A Technical Review on Digital Image Watermarking Techniques

Rohit Thanki¹
¹PhD Research Scholar,
C U Shah University,
Wadhwan City, Gujarat, India

Dr. Komal Borisagar²
²Guide and Assistant Professor,
Atmiya Institute of Technology & Science,
Rajkot, Gujarat, India

Abstract— With the use of digital technologies and distribution of digital contents like images, videos, audios and reports over the internet then digital right management has become a critical issue. The need of copyrights, intellectual property protection and authentication of the original data over the internet or communication channel is required. The solution of this problem is digital watermarking. Digital watermarking is a new and highly interdisciplinary research area for data protection over the internet. Digital watermarking is the process of inserting a digital signal or pattern into digital content. Digital watermarking system can be viewed as a communication system consisting of three main blocks an embedder, a communication channel which is optional and a detector. This paper provides a complete review of digital watermarking techniques in spatial, transform and sparse domain for different applications like copyright protection, tracking of original data, broadcast monitoring, authentication and medical safety. Nowadays digital watermarking techniques are used for security and authenticity of biometric templates and provide more accuracy in authentication system. Also, paper gives all evolution parameters for digital watermarking techniques.

Keywords— Digital Watermarking, Compressive Sensing, Content Protection, Correlation, Robustness, Parameters

I. INTRODUCTION

Digital watermarking is a relatively new research area that has attracted the interest of numerous researchers both in academia and industry and has become a hottest research areas in the multimedia signal processing and data protection. Digital watermarking is a process of inserting a digital signal or pattern into digital content. The digital signal known as a watermark or message can be used to identify owner work, authenticate content and to trace illegal copies. A 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 term ‘watermark’ was probably original from the German term ‘watermarke’. More than 700 years ago, watermarks were used in Italy to indicate the paper brand and mill that produced. By the 18th century, watermarks began to be used as anti-counterfeiting measures on money and other documents [1]. The first example of a technology similar to digital watermarking is a patent filed in 1954 by Emil Hembrooke for identifying work. In 1988, Komatsu and Tominaga appear to be the first to use the term “digital watermarking”. About 1995, interest in digital watermarking began to mushroom.

Digital watermarking system can be viewed as a communication system consisting of three main elements like embedder, a communication channel which is optional and a detector [2]. The block diagram of a general watermarking system is shown in figure 1.

![Fig. 1 Block Diagram of Watermarking System](image)

The embedder is embedded the watermark signal or image into the cover image or signal. The input of embedded is cover signal or image, watermark signal or image and private key. The output of embedder is watermarked image. This watermarked image is transmitted on communication channel which contains the message into cover or original image. Over communication channel, watermarked content will be affected by noise or different attacks which is given next section. At detector end, the detector is detected watermarked image and retrieval the watermark image from them. The input of detector is watermarked image and private key. The output of detector is the original watermark or message.
II. TYPES OF DIGITAL WATERMARKING

With images widely available on the web, watermarks could be used to provide authentication in terms of a secondary image which is overlaid on the primary image and provides a means of protecting the image [3]. The concept of watermarking use in various ways is described below:

A. Types of Medium Watermarking

Watermarking techniques can be divided into five categories according to the type of document to be watermarked as follows: Text watermarking, Image Watermarking, Video Watermarking, Audio Watermarking and Biometric Watermarking. The digital watermark can be divided into three different types according to inserted watermark type: Noise, Image and Biometric Templates.

