed pixels to 37.5% and 33%. In this Generalized LSB Matching Method (GLSBM), it was shown that increasing the size of pixels used as a group for embedding algorithm, the probability of changing the LSB of each sample can be further decreased up to about 22%.

The embedding rate is not limited to one bit per pixel, unlike the above examples, for all the methods categorized as LSB coding. The capacity of LSB coding can be increased by replacing more LSBs of each sample with more secret bits. For instance, suppose that the sample value of seven (0111 in binary) is supposed to carry two secret bits which are both zero. Replacing the last two bits of the sample with secret bits, one gets the output sample value of four (0100 in binary). But one can reduce the distortion in the host sample by changing the sample value to eight (1000 in binary) carrying the same message bits.

Rounding the sample value to the nearest value with LSBs is called the Optimum Pixel Adjustment Process (OPAP) [73, 74].

The OPAP method increases the capacity of the LSB technique, but it does not consider the HVS structure. Human vision is more sensitive to changes in smooth areas of an image thus the edges and the textured regions of image are more suitable to embed secret data. The Pixel Value Difference (PVD) method, embed the capacity of more than one bit per pixel in the host image in textured or edge regions of [30–33]. The criterion for detecting the edges of the host image is to compare the difference of adjacent pair of pixels with a certain threshold. The further the value of the difference between adjacent pairs of pixels, the more secret bits can be embedded.

Another group of recently proposed LSB techniques are those based on Exploiting Modification Direction (EMD) method introduced in 2006 [75]. While working on a group of n pixels in the previously discussed method, all of n pixels could be increased or decreased by one to embed n secret bits. Now suppose that only one pixel can be added or subtracted by one at the most. We have n pixels and two possibilities for changing each one which results in 2n+1different cases. Therefore, a digit from 2n+1 notationary system (from 0 to 2n) can be embedded in a set of n pixels with changing at most one of them. EMD proposes a scheme to realize this purpose. Diamond encoding (DE) is an extension to EMD [76]. DE works only on a pair of pixels and these pixels can be increased or decreased by k at most to embed a digit in 2k2+2k+1 notationary system. Adaptive Pixel Pair Matching (APPM) improves the performance of DE by altering the pixels in pixel pair in an optimum way in terms of MSE [77].There are more proposed algorithms based on EMD to improve either its embedding rate [78, 79] or imperceptibility [80].

4.2 Quantization Index Modulation (QIM)

This method was first proposed by Chen and Wornell [81]. In this scheme, the host signal is quantized with two or more quantizers based on the watermark. Each quantizer has its own index. QIM watermarking can be expressed via the following equation:

$$S(x, m) = Q_m(x) \tag{9}$$

where m is the watermark and Qm(x) is the function which quantizes the original signal regarding the watermark m. As an illustration, when the watermark is binary, 2-level quantizer is employed, which is shown in Figure 5.

Each sample is quantized to one of the crossed or circled values, which respectively implicate a 1 or 0 watermark bit.

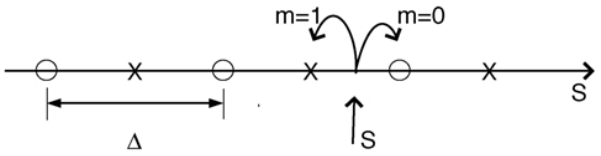

Figure 5. QIM Watermarking

This method, is not appropriate for watermarking, by itself because the attacker can realize the quantization pattern with a little effort. Below we will explain the embedding and the extraction algorithms of a QIM scheme called the Dither Modulation.

4.2.1 Watermark Embedding

Suppose that x is the original signal with N samples. Every L samples will carry one bit of the watermark. To this end, two vectors are defined as follows:

$$d[k,1] = \begin{cases} d[k,0] + \frac{\Delta}{2} & d[k,0] < 0 \\ d[k,0] - \frac{\Delta}{2} & d[k,0] > 0 \end{cases} \tag{10}$$

k=1,2,..,N/L

d[k,0] can be selected as a pseudo-noise with uniform distribution in $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$ where Δ is the quantization step. The length of dither vectors is $\frac{N}{L}$, which is equal to the number of the watermark bits. d[k,0] and d[k,1] are exploited to embed, respectively, zero and one bits in the host signal. The original signal is quantized using the above-mentioned two vectors and the following equation:

$$S(x; m_k) = q_\Delta(x + d[k, m_k]) - d[k, m_k] \tag{11}$$

where $q_\Delta(.)$ is the quantization function with step size of Δ, which is defined as follows:

$$q_\Delta(x) = \Delta * round(\frac{x}{\Delta}) \tag{12}$$

Function round(.) in this equation rounds its argument into the nearest integer.

4.2.2 Detection

In the detector, two vectors $S_0(k)$ and $S_1(k)$ are obtained by embedding respectively zero and one in the received vector $\hat{S}(k)$ and their Euclidean distances from $\hat{S}(k)$ are calculated. The index of the vector with smaller distance is considered as the embedded bit. In case where each bit of the watermark is embedded

in one sample of the host signal (L=1), the following equation states the detection procedure:

$$\hat{m}_k = \arg \min_{i\in\{0,1\}} (\hat{S}(k) - S_i(k))^2 \qquad (13)$$

when the embedder inserts each bit of the message in L samples of the host signal, the detection is based on the sum of the L samples of the vectors $S_0(k), S_1(k)$ and $\hat{S}(k)$; i.e.,

$$\hat{m}_k = \arg \min_{i\in\{0,1\}} \sum_{n=(k-1)L+1}^{kL} (\hat{S}(n) - S_i(n))^2 \qquad (14)$$
$$k = 1, 2, , N/L$$

where $\hat{m}_k$ is the k^{th} extracted bit of the message.

4.2.3 Projection Quantization

Geometric representation of this method is depicted in Figure 6 [81], [82]. Suppose that $\tilde{v} = [\tilde{v}_1, , \tilde{v}_N]^T$ is the vector in which the watermark has been embedded and v is the vector of the original signal. The relation between $\tilde{v}$ and v is expressed as:

$$\tilde{v} = v - Pv + B\Delta Q_i(\frac{B^H v}{\Delta})) \qquad (15)$$

In this expression, $Q_i(p) = \Delta Q_i(B^H v/\Delta)$ and is equal to $p+ \in_p$,where $\in_p$ is a random variable uniformly distributed within $\lfloor -\Delta, \Delta \rfloor$. i is the watermarking index and Q_0 and Q_1 are even and odd quantizers. B is the unitary projection vector and P is the projection matrix for the subspace of B. Inserting $p+ \in_p$ into (15), we will have:

$$\tilde{v} = (I - P)v + B(p+ \in_p) = v+ \in_p B \qquad (16)$$

As implicated in Figure 6, in order to detect the watermark, the projection procedure is applied on the received vector and the embedded bits are extracted in a way similar to (14).

This block method has priority over the one-sample scheme, since in the former method; the distortion is distributed among the samples and becomes imperceptible. Moreover, it gets difficult for the attacker to recognize the quantization pattern in the samples of the signal. Indeed, the projection vector plays the role of secret key here.

QIM scheme, especially its distortion compensated version [83], is one of the most popular watermarking methods. It has been proved in [84] that dither quantization can achieve the capacity by employing lattice

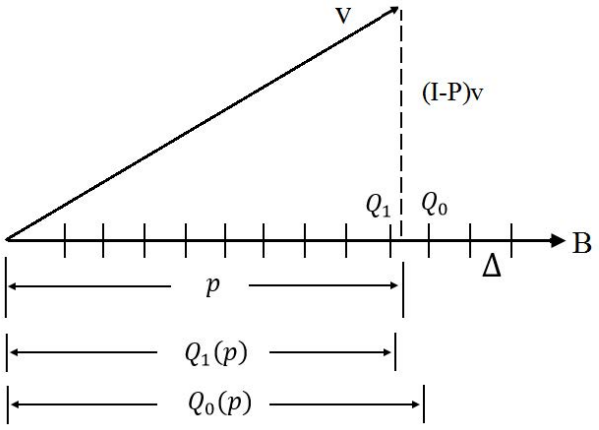

Figure 6. Geometric representation of projection quantization method

quantization in higher dimensions [84–87]. Furthermore, QIM method can be applied to the angle of the vectors [88], or dithers adapted to the host signal can be exploited [89]. A weakness of QIM scheme is its vulnerability to gain attack, which may occur in communication channels. Several solutions have been proposed to tackle this problem; some of which use pilot signal [90], employing conical codes with similarity-based receivers [91], and exploiting Rational Dither Modulation (RDM) [92]. There are several versions of RDM Scheme. For instance, by combining channel coding methods and choosing appropriate code words, one may achieve very good robustness [93]. In addition, by properly applying this scheme to transform domains, the QIM method can become resistant against exponential and filtering attacks [94, 95]. The main weakness of RDM method is the sub optimality of its rational function. In [96], the optimal rational function for specific conditions has been given. Besides, the QIM is not well compatible with HVS and HAS, athough several attempts are made in this direction [97-98]. The capacity of QIM is also discussed in [99–102].We now present a recent work in the field of QIM data hiding.

4.2.4 Logarithmic QIM Using μ-law Standard

A novel QIM data hiding framework called LQIM using the μ-law standard is proposed in [103]. In this work, the arrangement of the quantization levels has been designed using a novel approach. The logarithmic quantization is chosen due to its perceptual advantages; while the compression function of the μ-Law standard is applied in order to solve the problems of a previous logarithmic quantization-based method. For this purpose, the μ-Law compression function is applied to the host signal at first to transform it

into the logarithmic domain. After data embedding in this domain, the watermarked and transformed data is quantized uniformly and the inverse compression function is applied to transform the result back to the original domain. Due to the use of logarithmic function, smaller step sizes are devoted to smaller amplitudes and larger step sizes are associated with larger amplitudes. Therefore, in comparison with the previous logarithmic QIM techniques such as UQIM, this method poses perceptual advantages that lead to stronger watermark insertion. The scalar method is then extended to a vector quantization scheme, by quantizing the magnitude of each host vector on the surface of hyper-spheres with logarithmic radii. the optimum parameter μ for both the scalar and vector quantization is found according to the host signal distribution. Furthermore, in order to increase the security of algorithm, a secret key similar to the dither modulation in QIM is included. For the designed framework, the performance is analyzed and analytical results are verified through simulations on artificial and real signals. The performance of this μ-law-based logarithmic QIM is also compared to equivalent QIM techniques. The results demonstrate that this algorithm outperforms the conventional and logarithmic QIM techniques by removing their drawbacks such as vulnerability against scaling attacks and the performance drop for small amplitudes.

4.3 Patchwork Method

4.3.1 Primary Patchwork Algorithm

Patchwork is a statistical method proposed in 1996 [104]. The major application of this scheme is in audio watermarking for copyright protection, but later was extended to the audio applications [105]. In the primary patchwork method, two sets of samples, named A and B, are randomly selected from one block. The constant value d is added to the samples of one set and is subtracted from those of the other one.

$$a_i^* = a_i + d. \qquad b_i^* = b_i - d \tag{17}$$

In this way, the original signal is not required to detect the message and by knowing the location of the two sets, which can be determined by a shared key and a pseudo-random sequence generator, the existence of the hidden message can be revealed.

$$E\lfloor \bar{a^*} - \bar{b^*} \rfloor = E\lfloor (\bar{a}+d) - (\bar{b}-d) \rfloor = E\lfloor \bar{a} - \bar{b} \rfloor + 2d \tag{18}$$

Supposing the distribution of the chosen coefficients to be Gaussian (which is true about DCT coefficients used in modified Patchwork mentioned successively), the difference between the averages of the two sets has zero mean Gaussian distribution before embedding, while the distribution is shifted by $2d$ after Patchwork watermarking.

Setting the detection threshold to zero, the probability of missing the watermark is equal to the area under the shifted Gaussian curve in the left half-plane. The first solution to reduce this error is increasing d; however, d affects the imperceptibility directly and its increase decreases the quality of the signal.

Various improvements have been made to the classic Patchwork method. Among these improved schemes are the application of the Patchwork to the DCT coefficients, the use of the variance besides the mean in the detection procedure, and adaptive watermarking. In adaptive watermarking, the value of d is not the same for all the samples of the signal and is proportional to their magnitude. That is, the more the energy of a sample, the more capacity it has for embedding data. Consequently, $d$times of each sample magnitude is added to or subtracted from the sample.

$$a_i^* = a_i(1+d). \qquad b_i^* = b_i(1-d) \tag{19}$$

And in detection stage, we have:

$$E\lfloor \bar{a^*} - \bar{b^*} \rfloor = E\lfloor \bar{a}(1+d) - \bar{b}(1-d) \rfloor = E\lfloor \bar{a} - \bar{b} \rfloor + dE(\bar{a}+\bar{b}) \tag{20}$$

In this case, the center of the distribution for the difference of the two averages is shifted by $dE(\bar{a}+\bar{b})$ and consequently, the distribution apexes will remain close only when $E(\bar{a}+\bar{b})$ is small, which rarely happens.

