Digital Watermarking and Steganography

Name of Authors: Monal Mehta, Abhishek Shetty

Name of Institute: D.J. SANGHVI COLLEGE OF ENGINEERING.

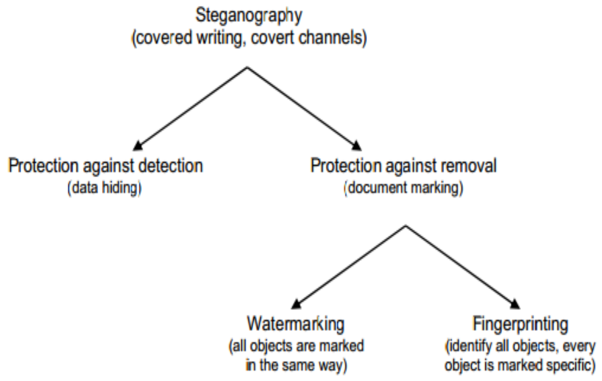
ABSTRACT:

Watermarking, which belongs to the data hiding field has seen plenty of research interest recently. There is a lot of work being conducted in several branches in this field. Steganography is used for secret communication, whereas Watermarking is used for content protection, copyright management, content authentication and tamper detection. In this paper we tend to present an in depth survey of existing and recently proposed steganographic and watermarking techniques. We classify the techniques based on completely different domains in which data is embedded.

INTRODUCTION:

Watermarking is a branch of data concealing that is used to hide proprietary data in digital media like photographs, digital music, or digital video. The benefit with which digital content is changed over the web has created infringement of copyright problems. Proprietary material can be simply changed over peer-to-peer networks, and this has caused major issues to those content suppliers. Hence to shield the interest of these suppliers, the digital content is watermarked.

The figure shows how data hiding will be broken down into totally different areas.
Steganography will be used to hide a message supposed for later retrieval by a selected individual or group. In this case the aim is to stop the message from being detected by any other party.
The other major area of steganography is copyright marking, where the message to be inserted is used to claim copyright over a document. This will be further divided into watermarking and which will be mentioned later.

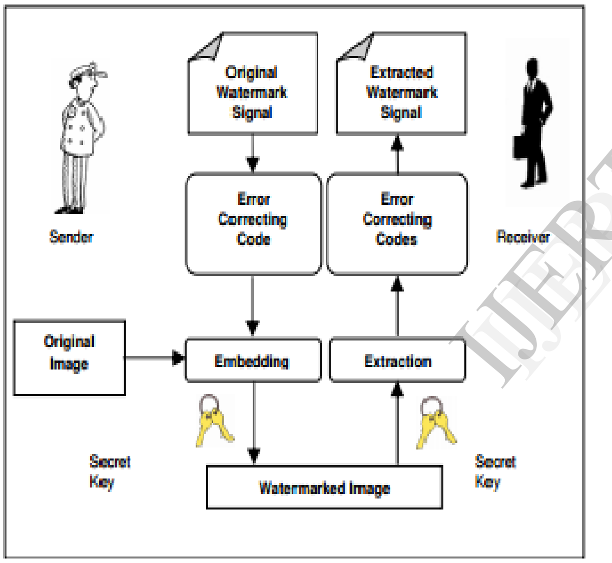

Steganography
(covered writing, covert channels)
Protection against detection
(data hiding)
Protection against removal
(document marking)
Watermarking
(all objects are marked in the same way)
Fingerprinting
(identify all objects, every object is marked specific)

KEY WORDS:

Watermarking, Watermark Detection, Spatial Domain, Image Transforms, DWT, DCT, DFT.

DIGITAL WATERMARKING TECHNIQUE:

The process of embedding a watermark during a multimedia object is termed as watermarking. Watermark is considered as a sort of a signature that reveals the owner of the multimedia object. Content suppliers want to enter watermarks in their multimedia objects (digital content) for several reasons like copyright protection, content authentication, tamper detection etc. A watermarking algorithmic program embeds a visible or invisible watermark in a given multimedia object



The embedding method is guided by use of a secret key that determines the locations inside the transmission object (image) where the watermark would be embedded.

Once the watermark is embedded it will experience many attacks as a result of the multimedia object being digitally processed. The attacks are often unintentional (in case of pictures, low pass filtering or gamma correction or compression) or intentional (like cropping) therefore the watermark needs to be very robust against these possible attacks. When the owner desires to see the watermarks within the probably attacked and distorted multimedia object, s/he relies on the key that was used to embed the watermark. Using the key, the embedded watermark sequence is often extracted. This extracted watermark may or may not resemble the first watermark as a result of the attack.

Hence to validate the existence of watermark, either the original object is used to match and find out the watermark signal (non-blind watermarking) or a correlation measure is used to detect the strength of the watermark signal from the extracted watermark (blind watermarking). In the correlation based detection the first watermark sequence is compared with the extracted watermark sequence.

