A Survey of Digital Watermarking Technologies

Lin Liu
linliu@ece.sunysb.edu

Abstract: This report introduces the basic concepts in digital watermarking. Common watermarking technologies are reviewed. Some experiment results are provided as well.

Index Terms: Digital watermarking, copyright protection

1. INTRODUCTION

During the past decade, with the development of information digitalization and internet, digital media increasingly predominate over traditional analog media. However, as one of the concomitant side-effects, it is also becoming easier for some individual or group to copy and transmit digital products without the permission of the owner. The digital watermark is then introduced to solve this problem. Covering many subjects such as signal processing, communication theory and Encryption, the research in digital watermark is to provide copyright protection to digital products, and to prevent and track illegal copying and transmission of them. Watermarking is embedding information, which is able to show the ownership or track copyright intrusion, into the digital image, video or audio. Its purpose determines that the watermark should be indivisible and robust to common processing and attack.

Currently the digital watermarking technologies can be divided into two categories by the embedding position——spatial domain and transform domain watermark. Spatial domain techniques developed earlier and is easier to implement, but is limited in robustness, while transform domain techniques, which embed watermark in the host's transform domain, is more sophisticated and robust. With the development of digital watermarking, spatial techniques, due to their weakness in robustness, are generally abandoned, and frequency algorithm based on DCT or DWT becomes the research focus. Another tendency in watermarking is blind extraction, which means the host is not need when extracting the watermark; otherwise it is hard to avoid the multiple claims of ownerships.

A. The Foundation of Digital Watermarking

It should be noted that he reason why digital watermarking is possible is that human vision system (HVS) is not perfect. Digital watermark utilizes the limitation of HVS to make itself invisible, thus avoiding to degrade original digital products, as well being hard to get identified or destroyed.

B. Properties and Requirements of Digital Watermarking

Invisible A watermarking system is of no use if it distorts the cover image to the point of being useless, or even highly distracting. Ideally the watermarked imaged should look indistinguishable from the original even on the highest quality equipment.
Robust The watermark should be resistant to distortion introduced during either normal use (unintentional attack), or a deliberate attempt to disable or remove the watermark present (intentional, or malicious attack). Unintentional attacks involve transforms that are commonly applied to images during normal use, such as cropping, resizing, contrast enhancement…etc.
Unambiguous Retrieval of the watermark should unambiguously identify the owner. Furthermore, the accuracy of owner identification should degrade gracefully in the face of attack.

C. Classification of Digital Watermarking

According to the domain for watermark embedding

Spatial-domain watermarking technologies change the intensity of original image or gray levels of its pixels. This kind of watermarking is simple and with low computing complexity, because no frequency transform is needed. However, there must be tradeoffs between invisibility and robustness, and it is hard to resist common image processing and noise. Frequency-domain watermarking embeds the watermark into the transformed image. It is complicated but has the merits which the former approach lacks.

According to how watermark is detected and extracted

Blind-extracting watermarking means watermark detection and extraction do not depend on the availability of original image. The drawback is when the watermarked image is seriously destroyed, watermark detection will become very difficult. Nonblind-extracting watermark can only be detected by those who has a copy of original image. It guarantees better robustness but may lead to multiple claims of ownerships.

According to the ability of watermark to resist attack

Fragile watermarks are ready to be destroyed by random image processing methods. The change in watermark is easy to be detected, thus can provide information for image completeness. Robust watermarks are robust under most image processing methods and can be extracted from heavily attacked watermarked image. Thus it is preferred in copyright protection.

D. The Classic Process of Digital Watermarking

Common watermarking algorithms usually include two steps: watermark embedding and watermark detection (extraction).

Let f() denote the embedding function, I the original watermark, W the watermark

to be embedded, then the watermarked image I' can be expressed as:

$$I' = f(I, W)$$

Common approach is as follows:

Extract a property sequence from original image V = v1, v2… vn, corresponding watermark sequence is X = x1, x2… xn. Embed X into V according to certain model to obtain the adjusted sequence V' = V + X = v'1, v'2…v'n. Put V' back and take the place of V, then we get the watermarked image I'.