4.3.2 Modified Patchwork Algorithm (MPA)

Here we briefly discuss the MPA [106]. First of all, two sets of indices are selected, one for embedding 1 (I^1) and the other for embedding 0 (I^0). Each set of indices represent $2n$ of DCT coefficients. The indices are chosen from $[K_1, K_2]$, which affects the robustness and imperceptibility of the watermark. Each I^i is broken in half producing two subsets of indices, which give two sets of n coefficients; A^i and B^i. The pooled standard deviation A^i and B^i is computed as follows:

$$S = \sqrt{\frac{\sum_{i=1}^{n}(a_i - \bar{a})^2 + \sum_{i=1}^{n}(b_i - \bar{b})^2}{n(n-1)}} \tag{21}$$

The greater the value of S, the more dispersed are the samples of A^i and B^i and consequently, the more is the difference of the averages and potential for data hiding. The two sets are modified by a multiple of S in a way that the samples of the set with smaller

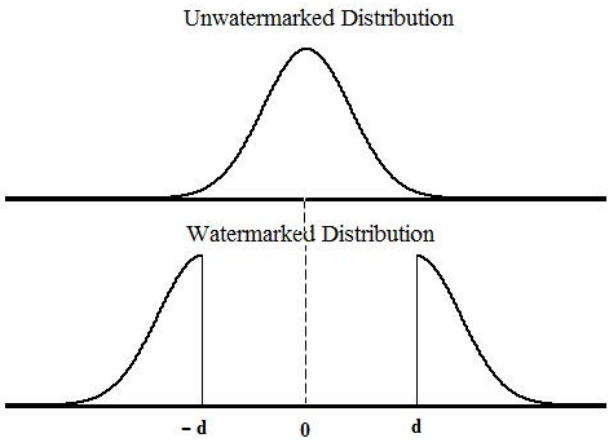

Figure 7. Modified patchwork watermarking algorithm

average are decreased and the samples of the other set are increased such that the difference of the averages in retrieval stage is always positive and the left half of the distribution is placed on its right half. Figure 7 states the embedding algorithm.

A required feature of a watermarking system is its capability to re-mark a watermarked signal. A point in case is when the allowed times of copying has been implicated to be one through a watermark. When the signal is copied one time, the watermark should be changed to zero. The previous watermark is removed via the following equation:

$$a_i = a_i^* - sign(\bar{a^*} - \bar{b^*})\sqrt{C}\frac{S^*}{2} \qquad (22)$$
$$b_i = b_i^* + sign(\bar{a^*} - \bar{b^*})\sqrt{C}\frac{S^*}{2}$$

which is the inverse of the embedding procedure. Regarding the fact that watermarking has been performed by adding and subtracting, the pooled standard deviation of the two sets remains unchanged, i.e., S and S^* are the same. In order to retrieve the watermark, the two sets of indices are re-generated using the shared key and the coefficients belonging to A^0, B^0, A^1, and B^1 are extracted. Then, the two pooled standard deviations S^0 and S^1 are computed and the following statistics are calculated:

$$T_0^2 = \frac{(\bar{a_0} - \bar{b_0})^2}{S_0^2} \qquad T_1^2 = \frac{(\bar{a_1} - \bar{b_1})^2}{S_1^2} \qquad (23)$$

and T^2 is considered as $\max(T_0^2, T_1^2)$. If T^2 is greater than the threshold M, the watermark exists and the index of the greater statistic is the embedded bit.

The GPA method is a more general scheme which exploits additive and multiplicative embedding rules in a way similar to MPA [107]. Overall, the weakness of all the above mentioned methods is their assumption on the two sets of coefficients to have the same statistical behavior, which is compensated by allocating a large number of samples to each set; however, this reduces the data rate of the watermarking system. Moreover, in MPA, the amplifying factor has been chosen regardless of HAS.

4.3.3 Multiplicative Audio Patchwork Watermarking

In order to embed the watermark data, two index sets are randomly generated using a secret key [38]. In this manner, two non-overlapping index sets, $I_A = i_A1, , i_{AM}$ and $I_B = i_B1, , i_{BM}$, which define two equally sized sub-sets of the host signal features are generated. Here, DWT is applied to each frame of the audio signal and coefficients of the approximation band are used as the host signal features. Generally, 2M can be equal or less than the number of host signal features (N). We have used 2N in all of our experiments. By using these two subsets, the watermark signal is embedded through the following multiplicative rule:

$$s'_{Ai} = \gamma s_{Ai} \qquad to\ embed\ 1 \qquad (24)$$
$$s'_{Ai} = \gamma^{-1} s_{Ai} \qquad to\ embed\ 0$$

where s_Ai is the sample of the subset which is represented by I_A, and γ is the watermark strength factor which has a value slightly larger than one. In this way, the watermark data is embedded by modifying the amplitude of one subset and leaving the other subset unchanged. The watermark strength factor is adaptively changed with an iterative scheme in order to reach a desired quality. Quality assessment is carried out using PEAQ algorithm.

The decoder extracts the watermark data by comparing the ratio of energies of two subsets with a defined threshold. According to the embedding function, presented in (24), the decoder can extract the watermark data perfectly provided that the energies of the two subsets are equal. Unfortunately this assumption is not correct in real audio signals. Therefore, we need to embed the watermark data in the selected frames of the host signal which satisfy the following condition:

$$((\gamma_{min} - 1)(1 - G) + 1)^{-2} < f \qquad (25)$$
$$= \frac{\sum\limits_{i \in I_A} s_i^2}{\sum\limits_{i \in I_B} s_i^2} < ((\gamma_{min} - 1)(1 - G) + 1)^2$$

where $\gamma_m in$ is the minimum allowable value for γ. G is named as Confidence Guard Factor (CGF) which varies between zero and one. CGF affects the performance of the system when the attack is low in power. By using this structure, the decoder can perfectly extract the watermark data in the absence of attacks even for small subsets.

According to the embedding functions presented in (24), and due to multiplication, the variance of the subset is changed. Thus, the variance ratio of two subsets should be compared with a threshold in order to extract the watermark data. In this regard, two index sets are generated using the same secret key used in the embedding process. Then, DWT is applied and from the coefficients of the approximation band, two subsets according to the two index sets are formed for each frame. The watermark data is extracted from the received signal using the following hypothesis test:

$$r = \frac{\sum\limits_{i \in I_A} v_i^2}{\sum\limits_{i \in I_B} v_i^2} ?T \qquad (26)$$

where v is the received noisy signal and T is the threshold value which is set to be equal to 1. Data is embedded in selected frames according to the condition presented in (24). Thus, in the decoding process we need to know which frames contain the watermark bits.

In comparison with previous patchwork methods, this algorithm provides some advantages. First, due to the use of multiplicative rule for embedding, stronger watermark insertion with less audibility is achieved. Second, as a result of data embedding in selected frames, the watermark data can be extracted without error in the case where the watermarked signal has not undergone any attack, whereas previous methods suffered from this error. This allows decreasing the frame length which results in higher data rate (approximately 2.5 times greater than the data rate of MPA). Third, the watermark strength is controlled locally to reach a desired quality of the watermarked audio. To control the quality an iterative approach is employed which evaluates the quality of the watermarked audio in each iteration aided by the PEAQ algorithm. Using this approach, more robustness is achieved whereas the quality of the watermarked audio is kept at an acceptable level. Probability of error for this method is also derived by modeling the host signal distribution with GGD and is verified by simulations. Simulation results show that this decoder is robust to the common audio watermarking attacks such as noise addition, MP3 compression, and resampling; this decoder is more robust than the previous patchwork [38].

4.4 Phase Coding

In this method [104], the signal is divided into blocks of N samples; thus, there are $M = \lfloor \frac{L}{N} \rfloor$ blocks, where L is the number of the original samples. Each block X_j is transformed into frequency domain via N-points DFT and the phase matrix $\phi_j[\omega_k]$ and magnitude matrix $A_j[\omega_k]$ are constructed $(0 \leq k \leq N-1)$. Then, the difference of the phase matrices of adjacent blocks is calculated. This matrix will be used at the retrieval stage.

$$\Delta\phi_{j+1}[\omega_k] = \phi_{j+1}[\omega_k] - \phi_j[\omega_k] \qquad (27)$$

If we show the phase of the j^{th} block in the watermarked signal with ϕ_{sw_j}, the watermark m is embedded into the phase of the first block as shown below:

$$\begin{aligned} \phi_{w0}[\omega_k] &= (-1)^{m[k]+1}.\pi/2 \\ m[k] \in \{0,1\}, \qquad & 0 \leq k \leq \frac{N}{2} - 1 \end{aligned} \qquad (28)$$

To guarantee the inaudibility of the modification made to the phase of the first block, the phase of other blocks should be altered consecutively in the following manner:

$$\phi_{\omega j+1}[\omega_k] = \phi_{\omega j}[\omega_k] - \Delta\phi_{j+1}[\omega_k] \qquad \forall j,k \qquad (29)$$

By applying the inverse DFT to the original magnitudes and the new phases, the watermarked signal is obtained in the time domain. To extract the watermark, the received signal is divided into blocks of N samples. Applying DFT on the first block, the phase of this block is computed, which contains the bits of the watermark. This watermarking scheme takes HAS characteristics into consideration and makes use of its insensitivity to the phase of the audio signal. A disadvantage of this method is that it embeds the watermark in the first block and in case of cropping attack, the watermark is easily lost. Phase coding watermarking can also be done in other ways. In [108], speech signals are watermarked using all-pass filters. This scheme has very good robustness against noise. To get more information on phase coding methods, you may refer to Takahashis paper [109].

4.5 Echo Watermarking

In this method, data is embedded by adding echo to the audio signal [110]. That is, different delays are used to embed watermark bits.

$$X_w(t) = X_0(t) + \alpha X_0(t - \Delta t) \qquad (30)$$

In this equation, Δt is the used data insertion. Δt and α can be adjusted in a way that the inaudibility of the echo is guaranteed. The above equation can be generally written as follows:

$$X_w(t) = \sum_{k=0}^{N} \alpha_k X_0(t - \Delta t_k) \tag{31}$$

where $\Delta t_0 = 0$, $\alpha_0 = 1$, and N is the number of added echoes. By using weighted impulse sequence $h(t) = \sum_{k=0}^{N} \alpha_k \delta(t - \Delta t_k)$, equation (31) can be stated as:

$$X_w(t) = X_0(t) * h(t) \tag{32}$$

Transforming this equation into the frequency domain, we get:

$$X_w(\omega) = X_0(\omega) H(\omega) \tag{33}$$

In the detection stage, $h(t)$ should be determined in order to obtain Δt_k and successively, extract the embedded bits. According to equation (33), $H(\omega)$ is obtained by dividing $X_w(\omega)$ by $X_0(\omega)$ and (ω); by applying the inverse Fourier transform yields $h(t)$. In this way, the original signal is required for the watermark detection. The detection procedure can also be performed via homomorphic de-convolution so that the original signal is not needed to separate the echo. This method converts the multiplicative equation (33) into an additive equation, which is a function of quefrency, by employing logarithmic multiplication function:

$$X_w(q) = IFFT(\log|X_0(\omega)H(\omega)|) = X_0(q) + H(q) \tag{34}$$

According to this equation, the original signal and the echo have been separated on the quefrency axis.

To embed the watermark, at the first step, the original signal is divided into M blocks. In each block $X_j, 0 \le j \le M-1$, one bit of the watermark is embedded. To this end, for each block, an echo is generated; the delay and decay rate of which is determined by the watermark bit:

$$w_k(t) = \alpha_k X_0(t - \Delta t_k) \qquad k = 0, 1 \tag{35}$$

The two modulating signals are produced for bits zero and one:

$$\begin{aligned} m_1(t) &= \sum_{j=0}^{M-1} b_j rect_j(t) \\ m_0(t) &= \sum_{j=0}^{M-1} (1 - b_j) rect_j(t) \\ b_j(t) &= \begin{cases} (m(j) & t_j < t < t_{j+1} \\ 0 & otherwise \end{cases} \\ rect_j(t) &= \begin{cases} 1 & t_j < t < t_(j+1) \\ 0 & otherwise \end{cases} \end{aligned} \tag{36}$$

Using the above signals, the watermarked signal is obtained as follows:

$$X_w(t) = X_0(t) + m_0(t)w_0(t) + m_1(t)w_1(t) \tag{37}$$

To detect the watermark, each block is transformed to the Cepstrum domain at first:

$$C_w = IFFT(log|FFT(X_w)|) \tag{38}$$

The autocorrelation, C_w, is calculated in this domain. Δt and the embedded bit is determined using the apex of C_w. Various methods have been devised based on echo hiding [111–114]. By combining spread spectrum and echo hiding schemes, a secure watermarking method is obtained [115]. In addition, a robust and secure watermarking algorithm has been designed using the hearing characteristics and filter bank structure of HAS [114].

