CHAPTER ONE

INTRODUCTION

With the rapid growth of Internet and networks technique, multimedia data transforming and sharing is common to many people. Multimedia data is easily copied and modified, so necessity for copyright protection is increasing. It is the imperceptible marking of multimedia data to "brand" ownership. Digital watermarking has been proposed as technique for copyright protection of multimedia data. Digital watermarking invisibly embeds copyright information into multimedia data. Thus, digital watermarking has been used for copyright protection, finger printing, copy protection and broadcast monitoring. Indeed, a watermarking algorithm requires both invisibility and robustness, which exist in a trade-off relation. Thus good watermarking algorithm must be satisfied the requirements.

The process of digital watermarking involves the modification of the original multimedia data to embed a watermark containing key information such as authentication or copyright codes. The embedding method must leave the original data perceptually un-changed, yet should impose modifications which can be detected by using an appropriate extraction algorithm. Common types of signals to watermark are images, music clips and digital video. The application of digital watermarking to still images is concentrated here. The major technical challenge is to design a highly robust digital watermarking technique, which discourages copyright infringement by making the process of watermarking removal tedious and costly.

The advent of the Internet has resulted in many new opportunities for the creation and delivery of content in digital form. Applications include electronic advertising, real-time video and audio delivery, digital repositories and libraries, and Web publishing. An important issue that arises in these applications is the protection of the rights of all participants.

It has been recognized for quite some time that current copyright laws are inadequate for dealing with digital data. This has led to an interest towards developing new copy deterrence and protection mechanisms. One such effort that has been attracting increasing interest is based on digital watermarking techniques.

Digital Watermarking describes methods and technologies that hide information, for example a number or text, in digital media, such as images, video. The embedding takes place by

manipulating the content of the digital data, which means the information is not embedded in the frame around the data. The hiding process has to be such that the modifications of the media are imperceptible. For images this means that the modifications of the pixel values have to be invisible.

A digital watermark is a message which is embedded into digital content (video, images or text) that can be detected or extracted later.

Moreover, in image the actual bits representing the watermark must be scattered throughout the file in such a way that they cannot be identified and manipulated. Watermarking is the insertion of imperceptible and inseparable information into the host data for data security & integrity.

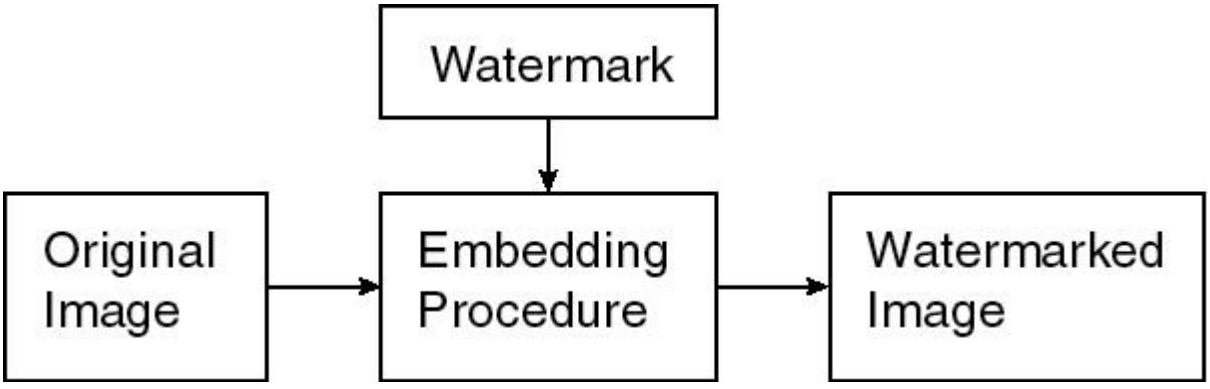
They are characterizing patterns, of varying visibility, added to the presentation media as a guarantee of authenticity, quality, ownership, and source.

However, in digital watermarking, the message is supposed **not** to visible (or at least not interfering with the user experience of the content), but (only) electronic devices can retrieve the embedded message to identify the code.

Another form of digital watermarking is known as steganography, in which a message is hidden in the content without typical citizens or the public authorities noticing its presence.

