A Review on Digital Image Watermarking

1Vivek Soni  
Research Scholar,  
Dept. of ECE, SRIT, Jabalpur

2Rashika Gupta  
Assistant Professor,  
Dept. of ECE, SRIT, Jabalpur

3Ravimohan  
Professor & H.O.D.  
Dept. of ECE, SRIT, Jabalpur

Abstract - There has been an increase in broadcasting media since the beginning of this century, because many techniques had been developed to solve this problem. Watermarking is the greatest bet from many researches around the world. Digital watermarks can be used by a lot of applications like: copyright protection, broadcast monitoring and owner identification. In this paper, we will show a classification of watermarks, propose a basic model for watermarking and explain some recent algorithms for image watermarking and their features, citing examples applicable to each category.

Keywords: Digital watermarking; wavelets; security; DWT, DCT, LWT, SVD.

I. INTRODUCTION

The increased Internet usage has turned a technique that is able to protect the copyright of published medias into a necessity. The easy of distribution of these documents through the web may transgress protection laws against unauthorized copies and make fidelity questionable. Digital watermarking has been proposed as a solution against these practices.

Digital watermark is a labeling technique of digital data with secret information that can be extracted in the receptor. The image in which this data is inserted is called cover image or host [1]. The watermarking process has to be resilient against possible attacks, keeping the content of the watermark readable in order to be recognized when extracted. Features like robustness and fidelity are essentials of a watermarking system, however the size of the embedded information has to be considered since data becomes less robust as its size increases. Therefore a trade-off [2] of these features must be considered.

The paper is organized as follows. In Section II, we described the classification and each feature. In Section III, we explain the main applications of watermarking. A basic model and the discussion about each block of the process that is proposed in Section IV. Section V is the conclusion and Section VI the acknowledgments.

Watermarking and cryptography are nearly related techniques but watermarking is discrete from encryption. In the digital watermarking system, it is containing information carrying the water is embedded in an original image. The watermarked image is transmitted or stored and then decoded to be determined by the receiver.

Cryptography scrambles the image so that it cannot be implicit. In Figure 1 explain the principal of cryptograph, in which plain text encrypted in to cipher text which is then decrypted into plain text.

II. CLASSIFICATION OF WATERMARKING

Digital watermarking rapidly growing research area of digitized images, video and audio has urged the need of copyright protection, which can be used to produce verification against any illegal attempt to either reproduce or manipulate them in order to change their identity. Digital watermarking is technique providing embedded exclusive rights information in
Digital watermarking is a collection of emerging of technology, such as signal processing, cryptography, probability theory, and stochastic theory, network technology, algorithm design, and other techniques. The number of bits that can be inserted through watermarking varies with each application. In case of images, a mark will be a static set of bits. In videos, capacity will be gauged by the quantity of inserted bits per second [1].

D. Detection Types:

This classification determines which resources are necessary for the analysis to extract the watermark from the cover image.

1) **Blind**: In this detection type the original image and the mark data is not available to the receiver. For example: Copy control applications must send different watermarks for each user and the receiver must be able to recognize and interpret these different marks [1].

2) **Non-Blind**: In this case, the receiver needs the original data, or some derived information from it, for the detection process [1]. This data will also be used in the extraction algorithm.

E. Embedding:

The method used to embed the watermark influences both the robustness against attacks and the detection algorithm, but some methods are very simple and cannot meet the application requirements. El-Gayyar and von zur Gathen [2] showed that designing a watermark should consider a trade-off among the basic features of robustness, fidelity, and payload.

There are two approaches for the embedding process:

1) **Spatial Domain**: These watermarks insert data in the cover image changing pixels or image characteristics [4]. The algorithms should carefully weight the number of changed bits in the pixels against the possibility of the watermark becoming visible [2]. These watermarks have been used for document authentication and tamper detection.

2) **Transform Domain**: These algorithms hide the watermarking data in transform coefficients, therefore spreading the data through the frequency spectrum [1] making it hard to detect and strong against many types of signal processing manipulations. The most used transforms are: Discrete cosine transform (DCT) [1], discrete wavelet transform (DWT) [6] and discrete lifting transform (LWT) [7].