4.6 Spread Spectrum Watermarking

4.6.1 Additive Method

A simple method of watermarking is the additive scheme. This approach was first introduced by Cox [115]:

$$f_{w,i} = f_i + \gamma w_i \tag{39}$$

where f_i is the i^{th} sample of the host signal, γ is the power factor, and w_i is the i^{th} sample of the watermark signal. w_i may be the samples of an arbitrary vector or a pseudo-random sequence. As a matter of fact, the additive method was first proposed in the block form where the spectrum of the watermark signal was spread using a pseudo-random sequence, PN, and was added to the host signal with a very small coefficient. The longer the sequence, the more the spectrum is spread and the less the summation coefficient. To detect the watermark, it is enough to compare the received signal

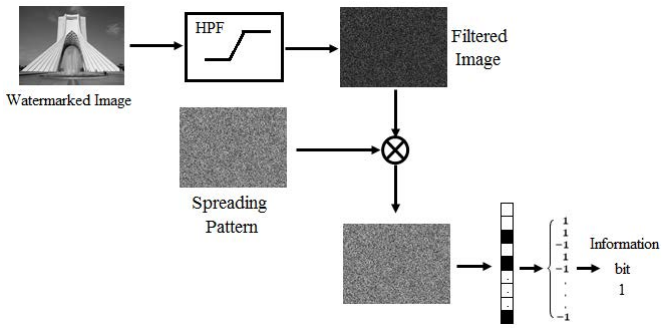

Figure 8. Additive spread spectrum watermark embedding

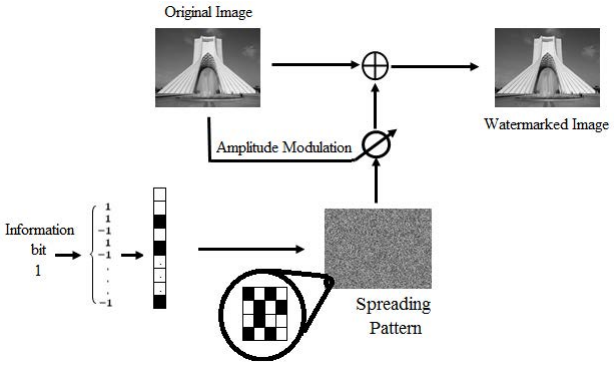

Figure 9. Additive spread spectrum watermark extraction

with the pseudo-random code. When all the bits of the code are the same as their corresponding bits embedded in the watermarked signal, there is a peak, which can be detected using a proper threshold. In fact, the receiver is a match filter. Generally, when the samples of the signal and the noise are modeled as Gaussian, all error expressions and the capacity can be obtained via the equations of information theory and communications. Figures 8 and 9 show a simple additive watermarking scheme.

To improve the additive method, the amplification factor can be adaptively selected according to the signal characteristics such that the watermark power can be maximized without losing transparency. As a case in point, equation (39) may be modified as follows:

$$f_{w,i} = f_i + \gamma_i(f) w_i \tag{40}$$

where γ is a function of f and is determined based on the cover features. It has been selected sub-optimally according to [116] and [117] for the DCT coefficients and the DWT of the image, respectively. The optimal receiver of the additive watermarking is implemented as follows:

$$\hat{b} = \arg \max_{b\in\{-1,1\}} \prod_{i=1}^{r} \frac{1}{\sqrt{2\pi(\sigma_f^2+\sigma_n^2)}} \times exp(-\frac{(f_i' - \mu_f - \gamma b w_i)^2}{2(\sigma_f^2+\sigma_n^2)}) \tag{41}$$

where f_i's are the watermarked coefficients, μ_f and σ_f^2 are the mean and variance of the host signal coefficients, respectively.

Additive watermarking can be easily implemented. The only problem is that in case the sequence is not long enough, the interference terms resulted from the overlapping of the original signal and the sequence will become large and may go beyond the threshold. To solve this problem, the received signal is passed through a high-pass filter so that the majority of its power is removed. Since the pseudo-random sequence is intrinsically high-pass, it passes through the filter without any modification; thus, there will be no loss in detecting the watermark. Figure 9 implicates the application of pre-processing filter before the overlapping.

Different implementations of additive watermarking have been described in [118–121]. In [118], Modulated Complex Lapped Transform (MCLT) has been used and the embedding has been done by altering the amplitude of MCLT coefficients in dB. The spread spectrum technique can also be exploited to synchronize the watermarked and original signals. In some methods, it is necessary to specify the beginning of the frame or the block. Regardless of the watermarking scheme, the synchronization can be done by adding a pseudo-random code with a small gain factor to the watermarked signal in spatial or temporal domain. When the pseudo-random code matches its peer in the watermarked signal, an apex appears, and that time is the beginning of the frame. Fore more information, please refer to [118–123].

4.6.2 Multiplicative Method

It seems that greater and more substantial coefficients have more capacity to carry the watermark. Indeed, several experiments on HAS and HVS have proved that the largest threshold of alteration imperceptibility is obtained when the interfering signal (the watermark signal in our discussion) has the same frequency as the host signal. In order to achieve this goal, multiplicative method has been proposed:

$$f_{w,i} = f_i(1 + \gamma w_i) \tag{42}$$

By this approach, we can make more use of the characteristics of HAS and HVS in DCT, DWT, and DFT domains and better adjust the power of the watermark

[104]. This has made multiplicative algorithm more successful than its additive counterpart. The only negative point about multiplicative method is that optimizing the parameter γ and obtaining equations for probability of error and capacity is not as easy as the additive scheme, and we can hardly make use of the information and communication theories. However, it leads to more open research topics on this algorithm.

In [3] and [119], similarity-based receivers have been proposed for multiplicative watermarking. Since these receivers are suboptimal for transform domains, various locally optimal detector have been investigated in [124–129]. Optimal ML receiver has been suggested in [125] for DCT, DWT, and DFT domains. In that paper, generalized Gaussian distribution has been considered for high-frequency DCT and DWT coefficients, while the amplitude of DFT coefficients has been assumed to meet Weibull distribution. It should be noted that in all cases except for the Gaussian one, the receivers have been examined without noise. Analyzing the performance of such receivers in the presence of noise are open research problems. For instance, the ML receiver with generalized Gamma distribution for multiplicative methods has been partly investigated by [127]. The authors of [128] have carefully computed the distribution of DFT coefficients and have shown that it does not exactly obey the Weibull model. Moreover, by applying HVS to the high frequency wavelet coefficients, locally optimal detector for Barnie multiplicative method discussed in [128, 130] is designed and analyzed in [129].

Multiplicative watermarking method has been extended to various transform domains other than the wavelet transform. One of the advantages of the multiplicative watermarking is its improved performance for the transforms which exhibit a sparser presentation of the host image. Therefore, the performance of the multiplicative watermarking is significantly improved when is applied to the multiresolution nonseparable transforms such as contourlet [130] and discrete ridgelet transforms [131].

4.7 Comparison among Various Data Hiding Techniques

In this part, we wish to briefly compare the various watermarking schemes. This comparison is summarized in Table 1. It is noteworthy that comparing various categories is not completely fair because of the differences in the type of signal and target application. Moreover, “low” or “high” do not mean “bad” or “good” here; both can be “good” depending upon the required application.

In this table, “open problems” column is an estimation of how much developed the theories are in each field and how many unsolved problems are left to be worked on. It is inferred from the Table that the LSB coding provides the most possible capacity. This feature makes it the choice of interest for the steganography applications, where high capacity algorithms are required to fulfill the covert communication necessities. The LSB schemes are also imperceptible enough and their robustness and security issues high vulnerability to the noise, and compression are not of much importance in the steganography application. These properties along with low complexity have attracted a lot of interest to the LSB steganography. By the way, it seems there is no significant improvement left to be made in this field. The QIM data hiding algorithms show similar properties to the LSB ones. However, the security issues are improved in QIM techniques considering the dither modulation idea.

Capacity decreases for the other five categories in the Table which make them applicable for the watermarking demands. Among them, phase coding and echo hiding are especially designed for the audio watermarking. By decreasing the capacity, these methods are expected to offer significant performance in other aspects, as a consequence of the principal data hiding trade-off. By the way, it is inferred from this table that the patchwork and the echo hiding techniques suffer from security problems, while the phase coding has the robustness issues.

Additive algorithms are outperformed by multiplicative ones, which efficiently improve the other data hiding issues in exchange of losing capacity. Multiplicative algorithms are also compatible to human visual and auditory systems. According to these properties, this group of data hiding techniques seems to be a suitable choice to answer the watermarking demands. Thus; in the remaining of this section, we will focus on the multiplicative schemes in particular the scaling based algorithms as one of the most efficient watermarking tools and investigate the performance of these algorithm from this category in more details.

4.8 State of the Art Multiplicative Watermarking

In the previous section, different categories of data hiding schemes were compared. Among them, multiplicative schemes were shown to be very efficient in terms of imperceptibility, robustness, complexity and security requirements. The only drawback of multiplicative techniques is their low capacity that makes them not to be a proper choice for the steganography demands. However, these methods seem to be the best choice for setting up a watermarking system. Here, we will discuss about a special case of multiplicative schemes which is called the scaling-based method. Al-

Table 1. Comparison of primitive watermarking schemes

Factor / Scheme	Imperceptibility	Bit Rate	Robustness	Complexity	Security	Open Problems
LSB	Good	Very high	Very low	Low	Low	Few
QIM	Good	High	Low	Low	High	Some
Patchwork	Moderate	Low	High	Low	Low	Few
Phase Coding	Very good	Average	Low	Average	—	Few
Echo	Moderate	Low	Average	Average	Low	Few
Additive	Good	Low	High	Low	High	Few
Multiplicative	Very good	Low	Very high	High	High	Many

though the method was first developed for image signals, similar concepts are extended to the audio and speech signal in [38, 132], and [133].

In [134], a scaling-based semi-blind image-adaptive watermarking system has been presented, which exploits HVS for adapting the watermark data to local properties of the host image. To have a better robustness, this algorithm is implemented on the low-frequency coefficients of the wavelet transform. The scaling-based embedding process is implemented by the following rule:

$$W_i' = \begin{cases} W_i.\alpha & For\ embedding\ 1 \\ \frac{W_i}{\alpha} & For\ embedding\ 0 \end{cases} \quad (43)$$

W_i and W_i' are the wavelet coefficients before and after embedding. α is the strength factor which controls the watermark power and is optimally selected, regarding the invisibility of the algorithm. To extract the watermark data, the Maximum Likelihood (ML) estimator is used. Experimental results confirm the imperceptibility of this method and its high robustness against various attacks such as JPEG compression, noise addition, and filtering compared to other similar methods [134].

The idea of multiplicative data hiding in the transform domain is also extended to the contourlet domain [135]. The contourlet transform [130] is designed to model singularities, high-dimensional discontinuities and edges in a more efficient way than the wavelet transform. Since HVS is less sensitive to the image edges, the contourlet transform which represents the image edges more sparsely seems to be an interesting choice for watermarking application. In the presented scheme, watermark is embedded in the most energetic directional subband using the following multiplicative rule:

$$w_i = \begin{cases} x_i.f_1(x_i) & For\ embedding\ 1 \\ x_i.f_0(x_i) & For\ embedding\ 0 \end{cases} \quad (44)$$

where $f_1(x)$ and $f_0(x)$ are strength functions which are chosen to be monotonous exponentially functions. To achieve the best performance, these functions are defined as follows:

$$\begin{aligned} f_1(x) &= -0.3e^{-0.2|x|} + 1.65 \\ f_0(x) &= 0.15e^{-0.2|x|} + 0.65 \end{aligned} \quad (45)$$

These functions are chosen exponentially in order that larger coefficients changes more than smaller ones during the watermarking process since the larger coefficients are related to the strong edges in the supposed directional subband. These functions are also defined in a way that the monotony of $xf(x)$ is satisfied for the practical range of x.

In this work it is also shown that the Generalized Gaussian Distribution (GGD) efficiently models the histogram of the contourlet coefficients:

$$\begin{aligned} GG_{\sigma_x,\beta}(x) &= C(\sigma_x,\beta)e^{-[\alpha(\sigma_x,\beta)|x|]^\beta} \\ -\infty < x < \infty, \sigma_x > 0, \beta > 0 \\ C(\sigma_x,\beta) &= \frac{\beta\alpha(\sigma_x,\beta)}{2\Gamma(\frac{1}{\beta})} \\ \alpha(\sigma_x,\beta) &= \sigma_x^{-1}\sqrt{\frac{\Gamma(\frac{3}{\beta})}{\Gamma(\frac{1}{\beta})}} \end{aligned} \quad (46)$$

where σ_x is the standard deviation of x, β is the shape parameter and Γ is the Gamma function.

By modeling the General Gaussian Distribution (GGD) for the contourlet coefficients, the distribution of the watermarked noisy coefficients is analytically calculated. At the receiver end, based on the ML decision rule, the optimal detector is proposed. Since the contourlet transform concentrates the image energy in the limited number of edge coefficients, using multiplicative approach in this domain yields high robustness accompanied by good transparency. Experimental results show the imperceptibility and high robustness of this method against Additive White Gaussian Noise (AWGN) and JPEG compression attacks [135]. The same idea of multiplicative embedding in the transform domain has been extended to the Ridgelet transform [131], where universally optimal decoder has been presented based on ANOVA (Analysis of Variances) [136].