Only a limited number of recipients can retrieve and decode the hidden message. Unlike a traditional watermark on paper, which is generally visible to the eye, digital watermarks can be made invisible or inaudible. They can, however, be read by a computer with the proper decoding software.

Figure shows the general watermarking embedding procedure. In an original image with the help of embedding procedure watermark is embedded and then we get a watermarked image.



The most common example of watermark is an Indian currency.

CHAPTER TWO
LITERATURE REVIEW

HISTORY

More than 700 years ago, watermarks were used in Italy to indicate the paper brand and the mill that produced it. By the 18th century watermarks began to be used as Anti-counterfeiting measures on money and other documents.

The term watermark was introduced near the end of the 18th century. It was probably given because the marks resemble the effects of water on paper.

The first example of a technology similar to digital watermarking is a patent filed in 1954 by Emil Hem Brooke for identifying music works .In 1988, Komatsu and Tominaga appear to be the first to use the term “digital watermarking”.

Digital watermarking is a technology for embedding various types of information in digital content. In general, information for protecting copyrights and proving the validity of data is embedded as a watermark. A digital watermark is a digital signal or pattern inserted into digital content. The digital content could be a still image, an audio clip, a video clip, a text document, or some form of digital data that the creator or owner would like to protect. The main purpose of the watermark is to identify who the owner of the digital data is, but it can also identify the intended recipient.

Why do we need to embed such information in digital content using digital watermark technology? The Internet boom is one of the reasons. It has become easy to connect to the Internet from home computers and obtain or provide various information using the World Wide Web (WWW). All the information handled on the Internet is provided as digital content. Such digital content can be easily copied in a way that makes the new file indistinguishable from the original. Then the content can be reproduced in large quantities.

For example, if paper bank notes or stock certificates could be easily copied and used, trust in their authenticity would greatly be reduced, resulting in a big loss. To prevent this, currencies and stock certificates contain watermarks. These watermarks are one of the methods for preventing counterfeit and illegal use. Digital watermarks apply a similar method to digital content. Watermarked content can prove its origin, thereby protecting copyright. A watermark

also discourages piracy by silently and psychologically deterring criminals from making illegal copies.

PRINCIPLE OF DIGITAL WATERMARKS

A watermark on a bank note has a different transparency than the rest of the note when a light is shined on it. However, this method is useless in the digital world. Currently there are various techniques for embedding digital watermarks. Basically, they all digitally write desired information directly onto images or audio data in such a manner that the images or audio data are not damaged. Embedding a watermark should not result in a significant increase or reduction in the original data. Digital watermarks are added to images or audio data in such a way that they are invisible or inaudible unidentifiable by human eye or ear. Furthermore, they can be embedded in content with a variety of file formats. Digital watermarking is the content protection method for the multimedia era.

CHAPTER THREE

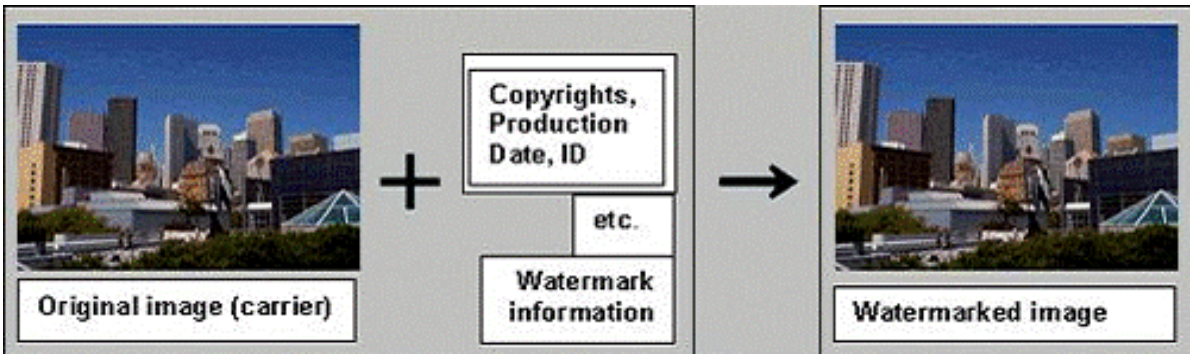
METHODOLOGY

MATERIALS SUITABLE FOR WATERMARKING

Digital watermarking is applicable to any type of digital content, including still images, animation, and audio data. It is easy to embed watermarks in material that has a comparatively high redundancy level ("wasted"), such as color still images, animation, and audio data; however, it is difficult to embed watermarks in material with a low redundancy level, such as black-and-white still images. To solve this problem, we developed a technique for embedding digital watermarks in black-and-white still images and a software application that can effectively embed and detect digital watermarks.