REQUIREMENTS OF DIGITAL WATERMARKING:

There are many various protocols and embedding techniques that enable us to hide data in a given object. However, all of the protocols and techniques should satisfy variety of requirements so steganography will be applied properly. The subsequent may be a list of main requirements that steganography techniques should satisfy:

- The integrity of the hidden information once it's been embedded within the stego object must be correct. The key message should not change in any manner, such as further information being added, loss {of data} or changes to the key information once it has been hidden. If secret information is modified throughout steganography, it'd defeat the whole purpose of the method.

- The stego object should stay unchanged or virtually unchanged to the naked eye. If the stego object changes considerably and might be detected, a 3rd party might even see that information is being hidden and so may conceive to extract or to destroy it.

• In watermarking, changes within the stego object should not have any impact on the watermark.

Imagine if you had an illegal copy of a picture that you just would like to manipulate in varied ways. These manipulations will be straightforward processes like resizing, trimming or rotating the image. The watermark within the image should survive these manipulations, otherwise the attackers will terribly simply take away the watermark and also the purpose of steganography will be broken.

• Finally, we have a tendency to always assume that the offender is aware of that there's hidden information within the stego object.

WATERMARKS AND WATERMARK DETECTION:

Basically there are 3 main sorts of watermarks which will be embedded within a picture.

A. Pseudo-Random mathematician Sequence

A Gaussian sequence watermark could be a sequence of numbers comprising of equal number of 1and -1s. It termed as a watermark with zero mean and one variation. Such watermarks are used for objective detection employing a correlation measure.

B. Binary Image or gray Scale Image Watermarks

Some watermarking algorithms introduce important data in form of a logo image rather than a pseudo-random Gaussian sequence. Such watermarks are termed as binary image watermarks or gray scale watermarks. Such watermarks are used for subjective detection.

Based on the type of watermark embedded, an appropriate decoder has to be designed to detect the presence of watermark. If it's a pseudo random Gaussian sequence hypothesis, testing is done to detect the presence of watermark. Suppose W is the original watermark bit sequence and W' is the extracted watermark bit sequence, then we can calculate bit error rate (BER) to detect the presence of watermark. If the BER is zero it indicates the presence of watermark; however, if it is one, it indicates absence of watermark. BER is calculated as follows. Suppose D is the retrieved signal and N is the number of bits in watermark then:

$$D = \begin{Bmatrix} 1 \text{ if } W_i \neq W_i^{'} \\ 0 \text{ if } W_i = W_i^{'} \end{Bmatrix} \quad BER(W, W') = \frac{\sum D}{N}$$

Normalized Correlation Coefficient can also be used to detect the presence of watermark.

$$NC(W, W') = \frac{\sum W W'}{\sqrt{\sum W_i^2} \sqrt{\sum W_i'^2}}$$

Images may be pictured in spatial domain and transform domain. The transform domain image is represented in terms of its frequencies; but, in spatial domain it's pictured by pixels. In simple terms transform domain means that the image is segmented into multiple frequency bands. To transfer a picture to its frequency representation we will use many reversible transform like discrete cosine transform (DCT), discrete wavelet Transform (DWT), or discrete Fourier transform (DFT). Each of those transforms has its own characteristics and represents the image in numerous ways that.

Watermarks may be embedded inside pictures by modifying these values, i.e. the constituent values or the transform domain coefficients. Simple watermarks may well be embedded in the spatial domain of pictures by modifying the constituent values or the least significant bit (LSB) values; but, more robust watermarks may well be embedded within the transform domain of pictures by modifying the transform domain coefficients. Since robust watermarking has several applications we'd limit this survey towards robust watermarking algorithms

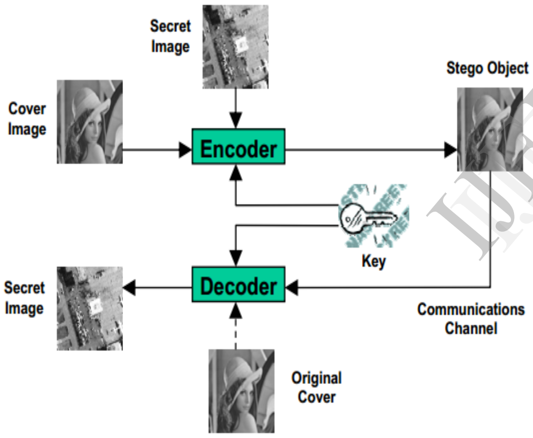


Figure 3. Generic process of encoding and decoding.

DCT DOMAIN WATERMARKING:

The DCT algorithm is one of the main components of the JPEG compression technique. This works

as follows:

1. 1st the image is split up into eight x eight squares.