B. Spatial Domain Watermarking

In this method the pixel information of the two dimensional image is altered so as to embed the hidden data. Three different techniques are defined in the spatial domain watermarking [5, 9-10]:

1. Least Significant Bit Substitution Technique
2. Correlation Based Watermarking Technique

C. Transform Domain Watermarking

Transform domain watermarking implies the used of various transforms to be applied on the cover medium so as to find out the frequency coefficients and then changing these coefficients according to the watermark information [5, 6]. Most powerful transforms used for the purpose of watermarking are:

1. Discrete Fourier Transform / Discrete Cosine Transform Based Technique [5, 6, 9, 17 and 18]
2. Discrete Wavelet Transform Based Techniques [5, 9-10]

D. Sparse Domain Watermarking

Last few years, a new signal processing theorem is introduced for sampling of signal or image which is called as compressive sensing [7, 8]. In this theorem show that super-resolved signals and images can be reconstructed from far fewer data/measurements than what is usually considered necessary [7]. Sparse domain watermarking use of compressive sensing framework applied to find out non sparse coefficients of message image and then add these random projections into host image. Few techniques using compressive sensing for purpose of tampering identification are:

2. Watermarking scheme with flexible self-recovery quality based on Compressive Sensing for tampering identification [12-13]

E. Types of Watermarking

The watermarked content is divided into two type’s base on human visual capacity which is given below:

1. Invisible Watermarking: An invisible watermark is an overlaid image which can’t be seen, but which can be detected algorithmically. Invisible watermarks do not change the signal to a perceptually great extent, i.e., there are only minor variations in the output signal [1]. The example in the figure 2 shows the invisibly watermarked image.

![Fig. 2 Invisible Watermarking](image)

2. Visible Watermarking: A visible watermarking is a visible translucent image that is overlaid on the primary image. Visible watermarking changes the signal altogether such that the watermarked signal is totally different from the actual signal, for example adding an image as a watermark into another image [2]. Consisting of the logo or seal of the organization allows the primary image to be viewed, but still marks it clearly as property of the owning organization. The watermark doesn’t totally obscure the primary image, but it does identify the owner and prevents the image from being used without that identification attached. The example in the figure 3 shows an image with the overlaid watermark.
F. Source Based Watermarking
This technique is used when owner of a document wants to distribute the document to multiple destinations with the same authentication information. With this method one can identify whether the received document is tempered or not [5].

G. Destination Based Watermarking
The purpose of this kind of technique is same as source based scheme but here each receiver gets unique watermark information that is embedded behind the document. Only that receiver can open that document. This method can prevent illegal reselling of the document [5].

III. REQUIREMENT OF DIGITAL WATERMARKING
There are various requirements that the watermarking techniques should have follow. They are given below [3, 5]:
A. Imperceptibility
The embedded watermarks are imperceptible both perceptually as well as statistically and do not alter the aesthetics of the multimedia content that is watermarked. The watermarks do not create visible artifacts in still images, alter the bit rate of video of introduce audible frequencies in audio signal.

B. Robustness
Depending on the application, the digital watermarking techniques can support different levels of robustness against changes made to the watermarked content. If digital watermarking is used for ownership identification then the watermark has to be robust against any modifications. The watermarks should not get degraded or destroyed as a result of unintentional or malicious signal and geometric distortions like analog to digital conversion, digital to analog conversion, cropping, re sampling, rotation, quantization, scaling and compression of content. On the other hand, if digital watermarking is used for content authentication, the watermarks should be fragile, i.e. the watermarks should get destroyed whenever the content is modified so that any modification to content can be detected.

C. Payload Capacity
It is the size of the message that can be embedded inside a cover medium. It depends on the method used for the watermarking.

Above three requirements carries a trade-off triangle as shown in figure 4 which states that to achieve two of the three requirements, third one should be traded off.

Fig. 4 Trade-off between Imperceptibility, Robustness and Payload Capacity

IV. APPLICATIONS OF DIGITAL WATERMARKING
Digital watermarking techniques have wide ranging applications. Some of the applications are listed below [1, 5]:

A. Copyright Protection
Digital watermarks can be used to identify and protect copyright ownership.
B. Copy Protection

Digital content can be watermarked to indicate that the content cannot be illegally replicated. Devices capable of replication can then detect such watermarks and prevent unauthorized replication of the content.