STRUCTURE OF A DIGITAL WATERMARK

The structure of a digital watermark is shown in the following figures.



The material that contains a digital watermark is called a carrier. A digital watermark is not provided as a separate file or a link. It is information that is directly embedded in the carrier file. Therefore, simply viewing the carrier image containing it cannot identify the digital watermark. Special software is needed to embed and detect such digital watermarks. Kowa 's SteganoSign is one of these software packages. Both images and audio data can carry watermarks. A digital watermark can be detected as shown in the following illustration.

THE IMPORTANCE OF DIGITAL WATERMARKS

The Internet has provided worldwide publishing opportunities to creators of various works, including writers, photographers, musicians and artists. However, these same opportunities provide ease of access to these works, which has resulted in pirating. It is easy to duplicate audio and visual files, and is therefore probable that duplication on the Internet occurs without the rightful owners' permission. An example of an area where copyright protection needs to be enforced is in the on-line music industry. The Recording Industry Association of America (RIAA) says that the value of illegal copies of music that are distributed over the Internet could reach $2 billion a year.

Digital watermarking is being recognized as a way for improving this situation. RIAA reports that "record labels see watermarking as a crucial piece of the copy protection system, whether their music is released over the Internet or on DVD-Audio". They are of the opinion that any encryption system can be broken, sooner or later, and that digital watermarking is needed to indicate who the culprit is. Another scenario in which the enforcement of copyright is needed is in newsgathering. When digital cameras are used to snapshot an event, the images must be watermarked as they are captured. This is so that later, image's origin and content can be verified. This suggests that there are many applications that could require image watermarking, including Internet imaging, digital libraries, digital cameras, medical imaging, image and video databases, surveillance imaging, video-on-demand systems, and satellite-delivered video.

TYPES OF WATERMARK

Digital watermarks are of four types:

- Visible
- Invisible
- Public, and
- Fragile

A **visible watermark** typically consists of a conspicuously visible message or a company logo indicating the ownership of the image. Any removal or tampering with the logo would break the copyright agreement. Another way is to write the copyright notice and other information into an

extra couple of lines within the image file. The extra lines can be removed from the image, without detriment to the image quality and content, but this again would break the copyright agreement of the image.

A visible watermark was added to the image to create this image.

The watermark is a repeating image of a bird in flight. Visible watermarks often look as if they were "embossed" onto the original image, as illustrated here.

An **invisible watermarked** image appears very similar to the original. The existence of an invisible watermark can only be determined using an appropriate watermark extraction or detection algorithm. It can be detected by an authorized agency only. Such watermarks are used for content and/or author authentication and for detecting unauthorized copier.

To insert the invisible watermark, we first supplied a special password, or a "key," for security. The key may be used to recover the message contained in the invisible watermark, and to determine whether the image was altered since the invisible watermark was inserted.

Public watermark such a watermark can be read or retrieved by anyone using the specialized algorithm. In this sense, public watermarks are not secure.

Fragile watermarks are also known as tamperproof watermarks. Such watermarks are destroyed by data manipulation. Fragile watermark is a mark which is (highly) sensitive to a modification .A fragile watermarking scheme should be able to detect any change in the signal and identify where it has taken place and possibly what the signal was before modification. It serves at proving the authenticity of a document.

DIFFERENT ATTRIBUTES ASSOCIATED WITH WATERMARKING

The characteristics of a watermarking algorithm is normally tied to the application is designed for. The most important properties of any digital watermarking techniques are robustness, security, imperceptibility, complexity, and verification. The following merely explain the words used in the context of watermarking.

