International Journal of Mathematical Archive-5(8), 2014, 85-91
Available online through www.ijma.info ISSN 2229 – 5046

COMPREHENSIVE SURVEY ON DIGITAL WATERMARKING TECHNIQUES

Govind N Sarage*
Dept. of Computer Science, National Defence Academy, Khadakwasla Pune-23, India.

(Received On: 18-08-14; Revised & Accepted On: 31-08-14)

ABSTRACT

Due to high speed computer networks, the use of digitally formatted data has increased the need for the protection of digital media. Digital media includes text, digital audio, images, video and software. Many approaches are available for protecting digital data; these include encryption, authentication and time stamping. This creates a high demand for content protection technique like watermarking, which is one of the most efficient ways to protect the digital properties in recent years. "Watermarking" is the process of hiding digital information in a carrier signal. Embedding a digital signal (audio, video or image) with information which cannot be removed easily is called digital watermarking. Digital watermarks may be used to verify the authenticity or integrity of the carrier signal or to show the identity of its owners. In this paper, we present a comprehensive survey on various digital watermarking techniques. A number of watermarking techniques have been discussed. This paper is a complete survey on general watermarking principles and focuses on describing various watermarking techniques and applications.

Keywords— *Digital Watermarking, Steganography, Encryption, Decryption, LSB.*

1. INTRODUCTION

The advent of the Internet has resulted in many new opportunities for creating and delivering 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 protection of the rights of content owners. In recent times, due to great developments in computer and internet technology, multimedia data i.e. audio, images and video have found wide applications. Digital watermarking [1] is one of the best solutions to prevent illegal copying, modifying and redistributing multimedia data. Encryption of multimedia products prevents an intruder from accessing the contents without a proper decryption key. But once the data is decrypted, it can be duplicated and distributed illegally. 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 in developing new copy deterrence and protective mechanisms. One approach that has been attracting increasing interest is based on *digital watermarking* techniques [1, 2]. Digital watermarking is the process of embedding information into digital multimedia content such that the information can later be extracted or detected for a variety of purposes, including copy prevention and control. To enforce IP rights and to prevent illegal duplication, interpolation and distribution of multimedia data, digital watermarking is an effective solution. Copyright protection, data authentication, covert communication and content identification can be achieved by Digital watermarking. Digital watermarking has become an active and important area of research, and development and commercialization of watermarking techniques is deemed essential to help address some of the challenges faced by the rapid proliferation of digital content.

Watermarking is a branch of information hiding which is used to hide proprietary information in digital media like photographs, digital music, or digital video. A watermark is a secret code or image incorporated into an original image. The use of perceptually invisible watermarks is one form of image authentication. Digital watermarking is the process of embedding information into digital multimedia content such that the information can later be extracted or detected for a variety of purposes, including copy prevention and control.

Digital watermarking is the act of hiding a message [3] related to a digital signal (i.e. an image, song, and video) within the signal itself. It is a concept closely related to steganography, in that they both hide a message inside a digital signal. Watermarking tries to hide a message related to the actual content of the digital signal, while in steganography the digital signal has no relation to the message, and it is merely used as a cover to hide its existence. Watermarking has been around for several centuries, in the form of watermarks found initially in plain paper and subsequently in paper bills. However, the field of digital watermarking was only developed during the last 15 years and it is now being used for many different applications.

*Corresponding author: Govind N Sarage**

2. TECHNICAL DETAILS

Digital watermarking technology makes use of the fact that the human eye has only a limited ability to observe differences. Minor modifications in the color values of an image are subconsciously corrected by the eye, so that the observer does not notice any difference. While vendors of digital watermarking schemes do not publicly release the exact methods used to create their watermarks, they do admit to using the following basic procedure (with obvious variations and additions by each vendor).

A secret key (string or integer) [3] produces a random number which determines the particular pixels, which will be protected by the watermarking. The watermark is embedded redundantly over the whole image, so that every part of the image is protected.