C. Tracking

Digital watermarks can be used to track the usage of digital content. Each copy of digital content can be uniquely watermarked with data specifying the authorized users of the content. Such watermarks can be used to detect illegal replication of content by identifying the users who replicated the content illegally. The watermarking technique used for tracking is called as fingerprinting.

D. Broadcast Monitoring

Digital watermarks can be used to monitor broadcasted content like television and broadcast radio signals. Advertising companies can use systems that can detect the broadcast of advertisement for billing purposes by identifying the watermarks broadcast along with the content.

V. ATTACKS ON DIGITAL WATERMARKING

There are different attacks against which watermarking system could be vulnerabilities which are given below [24, 26].

A. Removal Attacks

Removal attacks aim at the complete removal of the watermark information from the watermarked data without cracking the security of the watermarking algorithm. In this category of attacks includes denoising, quantization (e.g. for compression), remodulation and collusion attacks.

B. Geometric Attacks

In contrast to removal attacks, geometric attacks do not actually remove the embedded watermark itself, but intend to distort the watermark detector synchronization with the embedded information. In this category of attacks includes horizontal flip, rotation cropping, scaling, deletion of lines or columns, random geometric distortions, geometric distortions with JPEG.

C. Cryptographic Attacks

This attacks aim at cracking the security methods in watermarking schemes and thus finding a way to remove the embedded watermark information or to embed misleading watermarks. One such technique is the brute-force search for the embedded secret information. Another attack in this category is the so-called Oracle attack, which can be used to create a non-watermarked signal when a watermark detector device is available. In this category of attacks includes low pass filtering, sharpening, histogram modification, gamma correction, colour quantization, restoration, noise addition, printing-scanning, over-marking.

D. Protocol Attacks

This attacks aim at attacking the entire concept of the watermarking application. One type of protocol attack is based on the concept of invertible watermarks [27]. The idea behind inversion is that the attacker subtracts his own watermark from the watermarked data and claims to be the owner of the watermarked data.

VI. DIGITAL IMAGE WATERMARKING TECHNIQUES

Most watermarking research and publication are focused on images and videos. The reason might be that there is a large demand for image watermarking products due to the fact that there are many images and videos available at no cost on the World Wide Web, which need to be protected. To insert a watermark, research can use different domain likes spatial domain, transform domain and sparse domain. There are different watermarking techniques in spatial domain which are described below:

A. Spatial Domain Watermarking

Digital watermarking techniques in spatial domain class generally have following characteristics [9, 10]:

1. The watermark is applied in the pixel domain.
2. No transforms are applied to the host signal during watermark embedding.
3. Combination with the host signal is based on simple operations, in the pixel domain.
4. The watermark can be detected by correlating the expected pattern with the received signal.

The main strengths of pixel domain methods are that they are conceptually simple and have very low computational complexities. There are different watermarking techniques in spatial domain is given below:

1. Least Significant Bit (LSB) Substitution Technique

C. K. Chan and L. M. Cheng et al. [13] proposed a simple data hiding technique by simple LSB substitution. In this technique last bit of host data is chance and produce the watermarked data at output. The last bit of host image is randomly changed [1].
This technique is given extraordinarily high channel capacity of using the entire cover of image for transmission. LSB substitution however despite its simplicity brings a host of drawbacks. Although it may survive transformation such as cropping of image, any addition of noise or lossy compression is likely to defect the watermark image. An even better attack would be to simply set the LSB bits of each pixel to one. Furthermore, once the algorithm is designed, the embedded watermark could be easily modified by an intermediate stage [14].

2. Correlation Based Watermarking Technique

Langelaar G., Setyawan I. and Lagendijk et al. [5] proposed correlation based watermarking technique using PN sequence. This technique is use for generation of invisible watermarking. In this techniques, add a watermark to an image in the spatial domain is to add a pseudorandom noise pattern to the luminance values of its pixels. In general, the pseudorandom noise pattern consists of the integers {-1, 0, 1}, however also floating-point numbers can also be used. The pattern is generated based on a key using seeds, linear shift registers or randomly shuffled binary images.