Imperceptibility:-In watermarking, we traditionally seek high fidelity, i.e. the watermarked work must look or sounds like the original. Whether or not this is a good goal is a different discussion. Imperceptibility means the watermark is not seen by the human visual system.

Robustness:- By "robust" we mean the capability of the watermark to resist manipulations of the media, such as lossy compression (where compressing data and then decompressing it retrieves data that may well be different from the original, but is close enough to be useful in some way), scaling, and cropping, just to enumerate some. Robustness is defined as if the watermark can be detected after media (normal) operations such as filtering, lossy compression, color correction, or geometric modifications. In some cases the watermark may need to be fragile. "Fragile" means that the watermark should not resist tampering, or would resist only up to a certain, predetermined extent.

It is more a property and not a requirement of watermarking. The watermark should be able to survive any reasonable processing inflicted on the carrier (carrier here refers to the content being watermarked).

Security:-The watermarked image should not reveal any clues of the presence of the watermark, with respect to un-authorized detection, or (statistical) indefectibility or unsuspicious (not the

same as imperceptibility). Security means the embedded watermark cannot be removed beyond reliable detection by targeted attacks.

Complexity is described as the effort and time required for watermark embedding and retrieval**.**

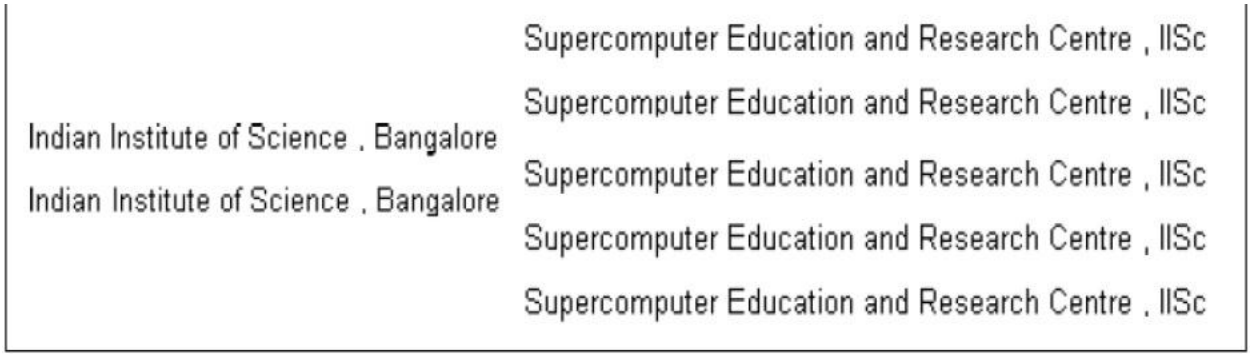
Verification is a procedure whereby there is a private key or public key function

WATERMARKING TECHNIQUES

Numerous methods for watermarking exist and they can be classified based on various parameters like the embedding algorithms and the detection algorithms used. We shall study them based on the data they watermark.

Watermarking for text: Two methods have been proposed for watermarking text, namely –

Word space coding: In this method, the spacing between words is varied by horizontally shifting the locations of the words within text lines, thus watermarking the document uniquely. This however is applicable only to documents in which variable spacing between adjacent words is possible. Also because interword spacing is often varied to format the document, the original document is necessary to verify the watermark. An example of word space coding is given below.