In a similar work, the performance of the multiplicative scheme is investigated when the host signal is assumed to be stationary Gaussian with first-order autoregressive (AR) model [137]. Partitioning the host signal into two separate parts, the watermark is embedded in one part using a multiplicative rule similar to (5-16) and the other is kept unchanged for blind parameter estimation. To drive the distribution of the decision variable, the ML decoding algorithm is suggested which is independent of the host signal distribution. This makes the algorithm suitable for any transform domains. The detector uses a decision variable resulting from the sum of samples, which due to the Central Limit Theorem converges to a Gaussian variable. Under this assumption, a Distribution Independent Optimum Decoder (DIOD) is introduced that works for any kind of distribution model such as Gaussian and GGD and in any transform domain such as wavelet, contourlet, ridgelet, and FFT. The proposed algorithm is applied to both artificial Gaussian autoregressive signals as well as various test images. Experimental results confirm the independence of the decoder performance to the host signal distribution and its great robustness against common attacks [137].

While in the previous schemes the samples of the host signals were assumed to be uncorrelated, another research is performed in [138] to investigate the scaling-based data embedding and the design of the optimal detector for correlated signals [138]. In this work, the host signal is assumed to be stationary Gaussian modeled with a first-order autoregressive process. It can be shown that this model is useful for representing the low frequency components of the natural images.

Let, u be a first order markov sequence of normally distributed random variables with mean μ, variance σ^2, and correlation coefficient ρ. If u contains N variables $u_1, u_2, , u_N$, let x and y sequences represent the samples of u in odd and even positions, respectively. That is, $x_i = u_(2i-1)$, and $y_i = u_2i$. Dividing the signal into these two categories satisfy the blindness of the watermarking approach. Now, consider the new sequence z, defined as:

$$z_i = \frac{x_i}{y_i} = \frac{u_{2i-1}}{u_{2i}} \tag{47}$$

This variable is the ratio of two correlated normal variables $x = N(\mu_x, \sigma_x^2)$ and $y = N(\mu_y, \sigma_y^2)$. By using some computation, it can be shown that for the case of non zero μ_x, μ_y, and $\sigma_y << \mu_y, \sigma_x << \mu_x$, the distribution of $z = x/y$ can be well approximated by a Gaussian distribution. More calculations result in finding the correlation coefficient between the samples of the new Gaussian random variable z.

For data embedding, the host signal is divided into two parts. One part is manipulated while the other part is kept unchanged for the parameter estimation. Therefore, the odd and even numbered samples of x_i and y_i are extracted at first. The data is embedded only in the x sequence based on a scaling rule similar to (24) to obtain the watermarked signal u'. The y part will be later used to estimate the original signal parameter at the decoder side. As we mentioned before, the ratio between these two parts of signal can be modeled using a Gaussian distribution. Therefore, a decoding scheme using the ratio of samples is suitable for this highly correlated signal. By calculating the distribution of the ratio, the performance of the maximum likelihood decoder is analytically studied. The unchanged part of the signal is exploited to help the estimation of the Gaussian parameters. The main characteristic of this decoder is that it can be easily implemented for highly correlated signals.

The error probability of the ML detector at the presence of AWGN is analytically calculated. This algorithm is applied to several artificial Gaussian autoregressive signals to verify the validity of the results. Simulation results show a great robustness of this method for low watermark to noise ratios (WNR) [138].

5 Steganalysis

Steganalysis is the art of detecting the presence of watermarked secret message within a digital signal. There are several ways to classify steganalysis methods, but in our brief review, we consider the "method-specific" and "universal" steganalysis approach. Method specific or non-blind steganalysis schemes are designed to attack a certain type of steganography algorithms based on its particular features and weaknesses. On the other hand, the applied algorithm is not important in blind or universal steganalysis which aim at extracting general features that are the most sensitive to malicious modifications. Although the steganalysis was initiated by attacking certain algorithms such as the LSB using limited number of features, the current trend of steganalysis is toward the development of universal steganalysis frameworks for large number of features that are selected efficiently.

5.1 Method-specific Steganalysis

There exists a variety of steganalysis methods designed to detect the existence of secret message embedded through a specific data hiding method. These algorithms find a weak point in a proposed steganography algorithm and concentrate on it to suggest a tool or measure for detecting the existence of secret message. Since LSB coding is almost the simplest and most

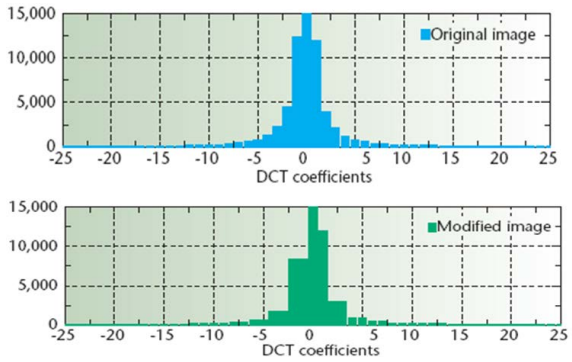

Figure 10. PoV appearance of histogram after LSB replacement

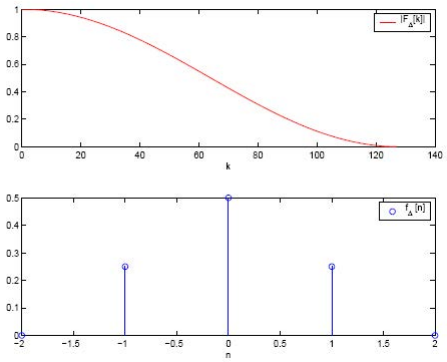

Figure 11. The Fourier transform (the upper curve) and the noise distribution of LSBM (the lower figure)

prevalent watermarking technique, in this section, we only consider some of steganalysis methods proposed specifically for detecting this kind of algorithms.

The LSB coding ideas developed after LSB replacement, where the least significant bit of signal sample was simply replaced by the secret bit. This simple technique results in a statistical drawback in the stego signal, called Pair of Values (PoV). Here we explain the PoV using a numerical example. Suppose that the number of samples with value of four (100) and five (101 in binary) in a cover signal before embedding are n4 and n5 respectively. Since the secret message is supposed to be encrypted and have equal number of ones and zeros, all the LSBs of samples with values equal to $(100)_b$ and $(101)_b$ are firstly removed and then randomly replaced by zero or one. It is simple to verify that as a consequence of this process, the expectation of the number of samples with values equal to four and five after data embedding are both $(n_4 + n_5)/2$. As a result, every two adjacent values in the histogram of digital signal tend to same values. This statistically drawback of LSB replacement is called PoV and is recognizable from the histogram of stego signal, as seen in Figure 10.

Two most famous steganalysis methods that exploit the PoV drawback to detect LSB replacement embedding are RS [67] and Chi-square [68] analyses. In the following, we briefly review the Chi-square analysis to present a method-specific steganalysis technique.

Suppose that the number of samples with value $2i$ and $2i + 1$ after embedding is n2i and n2i+1 respectively. We define the statistics of y_i and y_i* as:

$$y_i = n_{2i}, y_i^* = \frac{n_{2i} + n_{2i+1}}{2} \tag{48}$$

We know that after embedding, y_i* is the expectation of y_i, hence these values are expected to be almost the same in case of stego signal. This fact makes the χ^2 statistics defined below to tend to zero:

$$\chi^2 = \sum_{i=1}^{v} \frac{(y_i - y_i^*)^2}{y_i^*} \tag{49}$$

where v is the number of possible pair of values in the histogram of samples. the closeness of χ^2 statistics to zero value results in the following probability to tend to one in case of stego signal. The closer the p-value to zero, the more probable is the clean. $\Gamma(x)$ represents the gamma function of x.

$$p = 1 - \int_0^{\chi^2} \frac{t^{\frac{v-2}{2}} e^{\frac{-t}{2}}}{2^{\frac{v}{2}}\Gamma(\frac{v}{2})} dt \tag{50}$$

The LSB matching (LSBM) solves the problem of PoV using a simple idea. In the LSBM a sample value randomly increases or decreases by one in case of not being matched to the secret bit. As a result of this minor change, despite of LSB replacement where only $2i$ to $2i + 1$ and reverse changes are allowed, here both $2i$ to $2i + 1$ and $2i$ to $2i - 1$ changes are possible. Therefore, the PoV appearance of histogram does not happen anymore. As a consequence, more sophisticated steganlysis methods were required. For this purpose, Harmsen proposed his analysis based on center of mass of the histogram of the characteristic function (HCF COM) [139]. The main idea of HCF-COM analysis establishes the LSBM as a lowpass noise convolved with the original signal. Note that in LSBM, both the probability of incense and decrease of the value of a sample in the host signal equals 0.25, where the sample may remain unchanged by the probability of 0.5. The noise-like distribution of the LSBM embedding effect and its Fourier transform is shown in Figure 11.

As seen in Figure 11, the LSBM acts as a lowpass filter. A noise with this distribution is added to the host signal. We know that if two uncorrelated random variables are added, the distribution of the result would be the convolution of the variables distribution.

Assume the histogram of the original signal as the estimation of its distribution. Then we will have the histogram and the noise distribution convolved together after embedding. Since the embedding noise was found to behave like a lowpass filter, higher frequencies of the histogram function are attenuated comparing to lower ones. We refer to the Fourier transform of the histogram as the Histogram Characteristic Function (HCF). Having higher values of HCF, the center of mass of HCF (HCF-COM of stego signal) stands lower than that of the original image. Center of mass of HCF for 256-point DFT is defined as:

$$C(H[k]) = \frac{\sum_{k=0}^{1} 27k|H[k]|}{\sum_{k=0}^{1} 27|H[k]|} \qquad (51)$$

where $H[k]$ is the k^{th} sample of HCF (Fourier transform of the histogram). Generally, classifiers work on a set of features similar to the introduced HCF-COM. A number of stego and clean signals are generated at first. Then the classifier is trained using these features to distinct stego signals from the clean ones in the best possible way. After finding the decision rule, the classifier can decide either every new unknown signal is stego or clean. For only HCF-COM feature, the decision rule reduces to a simple thresholding.

Although Harmsen HCF-COM analysis showed an acceptable performance for colored images, it encountered some difficulties in case of grey-scale ones. Ker modified the HCF-COM feature and proposed adjacency HCF-COM and calibrated adjacency HCFCOM features [140]. The histogram function used for simple HCF-COM was a one-dimensional histogram over possible values from minimum to maximum value, for instance 0 to 255 for 8-bit signals. But adjacency histogram is a 2-dimensional histogram which its (i, j) entry is the frequency of pair of adjacent samples with values equal to i and j. Center of mass of this histogram also decreases as a consequence of data embedding. Therefore, Ker proposed adjacency HCF-COM as another steganalysis feature and showed that it outperforms normal HCF-COM in grey-scale images.

In addition, it is impossible to impose an absolute threshold for distinction between stego and clean signals. HCF-COM of one signal after data embedding might be still higher than the other one prior to embedding. This depends upon the nature of signal. To make analysis more independent of the nature of the signal, Ker proposed calibrated HCF-COM. Calibrated signal is the received signal down-sampled by the rate of two (for images it results in the decrease of the size of image by four). This signal is considered as an estimate for the original signal which is not available at the receiver side. Then thresholding is done on the ratio of HCF-COM of the received signal to that of the calibrated one. Ker showed that this improvement makes the analysis more independent of the original signal and results in better discrimination between stego and clean signals.

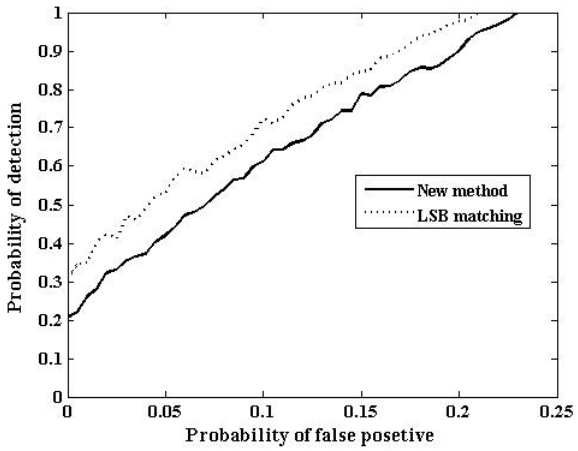

Figure 12. The ROC performance of the calibrated adjacency HCF-COM

Even with these improvements, the classifier still has errors in detection. Every threshold yields a number of true detections (probability of detection) and a number of false detections of clean signals as stego ones (probability of false alarm). Plotting different pairs of detection and false alarm probabilities versus each other for different thresholds, we get the Receiver Operation Characteristic (ROC) results. Figure 12 shows the ROC performance of the calibrated adjacency HCF-COM against the LSB matching method, and the one in [35] which offers the probability of change per pixel of 37.5% for data embedding at the rate of the one bit per pixel (1bpp). It is obvious that among these methods, one with the least probability of change, has the curve closer to the $y = x$ line, which means the best performance in terms of not being detected by the steganalyzer.