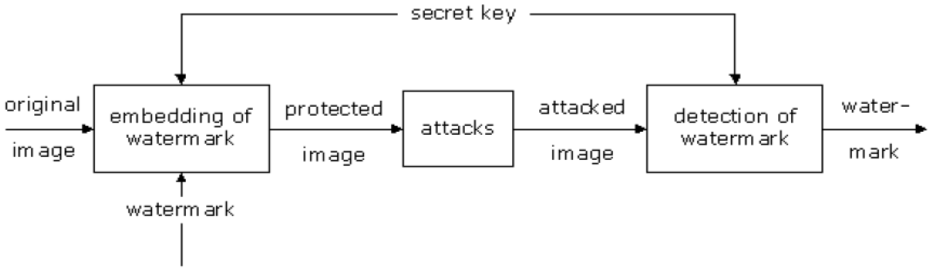

Fig: procedure for watermarking

One way of doing this is by “Patchwork”. This technique uses a random number generator to select n pairs of pixels and slightly increases or decrease their luminosity (brightness level). Thus the contrast of this set is increased without any change in the average luminosity of the image. With suitable parameters, Patchwork even survives compression using JPEG.

Although the amount of secret information has no direct impact on the visual fidelity of the image or the robustness of the watermark, it plays an important role in the security of the system. The key space, that is the range of all possible values of the secret information, x must be large enough to make exhaustive search attacks impossible.

In the process of extracting the watermark, the secret key is used to identify the manipulated pixels and finally to decode the watermark.

3. WATERMARKING TECHNIQUES

The various watermarking techniques are:

Spatial Domain Techniques

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. The spatial domain is the normal image space, this mean any altering in any location in the image is reflected in the corresponding scene which the image is formed by its projection. However, this technique is not reliable when subjected to normal media operations such as filtering or lossy compression. Various spatial domain techniques are as follows:-

Least Significant Bit Coding (LSB)

LSB is the most familiar technique in hiding a watermark in an image it depends on modifications done to the least significant bits of certain pixels in the image. Some algorithms use a sequence of prime numbers instead of the LSB, or add a sequence of prime number to the LSB, other algorithms embeds check sum of the image data into the LSB. [13]
In this method the LSB of the carrier signal is substituted with the watermark. The bits are embedded in a sequence which acts as the key. In order to retrieve it back this sequence should be known. The watermark encoder first selects a subset of pixel values on which the watermark has to be embedded. It then embeds the information on the LSBs of the pixels from this subset. LSB coding is a very simple technique but the robustness of the watermark will be too low. With LSB coding almost always the watermark cannot be retrieved without a noise component. *least significant bit (LSB) substitution* (or *overwriting*). The "least significant bit" term comes from the numeric significance of the bits in a byte. The high-order, or most significant, bit is the one with the highest arithmetic value (i.e., 2^7 =128) while the low-order, or least significant, bit is the one with the lowest arithmetic value (i.e., 2^0 =1)

Information hiding inside the hard drive:

Information can also be hidden on a hard drive in a secret partition. A hidden partition will not be seen under normal circumstances although disk configuration and other tools might allow complete access to the hidden partition [4]. This theory has been implemented in a steganographic ext2fs file system for Linux. A hidden file system is particularly interesting because it protects the user from being inextricably tied to certain information on their hard drive. This form of *plausible deniability* allows a user to claim not to be in possession of certain information or to claim that certain events never occurred. Under this system, users can hide the number of files on the drive, guarantee the secrecy of the files' contents, and not disrupt non-hidden files by the removal of the stego file driver [5, 6, 7].

Spread spectrum steganography:

Spread spectrum steganography methods are analogous to spread spectrum radio transmissions (developed in World War II and commonly used in data communications systems today) where the "energy" of the signal is spread across a wide frequency spectrum rather than focused on a single frequency, in an effort to make detection and jamming of the signal harder. Spread spectrum stego has the same function; avoid detection. These methods take advantage of the fact that little distortions to image and sound files are least detectable in the high energy portions of the carrier; i.e., high intensity in sound files or bright colors in image files. Even when viewed side-by-side, it is easier to fool human senses when small changes are made to loud sounds and/or bright colors [10].

Predictive Coding Schemes

In this method the correlation between adjacent pixels are exploited. A set of pixels where the watermark has to be embedded is chosen and alternate pixels are replaced by the difference between the adjacent pixels. This can be further improved by adding a constant to all the differences. A cipher key [4] is created which enables the retrieval of the embedded watermark at the receiver. This is much more robust as compared to LSB coding.