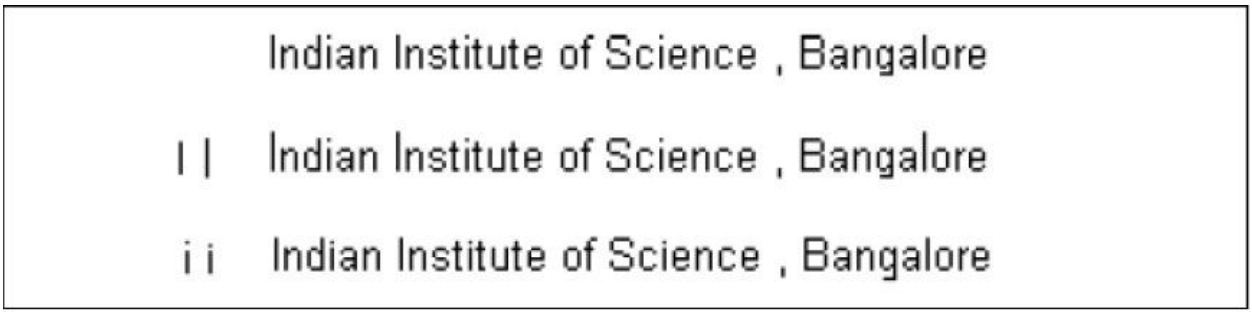


This figure illustrates word shift encoding. Note the additional space of one pixel between ‘n’ of ‘Indian’ and ‘I’ of ‘Institute’. Such minor variations are not perceptible to the human eye. They are recognized only on close comparison**.**

Line Shift Coding: The same concept of word space coding is used, only it is applied in the vertical domain. Text lines are shifted vertically to watermark the image uniquely. If the original image has uniform line spacing, then verification of the watermark can be accomplished without the original. An example of line shift coding is shown in above figure. This illustrates the

technique of line shift encoding (see lines 2 and 3). Notice that line shift encoding is more evident than word shift encoding because of the length of the lines.

Feature Coding: This method alters the specifications of particular characters by either lengthening or shortening them. It provides numerous possibilities of watermarking. However, for decoding, the original image is necessary, or rather, the original characteristics of the characters are required.



In the second line the length of the letter 'I' has been imperceptibly increased; it is evident only on close comparison. In the third line the distance between the dot and the line in the character 'i' has been reduced. Such features are not noticeable until and unless specifically looked for.

ii) **Watermarking for Images:** Image data is binary in nature i.e. all image files are a combination of zeros and ones. Thus they are easily manipulated, processed, and tampered with. Hence, robust and standard watermarks for image files are a challenge. Images, being digital in nature, can be visualized in two forms – either they can be thought of as a two-dimensional array of zeros and ones or they can be considered to be the digital representation of an analog signal. Watermarking techniques for images are based on these methods of representation.

Spatial Domain Watermarking: The image is considered to be a two-dimensional array and manipulating certain pixels based on their spatial locations in the array embeds the watermark. Spatial domain watermarking slightly modifies the pixels of one or two randomly selected subsets of an image. Modifications might include flipping the low-order bit of each pixel. However, this technique is not reliable when subjected to normal media operations such as filtering or lossy compression. Techniques may be as simple as flipping the least significant bit (LSB) or may be a complex superposition of watermarking symbols over an area of the image. In the latter technique, a lot of flexibility exists in terms of placement, size, and intensity of the watermark.

FREQUENCY DOMAIN WATERMARKING: Frequency domain watermarking technique is also called transform domain. The image is considered to be a sampled-digitized data of an analog signal. The analog signal can be obtained by various transforms like the DCT (Discrete Cosine Transform), DFT (Discrete Fourier Transform), FFT (Fast Fourier Transform) etc. and hence represented as a series of signals of increasing frequencies. The watermark can now be embedded in the coefficients of the various frequency components. The watermark is not embedded in the high frequency components, as they are usually lost on compression or scaling. The watermark is applied to the whole image so as not to be removed during a cropping operation. However, it is more difficult to decode a watermark applied in the frequency domain. Verification can be difficult since this watermark is applied indiscriminately across the whole image.

DETECTION OF WATERMARKS

It needs to be emphasized that a watermark can be defeated in two ways – one, by removing the watermark from the original data and two, by proving it to be unreliable i.e. identifying a watermark when there is none. If the latter can be achieved then the watermark cannot be proved in a legal battle. Hence detection of watermarks needs to be even more reliable than their embedding. Detection algorithms are dependent on or are derived from embedding algorithms. Hence, a rigorous and detailed classification is ruled out.

