AIN SHAMS UNIVERSITY
FACULTY OF COMPUTER AND INFORMATION SCIENCES
COMPUTER SCIENCE DEPARTMENT

USING DIGITAL WATERMARKING FOR IMAGE VERIFICATION AND AUTHENTICATION

This thesis is submitted as a partial fulfillment of the requirements for the degree of **Master of Computer and Information Sciences**.

By

Applicant **Salma Hamdy Mohammad El-Sayed**
Computer Science Department
Faculty of Computer and Information Sciences
Ain Shams University

Under Supervision of

Prof. Dr. M. Hatem El-Ayadi
Professor of Signals and Systems,
Faculty of Computer and Information Sciences,
Ain Shams University.

Prof. Dr. Taha I. El-Areef,
Professor of Computer Science,
Faculty of Computer and Information Sciences,
Ain Shams University.

Cairo 2004

DEDICATION

To Hamdy and Hanaa.
To Wessam.
And to Sara.

For their patience and support during the elaboration of this thesis.

ACKNOWLEDGMENTS

I acknowledge my deep gratitude to ALLAH the most beneficent and most merciful, who helped me complete this work on a level I hope it will please the reader.

I must also thank *Prof. Dr. Mohamed Said Abd El-Wahab*, Dean of the faculty of Computer and Information Sciences, Ain shams University, for his support and concern.

Special thanks are due to *Prof. Dr. Mohamed Hatem El-Ayadi*, Professor of Signals and Systems, at the Faculty of Computer and Information Sciences, Ain Shams University, who constantly guided me in elaborating this thesis, assisted me in understanding and analyzing problems, and continuously provided support and valuable comments.

Also my thanks are due to *Prof. Dr. Taha Ibrahim El-Areef*, Professor of Computer Science, at the Faculty of Computers and Information Sciences, Ain Shams University, who supported me in my work.

I would like also to thank others who have helped me out in a variety of ways including *Prof. Dr. Said El-Ghonaimy*, at the Faculty of Computer and Information Sciences, Ain Shams University, *Prof. Dr. Laila El-Masry,* Dean of Graduate Studies and Professor of Urban Design, Faculty of Urban and Regional Planning, Cairo University, and my colleagues *Mohamed Abd El-Megeed*, *Safaa El-Sayed*, and *Islam Hegazy*.

ABSTRACT

The presented work focuses on using digital watermarking for image verification and authentication. This entails elaborate presentation, implementation, performance evaluation, and comparison of known watermarking techniques. Suggestion of possible improvement, enhancements, or upgrades of studied techniques falls within the scope of interest of the thesis. All investigations are based on extensive Monte-Carlo simulations.

We study closely four typical single-bit watermarking systems that differ in the embedding algorithm, the detection algorithm, or both. Through elaborate presentation, implementation, and performance evaluation, we show the positive and the negative merits of each system. We address important issues such as selection of an embedding strength satisfying the fidelity requirement, evaluation of potential detection performance and analysis of trade-offs among fidelity, robustness and detection performance of each considered system.

The thesis also investigates multibit-watermarking systems and considers methods of representing messages with watermarks. We elaborate on two approaches to modulating messages. These are direct message modulation, and multiplexing. We study an eight-bit code division multiplexing watermarking system with blind embedding and linear correlation detection. We show that there is a trade-off between achieving good performance and sustaining an acceptable distortion. We show also how error correction codes can much reduce the error probability of the system. Moreover, we investigate the effect of informed embedding on system performance. We also investigate a direct message modulation DMM watermarking system with blind embedding and linear correlation detection. We suggest the use of whitening to improve detection performance if the spatial

covariance of images is known. We also develop approaches to implement and to examine the effect of informed embedding and informed modulation on the performance of the system. Results show that informed embedding and modulation improves detection performance over informed embedding but is outperformed by whitening.

We also focus on the use of digital watermarking for exact content authentication. A blind embedder that uses ordinary addition, followed by rounding and clipping does not allow exact recovery of the original image. One solution is to let the embedder apply only rounding and modulo addition. However, this solution introduces a 'salt-and-pepper' noise that degrades the detection performance. We suggest a *gray-level unwrapping algorithm* to alleviate the bad impact of the modulo addition on the detection of the modulo-added erasable watermarks. Computer simulations demonstrate the efficacy of the suggested unwrapping algorithm.

We conclude the thesis with a discussion of watermark immunity, supported by an investigation of three algorithms for watermark removal, two published and one contributed by the author, to estimate a watermark, created using the single-bit blind embedder. Evaluation of the relative perceptual distortion of the image obtained after removal of the estimated watermark confirmed good estimation quality and a successful removal of the mark.