III. CHARACTERISTICS OF DIGITAL WATERMARKING

There are a number of important characteristics that watermark can exhibit, Jalil and Mirza (2010), Bandyopadhyay and Paul (2010). The main characteristics of digital watermarking are classified into major categories as follows.

- **Robustness**: The watermark should be capable to resist after normal image processing operations such as image cropping, transformation, compression etc.
- **Imperceptibility**: The watermarked image should appear like same as the original image to the ordinary eye. The observer cannot detect that watermark is embedded in it.
- **Security**: An unauthorized someone cannot detect, retrieve or change the embedded watermark.

A. Robustness:

This feature refers to the ability to detect the watermark after some signal processing operation [1]. Marks cannot survive all kinds of attacks, hence attacks resilience must be optimized according to application. For example: To verify data integrity a correlation between the received image and the signal is carried out when the watermark is extracted. If differences are found then manipulations must have occurred [3]. With that in mind the following classification can be made:

1) **Fragile**: These marks can be destroyed by small manipulations of the watermarked image [4]. Such marks have been used for authentication and integrity verification.

2) **Semi Fragile**: These behave as fragile watermarks against intentional modifications and as robust watermarks against casual manipulations [5] like noise. These marks have been used in image authentication and tamper control.

3) **Robust**: According to [4], these watermarks are designed to resist heterogeneous manipulations. They can be used in copy control and monitoring.

B. Fidelity:

This requirement could be called invisibility. It preserves the similarity between the watermarked object and the original image according to human perception [1]. The mark must remain invisible notwithstanding the occurrence of small degradations in image brightness or contrast.

C. Capacity or Data Payload:
Transparency: Transparency relates to the properties of the human sensory. A transparent watermark causes no artifacts or feature loss.

Capacity: Capacity describes how many information bits can be fixed. It addresses also the possibility of embedding multiple watermarks in one document in parallel. Capacity requirement always effort against two other important requirements, that is, imperceptibility and robustness (Fig 4). A higher capacity is usually obtained at the expense of either robustness strength or imperceptibility, or both.

![Fig. 4: characteristics of digital watermarking](image)

IV. APPLICATIONS

Before discussing watermarking algorithms let us review some common applications:

- **Broadcast Monitoring:** This type of monitoring is used to confirm the content that is supposed to be transmitted [1], [3], [8]. As an example, commercial advertisements could be monitored through their watermarks to confirm timing and count.

- **Owner Identification:** The conventional form of intellectual ownership verification is a visual mark. But, nowadays, this is easily overcome by the use of softwares that modify images. An example is images with a copyright registration symbol which have this mark removed by specialized softwares. In this case invisible watermarks are used in order to overcome the problem.

- **Fingerprinting:** A watermarked object contains information about the owner permissions. Several fingerprints can be hosted in the same image since the object could belong to several users [3], [8].

- **Publication monitoring and copy control:** The watermark contains owner data and specifies the corresponding amount of copies allowed. This presupposes a hardware and a software able to update the watermark at every use [3]. It also allows copy tracking of unauthorized distribution since owner data is recorded in the watermark.

V. BASIC MODEL

Liu and He [3] present a model with three stages: Generation & Embedding, Distribution & Possible Attacks and Detection. In this paper, we adapted this model dividing the first block in Generation and Embedding, because the both use different watermarking algorithms and can be studied independently. The general approach is presented below:

![Fig. 5. Basic Model](image)

The explanation about the Embedding and Detection stage will be presented together, because the algorithms are related. In Fig 1 the basic model can be divided in four stages:

A. Generation

In this stage the mark is created and its contents must be unique and complex in order to be difficult to extract or not to be damaged by possible attacks. Some algorithms that have been used for watermark generation will presented below:

1) **Images in grayscale or binary:** Many marking objects can be images or brands that represent some enterprise or have some data that identify the cover image. Depending on the application the marks can be binary images or grayscale containing a larger amount of data and even some intrinsic features that helps in the extraction process [8].

2) **Pseudo-Random Sequence:** This mark has a random seed that is used to generate the marked matrix. This seed must be stored like a secret key and will be used in the detection process to reconstruct the mark. The use of binary marks with this algorithm is rather common.