Correlation-Based Techniques

In this method a pseudo random noise (PN) with a pattern W(x, y) is added to an image. At the decoder the correlation between the random noise and the image is found out and if the value exceeds a certain threshold value the watermark is detected else it is not.

Replica Method

Original signal can be used as an audio watermark. Echo hiding is a good example. Replica modulation also embeds part of the original signal in frequency domain as a watermark. Thus, replica modulation embeds replica, i.e., a properly modulated original signal, as a watermark. Detector can also generate the replica from the watermarked audio and calculate the correlation. The most significant advantage of this method is its high immunity to synchronization attack.

Text Watermarking

Many paper documents (e.g., contracts, wills, etc.) are more valuable than multimedia like sound clips and images. Digital libraries and archives distribute copyrighted articles, journals, and books in electronic form. Watermarking of text documents provides a means of tracing documents that have been illegally copied, distributed, altered, or forged.
Raw text, such as an ASCII text file or computer source code, cannot be watermarked because there is no “perceptual headroom” in which to embed hidden information. However, final versions of documents are typically formatted (e.g., PostScript, PDF, RTF), and it is possible to hide a watermark in the layout information (e.g., word and line spacing’s) and formatting (e.g., serifs). Although *optical character recognition* (OCR) can theoretically remove any layout information, OCR is expensive, imperfect, and often requires manual supervision. Brassil *et al.* [11–12] have investigated text watermarking and proposed a variety of methods for embedding hidden messages in PostScript documents. The work of Brassil *et al.* currently does not use SS embedding, but it could be added to the system to strengthen robustness and security. In [11–12], the message is embedded by altering different parts of the document.
Line shifting moves entire lines of text up or down by a small amount, typically 1/150 or 1/300 inch (0.170 or 0.085 mm). Similarly, word shifting may horizontally shift individual words or blocks of words; words at the ends of a line are not shifted to preserve justification.

Recovery of the message from a printed or photocopied document requires a number of post-processing steps (scanning, skew correction, and noise removal). After post-processing, the message receiver automatically measures line shifts, word shifts, and/or feature alterations to detect the message. In experiments, these methods have shown promise. Line shifts could be correctly detected even after photocopying ten times. Word shifts on a single page were correctly detected 75 percent of the time, after photocopying four times or after fax transmission. With simple ECC, 26–30 of 30 embedded message bits per page could be decoded, depending upon the amount of degradation (e.g., photocopying multiple times).

Image Watermarking

Digital images can be produced from many sources, such as everyday photographs, satellite pictures, medical scans, or computer graphics. Watermarks for natural images typically modify pixel intensities or transform coefficients, although it is conceivable that a watermark could alter other features such as edges or textures. An image may be viewed for an extended period of time, and it may also be subject to a great deal of manipulation, such as filtering, cropping, geometric transformations, compression, and compositing with other images, and hostile attacks. Thus, imperceptibility, robustness, and security are usually the most important properties of image watermarks; speed and complexity are often secondary. Also, since many images are compressed (e.g., JPEG or GIF), watermarking algorithms that operate in the transform or wavelet domain may be useful.

Video Watermarking
Digital video is a sequence of still images, and many image watermarking techniques can be extended to video in a straightforward manner. In contrast to single images, the large video bandwidth means that long messages can be embedded in video. Speed is also an important issue because of the huge amounts of data that must be processed. Except for video production (which takes place before distribution), digital video is typically stored and distributed in compressed form (e.g., MPEG). Hence, it is often desired that the marked, compressed video should not require more bandwidth than the unmarked, compressed video. This bit-rate constraint could also be an issue for single images. Compressed-domain video watermarking is especially attractive. Operating on the compressed bit stream obviates the need for compute-intensive, time-consuming decompression and recompression, such that the watermark can be embedded at the time of distribution or reception.

4. WATERMARKING APPLICATIONS

One of the first applications for watermarking was *broadcast monitoring*. It is often crucially important that we are able to track when a specific video is being broadcast by a TV station. This is important to advertising agencies that want to ensure that their commercials are getting the air time they paid for. Watermarking can be used for this purpose. Information used to identify individual videos could be embedded in the videos themselves using watermarking, making broadcast monitoring easier.