Let E() denote the detection function and I' the image to be examined. Extract the watermark from I'

$W' = E(I')$ Blind-extracting watermarking, or

$W' = E(I', I)$ Nonblind-extracting watermarking

If the correlation function C(W, W') satisfies

$C(W, W') >= T$ (T is the threshold value)

Then we consider there is a watermark W in I'. Otherwise there is none.
The whole process is illustrated as in Figure 1.

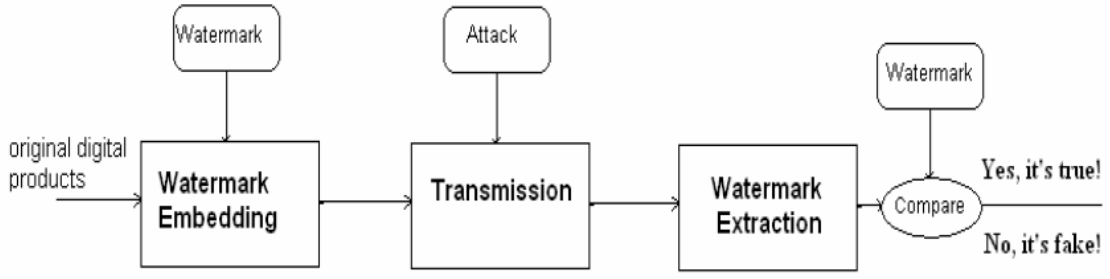

Figure 1.Digital watermarking process

E. Important Parameters in Digital Watermarking Systems

There are a lot of parameters and variables in digital watermarking systems. Tradeoffs must be made between some of them. The most important ones are listed here:

Quantum of information embedded: This important parameter is determined by the specific application and directly influences the robustness of the system. The more information inserted, the less robust the watermarking will be.

Watermark intensity: Also known as the power of the embedded watermark. To increase the robustness, one may increase this parameter, but at the cost of the degradation of original image.

Size of watermark: Similar to its intensity, the larger the size of watermark is, the robust the system will be. It should be noted that watermark that is too small tend to have little value in real application.

Control information: Though it has nothing to do with the invisibility or robustness of the watermarking system, the control information, for example, the key used to

rearrange the watermark before embedding it, plays an important role in system security.

F. Digital Watermarking V.S. Information Hiding

Though in many literatures, even some famous papers, digital watermarking and information hiding refer to the same technology, the two have some differences worth pointing out. Strictly speaking, digital watermarking should not be considered as a branch of the latter, neither. They are same in terms of embedding some secret information into hosts, as well taking imperceptibility as an important criterion. Many algorithms for information hiding can be moved to digital watermarking. However, fundamentally speaking, information hiding is to realize secrete communication. The host is easy to obtain and have little actual value. In contrast, the host for digital watermarking is the product to protect and may not have many available copies. Also in digital watermarking the information to be embedded is significantly less than that in information hiding, which thus has to pay more attention to the imperceptibility part.

G. Performance Evaluation of Watermarking Systems

Signal-to-noise ratio (SNR) is a common metric in signal processing industry. Suppose the original image is Im,n, the output image is Dm,n, then generally SNR is defined as:

$$SNR = 10\log_{10}[\sum_m\sum_n I(i,j)^2 / \sum_m\sum_n (I(i,j) - D(i,j))^2]$$

When SNR approaches infinity, the original image and output image are totally the same.

Another similar one is Peak SNR (PSNR). For images with 255 gray levels, the PSNR is defined as:

$$PSNR = 10\log_{10}[\sum_m\sum_n 255^2 / \sum_m\sum_n (I(i,j) - D(i,j))^2]$$

The similarity of extracted watermark W1 and original watermark W is computed by the following formula:

$$SM = \sum_m\sum_n W(i,j) * W1(i,j) / \sqrt{\sum_m\sum_n W(i,j)^2 * \sum_m\sum_n W1(i,j)^2}$$

If the result is larger than some determined threshold, we consider W1 = W.