3) **Chaotic Sequence:** The watermarks are prepared using maps of chaotic functions [9]–[11]. These sequences are easy to implement because there are predefined models to create them. Due to statistic features these watermarks resist several types of attacks, like simple attacks and distortion.

4) **Error Correcting Codes and Cryptography:** The insertion of redundancy in watermarks or in the cover image can improve the extraction process or the reconstruction of the watermark after attacks. However codes can cause collateral effects by increasing the amount of embedding data, which may in turn harm the watermark robustness or decrease its data payload. The most commonly used codes are: Hamming [12], Bose-Chaudhuri- Hocquenghen (BCH) [12]–[14], Reed Solomon [12], [13], Low density parity check (LDPC) [15], and Turbo [16].

B. Embedding and Detection:

The embedding is directly related with the extraction algorithm, in this section we will discuss how this has been done in recent algorithms. The embedding algorithm is basically a combination of the watermark with the chosen media [3], so the result is equivalent to:

\[ I_W = E(I,W) \]  

Where: I is the original media, W the watermark, E is the embedding function and I_W the watermarked media. The function depends on the algorithm and the analyzed domain.

VI. CLASSIFICATION

1) **Spatial Domain:** In this case the embedded watermark is equivalent to noise addition to the original media, thereby
influencing the watermarked object characteristics. Two following we will be presented below:

(i) LEAST SIGNIFICANT BIT (LSB)

This is the simplest approach, because the least significant bits carry less relevant information and their modification does not cause perceptible changes. Among these approaches there are types using only the salient points [17] or type, which use some kind of cryptography on the watermark message before the embedding process [18]. In this last case, a cipher called "datamark" is created, which is embedded in the cover image using a key. This key determines which points must be modified by the embedding process.

The extraction algorithm is the inverse of embedding. The marked object must be analyzed and its least significant pixel bits isolated. These extracted bits can be used together with the cryptography keys in decoding algorithms to recover the original watermark.

(ii) SINGULAR VALUE DECOMPOSITION (SVD)

Singular Value Decomposition (SVD) is a numeric analysis of linear algebra which is used in many applications in image processing. It is used to decompose a matrix with a little truncate error according to the equation below:

$$A = USV^T$$  (2)

Where A is the original matrix, U and V is orthogonal matrices with dimensions M x M and N x N respectively, S is a diagonal matrix of the Eigen values of A and T indicates matrix transposition. R. Liu and T. Tan [11] did the decomposition of the cover image and added the watermark using a scale coefficient α to get the following equation:

$$S + \alpha W = U_W S_W V_T^T$$  (3)

Multiplying matrices U, VT and SW result in the marked image Aw:

$$A_W = US_W V_T^T$$  (4)

This is possible due to the high stability of singular value of SVD. In another approach, the cover image is separated in blocks and the SVD applied to each block [46], in this case the dimension of watermark must be equal to the block size and a copy of the watermark is embedded in each block. This technique improves watermark robustness and resistance against many kinds of attacks.

Singular Value Decomposition technique is shown to be powerful methods for robust image watermarking [17], [18]. This can be attributed to the facts that:

- Singular value (SV) of a digital image is stable. The SVs remain intact when disturbances are added to an image.
- SVD preserves both one-way and non-symmetric properties, which are not obtainable using DCT or DFT transformations.
- SVs are able to represent intrinsic algebraic properties of a digital image.
- SVD can be performed on both square and rectangular matrices.

Chandra et al. [14] proposed a method based on the SVD of both the host image and visual watermark. The SVs of the watermark are multiplied by a scaling factor and added to the SVs of the host image. The attacks used are JPEG and low pass filter. However this method is non-blind in nature.

Sun et al. [12] proposed a method based on SVD watermarking scheme, wherein the D component with a diagonal matrix is explored for embedding. The basic mechanism used is the quantization of the largest component with a fixed constant integer, called quantization coefficient. A trade-off can be achieved between transparency and robustness by varying the quantization coefficient. However, this method is failed in extracting the watermark with zero error rates.