Another very important application is *owner identification*. Being able to identify the owner of a specific digital work of art, such as a video or image can be quite difficult. Nevertheless, it is a very important task, especially in cases related to copyright infringement. So, instead of including copyright notices with every image or song, we could use watermarking to embed the copyright in the image or the song itself.

Transaction tracking is another interesting application of watermarking [6]. In this case the watermark embedded in a digital work can be used to record one or more transactions taking place in the history of a copy of this work. For example, watermarking could be used to record the recipient of every legal copy of a movie by embedding a different watermark in each copy. If the movie is then leaked to the Internet, the movie producers could identify which recipient of the movie was the source of the leak.

In general, digital watermarking has various types of applications. Different applications employ and emphasize different properties of watermarking.

Copy Protection: Watermarking can be utilized to prevent or dissuade unauthorized copying and distribution of content. Individuals can use watermarks to prove authorship of a certain work they created if there is a dispute. It enables the identification of the copyright holder and thus protects his or her right in content distribution. Watermarking is embedded into an image to protect the rights of the owner. Broadcasting companies may use this technology to insert marks into their programs so that commercials can be identified and monitored. For these applications, robustness is critical in that the embedded watermark should not be easily removed or destroyed by small transmission imperfections [6].

Tamper Detection and Authentication: Tamper detection and authentication is needed when evidence is presented to an authority. Tamper detection is used to disclose alterations made into an image. It is closely related to authentication. If tampering is detected in an image, then the image is considered unauthentic. Applications include photo forensics [7], where photographs are presented as evidence in the court of law and photo journalism, where the integrity of the media has to be maintained. Also, in security monitoring systems, watermarks can be used to make sure that all video inputs are from authorized sources. In these applications, a watermark which describes the work is sometimes used. It is important that the description of the file is unique and hard to obtain by an attacker.

Fingerprinting. In applications where multimedia content is electronically distributed across a network, the content owner would like to discourage unauthorized duplication and distribution by embedding a distinct watermark (or a fingerprint) in each copy of the data. If, unauthorized copies of the data are found, at a later time, then the origin of the copy can be determined by retrieving the fingerprint. In this application, the watermark needs to be invisible and must also be invulnerable to deliberate attempts to forge, remove or invalidate it. The watermark, should also be resistant to collusion, that is, a group of users that have the same image but contains different fingerprints should not be able to collude and invalidate any fingerprint or create a copy without any fingerprint. Another example is in digital cinema, where information can be embedded as a watermark in every frame or sequence of frames to help investigators locate the scene of the piracy more quickly and point out security weaknesses in the movie's distribution. The information could include data such as the name of the theater and the date and time of the screening.

ID Card Security and Data Monitoring:
Information in a passport or ID (e.g., passport number or person's name) can also be included in the person's photo that appears on the ID. The ID card can be verified by extracting the embedded information and comparing it to the written text. The inclusion of the watermark provides an additional level of security in this application. For example, if the ID

card is stolen and the picture is replaced by a forged copy, failure in extracting the watermark will invalidate the ID card. These are a few examples of applications where digital watermarks could be of use. In addition, there are many other applications in digital rights management (DRM) and protection that can benefit from watermarking technology. Examples include tracking use of content, binding content to specific players, automatic billing for viewing content, and broadcast monitoring. [8]

Data monitoring is a way to save data of what is transmitted from for example a TV-channel. In 1997 it was discovered that some TV stations in Japan where over booking their air time for commercials. They got paid by advertisers for hours of commercials that were never aired [11]. Watermarking can be used as part of an automatic monitoring system that stores information about what have been aired by the TV stations. Data monitoring can also be used for statistical data collection and analysis [8, 9].

Copy and Usage Control.
Users can have different privilege (play/copy control) on the object due to different payment for that object. It is expected in some systems to have a copy and usage control mechanism to check illegal copies of the content or limit the number of times of copying. A watermark can be used for this purpose.

5. WATERMARKING MODELS

There are several ways in which we can model a watermarking process. These can be broadly classified in one of two groups. The first group contains models which are based on a communication-based view of watermarking and the second group contains models based on a geometric view of watermarking.