2. EXISTING WATERMARKING TECHNOLOGIES

A. Spatial-Domain technologies

Spatial-domain technologies refer to those embedding watermarks by directly changing pixel values of host images. Some common spatial-domain algorithms include Least Significant Bit (LSB) Modification, Patchwork, Texture Block Coding, etc. The most serious drawback of spatial-domain technologies is limited robustness.

It is difficult for spatial-domain watermarks to survive under attacks such as lossy compression and low-pass filtering. Also the information can be embedded in spatial domain is very limited. In recent years they are becoming generally abandoned. We introduce the most famous spatial-domain technology, LSB Modification, to keep the discussion complete.

The LSB is the most straight-forward method of watermark embedding. Given the extraordinarily high channel capacity of using the entire cover for transmission in this method, a smaller object may be embedded multiple times. Even if most of these are lost due to attacks, a single surviving watermark would be considered a success.

LSB substitution however despite its simplicity brings a host of drawbacks. Although it may survive transformations such as cropping, any addition of noise or lossy compression is likely to defeat the watermark. An even better attack would be to simply set the LSB bits of each pixel to one, which fully defeating the watermark with negligible impact on the cover object. Furthermore, once the algorithm is discovered, the embedded watermark could be easily modified by an intermediate party.

An improvement on basic LSB substitution would be to use a pseudo-random number generator to determine the pixels to be used for embedding based on a given “seed” or key. Security of the watermark would be improved as the watermark could no longer be easily viewed by intermediate parties. The algorithm however would still be vulnerable to replacing the LSB’s with a constant. Even in locations that were not used for watermarking bits, the impact of the substitution on the cover image would be negligible. LSB modification proves to be a simple and fairly powerful tool for stenography, however lacks the basic robustness that watermarking applications require.

B. Frequency-Domain Technologies

Compared to spatial-domain watermark, watermark in frequency domain is more robust and compatible to popular image compression standards. Thus frequency-domain watermarking obtains much more attention. To embed a watermark, a frequency transformation is applied to the host data. Then, modifications are made to the transform coefficients. Possible frequency image transformations include the Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) and others.

The first efficient watermarking scheme was introduced by Koch et al. In their method, the image is first divided into square blocks of size 8x8 for DCT computation. A pair of mid-frequency coefficients is chosen for modification from 12 predetermined pairs. Bors and Pitas developed a method that modifies DCT coefficients satisfying a block site selection constraint. After dividing the image into blocks of size 8x8, certain blocks are selected based on a Gaussian network classifier decision. The middle range frequency DCT coefficients are then modified, using either a linear DCT constraint or a circular DCT detection region. A DCT domain watermarking technique based on the frequency masking of DCT blocks was introduced by Swanson. Cox developed the first frequency-domain watermarking scheme. After that a lot of watermarking algorithms in frequency domain have been proposed.

Most frequency-domain algorithms make use of the spread spectrum communication technique. By using a bandwidth larger than required to transmit the signal, we can keep the SNR at each frequency band small enough, even the total power transmitted is very large. When information on several bands is lost, the transmitted signal can still being recovered by the rest ones. The spread spectrum watermarking schemes are the use of spread spectrum communication in digital watermarking. Similar to that in communication, spread spectrum watermarking schemes embed watermarks in the whole host image. The watermark is distributed among the whole frequency band. To destroy the watermark, one has to add noise with sufficiently large amplitude, which will heavily degrade the quality of watermarked image and be considered as an unsuccessful attack.

Figure 2 and Figure 3 illustrate the watermark embedding and detection /extraction in frequency domain, respectively.