To create the watermarked image $I_{W}(x, y)$, the pseudorandom pattern $W(x, y)$ is multiplied by a small gain factor $k$ and added to the host image $I(x, y)$, as using below equation [5]:

$$I_{W}(x, y) = I(x, y) + k \times W(x, y)$$  \hspace{1cm} (1)

To detect a watermark in possible watermarked image, we calculate the correlation between the image $I_{W}(x, y)$ and pseudorandom noise pattern $W(x, y)$. In general, $W(x, y)$ is normalized to a zero mean before correlation. Pseudorandom patterns generated using different keys have very low correlation with each other. Therefore, during the detection process the correlation value will be very high for a pseudorandom pattern generated with the correct key and would be very low otherwise [5]. During the detection process, it is common to set a threshold $T$ to decide whether the watermark is detected or not. If the correlation exceeds a certain threshold $T$, the watermark detector demines that image $I_{W}(x, y)$ contains watermark $W(x, y)$.

$$I_{W}(x, y) \circledast W(x, y) > T \rightarrow W(x, y) \text{det ected}$$  \hspace{1cm} (2)

$$I_{W}(x, y) \circledast W(x, y) < T \rightarrow \text{NoW}(x, y) \text{det ected}$$  \hspace{1cm} (3)

Thanki R., Trivedi R., Kher R. and Vyas D. et al. [15] proposed correlation based watermarking techniques using WGN Sequence. In this technique, we are generated watermarked image using different Gaussian noise sequence and exploited correlation properties of Gaussian noise at detector size for detection of secure watermark. This technique is use for generation of visible watermarking. This technique is shown in figure 7. In this technique we use block processing of image and generate watermarked image. The limitation of this technique is that payload capacity is around 7 percentages.
3. Spread Spectrum Based Watermarking Technique

Cox I., Kilian J., Leighton T. and Shamoon T. et al. [6] proposed spread spectrum based watermarking technique. In this technique, determining the values of the watermark from blocks in the spatial domain, employ spread spectrum technique [16] to scatter each of the bits randomly throughout the host image, increasing capacity and improving resistance to cropping. The watermark is first formatted as a long string rather than a 2D image. For each value of the watermark, a PN sequence is generated using an independent state. These sates could either be stored, or themselves generated through PN methods. The summation of all these PN sequences represents the watermark, which is then scaled and added to the host image [6, 9]. To detect the watermark, each state is used to generate its PN sequence, which is then correlated with the entire watermarked image. If the correlation is high, that bit in the watermark is set to “1”, otherwise a “0”. The process is then repeated for all the values of the watermark [5]. This technique improves on the robustness of the watermark significantly, but requires several orders more of calculation.

B. Transform Domain Watermarking

An advantage of the spatial techniques can be easily applied to any image; regardless of subsequent processing. A possible disadvantage of spatial Techniques is they do not allow for the exploitation of this subsequent processing in order to increase the robustness of the watermark. Both the robustness and quality of the watermark could be improved if the properties of the cover image could similarly be exploited. For instance, it is generally preferable to hide watermarking information in noisy regions and edges of images, rather than in smoother regions. The benefit is two-fold; degradation in smoother regions of an image is more noticeable to the HVS, and become a prime target for lossy compression schemes [5, 9]. Taking these aspects into consideration, working in transform domain of some sort becomes very attractive. The classic and still most popular domain for image processing is Frequency domain transform like Discrete Fourier Transform Discrete Cosine Transform and Discrete Wavelet Transform.

1. Frequency Domain Watermarking Technique

The DFT/DCT allows an image to be broken up into different frequency bands, making it much easier to embed watermarking information into the middle frequency bands of an image. The middle frequency bands are chosen such that they have minimize they avoid the most visual important parts of the image (low frequencies) without over-exposing themselves to removal through compression and noise attacks (high frequencies) [5]. One such technique utilizes the comparison of middle-band DCT coefficients to encode a single bit into a DCT block. To begin, we define the middle-band frequencies (F_M) of an 8*8 DCT block as shown below in figure 8 [5].