Communication-based models

Communication-based models describe watermarking in a way very similar to the traditional models of communication systems. Watermarking is in fact a process of communicating a message from the watermarking embedder to the watermarking receiver. Therefore, it makes sense to use the models of secure communication to model this process.
In a general secure communication model we would have the sender on one side, which would encode a message using some kind of encoding key to prevent eavesdroppers to decode the message if the message was intercepted during transmission. Then the message would be transmitted on a communications channel, which would add some noise to the noise to the encoded message. The resulting noisy message would be received at the other end of the transmission by the receiver, which would try to decode it using a decoding key, to get the original message back. This process can be seen in Figure 1.

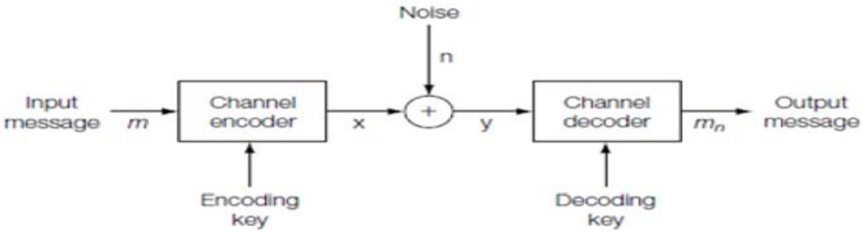

Figure-1: Standard model of a communications channel with key-based encoding

Communication-based watermarking models can be further divided into two sub-categories. The first uses side-information to enhance the process of watermarking and the second does not use side-information at all. The term side-information refers to any auxiliary information except the input message itself that can be used to better encode or decode it. The best example of this is the image used to carry the message, which can be used to provide useful information to enhance the correct detection of the message at the receiver.

Geometric models

It is often useful to think of watermarking in geometric terms. In this type of model, images, watermarked and unwatermarked, can be viewed as high-dimensional vectors, in what is called the media space. This is also a high-dimensional space that contains all possible images of all dimensions. For example a 512 X 512 image would be described as a 262144 elements vector in a 262144-dimensional space.

Geometric models can be very useful to better visualize the watermarking process using a number of regions based on the desirable properties of watermarking. One of these regions is the embedding region, which is the region that contains all the possible images resulting from the embedding of a message inside an unwatermarked image using some watermark embedding algorithm. Another very important region is the detection region, which is the region containing all the possible images from which a watermark can be successfully extracted using a watermark detection algorithm. Lastly, the region of acceptable fidelity contains all the possible images resulting from the embedding of a message into an unwatermarked image, which essentially looks identical to the original image. The embedding region for a given watermarking system should ideally lie inside the intersection of the detection region and the region of acceptable fidelity, in order to produce successfully detected watermarks that do not alter the image quality very much.

An example of a geometric model can be seen in Figure 2. Here we can see that if mean square error (MSE) is used as a measure of fidelity, the region of acceptable fidelity would be an n-dimensional sphere centered on the original unwatermarked image (co), with a radius defined by the largest MSE we are willing to accept for images with acceptable fidelity. The detection region for a detection algorithm based on linear correlation would be defined as a half space, based on the threshold used to decide whether an image has a watermark embedded or not. Note that the diagram is merely a projection of an n-dimensional space into a 2d space.

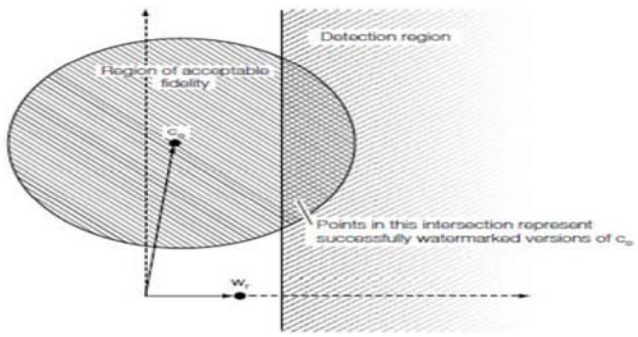
Detection region
Region of acceptable fidelity
c_o
Points in this intersection represent successfully watermarked versions of c_o
w_r

Figure-2: The region of acceptable fidelity (defined by MSE) and the detection region (defined by linear correlation)

When thinking about complex watermarking systems, it is sometimes more useful to consider a projection of the media space into a possibly lower-dimension marking space in which the watermarking then takes place as usual. This projection can be handled more easily by computers because of the smaller number of vector elements and can be possibly expressed by block-based watermarking algorithms which separate images into blocks instead of operating on a pixel basis.