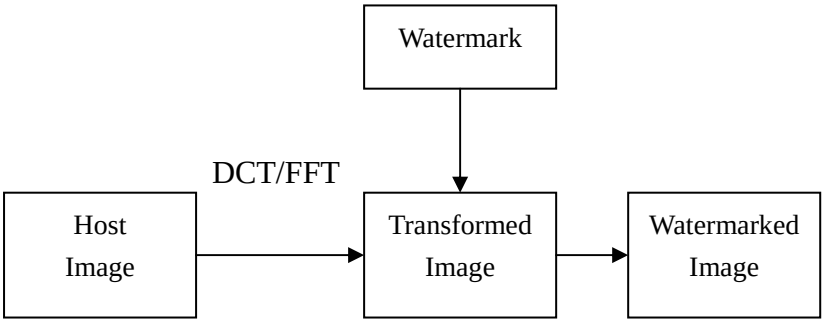

Figure 2. Watermark embedding in frequency domain

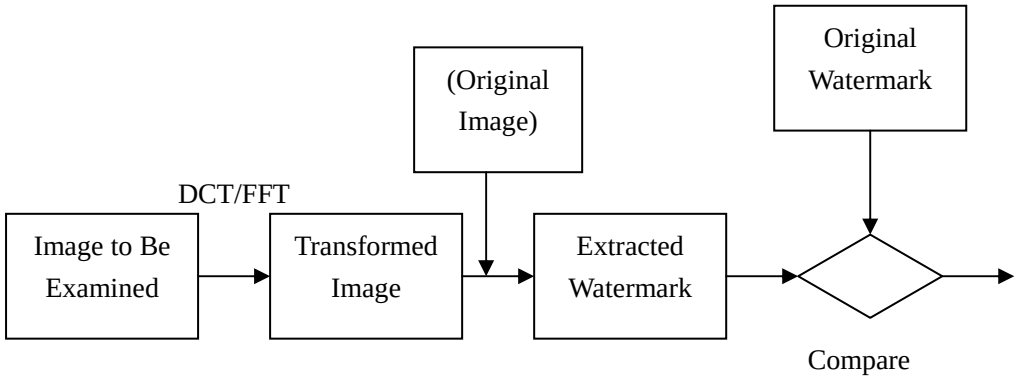

Figure 3. Watermark detection/extraction in frequency domain

One major reason why frequency domain watermarking schemes are attractive is their compatibility with existing image compression standards, in particular, the JPEG standard. The compatibility ensures those schemes a good performance when the watermarked image is subject to lossy compression, which is one of the most common image processing methods today. In consequence, those schemes become particularly useful in practical applications on the Internet.

A widely accepted point now is the frequency-domain watermark should be embedded into the mid-band of the transformed host image. Watermarks in high frequency band tend to have less influence on the quality of original image, while

watermarks in low band will achieve a better robustness (since a large portion of high frequency components may be quantized to zero under JPEG compression, as shown in figure 4). And the mid-bind scheme is right a tradeoff between the imperceptibility and robustness.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	69	56
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

figure 4. JPEG quantization table for intensity (left) and hue(right)

C. Wavelet-domain Domain Technologies

The new JPEG2000 standard has adopted a new technique, the wavelet transform. Though this standard has not been widely used yet, any new watermarking algorithm that intends to survive in the future should get along with it. Here come the watermarking schemes based on wavelet transform. The difference between different wavelet domain methods depends on the way the watermark is weighted. The reason for this is to reduce the presence of visual artifacts.

The DWT (Discrete Wavelet Transform) separates an image into a lower resolution approximation image (LL) as well as horizontal (HL), vertical (LH) and diagonal (HH) detail components. The process can then be repeated to computes multiple “scale” wavelet decomposition, as in the 2 scale wavelet transform shown below in figure 4.

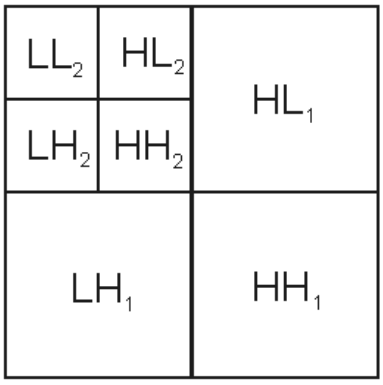

Figure 5. 2 Scale 2-Dimensional Discrete Wavelet Transform

One of the many advantages over the wavelet transform is that that it is believed to more accurately model aspects of the HVS as compared to the FFT or DCT. This allows us to use higher energy watermarks in regions that the HVS is known to be less sensitive to, such as the high resolution detail bands {LH,HL,HH). Embedding