More recently, steganography algorithms with better performance against the calibrated adjacency HCF-COM were introduced. These algorithms require more complicated detection features. For instance, the HCF-COM was proposed for the difference image, i.e., the image generated from the difference of adjacent pixels [141, 142]. This competition between steganography and steganalysis is still continuing. The invention of new analysis methods, trigger the design of more developed data hiding algorithms. As a result of this competition, more and more sophisticated algorithms are available both in steganography and steganalysis. Nowadays, steganography algorithms are sophisticated enough not be discovered by limited number of

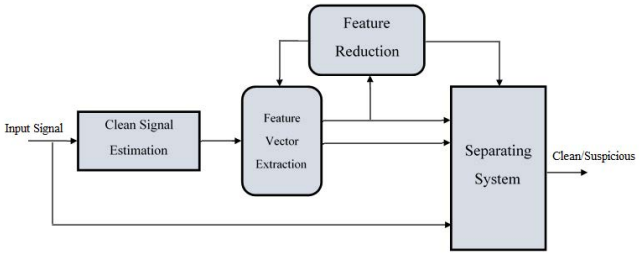

Figure 13. Generic block diagram of a universal steganalysis system

features. As a consequence, current trends of steganography algorithms are to design universal steganalysis frameworks that work on the large number of features independent of the applied steganography scheme.

5.2 Universal Steganalysis Methods

Analyses techniques introduced in the previous section, concentrates on a specific statistical drawback of certain steganography methods. However, one can provide a set of statistical features that is large enough to analyze signals generated through a variety of data hiding methods. Therefore, universal analysis is an approach independent of the embedding which requires classifiers much more complicated than simple thresholding techniques. Figure 13 shows the overall performance of a universal steganalysis system.

Feature reduction is necessary to ensure that the best possible analysis performance is guaranteed using the least size of selected features. As an example for the clean signal estimation, we can apply the Ker's method for calibrated signal at the receiver end. The way we can discriminate different universal steganalysis methods is through their selected feature set. One of the first universal steganalysis schemes is Farid's universal feature set, which exploits the high order moments of the image wavelet coefficients [143]. Some well-known image quality metrics are used as a feature set by Avcibas to form another steganalysis system [144]. Moulin [145] and Huang [146] feature sets are another examples of the universal feature sets introduced for the purpose of steganalysis.

Ker employed the histogram of the adjacency for the sake of steganalysis [140]. This feature counts the number of adjacent pixel pairs with a certain pair of values. The adjacency feature is a second order statistic, because it considers the co-occurrence of two certain values. One can consider the occurrence probability of certain triples as the third order statistics. Although higher order statistics can be defined as well, the complexity imposed by large number of possibilities in the higher order statistics make them impractical. Therefore, the second and third order

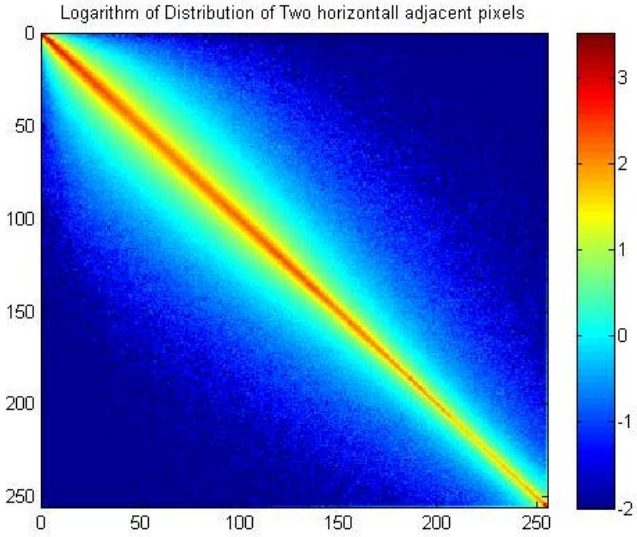

Figure 14. average of 2D co-occurrence histogram of 1000 images

statistics are much more common in the literature of the steganalysis. Figure 14 shows the average of two dimensional co-occurrence histogram for a set of 1000 images:

As it is illustrated in Figure 14, close pixel values are more probable to occur. This fact, leads to the idea of applying the co-occurrence to the difference images. In the next equations, we denote with F the original matrix which can be the image itself, or its transformed version in quantized DCT or wavelet subbands. We also define the difference matrix D whose entries are the difference between the consecutive entries of the original image. We define four types of difference matrices in this paper. Suppose that $F^{i,j}$ is the F matrix when shifted by i rows to the right and j columns to the bottom. The horizontal, vertical, diagonal and minor diagonal difference matrices are defined based on (52):

$$\begin{aligned} CD^h &= F - F^{1,0} \\ D^v &= F - F^{0,1} \\ D^d &= F - F^{1,1} \\ D^{md} &= F - F^{1,-1} \end{aligned} \tag{52}$$

The second and third order statistic can be extracted from the difference images described by (52). As seen before, the pixel pairs with limited difference are most probable. Therefore, we can restrict the values of the statistics to the range of $[-T, T]$ to limit the number of second order statistics. The information about the number of co-occurrences is recorded in NJ matrices as features of higher order statistics. For instance, NJ is a 7×7 matrix for $T = -3$ and

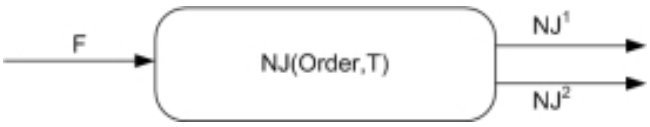

Figure 15. The co-occurrence matrix extraction

$order = 2$. This procedure is shown in Figure 15.

One of the most important feature sets extracted from the difference images is 686-D Subtractive Pixel Adjacency Matrix (SPAM) [147]. In this case, the third order statistics only at the range of $[-3, 3]$ are considered. All the features are extracted from spatial domain signal. In order to decrease the number of features, horizontal and vertical feature sets are reduced to one by averaging. The same procedure is performed to the diagonal and minor diagonal feature sets. Therefore, the final number of the features in the SPAM feature set equals $2 \times (2 \times 3 + 1)^3 = 686$.

The second and third order statistics can also be extracted from the transform domain. Since DCT is the transform used for JPEG compression and is common in steganalysis, there exists a trend to extract features in the DCT domain. In the JPEG compression, the original image is divided into a block of size 8×8 and then the DCT is applied. Since the dependencies are decreased in the transform domain, the second order statistics seem to be sufficient.

In order to extract these features in the JPEG domain, there are several approaches considering 8×8 blocks. One can extract the features among the coefficients of each block. This is called inter-block feature extraction, while the term intra-block is referred to the case where coefficients in the same place of different blocks are considered. Here, we discuss these features in more details.

1) Features Extracted from Inter-block Correlation: These features are extracted through considering the DCT 8×8 blocks as the F matrix. For each block, D and NJ matrices are calculated separately. Apparently, this process generates a very large number of matrices. In order to produce manageable results, we average over all matrices. Since dependencies are considered only inside blocks, this method is called inter-block correlation and is summarized and illustrated in Figure 16.

2) Features Extracted from Intra-block Correlation: In this case, we combine certain entries (frequencies) of all 8×8 blocks excluding the first entry (DC coefficient), together to form a set of 63 new F matrices. The D and NJ matrices are calculated similar to the former sections. The number of matrices is reduced by averaging over all of them as discussed in the previous section. Since similar frequencies in different blocks

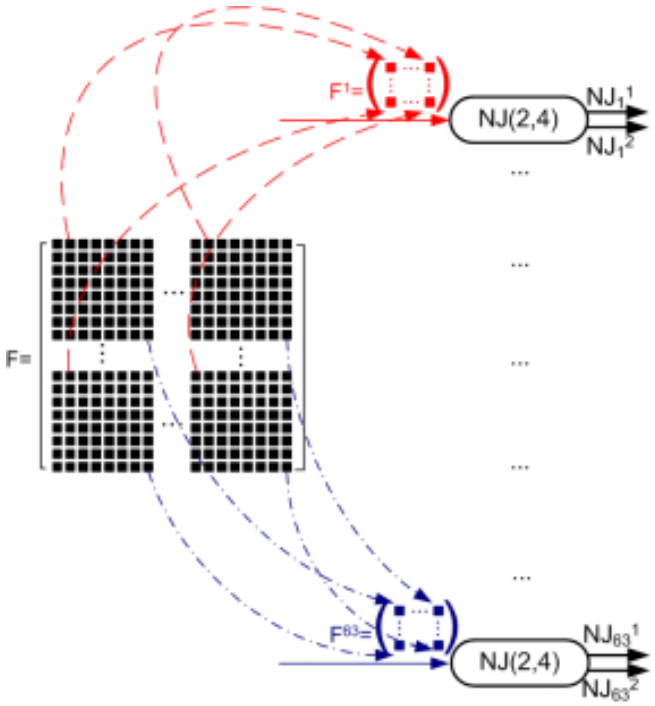

Figure 16. The feature extraction in DCT domain using inter-block correlation

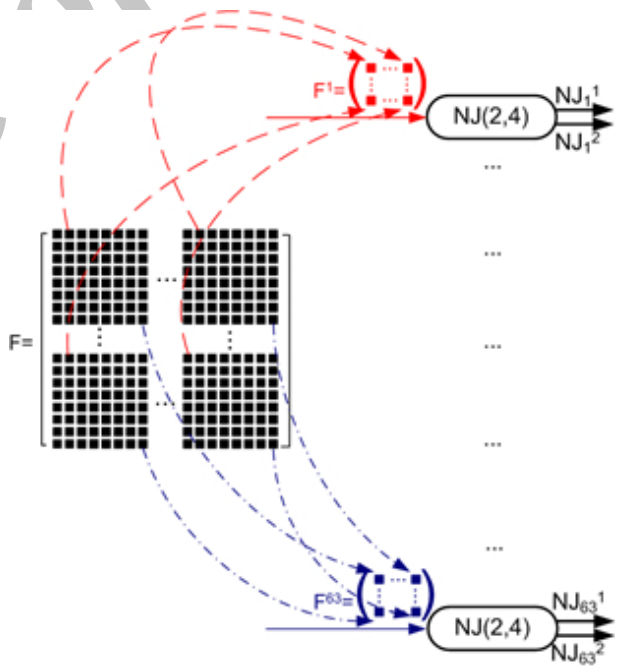

Figure 17. Feature extraction in the DCT domain using intra-block correlation

are considered in this method, it is called the intra-block correlation which is presented in Figure 17.

The inter-block and intra-block features are exploited in some recent and efficient steganalyzers such as 648-D CC-Shi [148, 149], and 548-D CC-Pev [150]

feature sets. However, there might be still more weak dependencies not captured by these features; such as the dependency between the second coefficient from one of the blocks and the third one from another block. Including these dependencies, the 48600-D CC-C [151] feature set is designed to capture even the weakest dependencies. Although this large feature set reflects even the weakest dependencies, one might ignore its performance improvement in the exchange of much less complexity and realization time.

It was shown that recent steganalyzers work on large feature sets, because a limited number of features are not capable to discover the sophisticated current steganography schemes. As the number of features grows, the other important part of the steganalysis frameworks the classifier- becomes more and more important. Although several classifiers such as neural networks [152] have been applied in the field of steganalysis, the support vector machines (SVM) [153] are of the most interest among steganalyzers. Although the SVM classifier performs efficiently for steganalysis purposes, its complexity increases exponentially by growing size of the current feature sets. On the other hand, the efficient analyzing of the recent sophisticated and diverse steganographic algorithms necessitates applying a large number of features to reflect even the weakest dependencies in the images. However, the implementation time and the complexity of these classifiers such as SVM restrict the steganalyzer to use only limited number of features. In order to tackle this challenge, the ensemble classification structure is very recently investigated in the steganalysis literature [154].

Actually, an ensemble structure consists of application of several classifiers on a limited number of features, whose decisions are integrated by an efficient rule to make the final decision. The base classifiers work on a limited number of features (100 for instance), while a feature can be repeated in several classifiers. The aggregation of the results from simple classifiers leads to a simple and efficient steganalyzer. In this way, the ensemble classifier decreases the complexity of the classification algorithm significantly, while it keeps the ability of dealing with a large number of features with offering almost the same performance. Moreover, the ease of dealing with a large feature sets, makes it possible to avoid the need of conducting many experiments to pick the optimized features for each specific embedding method. Trying several classifiers and decision rules, the authors have come to the result that the application of complex classifiers or decision rules does not improve the performance of the final classifier. Therefore, the majority rule is exploited to aggregate the decisions that come from a simple base classifiers [154].