CONTENTS

Chapter I
INTRODUCTION

Usage of digital media has witnessed a tremendous growth during the last decades, as a result of their notable benefits in efficient storage, ease of manipulation and transmission. In particular, the Internet has become an excellent distributing system that is inexpensive, eliminates warehousing and stock, and delivery is almost instantaneous [1]. All of these features make digital media vulnerable to copyright infringement, tampering and unauthorized distribution. The original information creators and distributors may prohibit their products to be copied at all. And if distributing copies is allowed, they want to be compensated every time their work is used. Furthermore, they want to make sure these products are not used in an improper way (e.g. modified without their permission). These concerns about protecting digital media copyright and proving ownership of distributed products have triggered wide researches to find ways of hiding copyright notices into digital data. This led to the so-called digital watermarking; a branch of information-hiding technology. First publications, related to digital watermarking date back to 1979. However, it was only in 1999 that it started to gain large international interest.

In information hiding technology, we refer to the media in which we hide information (song, video, or picture) as a *Work* and the set of all possible Works of a particular type as *content*. The original object (with no hidden messages) is called the *Cover Work*. Information hiding is a general term that encompasses a wide range of problems beyond that of embedding messages in content. The term “hiding” here can refer to either making the information imperceptible (as in watermarking) or keeping the existence of the information secret. Information hiding has been receiving significant attention in the digital media community, and a number of

techniques that try to address the problem by hiding appropriate information (e.g. copyright or authentication data) within digital media have been proposed. Information hiding is like an umbrella that covers several branches; steganography, watermarking, cryptography, anonymity, to mention some.

Steganography [1-2] is the art of concealed communication by hiding messages in seemingly innocuous objects. The term is derived from the Greek words *steganos*, which means "covered", and *graphia*, which means "writing". The main requirement in steganographic applications is the secrecy of the message which is of prime value to the recipient. The hidden information is not necessarily related to the object in which it is hidden.

On the other hand, watermarking is the act of imperceptibly embedding in a Work, a message about that Work [1, 3]. As an information hiding subdiscipline, watermarks (messages) are meant to be hidden within some Cover Work. But as opposed to steganography, the existence of the mark need not be secret and the watermark carries some information about the Work it is embedded within (e.g. owner or source of data). However, differences between steganography and watermarking are rather philosophical than materialistic. In steganography, the embedded message is of prime interest to the recipient, while in watermarking applications, it serves the owner (sender) of the Work more. Thus, we can conclude that both technologies do not differ in implementation as much as they do in objective and hence in requirements and design of technical solutions.

The concept of digital watermarking came up while trying to solve the problems related to the management of intellectual property in digital media. It was raised as an alternative or complement to cryptography, as a technology that can protect content even after it is decrypted. Watermarking has the potential to fulfill this need because it places information within the content where it is never removed during normal usage. Decryption, re-encryption, compression, digital-to-

analog conversion, and file format changes—a watermark can be designed to survive all of these processes [1]. Thus it is used to give proof of ownership of digital data by embedding copyright statements and hence serving for verification of the data. In short, watermarking has three important features that distinguish it from other techniques. Watermarks are imperceptible, inseparable from the Work, and they undergo the same transformations as the Work.

1.1 Applications of Watermarking

Watermarking is used in a wide variety of applications. Below, we highlight the most important ones.

1. *Broadcast Monitoring [4-5]:* It is an application of watermarking in which a detector monitors radio or television broadcasts, searching for watermarks. This can be used to verify that advertisements are properly broadcast, and that royalties are properly paid.
2. *Owner Identification and Proof of Ownership [6]:* For the copyrighted Work, the exact form of copyright notices is important. For visual Works, it must be either "Copyright *data owner*", "© *data owner*", or "Copr. *Data owner*". For sound recordings, the copyright notice takes the similar form "(p) *data owner*" and must be placed on the surface of the physical media, the label, or on the packaging. These textual copyright notices are very weak to be used to identify the owner of the Work as they are easily removable from documents, and thus newly made copies of the Work will not be copyrighted anymore. Even if they are assumed to be protected by copyright, it is difficult to know the identity of the owner whom we must contact for permission of copying. Furthermore, text copyright notices may cover part of the image and reduce its aesthetics. In the case of audio data, the copyright notice is never copied to the newly made copy as it is marked on the physical media (tape or CD). Therefore, a means of identifying or verifying the owner and maintaining the copyright of the

Work is required to be visually ineffective and inseparable from the Work. Because watermarking places an imperceptible mark within the data, it seems obvious to use it for owner identification. In addition, we may also use watermarks to prove ownership not just identify the owner. Because textual notices are easily forged or removed, they are not reliable for proving ownership of data.