watermarks in these regions allow us to increase the robustness of our watermark, at little to no additional impact on image quality. One of the most straightforward techniques is to use a similar embedding technique to that used in the DCT, the embedding of a CDMA sequence in the detail bands according to the equation

$$I_{Wu,v} = \begin{cases} W_i + \alpha |W_i| x_i, & u,v \in HL, LH \\ W_i & u,v \in LL, HH \end{cases}$$

where Wi denotes the coefficient of the transformed image, xi the bit of the watermark to be embedded, and a scaling factor. To detect the watermark we generate the same pseudo-random sequence used in CDMA generation and determine its correlation with the two transformed detail bands. If the correlation exceeds some threshold T, the watermark is detected. This can be easily extended to multiple bit messages by embedding multiple watermarks into the image. As in the spatial version, a separate seed is used for each PN sequence, which is then added to the detail coefficients as per figure 10. During detection, if the correlation exceeds T for a particular sequence a "1" is recovered; otherwise a zero. The recovery process then iterates through the entire PN sequence until all the bits of the watermark have been recovered. Furthermore, as the embedding uses the values of the transformed value in embedded, the embedding process should be rather adaptive; storing the majority of the watermark in the larger coefficients.

3. EXPERIMENTS AND RESULTS

In this section, we provided some experiment results for a typical DCT-based watermarking algorithm. The host images in this experiment are both 512 x 512 JPEG images with 256 gray levels. The watermark is 81 x 80 JPEG image with 256 gray levels. They are shown in figure 6 and 7.

Figure 6. Original images lena and baboon

Figure 7. Original watermark

A. Attack-Free

.

Figure 8. Watermarked Lena and extracted watermark (no attack)

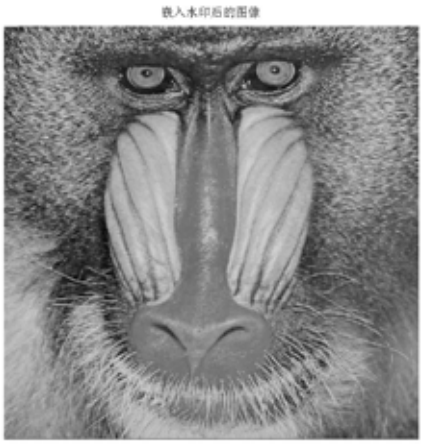

Figure 9. Watermarked Baboon and extracted watermark (no attack)

Table 1. Performance under ideal condition

	SNR(dB)	PSNR(dB)	SM(W,W1)
lena	39.9075	45.6740	0.9993
baboon	38.7344	44.1982	0.9993

From figure 8, 9 and table 1, it can be seen that the watermarked images and original ones are almost identical. The similarity between original and extracted watermark is very close to 1.

B. Experiment results under image cropping

In this experiment we crop the watermarked lena by different ratio, then try to extract the watermark and compute the similarity. Results are shown in figure 10, 11 and table 2.

Figure 10. Watermarked Lena after 1/8 cropping and the extracted watermark.

Figure 10. Watermarked Lena after 1/4 cropping and the extracted watermark.

Table 2. Performance under cropping attack

	SNR	SM
1/8 Cropped	8.5885	0.9983
1/4 Cropped	5.1837	0.9937

The results show this algorithm deals with cropping excellently. The extracted watermark can maintain a good similarity with the original one even after the watermarked image is cropped 1/4.

C. Experiment results under JPEG compression

Image files on the Internet are usually compressed by JPEG standard in order to reduce the file size and save limited bandwidth. As a result, digital watermarking algorithms should be robust under JPEG compression.

In this experiment, the watermarked image, baboon, is compressed with ratio 1.7 and 2.3, respectively. The corresponding SNRs are 34.7300 and 30.8038, respectively, and the similarities 0.9922 and 0.9580 respectively. So we come to the conclusion that under JPEG compressions with relatively small ratios, the watermark can be well detected and extracted.