6 Conclusion

A survey on data hiding concepts and algorithms is given in this paper. The main attributes of data hiding systems are briefly discussed and a variety of data hiding applications are presented. Transparency which means the similarity between the original and stego/watermarked signal is almost required for every data hiding algorithm. In watermarking applications where a low-rate data transmission is needed, the watermark is required to be robust against some certain attacks; while in the steganography applications that requires a high rate of embedding, the robustness issue is less important. On the other hand, the capacity considerations are critical for steganography schemes, while they are not important in watermarking applications. The data hiding tradeoff declares that one cannot satisfy all these three concepts simultaneously. The security of algorithm and the computational complexity are the other attributes of data hiding schemes. Since the knowledge of human auditory and visual system helps to design the data hiding algorithms more efficiently, they are also briefly reviewed in this paper.

Next in this paper, data hiding algorithms were categorized and discussed separately. It was shown that the LSB algorithms provide the best capacity without acceptable robustness, which makes them a proper choice for steganography applications. These parameters are more or less compromised in the other methods; while the multiplicative embedding was seen to present the best properties to satisfy the watermarking requirements. Therefore, this group of data hiding schemes is discussed in more details. Finally, the analysis of data hiding algorithm was reviewed by being divided into blind and non-blind steganalyzers.

The main target of this work is to help the newcomers in the field of data hiding to find their right path of research by getting familiar with the main concepts and algorithms of this field. However, after about two decades from the inception of data hiding concepts, it is not as simple to design a new data hiding scheme or improve an elder one to achieve a significant gain in performance. Somehow, it seems that the problem of data hiding has been approached from all possible aspects, especially for images. On the other hand, the development of steganalyzers has also been slowed down and newly introduced features that are usually supported with the simulation results rather than solid mathematical frameworks, and result in a very slight performance achievement.

Nowadays, it seems that the future of the data hiding field is focused on the development of the novel applications rather than designing a new schemes. For instance, the generation of the tamper-proof images

with the self-recovery capability is introduced as an interesting novel application of the data hiding systems. It was shown that the self-recovery problem can be modeled by applying a proper source code to generate the "image digest" and a channel code to protect it against the tampering. The source and channel coded image is embedded in itself as the tamper-proof watermark. Although in a recent work in this field, these source and channel codes are selected to be optimal separately; one can inject the joint source channel coding concepts to the problem and optimize both source and channel codes simultaneously to achieve a performance gain.

Recently, the concept of data hiding has also been extended to the fields other than multimedia. For instance, some fragile watermarks can be embedded in the printed circuit boards when they are designed and ordered to be implemented. In this case, the owner will notice every manipulation when the implemented board is returned back to him. These manipulations might be done for example to disrupt the board performance in a particular geographical area. This is called hardware watermarking. The same concept can be extended to the software watermarking, when the designer of the software desires to notice every malicious manipulation on the source code or the library files. In the field of cryptography, a new problem of interest is to embed the secret key of a conversation into the encrypted message itself, instead of communicating it through a separate secure channel which may not exist.

Although these applications need to leave a watermark to find the manipulations later, there is another new similar trend called forensics, in which the target tries to find any additional useful information from the media. For example, one can find that an object is artificially repeated in an image via investigating the correlations between blocks; or as another interesting application of forensics one may try to find the location in which a photo is taken, by analyzing the signatures of the city power system on the image. In the computer networks field, it will be interesting to find any forbidden application which is run on the network through analyzing the signatures of its header and information using forensics tools.

All in all, by considering the tremendous size of works accomplished in the field of data hiding, to achieve a the real performance gain in this field, one requires the application of very sophisticated ideas. There exists another hot topic in data hiding to find its novel applications which may be extended to the fields such as multimedia protection, self-recovery, hardware, software protection and network monitoring. The application of the watermarks to find malicious modifications is a sub-discipline of a general concept called forensics in which, the aim is to extract maximum possible information from multimedia to combat malicious actions.

References

[1] C. S. Lu, Multimedia security: steganography and digital watermarking techniques for protection of intellectual property, Idea Group Publishing, 2004.

[2] J. Seitz, Digital watermarking for digital media, Information Science Publishing, 2005.

[3] G. C. Langelaar, I. Setyawan, and R.L. Lagendijk, Watermarking digital image and video data: A state-of-the-art overview, IEEE Trans. Signal Process. Magazine, vol. 17, no. 5, pp. 20-46, 2000.

[4] S. Katzenbeisser, and F. A. Petitcolas, Information hiding techniques for steganography and digital watermarking, Artech House, Boston, 2000.

[5] I. J. Cox, M. L. Miller, and J. A. Bloom,. Digital watermarking, first edition, San Francisco: Morgan Kaufmann, 2002.

[6] M. Barni and F. Bartolini, Watermarking systems engineering: Enabling Digital Assets Security and Other Applications, CRC, 2008.

[7] A. B. Watson, Handbook of human perception and performance, in Temporal Sensitivity, K. Boff, L. Kaufmann, and J. Thomas, Eds. New York: Wiley, 1986.

[8] A. B. Watson, M. Taylor, and R. Borthwick, Image quality and entropy masking, Proc. SPIE , Human Vision, Visual Processing, and Digital Display VIII, 1997, vol. 3016, pp. 2-12.

[9] A. B. Watson, J. Y. Yang, J. A. Solomon, and J. Villasenor, Visibility of wavelet quantization noise, IEEE Trans. on Image Process., vol. 6, no. 8, pp. 1164-1175, Oct. 1997.

[10] J. R. Deller, J. H. L. Hansen, and J. G. Proakis, Discrete- Time Processing of Speech Signals, 2nd edition, IEEE Press, 2000.

[11] SQAM - Sound Quality Assessment Material, http://sound.media.mit.edu/mpeg4/audio/sqam/, 2006.

[12] K. Brandenburg, T. Sporer, NMR and masking flag: Evaluation of quality using perceptual criteria, Proceedings of the International Audio Engineering Society Conference on Audio Test and Measurement, pp.169-179, Sept, 1992.

[13] Z. Wang, and A. C. Bovik, Image quality assessment: from error visibility to structural similarity, IEEE Trans. on Image Process., vol. 13, no. 4, pp. 600-612, 2004.

[14] Z. Wang, and A. C. Bovik, A universal image quality index, IEEE Signal Processing Letters, vol. 9, no. 3, pp. 81-84, 2002.

[15] P. Kabal, An Examination and Interpretation of ITU-R BS.1387: Perceptual evaluation of audio quality, Technical Report of Telecom. Signal Process. Lab., version 2, (http://www.tsp.ece.mcgill.ca/), McGill University, 2003.
[16] Q. Cheng, and T. S. Huang, Robust optimum detection of transform domain multiplicative watermarks, IEEE Trans. signal Processing, vol. 51, no. 4, pp. 906-924, 2003.
[17] S. Wu, J. Huang, D. Huang, Y. Q. Shi, Efficiently self-synchronized audio watermarking for assured audio data, IEEE Transmissions on Broadcast., vol.51, no. 1, pp. 69-76, Mar. 2005.
[18] E. T. Lin and E. J. Delp, A review of fragile image watermarks, Proc. Multimedia and Security Workshop on Multimedia Contents, Orlando, pp. 25-29, Oct. 1999.
[19] L. M. Marvel, G. W. Hartwig, and C. Boncelet, Compression compatible fragile and semi fragile tamper detection, Proc. SPIE, vol 39, no 71, 131-139 ,2002.
[20] O. Ekici, B. Sankur, B. Coskun, U. Naci, M. Akcay, Comparative assessment of semi fragile watermarking methods, Journal of Electronic Imaging, vol. 13, no. 1, pp. 209-216, Jan. 2004.
[21] C. Lu and H. M. Liao, Multipurpose watermarking for image authentication and protection, IEEE Trans. Image Process., vol. 10, no. 10, pp. 1579-1592, Oct., 2001.
[22] J. Fridrich, Security of fragile authentication watermarks with localization, Proc. SPIE, vol. 46, no. 75, 691-700, 2002.
[23] G. W. Yu, C. S. Lu, and H. Y. M. Liao, Mean quantization-based fragile watermarking for image authentication, Opt. Eng. vol. 40, no. 7, 1396-1408, 2004.
[24] H. Yuan, and X. P. Zhang, Multiscale fragile watermarking based on the Gaussian mixture model, IEEE Trans. on Image Process., vol. 15, no. 10, pp. 3189-3200, Oct. 2006.
[25] E. T. Lin, C. I. Podilchuk, and E. J. Delp, Detection of image alterations using semi-fragile watermarks, Proc. SPIE, vol. 39, no. 71, pp. 152-163, 2000.
[26] Z. M. Lu, C. H. Liu, D. G. Xu, and S. H. Sun, Semi-fragile image watermarking method based on index constrained vector quantization, Electronic Letter, vol. 39, no. 7, pp. 35-36. Jan. 2003.
[27] D. Zou, Y. Q. Shi, Z. Ni, and W. Su, A semi-fragile lossless digital watermarking scheme, IEEE Trans. on Circuit and Systems for Video Tech., vol. 16, no. 10, pp. 1294-1300, Oct. 2006.
[28] J. Chou, K. Ramchandran, and A. Ortega, High capacity audio data hiding for noisy channels, Proc. of the International Conference on Information Technology: Coding and Computing, pp. 108-111, 2001.
[29] K. Hofbauer and G. Kubin, High-rate data embedding in unvoiced speech, Proc. International Conference on Spoken Language Processing, pp. 176-180, 2006.
[30] D. C. Wu and W. H. Tsai, A steganographic method for images by pixel-value differencing, Pattern Recognition Letters, vol. 24, no. 910, pp. 1613 - 1626, 2003.
[31] X. Zhang and S. Wang, Vulnerability of pixel-value differencing steganography to histogram analysis and modification for enhanced security, Pattern Recognition Letters, vol. 25, no. 3, pp. 331 - 339, 2004.
[32] C. H. Yang, C. Y. Weng, S. J. Wang, and H. M. Sun, Adaptive data hiding in edge areas of images with spatial lsb domain systems,IEEE Trans. on Info. Forensics and Security, vol. 3, no. 3, pp. 488 -497, Sept. 2008.
[33] H. C.Wu, N. I.Wu, C. S. Tsai, and M. S. Hwang, Image steganographic scheme based on pixel-value differencing and lsb replacement methods, Vision, Image and Signal Processing, IEE Proc., vol. 152, no. 5, pp. 611- 615, Oct. 2005.
[34] H. Hering and M. Hagmuller, Safety and security increase for air traffic management through unnoticeable watermark aircraft identification tag transmitted with the VHF voice communication, Proc. of International Conference on Digital Avionic Systems, pp. 202 - 206, 2003.
[35] J. Mielikainen, LSB Matching Revisited, IEEE Signal Processing Letters, vol. 13, no. 5., May, 2006.
[36] S. Sarreshtedari, M. Ghotbi, and S. Ghaemmaghami, One-third probability embedding: Less detectable LSB steganography, Proc. of International Conference on Multimedia and Expo (ICME), pp. 1002-1005, 2009.
[37] X. Li, B. Yang, D. Cheng, and T. Zeng, A generalization of lsb matching, Signal Processing Letters, IEEE, vol. 16, no. 2, feb. 2009.
[38] N. Khademi-kalantari, M. A. Akhaee, and S. M. Ahadi, and S. M. R. Amindavar, Robust multiplicative patchwork method for audio watermarking, IEEE Trans. on Audio, Speech, and Language Processing, vol. 17, no. 6, pp. 1133-1141, 2009.
[39] A. Westfeld, F5-A Steganographic algorithm, in Lecture Notes in Computer Science. Springer, 2001, vol. 2137, pp. 289-302.
[40] J. Fridrich, T. Pevny, and J. Kodovsky, Statistically undetectable JPEG steganography: dead ends challenges, and opportunities, Proc. of 9th workshop on Multimedia & security, New York, NY, USA: ACM, 2007, pp. 3-14.