6. CONCLUSION

Digital watermarking is a rapidly evolving area of research and development. Digital Watermarking defines methods and technologies that hide information, for example a number or text, in digital media, such as images, copyright protection, tamper proofing, video or audio. In this paper we are briefly defining the concepts of watermarking, history of watermarks and the properties of a watermarking system as well as an application. In this paper, we have surveyed many of the techniques proposed for watermarking. Thus this paper may serve as a ready reference for any new researcher willing to explore the basics and foundation works in the area of digital watermarking since its evolution to the point where it started gaining prominence in the area of digital media control. This paper gives a base to understand the recent advances in digital watermarking which may have happened after the previous works described in this paper but not covered in this paper.

7. REFERENCES

[1] Jian Liu, Xiangjian He; “A Review Study on Digital Watermarking”, Information and Communication Technologies, 2005. ICICT 2005. First International Conference, Page(s):337 – 341, 27-28 Aug. 2005.

[2] Cox, I.J., M.L. Miller, and J.A. Bloom, “Digital Watermarking.”, 1st edition 2001, San Francisco: Morgan Kaufmann Publisher.

[3] Johnson, N.F., Duric, Z. and Jajodia, S.G. *Information Hiding: Steganography and Watermarking - Attacks and Countermeasure*s.Norwell (MA): Kluwer Academic Publishers, 2001.

[4] Johnson, N.F., Duric, Z. and Jajodia, S.G. *Information Hiding:Steganography and Watermarking - Attacks and Countermeasure*s. Norwell (MA): Kluwer Academic Publishers, 2001.

[5] Anderson, R., Needham, R., and Shamir, A. "The Steganographic File System." In: Aucsmith, D. (ed.). *Proc. of the Second International Workshop on Information Hiding (IH '98*), Portland, OR, April 1998. *Lecture Notes in Computer Scienc*e, Vol. 1525. New York: Springer-Verlag, 1998.

[6] Artz, D. "Digital Steganography: Hiding data within Data." *IEEE Internet Computin*g, May/June 2001. URL: http: //www.cc.gatech.edu/classes/AY2003/cs6262_fall/digital_steganography.pdf. Last accessed: 2003.

[7] McDonald, A.D. and Kuhn, M.G. "StegFS: A Steganographic File System for inux." In: Pfitzmann, A. (ed.). *Proc. of the Third International Workshop on Information Hiding (IH '99*), Dresden, Germany, Sept.-Oct. 1999. *Lecture Notes in Computer Scienc*e, Vol. 1768. New York: Springer-Verlag, 2000. URL: http://www.cl.cam.ac.uk/~mgk25/ih99-stegfs.pdf. Last accessed: 2003.

[8] Fridrich, J. and Du, R. "Secure Steganographic Methods for Palette Images." In: *Proc. of The 3rd Information Hiding Worksho*p, September 1999, Dresden, Germany. *Lecture Notes in Computer Scienc*e, Vol. 1768. New York: Springer-Verlag, 2000.

[9] Wayner, P. *Disappearing Cryptography - Information Hiding: Steganography & Watermarkin*g, 2nd. ed. San Francisco: Morgan Kaufmann; 2002.

[10] Wayner, P. *Disappearing Cryptography - Information Hiding: Steganography & Watermarkin*g, 2nd. ed. San Francisco: Morgan Kaufmann; 2002.

[11] Brassil, J., Low, S., Maxemchuk, N., and O' Gorman, L., Electronic marking and identification techniques to discourage document copying. *Proceedings of IEEE INFOCOM '94*, 1994 3, pp. 1278–1287.

[12] Brassil, J., Low, S., Maxemchuk, N., and O' Gorman, L. Hiding information in document images. *Proceedings of the 29th Annual Conference on Information Sciences and Systems*, 1995, pp. 482-489.

[13] Audio Watermarking Technology To Protect Digital Audio Copyrights Last Updated March 1, 2009 By Johnston Yoon Markany, Inc.

Source of support: Nil, Conflict of interest: None Declared