Figure 11. Extracted watermarks after JPEG compression

of watermarked baboon with ratio 1.7 (left) and 2.3 (right)

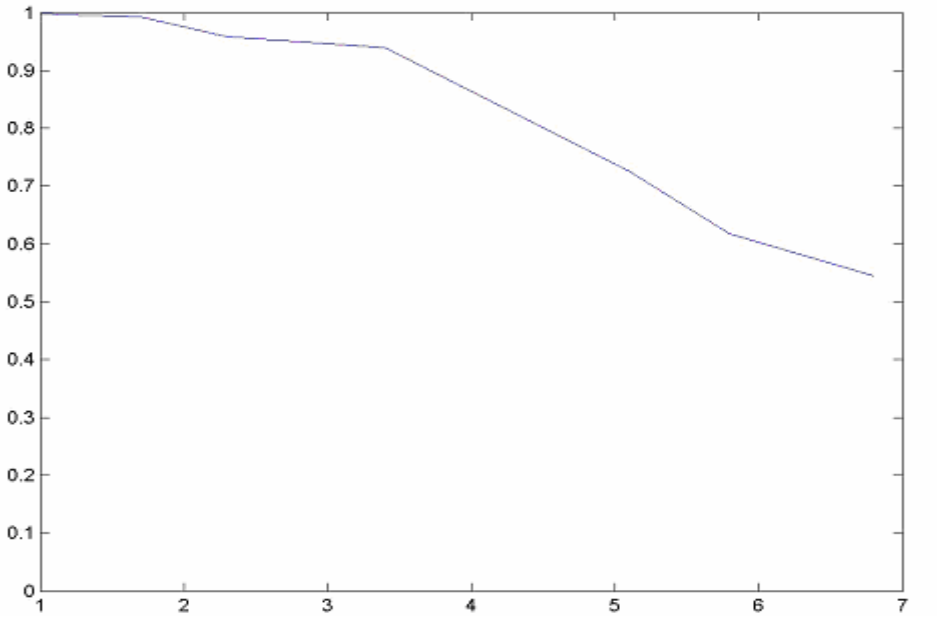

Figure 12. The relationship between compress ratio and watermark similarity

This relationship figure between compress ratio and similarity between the original and extracted watermark is given as figure 12. We can observe that when the

compression ratio becomes large, to recognize the extracted watermark will be difficult by our naked eyes but still easy by the corresponding watermark detector.

4. CONCLUSIONS

In this report, we have introduced some important basic concepts in digital watermarking, including its foundation, properties, requirements and applications, as well as the comparison between digital watermark and information hiding. After that, common watermarking techniques are reviewed; schemes in spatial domain, frequency domain and wavelet domain are introduced with analysis of pros and cons, in terms of imperceptibility, robustness, implementation complexity etc., for each domain. Typical algorithms in all domains have been described in detail.

The last part of this report presents some experiment results, taking the typical frequency-domain, DCT-based watermarking approach as the underlying algorithm. The watermark is embedded into the mid-band of the host image to achieve a good tradeoff between the imperceptibility and robustness of the watermarking system. The results show that this kind of algorithms has a satisfactory performance under image cropping and JPEG lossy compression.

REFERENCES

[1] J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, pp. 1673-1687, Dec. 1997

[2] A. Bors and I. Pitas, “Image watermarking using DCT domain constraints.” in *Proc. IEEE. Int. Conf. Image Processing*, Lausanne, Switzerland, Sept. 1996, pp. 231-234

[3] R.C. Gonzalez, R.E. Woods, “Digital Image Processing”, Upper Saddle River, New Jersey, Prentice Hall, Inc., 2002

[4] Perez-Gonzalez, F.; Hernandez, J.R.;” A tutorial on digital watermarking” *Security Technology, 1999. Proceedings. IEEE 33rd Annual 1999 International Carnahan Conference on*. Oct. 1999 Page(s):286 – 292

[5] Ingemar J. Cox, Matthew L. Miller, Jeffrey A. Bloom “Digital Watermarking” Morgan Kaufmann Publishers ISBN: 1-55860-714-5