3. *Transaction Tracking [1]:* Watermarks can be used to indicate transactions that have occurred in the history of the Work. In a typical transaction tracking scenario, the owner places a different watermark in each copy, he wishes to distribute. These marks identify the recipient (first legal user) of the copy. Afterwards, if any of the distributed copies were misused somehow, the owner can easily track the traitor who is responsible of illegally re-distributing his copy. The owner obtains one of the illegal copies and simply extracts the watermark from it. Tracking adversaries can also be done using visible marks. For example, identification codes printed in the background of text documents with each copy of the document has a different code. Of course, visible marks can easily be forged or removed completely.
4. *Content Authentication [7-9]:* Because digital data is very easy to tamper with in ways that are difficult to detect, there is a strong need for a method that can inform of any modifications to which a Work may have been exploited. In other words, there is a strong need for a method that can verify the integrity of the cover Work. Cryptography is one of the common ways for authenticating a Work by using *digital signatures*. A digital signature is an encrypted summary of the message that is computed and transmitted along with the Work it verifies. Unfortunately, it is easy to lose the signature in normal usage and hence we might want to

embed the signature directly into the Work using watermarking.

Fragile watermarking is a technique to insert a mark for the purpose of authentication. The mark will be altered when the host data is manipulated. Detecting changes that may occur in digital data is important in application where any slight change in the data must be noticed, as in news reports, medical archiving, or malicious tampering of photographic evidence in a criminal trial. Another need for content authentication is in e-commerce where a customer buys digital images which are sent to her over the network. In such case, she wants to make sure she receives the original intact image sent to her by the seller. Here, we use watermarking to verify the integrity of the content plus checking for the owner too. We discuss content authentication in more detail in chapter IV.

5. *Device Control [10-12]:* In this category of applications, devices react to watermarks they detect in content. Copy control falls within this category. We aim to prevent people from making illegal copies of copyrighted content. One way of doing this is through cryptography. By encrypting Works and only providing the keys to legitimate users, we make these Works unusable for those who don't have a key for decryption. This method doesn't completely solve the problem because cryptography can protect a Work from illegal copying until it is decrypted, afterwards there is no way that encryption can prevent a legitimate user from making and distributing illegal copies. Because watermarks are embedded into the content itself, they are present in every representation of that content and therefore might provide a better method for implementing copy control. Even after the encrypted Work is decrypted, the watermark still exists and the recording device prohibits

recording whenever it detects a `never-copy` watermark.

1.2 Preliminary Technical Background and Terminology

A generic watermarking system is shown in Figure 1.1. The suitability of a given watermarking system for a given application may be judged in terms of the following features of the system.

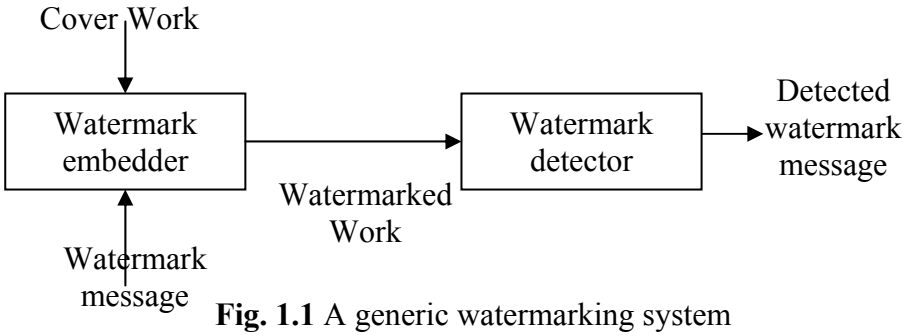

Fig. 1.1 A generic watermarking system

A. Data Payload

Data payload is the number of information bits, a watermark encodes within a unit of time or within a Work. In other words, it is the amount of information conveyed by the watermark. A watermark that encodes N bits is referred to as an N-bit watermark. Such a system can be used to embed any one of 2^N different messages. Different applications require different data payloads. For example, copy control applications require fewer bits of information than broadcast monitoring applications that require more bits to identify all commercials. According to the value of N, we can classify watermarking systems into single-bit and multi-bit systems.

B. Robustness

In most of the above cited applications, and as opposed to steganography, watermarking has the additional requirement of "robustness" which refers to how well the watermark survives usual signal processing operations (both in transmission and reception) as well as usual degradation effects.