[41] J. Fridrich and M. Goljan, Images with self-correcting capabilities, Proc. of International Conference on Image Processing, vol. 3, pp. 792-796, 1999.
[42] H. J. He, J. S. Zhang, and F. Chen, Adjacent-block based statistical detection method for self-embedding watermarking techniques, Signal Processing, vol. 89, no. 8, pp. 1557 - 1566, 2009.
[43] S. H. Liu, H. X. Yao, W. Gao, and Y.-L. Liu, An image fragile watermark scheme based on chaotic image pattern and pixel-pairs, Applied Mathematics and Computation, vol. 185, no. 2, pp. 869 - 882, 2007.
[44] V. Mall, K. Bhatt, S. Mitra, and A. Roy, Exposing structural tampering in digital images, Proc. International Conference Signal Processing, Computing and Control (ISPCC), pp. 1-6. 2012.
[45] R. Chamlawi, A. Khan, and I. Usman, Authentication and recovery of images using multiple watermarks, Computers and Electrical Engineering, vol. 36, no. 3, pp. 578 - 584, 2010.
[46] C. W. Yang and J. J. Shen, Recover the tampered image based on vq indexing, Signal Processing, vol. 90, no. 1, pp. 331 - 343, 2010.
[47] A. Cheddad, J. Condell, K. Curran, and P. M. Kevitt, A secure and improved self-embedding algorithm to combat digital document forgery, Signal Processing, vol. 89, no. 12, pp. 2324 - 2332, 2009.
[48] X. Zhang and S. Wang, Statistical fragile watermarking capable of locating individual tampered pixels, Signal Processing Letters, IEEE, vol. 14, no. 10, pp. 727-730, 2007.
[49] X. Zhang and S. Wang, Fragile watermarking with error-free restoration capability, IEEE Transactions on Multimedia, vol. 10, no. 8, pp. 1490-1499, 2008.
[50] X. Zhang and S. Wang, Fragile watermarking scheme using a hierarchical mechanism, Signal Processing, vol. 89, no. 4, pp. 675 - 679, 2009.
[51] X. Zhang, S. Wang, and G. Feng, Fragile watermarking scheme with extensive content restoration capability, in Digital Watermarking, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, vol. 5703, pp. 268-278. 2009.
[52] X. Zhang, S. Wang, Z. Qian, and G. Feng, Self-embedding watermark with flexible restoration quality, Multimedia Tools and Applications, vol. 54, no. 2, pp. 385-395, 2011.
[53] Z. Qian, G. Feng, X. Zhang, and S. Wang, Image self-embedding with high-quality restoration capability, Digital Signal Processing, vol. 21, no. 2, pp. 278 - 286, 2011.
[54] P. Korus and A. Dziech, A novel approach to adaptive image authentication, Proc. International Conference on Image Processing (ICIP), pp. 2765-2768. 2011.
[55] X. Zhang, Z. Qian, Y. Ren, and G. Feng, Watermarking with flexible self-recovery quality based on compressive sensing and composite reconstruction, IEEE Transactions on Information Forensics and Security, vol. 6, no. 4, pp. 1223-1232, 2011.
[56] P. Korus and A. Dziech, Efficient method for content reconstruction with self-embedding,IEEE Transactions on Image Processing, vol. 22, no. 3, pp. 1134-1147, 2013.
[57] D. J. C. MacKay, Fountain codes, Communications, IEE Proceedings, vol. 152, no. 6, pp. 1062-1068, 2005.
[58] S. Sarreshtedari, M. Akhaee, On source channel coding for image tampering protection and self-recovery, submitted to IEEE Transactions on Image Processing, 2013.
[59] A. Said and W. Pearlman, A new, fast, and efficient image codec based on set partitioning in hierarchical trees, IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 3, pp. 243-250, 1996.
[60] S. B. Wicker, Reed-Solomon Codes and Their Applications. Piscataway, NJ, USA: IEEE Press, 1994.
[61] J. Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications, 1st Edition, Cambridge University Press, NY, 2010.
[62] M. Parvaix, L. Girin, Informed Source Separation of Linear Instantaneous Under-Determined Audio Mixtures by Source Index Embedding, IEEE Transactions on Audio, Speech, and Language Processing, vol.19, no.6, pp.1721,1733, Aug. 2011.
[63] P. H. W. Wong, O. C. Au, A capacity estimation technique for JPEG-to-JPEG image watermarking, IEEE Transactions on Circuits and Systems for Video Technology, vol.13, no.8, pp.746,752, Aug. 2003.
[64] E. Zwicker and H. Fastl, Psychoacoustics: Facts and models, 2nd edition, Springer-Verlag, 1999.
[65] V. Schyndel, R. G., A. Z. Tirkel, and C. F. Osborne, A digital watermark, International Conference on Image Processing (ICIP), Austin, pp. 86-90. 1994.
[66] R. Crandall, Some Notes on Steganography, posted on Steganography Mailing List. 1998.
[67] A. Westfeld and A. Pfitzmann, Attacks on steganographic systems, in Proc. 3rd Int. Workshop on Information Hiding, vol. 1768, pp. 61-76. 1999.
[68] J. Fridrich, M. Goljan, and R. Du, Detecting lsb steganography in color, and gray-scale images, IEEE, Multimedia, vol. 8, no. 4, pp. 22-28, Oct.

2001.
[69] A. Latham, JPEG Hide and Seek. 1999. Available: linux01.gwdg.de/ alatham/stego.
[70] N. Provos, Outguess. [Online]. Available: www.outguess.org.
[71] J. Fridrich, M. Goljan, and D. Soukal, Perturbed quantization steganography, Multimedia Syst., vol. 11, no. 2, pp. 98-107, Dec. 2005.
[72] K. S. Wong, X. Qi, and K. Tanaka, A DCT-based Mod4 steganographic method, Signal Processing, vol. 87, pp. 1251-1263, 2007.
[73] C. K. Chan and L. Cheng, Hiding data in images by simple lsb substitution, Pattern Recognition, vol. 37, no. 3, pp. 469 - 474, 2004.
[74] C. H. Yang, Inverted pattern approach to improve image quality of information hiding by lsb substitution, Pattern Recognition, vol. 41, no. 8, pp. 2674 - 2683, 2008.
[75] X. Zhang and S. Wang, Efficient steganographic embedding by exploiting modification direction, Communications Lett., IEEE, vol. 10, no. 11, pp. 781 -783, Nov. 2006.
[76] R. M. Chao, H. C. Wu, C. C. Lee, and Y. P. Chu, A novel image data hiding scheme with diamond encoding, EURASIP J. Inf. Security, vol. 4, 2009.
[77] W. Hong and T. S. Chen, A novel data embedding method using adaptive pixel pair matching, IEEE Trans. on Info. Forensics and Security, vol. 7, no. 1, pp. 176 -184, Feb. 2012.
[78] C. F. Lee, C. C. Chang, and K. H. Wang, An improvement of EMD embedding method for large payloads by pixel segmentation strategy, Image and Vision Computing, vol. 26, no. 12, pp. 1670 - 1676, 2008.
[79] W. Hong, T. S. Chen, and C. W. Shiu, A minimal Euclidean distance searching technique for sudoku steganography, in International Symposium Info. Science and Engineering, vol. 1, pp. 515 -518. 2008.
[80] J. Wang, Y. Sun, H. Xu, K. Chen, H. J. Kim, and S. H. Joo, An improved section-wise exploiting modification direction method, Signal Processing, vol. 90, no. 11, pp. 2954 - 2964, 2010.
[81] B. Chen and G. Wornell, Quantization index modulation: A class of provably good methods for digital watermarking and information embedding, IEEE Trans. Inf. Theory, vol. 47, no. 4, pp. 1423-1443, May 2001.
[82] T. H. Lan, A. H. Tewfik, A novel high-capacity data-embedding system, IEEE Trans. On Image Process., vol. 15, no. 8, , pp. 2431-2440, Aug. 2006.
[83] J. J. Eggers, R. Buml, R. Tzschoppe, and B. Girod, Scalar costa scheme for information embedding, IEEE Trans. Signal Process., vol. 4, no. 51, pp. 1003-1019, Apr. 2003.
[84] R. Zamir, S. Shamai, and U. Erez, Nested linear/lattice codes for structured multi-terminal binning, IEEE Trans. Inf. Theory, vol. 48, no. 6, pp. 1250-1276, Jul. 2002.
[85] J. J. Eggers, J. K. Su, B. Girod, A blind watermarking scheme based on structured codebooks, In Secure Images and Image Authentication, Proc. IEE Colloquium, pp. 4/1-4/6, London, UK, Apr. 2000.
[86] Q. Zhang, and N. Boston, Quantization index modulation using E8 lattice, Proc. of 41th Annual Allerton Conf. on Communication, Control and Computing, Allerton, IL, USA, 2003.
[87] R. Fischer, R. Tzschoppe, and R. Buhamel, Lattice costa schemes using subspace projection for digital watermarking, European Trans. Telecommunications, vol. 15, no. 4, pp. 51-362, Aug. 2004.
[88] A. Abrardo and M. Barni, Informed watermarking by means of orthogonal and quasi-orthogonal dirty paper coding, IEEE Trans. Signal Processing, vol. 53, no. 2, pp. 824-833, 2005.
[89] M. A. Akhaee, M. J. Saberian, S. Feizi, and F. Marvasti, Robust audio data hiding using correlated quantization with histogram based detector, IEEE Trans. on Multimedia, vol 11, no. 5, pp. 834-842, Aug. 2009.
[90] J. J. Eggers, R. Buml, and B. Girod, Estimation of amplitude modifications before SCS watermark detection, Proc. SPIE Security Watermarking Multimedia Contents, vol. 46 no. 75, pp. 387-398, Jan. 2002.
[91] M. L. Miller, G. J. Doerr, and I. J. Cox, Applying informed coding and embedding to design a robust high capacity watermark, IEEE Trans. Image Process., vol. 13, no. 6, pp. 792-807, Jun. 2004.
[92] F. Perz-Gonzalez, C. Mosquera, M. Barni, and A. Abrado, Rational dither modulation: A high rate data-hiding method invariant to gain attacks, IEEE Trans. Signal Process., vol. 53, no. 10, pp. 3960-3975, Oct. 2005.
[93] A. Abrardo, M. Barni, F. Perez-Gonzallez and C. Mosquera, Improving the performance of RDM watermarking by means of trellis coded quantization, IEE Proc. Inf. Security, vol. 153, no. 3, pp. 107-114, Sept. 2006.
[94] P. Guccione, M. Scagliola, Hyperbolic RDM for nonlinear volumetric distortions, IEEE Trans. Inf. Forensics and Security, vol. 4, no. 2, pp. 25-35, March 2009.
[95] F. Perz-Gonzalez, C. Mosquera, Quantization-based data hiding robust to linear-time-invariant filtering, IEEE Trans. Inf. Forensics Security, vol. 3, no. 2, pp. 137-152, June 2008.
[96] M. A. Akhaee, A. Amini, G. Ghorbani, and F. Marvasti, A solution to gain attack on water-

marking systems: Logarithmic homogeneous rational dither modulation, Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1312-1316, 2010.
[97] P. Comesaa and F. Perez-Gonalez, Dither modulation in the logarithmic domain, Proc. of International Workshop in Digital Watermarking (IWDW'07), Guangzhou, China, Dec. 2007.
[98] P. Comesata and F. Perez-Gonzalez, On a watermarking scheme in the logarithmic domain and its perceptual advantages, Proc. of International Conference on Image Processing (ICIP'07), pp. 2036-2039, 2007.
[99] U. Erez and R. Zamir, Achieving (1=2log(1+SNR) on the AWGN channel with lattice encoding and decoding, IEEE Trans. on Inf. Theory, vol. 50, no. 10, pp. 2293-2314, Oct. 2004.
[100] P. Moulin and R. Koetter, Data-hiding codes, IEEE Trans. on Signal Process., vol. 93, no. 12, pp. 2081-2127, Dec. 2005.
[101] R. Tzschoppe, R. Bahml, R. Fischer, A. Kaup, and J. Huber, Additive non-Gaussian attacks on the scalar costa scheme, in Proc. SPIE, San Jose, CA, Jan. 2005.
[102] P. Moulin, and A. K. Goteti, Block QIM watermarking games, IEEE Trans. on Inf. Forensics and Security, vol. 1, no. 3, pp. 293-310, Sept. 2006.
[103] N. K. Kalantari, S. M. Ahadi, A logarithmic quantization index modulation for perceptually better data hiding, Image Processing, IEEE Transactions on , vol.19, no.6, pp.1504,1517, June 2010.
[104] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, Techniques for data hiding, IBM Systems, vol. 35, no. 3, pp. 313-336, 1996.
[105] M. Arnold, Audio watermarking: Features, applications and algorithms, IEEE International Conference Multimedia and Expo, vol. 2, pp. 1013-1016, 2008.
[106] I. K. Yeo and H. J. Kim, Modified patchwork algorithm: A novel audio watermarking scheme, IEEE Trans. on Speech, Audio, and Language Process., vol. 11, no. 4, pp. 381-386, Jul. 2003.
[107] I. K. Yeo, H. J. Kim Generalized patchwork algorithm for image watermarking, Multimedia Systems, vol. 9, no. 3, pp. 261-265, 2003.
[108] H. Malik, R. Ansari, and A. Khokhar, Robust data hiding in audio using allpass filters, IEEE Trans. on Audio, Speech, and Language Process., vol 15, no. 4, pp. 1296-1304, May 2007.
[109] A. Takahashi, R. Nishimura, Y. Suzuki, Multiple watermarks for stereo audio signals using phase-modulation techniques, IEEE Trans. on Signal Process., vol. 53, no. 2 , pp. 806-815, Feb. 2005.
[110] D. Gruhl and W. Bender, Echo hiding, Proc. of Information Hiding Workshop, pp. 295-315, 1996.
[111] H. O. Oh, J. W. Seok, J. W. Hong, and D. H. Youn, New echo embedding technique for robust and imperceptible audio watermarking, Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2011-2014, 2001.
[112] C. Xu, J.Wu, Q. Sun, and K. Xin, Applications of digital watermarking technology in audio signals, J. Audio Eng. Soc., vol. 47, no. 10, Oct. 1999.
[113] B. S. Ko, R. Nishimura, and Y. Suzuki, Time-spread echo method for digital audio watermarking, IEEE Trans. on Multimedia, vol. 7 , no. 2 , pp. 212-221, Apr. 2005.
[114] O. T. C. Chen, W. C. Wu, Highly Robust, Secure, and Perceptual- Quality Echo Hiding Scheme, IEEE Trans. on Audio, Speech, and Language Process., vol. 16, no. 3, pp. 629-638, Mar. 2008.
[115] I. J. Cox, M. L. Miller, and A. L. McKellips, Watermarking as communications with side information, Proceding of the IEEE, 87, pp. 1127-1141, July 1999.
[116] A. B. Watson, J. Hu, and J. F. McGowan, III, DVQ: A digital video quality metric based on human vision, Journal Electronic Imaging, vol. 10, pp. 20-29, Jan. 2001.
[117] A. B. Watson, J. Y. Yang, J. A. Solomon, and J. Villasenor, Visibility of wavelet quantization noise, IEEE Trans. on Image Process., vol. 6, no. 8, pp. 1164-1175, Oct. 1997.
[118] I. J. Cox, J. Kilian, F.T. Leighton, and T. Shamoon, Secure spread spectrum watermarking for multimedia, IEEE Trans. Image Process., vol. 6, no. 12, pp. 1673-1687, 1997.
[119] Q. Cheng and T.S. Huang, An additive approach to transform-domain information hiding and optimum detection structure, IEEE Trans. Multimedia, vol. 3, no. 3, pp. 273-284, 2001.
[120] P. Moulin and A. Ivanovic The zero-rate spread-spectrum watermarking game, IEEE Trans. on Signal Process., vol. 51, no. 4, pp. 1098-1117, Apr. 2003.
[121] H. O. Altun, A. Orsdemir, G. Sharma, and M. F. Bocko, Optimal spread spectrum watermark embedding via a multi-step feasibility formulation, IEEE Trans. Image Process., vol. 18, no. 2, pp. 371-386, Aug. 1999.
[122] L. M. Marvel, C. G. Boncelet, and C. T. Retter, Spread spectrum image steganography, IEEE Trans. on Signal Process., vol. 8, no. 8, pp. 1285-1293, Aug. 1999.
[123] S. P. Maity, and S. Maity, Multistage spread spectrum watermark detection, IEEE Signal Processing Lett., vol. 16, no. 4, Apr. 2009.