![Fig. 8 Definitions of DCT Regions](image)

F_L is used to denote the lowest frequency components of the block, while F_H is used to denote the higher frequency components. F_M is chosen as the embedding region as to provide additional resistance to lossy compression techniques, while avoiding significant modification of the cover image [17]. Next two locations $B_1(u_1, v_1)$ and $B_2(u_2, v_2)$, are chosen from the FM region for comparison. Rather than arbitrarily choosing these locations, extra robustness to compression can be achieved if we base the choice of coefficients on the recommended JPEG quantization table shown below in table 1. If two locations are chosen such that they have identical quantization values, we can feel confident that any scaling of one coefficient will scale the other by the same factor preserving their relative size.
Based on the table, we can observe that coefficients (4, 1) and (3, 2) or (1, 2) and (3, 0) would make suitable candidates for comparison, as their quantization values are equal. The DCT block will encode a “1” if \( B_1(u_1, v_1) > B_1(u_2, v_2) \); otherwise it will encode a “0”. The coefficients are then swapped if the relative size of each coefficient does not agree with the bit that is to be encoded [18]. The swapping of such coefficients should not alter the watermarking image significantly, as it is generally believed that DCT coefficients of middle frequencies have similar magnitudes. The robustness of the watermark can be improved by introducing a watermark strength constant \( k \), such that \( B_1(u_1, v_1) - B_1(u_2, v_2) > k \). Coefficients that do not meet these criteria are modified though so as to then satisfy the relation. Increasing \( k \) thus reduces the chance of detection errors at the expense of additional image degradation [18]. Another possible technique is to embed a PN sequence \( W \) into the middle frequencies of the DCT block.

We can modulate a given DCT block \( x, y \) of the image, the DCT for the block is first calculated, in that block, the middle frequency components \( F_M \) are added to the PN sequence \( W \), multiplied by noise factor \( N \). Coefficients in the low and middle frequencies are copied over to the transformed image unaffected. Each block is then inverse-transformed to give us our final watermarked image \( I_W \).

The watermarking procedure can be made somewhat more adaptive by slightly altering the embedding process to the techniques shown below in equation [5].

\[
I_{W,x,y}(u,v) = I_{x,y}(u,v) + K \times W_{x,y}(u,v), u,v \in F_M
\]

Otherwise, \( I_{W,x,y}(u,v) = I_{x,y}(u,v) \)

This slight modification scales the strength of the watermarking based on the size of the particular coefficients being used. Larger \( k \)’s can thus be used for coefficients of higher magnitude in effect strengthening the watermark in regions that can afford it; weakening it in those that cannot [5]. For detection, the image is broken up into those same 8×8 blocks, and a DCT performed. The same PN sequence is then compared to the middle frequency values of the transformed block. If the correlation between the sequences exceeds some threshold \( T \), a “1” is detected for that block; otherwise a “0” is detected. Again \( k \) denotes the strength of the watermarking, where increasing \( k \) increases the robustness of the watermark at the expense of quality [5].

2. Wavelet Domain Watermarking Technique

Another possible domain for watermark embedding is that of the wavelet domain. The DWT separates an image into a low 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 transforms shown in figure 9 [5].

![Wavelet Transform Diagram](Image)

Fig. 9 2 Scale 2 Dimensional Discrete Wavelet Transform

One of the many advantages over the wavelet transform is 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 increase the robustness of our watermark, at little to no additional impact on image quality [5]. 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 shown below:
To detect the watermark, we generate the same pseudorandom 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 equation. This technique uses seed for embedding multiple watermarks into one host image. The author [19] claims that this technique should prove resistant to JPEG compression, cropping and other typical attacks.