Detection algorithms can be divided into two broad categories –those which need the original unwater marked data, and those which do not. The former generally makes a byte-by-byte Comparison and arrives at a decision after allowing for a reasonable amount of error. Say, for example, the image has been watermarked by increasing the intensity of certain pixels in the Original unwater marked image by a known factor K, the average intensity of these pixels in the original image and the test image are compared. If they differ by more than 0.7K, the image is watermarked while if they differ by less than 0.3K, the image is not watermarked. The in-between range of 0.3K to 0.7K is a gray area and needs a more detailed analysis of the conditions undergone by the image. It should be said that the figures of 0.3K and 0.7K are an offhand estimate. They need to be arrived at after mathematical estimations.

APPLICATIONS OF WATERMARKING

The first applications that came to mind were related to copyright protection of digital media. In the past duplicating art work was quite complicated and required a high level of expertise for the counterfeit to look like the original.

However, in the digital world this is not true. Now it is possible for almost anyone to duplicate or manipulate digital data and not lose data quality. Similar to the process when artists creatively signed their paintings with a brush to claim copyrights, artists of today can watermark their work by hiding their name within the image.

Hence, the embedded watermark permits identification of the owner of the work.There are a number of possible applications for digital watermarking technologies and this number is increasing rapidly.

Security: In the field of data security**,** watermarks may be used for certification, authentication, and conditional access. Certification is an important issue for official documents, such as identity cards or passports.

Example on the left of a protected identity card. The identity number "123456789" is written in clear text on the card and hidden as a digital watermark in the identity photo. Therefore switching or manipulating the identity photo will be detected. Digital watermarking permits linking information on documents. That means that key information is written twice on the document. For instance, the name of a passport owner is normally printed in clear text. But it would also be hidden as an invisible watermark in the passport photo. If anyone tries to tamper with the passport by replacing the photo it would be possible to detect the change by scanning the passport and verifying the name hidden in the photo.

Tampering with images: Another application is the authentication of image content. The goal of this application is to detect any alterations and modifications in an image.

The three pictures below illustrate this application. The picture on the left shows an original photo of a car that has been protected with a watermarking technology. In the center, the same picture is shown but with a small modification: the numbers on the license plate have been changed. The picture on the right shows the photo after running the digital watermark detection program on the tampered photo. The tampered areas are indicated in white. We can clearly see that the detected area corresponds to the modifications applied to the original photo.

Using digital watermarks for integrity verification: the protected image is the image (a) above; a modified image is obtained by swapping the numbers 9 and 4 of the number plate (b); digital watermarking technology allows detecting and highlights the modified areas, as shown on (c).

Copy prevention or control. Watermarks can also be used for copy prevention and control. For example, in a closed system where the multimedia content needs special hardware for copying and/or viewing, a digital watermark can be inserted indicating the number of copies that are permitted. Every time a copy is made the watermark can be modified by the hardware and after a point the hardware would not create further copies of the data. An example of such a system is the Digital Versatile Disc (DVD).

ADVANTAGES

- Digital Watermarking allows embedding of arbitrary information (the watermark) into digital media (such as video or images) by applying imperceptible, systematic alterations to the media data.
- Higher level of security**:** Security and confidentiality of the embedded information is provided by a secret key. Without this key the watermark cannot be accessed or modified.

Watermarks can be designed in such a way that the embedded information is still retrievable even after the carrier medium changed.

- The advantage of digital watermarking is that the product of the embedding process is still a digital medium. Customers can do everything with a marked medium that they can do with an unmarked one. Watermarked media can be played or copied without any restrictions
- Digital Watermarking is non-restrictive – only misuse is detectable and traceable.

DISADVANTAGES

Digital watermarking is a recent research field; therefore its intrinsic limits are not well understood yet.

On the other hand, more insight into the technical possibility of satisfying the requirements imposed by practical applications is needed, if the practical possibility of using watermarking for copyright protection is to be evaluated. In the following, some of the most common limits shared by digital watermarking schemes are described.

- Visible watermark can be easily removed.

- A watermarking algorithm which is really robust does not exist yet. In the image case, robustness is still an open issue. More specifically, resistance to geometric manipulations such as cropping is recognized as a very difficult goal to achieve in a computationally efficient way.

- Owners can erase the mark: virtually all the watermarking schemes proposed so far are reversible according to the definition previously given.

In other words, by knowing the exact content of the watermark, and the algorithms used to embed and retrieve it, it is always possible to make it unreadable without any significant degradation of the data.