2. Next every of those squares is transformed via a DCT, that outputs a multi dimensional array of 63 coefficients.

3. A quantizer rounds each of those coefficients, primarily in the compression stage as this can be where data is lost.

4. Small unimportant coefficients are rounded to zero whereas larger ones lose a number of their precision.

5. At this stage you must have an array of streamlined coefficients, which are further compressed via a Huffman encoding scheme or similar.

6. Decompression is done via an inverse DCT.

Hiding via a DCT is beneficial as somebody who simply looks at the pixel values of the image would be unaware that something is amiss. Also the hidden data can be distributed more equally over the whole image in such the way on make it more robust.

One technique hides data within the quantizer stage. If you would like to encode the bit value zero in a specific eight x eight sq. of pixels, you'll be able to do that by ensuring all the coefficients are even, for example by tweaking them. Bit value one is often stored by tweaking the coefficients so they are odd. During this manner a large image will store some data that's quite tough to observe in comparison to the LSB methodology.

This is a very simple process and whereas it works well to keep down distortions, it is vulnerable to noise.

DWT DOMAIN WATERMARKING

While DCT transformations facilitate hide watermark data or general data, they don’t do a great job at higher compression levels. The blocky look of highly compressed JPEG files is due to the eight x eight blocks used in the transformation method. Wave transformations on the other hand are much better at high compression levels and therefore increase the amount of hardiness of the data that's hidden, one thing that is important in a section like watermarking.

This technique works by taking several wavelets to encode an entire image. They allow pictures to be compressed so highly by storing the high frequency "detail" in the image separately from the low frequency elements. The low frequency areas will then be compressed that is acceptable as they're most viable for compression. Quantization will then happen to compress things and also the whole method will start once more if required.

A simple technique using wavelets to hide information is precisely like one in all the techniques discussed in the previous section rather than altering the DCT coefficients with pseudo noise, instead the coefficients of the wavelets are altered with the noise within tolerable levels.

Embedding data into wavelets is an in progress research topic, that still holds a lot of promise.

DFT DOMAIN WATERMARKING

DFT domain has been explored by researches because it offers robustness against geometric attacks like rotation, scaling, cropping, translation etc. In this section we discuss some watermarking algorithms based on the DFT domain.

A. Characteristics of DFT

1) DFT of a real image is usually complex valued, which results in the phase and magnitude illustration of an image.

2) DFT shows translation invariance. Spatial shifts in the image affects the phase illustration of the image however not the magnitude representation, or circular shifts in the spatial domain don't have an effect on the magnitude of the Fourier transforms.

3) DFT is additionally resistant to cropping because impact of cropping results in the blurring of spectrum. If the watermarks are embedded in the magnitude, which are normalized coordinates, there is no need of any synchronization.

4) The strongest components of the DFT are the central components that contain the low frequencies.

5) Scaling of image results in amplification of extracted signal and can be detected by correlation coefficient.

Translation of image has no result on extracted signal.

6) Rotation of image leads to cyclic shifts of extracted signal and may be detected by complete search.

7) Scaling within the spatial domain causes inverse scaling in the frequency domain. Rotation within the spatial domain causes a similar rotation within the frequency domain.

B. Coefficient selection Criteria

1) Modification to the low frequency coefficients can cause visible artifacts in the spatial domain. Hence, low frequency coefficients ought to be avoided

2) High frequency coefficients are not appropriate because they are removed during JPEG compression

3) The most effective location to embed the watermark is the mid frequency.

C. Benefits of DFT over DWT and DCT

1) DFT is rotation, scaling and translation (RST) invariant.

Hence it can be used to recover from geometric distortions, whereas the spatial domain, DCT hence DWT are not RST invariant and hence it's difficult to overcome from geometric distortions.

CONCLUSION:

As steganography becomes more widely used in computing there are issues that need to be resolved. There are a wide variety of different techniques with their own advantages and disadvantages.

Many currently used techniques aren't robust enough to prevent detection and removal of embedded data. The use of benchmarking to evaluate techniques should become more common and a lot of standard definition of robustness is needed to assist overcome this.

As attacks are found that work against existing techniques, it's likely that new techniques can be developed that overcome these deficiencies. The continuing use of digital media will drive development of new techniques and standards for watermarking are likely to be developed.

Meanwhile techniques used by law enforcement authorities to detect embedded material can improve as they continue to try to prevent the misuse of steganography.

REFERENCES:

- www.wikipedia.org
- D. Artz, “Digital Steganography: Hiding Data within Data”
- S. Cacciaguerra and S. Ferretti, “Data Hiding: Steganography And Copyright Marking”
- www.authorstream.com
- www.seminarelectronicstopic.com
- www.seminaron.in
- www.techalone.com