C. Sparse Domain Watermarking

The recent theory of compressive sensing introduced by Candès, Romberg and Tao [7, 8 and 20] and Donoho [21] that a length $N$ signal that is $K$ Sparse in one basis (the Sparsity basis) can be recovered from $O (K \log (N/K))$ non-adaptive linear projections onto a second basis (the measurement basis), that is incoherent with the first. Researchers use the theory of compressive sensing for tamper identification and increase payload capacity.

Sheikh M. A. and Baraniuk R. G. et al. [11] proposed watermarking technique based on compressive sensing in transform domain. In this technique, embed a given confidential watermark $f$ by first encoding it as a sequence $p=Axf$ and then adding it to sparse image coefficients $e$; this will give watermarked image coefficients $y=p+e$. This setup is perfect for $L_1$ decoding [22] to accurately determine the sparse coefficients. This transform domain watermarking technique has been shown to be more robust and tamper proof as compared to straightforward spatial watermarking [5].

Tagliasacchi M., Valenzise G., Tubaro S., Cancelli G. and Barni M. et al. [12] proposed watermarking technique for tampering identification based on compressive sensing. Authors describe a robust watermarking technique for image tampering identification and localization. A compact representation of the image is first convert in different block and then convert this block value into random projection using random seed which same for encoder and detector. Then applying wyner-ziv encoding and generate hash value $H$ which is embed as a watermark into host image. In detector side, distortion estimation will be calculated using MSE value.

VII. PERFORMANCE EVALUATION PARAMETERS FOR DIGITAL WATERMARKING TECHNIQUES

The watermark robustness dependents directly on the embedding strength, which depend on visual degradation of image. For performance evaluation parameters, the visual degradation due to the embedding is an important. In this section we give most popular pixel based distortion criteria and introduce one metric which makes use of the effect in the human visual system (HVS) [24]. Most distortion measures or quality metrics use

$\text{Peak Signal to Noise Ratio (PSNR)} = 10 \times \log \frac{255^2}{\text{MSE}}$

<table>
<thead>
<tr>
<th>TABLE III</th>
<th>COMMONLY USED PIXEL BASED VISUAL DISTORTION METRICS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Difference</strong></td>
<td>$AD = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (I(x, y) - I'(x, y))$</td>
</tr>
<tr>
<td><strong>Maximum Difference</strong></td>
<td>$MD = \max \left(\left</td>
</tr>
<tr>
<td><strong>Mean Square Error</strong></td>
<td>$MSE = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (I(x, y) - I'(x, y))^2$</td>
</tr>
<tr>
<td><strong>Normalized Absolute Error</strong></td>
<td>$NAE = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \left</td>
</tr>
<tr>
<td><strong>Normalized Cross Correlation</strong></td>
<td>$NCC = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} I(x, y) \times I'(x, y)}{\sum_{x=1}^{M} \sum_{y=1}^{N} I^2(x, y)}$</td>
</tr>
</tbody>
</table>
In this paper numbers of techniques for the watermarking of digital images have been studied along with their limitations and applications. LSB substitution technique is not a very good technique for watermarking because of lack ability of survives against attacks. In the correlation based watermarking techniques, when increase power of noise then visibility of recover watermark image is increase. The limitation of spread spectrum is that low payload capacity and processing time require for embedding watermark and detecting are very large as compare to other watermarking techniques. The watermarking technique using WGN is used for generation of visible watermarking and that they cannot be applied for larger size of watermarks. All these watermarking techniques are used for different types grey scale as well as colour images. In this paper also provide new fragile watermarking algorithm using compressive sensing which provides protection against tempering of watermarking data and provides more payload capacity compare to traditional watermarking algorithms.

Now days so many organizations and institution are used biometric system for authentication. So research direction in watermarking area is gone toward developing watermarking algorithm for biometric data because of less standard algorithm is available. So research is required to design more watermarking algorithm for biometric embedding which can be used for biometric data protection and large scale applications.

**REFERENCES**


[23] L1-magic Library Website: http://users.ece.gatech.edu/~justin/l1magic/