[124] M. Barni, F. Bartolini, A. De Rosa, and A. Piva, A new decoder for the optimum recovery of non-additive watermarks, IEEE Trans. Image Process., vol. 10, no. 5, pp. 755-766, 2001.
[125] Q. Cheng, and T. S. Huang, Robust optimum detection of transform domain multiplicative watermarks, IEEE Trans. signal Processing, vol. 51, no. 4, pp. 906-924, 2003.
[126] M. Barni, F. Bartolini, A. De Rosa, and A. Piva, Optimum decoding and detection of multiplicative watermarks, IEEE Trans. on Signal Process., vol. 51, no. 4, pp.1118-1123, 2003.
[127] T. M. Ng, H. Garg, Maximum likelihood detection in image watermarking using generalized gamma model, Proc. of 39th Asilomar Conference on Signals, Systemsand Computer, pp. 1680-1684, 2005.
[128] V. Solachidis, and I. Pitas, Optimal detector for multiplicative watermarks embedded in the DFT domain of non-white signals, EURASIP Journal on Applied Signal Processing, vol. 16, pp. 522-532, 2004.
[129] J. Wang, G. Liu, Y. Dai, and J. Sun, Locally optimum detection for Barni multiplicative watermarking in DWT domain, Signal Processing, vol. 88, pp. 117-130. 2008.
[130] M. N. Do, and M. Vetterli, The contourlet transform: An efficient directional multiresolution image representation, IEEE Trans. on Image Process. vol. 14, no. 12, pp. 2091-2106, 2005.
[131] M. N. Do, and M. Vetterli, Framing pyramids, IEEE Trans. on Signal Process., pp. 23292342, Sep. 2003.
[132] M. A. Akhaee, N. Khademi-Kalantari, and F. Marvasti, Robust Multiplicative Audio and Speech Watermarking Using Statistical Modeling, Proc. of International Conference on Communications (ICC), 2009.
[133] M. A. Akhaee, N. K. Kalantari, F. Marvasti, Robust audio and speech watermarking using Gaussian and Laplacian modeling, Signal Processing, vol. 90, no. 8, pp. 2487-2497, August 2010.
[134] M. A. Akhaee, S. M. E. Sahraeian F. Marvasti, and B. Sankur, Robust scaling-based image multiplicative watermarking technique using maximum likelihood decoder with optimum strength factor, IEEE Trans. on Multimedia, vol 11, no 5, pp. 822-833, Aug. 2009.
[135] M. A. Akhaee, S. M. E. Sahraeian, F. Marvasti, Contourlet-based image watermarking using optimum detector in a noisy environment, IEEE Trans. on Image Process., vol.19, no.4, pp. 967 -980, Apr 2010.
[136] N. K. Kalantari, S. M. Ahadi, M. Vafadust, M., A robust image watermarking in the ridgelet domain using universally optimum decoder, IEEE Trans. on Circuits and Systems for Video Technology, vol.20, no.3, pp. 396 -406, March 2010.
[137] M. A. Akhaee, S. M. E. Sahraeian, F. Marvasti, Universal optimum blind scaling based Watermarking using maximum likelihood decoder, Proc. of International Conference on Image Processing (ICIP), pp. 765-768, 2009.
[138] S. M. E. Sahraeian, M. A. Akhaee, F. Marvasti, Information hiding with optimal detector for highly correlated signals, Proc. of International Conference on Communications (ICC), 2009.
[139] J. J. Harmsen and W. A. Pearlman, Steganalysis of additive-noise modelable information hiding, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conf., vol. 5020, pp. 131-142, 2003.
[140] A. D. Ker, Steganalysis of lsb matching in grayscale images, IEEE Signal Process. Lett., vol. 12, no. 6, pp. 441 - 444, Jun. 2005.
[141] X. Li, T. Zeng, and B. Yang, Detecting lsb matching by applying calibration technique for difference image, Proc. of the 10th ACM workshop on Multimedia and security, pp. 133-138, 2008.
[142] T. Pevny, P. Bas, and J. Fridrich, Steganalysis by subtractive pixel adjacency matrix,IEEE Trans. on Information Forensics and Security, vol. 5, no. 2, pp. 215 -224, Jun. 2010.
[143] H. Farid, Detecting hidden messages using higher-order statistical models, International Conference on Image Processing, vol. 2, pp. 905-908, 2002.
[144] I. Avcibas, N. Memon, and B. Sankur, Steganalysis using image quality metrics,IEEE Transactions on Image Process., vol. 12, no. 2, pp. 221-229, 2003.
[145] F. Huang, B. Li, and J. Huang, Attack lsb matching steganography by counting alteration rate of the number of neighborhood gray levels, Proc. of International Conference on Image Processing, pp. 401-404, 2007.
[146] Y. Wang and P. Moulin, Optimized feature extraction for learning-based image steganalysis, IEEE Trans. on Info. Forensics and Security, vol. 2, no. 1, pp. 31 -45, Mar. 2007.
[147] T. Pevny, P. Bas, and J. Fridrich, Steganalysis by subtractive pixel adjacency matrix, IEEE Transactions on Info. Forensics and Security, vol. 5, no. 2, pp. 215-224, 2010.
[148] Y. Q. Shi, C. Chen, and W. Chen, A markov process based approach to effective attacking jpeg steganography, in Information Hiding. Springer, pp. 249-264, 2007.
[149] C. Chen and Y. Shi, Jpeg image steganalysis utilizing both intrablock and interblock correlations, Proc. of International Symposium on

Circuits and Systems, pp. 3029-3032, 2008.
[150] T. Pevny and J. Fridrich, Merging Markov and DCT features for multi-class JPEG steganalysis, Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents, vol. 3, pp. 1117-1126 2007.
[151] J. Kodovsky , J. Fridrich, Steganalysis in high dimensions: fusing classifiers built on random subspaces, Proc. of SPIE Media Watermarking, Security, and Forensics III, pp. 23-26, 2011.
[152] Y. Shi, G. Xuan, D. Zou, J. Gao, C. Yang, Z. Zhang, P. Chai, W. Chen, and C. Chen, Image steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image, and neural network, Proc. of International Conference on Multimedia and Expo (ICME), 2005.
[153] C. C. Chang and C. J. Lin, Libsvm: a library for support vector machines, 2001, software available at http://www. csie.ntu.edu.tw/ cjlin/libsvm.
[154] J. Kodovsky, J. Fridrich, and V. Holub, Ensemble classifiers for steganalysis of digital media, IEEE Trans. on Info. Forensics and Security, vol. 7, no. 2, pp. 432-444, 2012.

Mohammad Ali Akhaee received his B.S. degree in both electronics and communications engineering from Amirkabir University of Technology, and the M.S. and Ph.D. degrees in communication systems from Sharif University of Technology in 2005 and 2009, respectively. He has been awarded governmental Endeavour research fellowship from Australia in 2010. He is an author/coauthor of more than 35 papers and holds an Iranian patent. He served as the technical program chair of EUSIPCO11-13. Dr. Akhaee serves as a faculty member at the College of Eng., University of Tehran, Tehran, Iran. His research interests include multimedia security, cryptography, network security, and statistical signal processing.

Farokh Marvasti received his B.S., M.S. and Ph.D. degrees all from the Renesselaer Polytechnic Institute in 1970, 1971 and 1973, respectively. He has worked, consulted and taught in various industries and academic institutions since 1972. Among which are Bell Labs, University of California Davis, Illinois Institute of Technology, University of London, Kings College. He was one of the editors and associate editors of IEEE Trans on Communications and Signal Processing from 1990-1997. He has about 60 journal publications and has written several reference books. His last book is entitled Nonuniform Sampling: Theory and Practice and published by Kluwer in 2001. He was also a guest editor of the Special Issue on Nonuniform Sampling for the Sampling Theory and Signal and Image Processing journal. Dr. Marvasti is currently a professor at Sharif University of Technology and the director of the Advanced Communications Research Institute (ACRI).

The ISC Int'l Journal of Information Security
ISeCure

January 2013, Volume 5, Number 1 (p. 111)
http://www.isecure-journal.org

Persian Abstract

بررسی اصول، روش‌ها و کاربردهای سیستم‌های پنهان‌سازی اطلاعات

محمدعلی اخایی[۱] و فرخ مروستی[۲]

[۱] عضو هیأت علمی دانشکده برق و کامپیوتر، پردیس دانشکده‌های فنی دانشگاه تهران

[۲] عضو هیأت علمی دانشکده برق و مدیر پژوهشکده مخابرات نظری، دانشگاه صنعتی شریف

با گسترش روزافزون دنیای رقمی و راه‌های آسان تبادل این گونه اطلاعات، پنهان‌نگاری رقمی مورد توجه بسیاری قرار گرفته است. در این مقاله سعی شده موضوع نهان‌نگاری و نشان‌گذاری مورد تحلیل و بررسی قرار گیرد. در این راستا ابتدا مفاهیم اولیه در پنهان‌سازی بیان گردیده و در ادامه نیازمندی‌ها و کاربردهای آن ارائه می‌شود. برای طراحی یک سیستم نهان‌نگاری کارا دانستن مفاهیم پایه‌ای از درج و استخراج نهان‌نگاره لازم و ضروری است. داشتن اطلاعات کافی از سیگنال میزبان (نظیر صحبت، صوت، تصویر و ویدئو) و آشنا بودن با ساختار گیرنده نهایی آن که چشم و گوش انسان است به طراح کمک می‌کند که عمل درج را به بهترین نحو انجام دهد. در این حالت نهان‌نگاره‌ی بیشتر با اثرات ظاهری/آماری کمتر در سیگنال میزبان جاسازی می‌گردد. در این خصوص روش‌های بسیاری معرفی و پیاده‌سازی شده است. البته شایان ذکر است که اگرچه الگوریتم‌های بسیار زیادی تاکنون ارائه گردیده اما به سادگی می‌توان این روش‌ها را در چند گروه کلی طبقه‌بندی کرد. در این مقاله پس از معرفی اجمالی سیستم بینایی و شنوایی انسان، روش‌های نخستین به همراه نسخه‌های پیشرفته و بهبودیافته آن‌ها معرفی می‌گردد. سپس یک مقایسه جامع بین این الگوریتم‌ها صورت پذیرفته تا یک طراح بتواند برحسب نیازها و امکاناتی که در اختیار دارد در مورد انتخاب الگوریتم و تنظیم پارامترهای آن تصمیم‌گیری کند. در ادامه با در نظر گرفتن مساله زندانی و زندانبان و سوء نیت زندانبان برای کشف یا خراب کردن ارتباط بین زندانیان، حملات به دو صورت عمدی و غیرعمدی تقسیم‌بندی شده است. این حملات که همانند یک سنگ محک برای ارزیابی کیفیت و عملکرد روش‌ها در نظر گرفته می‌شود در این‌جا معرفی و دسته‌بندی شده است.

واژه‌های کلیدی: امنیت، نهان‌نگاری، نشان‌گذاری، نهان‌کاوی، پایداری.