Chin-Chen Chang et al. [13] proposed a watermarking scheme based on the SVD domain. U matrix of SVD is used for the watermark embedding. The absolute difference between the two rows of U matrix is used for the watermark embedding. They explored the positive relationships between the rows of U and V matrices that are preserved after JPEG compression also.

Chung et al. [49] proposed two notes on the SVD based watermarking algorithm. From their method, if the watermark is embedded in the columns of U matrix and rows of VT, the perceptibility of the host image is improved. However, their method is not robust to many attacks since watermark embedding is in U and VT matrices.

Singular values represent the algebraic properties of an image [50]. Singular values possess the algebraic and geometric invariance to some extent.

The properties of SVD are very much desirable in image watermarking. When the watermarked image undergoes attacks like rotation, scaling and noise addition, the watermark can be retrieved effectively from the attacked watermarked image due to the above said properties.

2) Transform Domain: The mark is embedded into the cover image spectrum, thus not directly influencing the selected image quality. The following transforms are used, among others, in image spectral analysis: DCT, DWT. Some watermarking algorithms using these transforms are presented below:

(i) DISCRETE COSINE TRANSFORM (DCT)

The DCT makes a spectral analysis of the signal and orders the spectral regions from high to low energy. It can be applied globally or in blocks. When applied globally, the transform is applied to all parts of the image, separating the spectral regions according to their energy. When applied in blocks, the process is analogous, only the transform is applied to each block separately.

Below, we list the typical algorithm steps found in the literature [1], [8]:

1) Segment the image into non-overlapping blocks of 8x8;
2) Apply forward DCT to each of these blocks;
3) Apply some block selection criteria;
4) Apply coefficient selection criteria;
5) Embed watermark by modifying the selected coefficients;
6) Apply inverse DCT transform on each block.

The technique proposed by R. Mehul and R. Priti [15] provide the watermark is inserted in four different frequency ranges by selecting coefficients in zigzag order. This technique produced good results when attacks are applied but failed to achieve
robustness to both compression and image processing tasks simultaneously when only one copy of watermark is inserted.

The technique proposed by J.R Hernandez, M. Amado and F. Perez Gonzalez [18] presents that if the watermark is inserted in perceptually most significant components, i.e., low frequencies; the technique tends to be robust to attack but it is difficult to hide the watermark. On the other hand, if the watermark is inserted in perceptually insignificant components, i.e., high frequencies; it is easier to hide the watermark but the technique is then less resistant to attacks.

The technique proposed by Y. Yang, X. Sun, H. Yang, and C.T. Li [19] present a DCT domain based removable visible watermarking algorithm that moderately succeeds in defeating illegal removal and resisting compression. They intended to protect the multimedia content and to ensure that the reconstructed images are high quality for authorized user, or else, of low-quality for unauthorized users by embedding the visible watermark. Their technique enabled preventing the embedded visible watermark from being illegally removed by unauthorized users without correct user keys as their proposed scheme. In conclusion, the watermarked image is generated by adaptive addition of the significant DCT coefficients of the pre-processed watermark and the corresponding host image. The watermarking system is somewhat robust against compression. They show the performance of their proposed technique the success of the introduced scheme in preventing the inserted watermark from illegal removal is illustrated through the results.

The idea proposed by I.J. Cox, J. Kilian, F.T. Leighton and T. Shamon [20] present the host image and the watermark communication channel and a signal to be transmitted, respectively. The perceptually important part of signal spectrum is spectrum with the watermark message. Gaussian noise like watermarks is employed to accomplish security. The watermarked image will be damaged if an attempt is made to destroy the watermark.

A method called Optional differential energy watermarking of DCT encoded images and video is proposed by Langelaar and Langendijk [21]. A block which composes of several $8 \times 8$ DCT blocks is inserted with a watermark bit by dividing the block into two parts. In order to produce an energy difference in the two parts of the same block, where the energy difference is determined by the watermark bit, the high frequency DCT coefficients in the compressed bit stream are selectively discarded. The number of $8 \times 8$ DCT blocks in a block, JPEG quantization, step size, and a minimal cut-off index for watermarking are the three parameters in this technique.

In this [16] techniques, they proposed technique that inserts the watermark into image and extracts the watermark from the watermarked image more efficiently by exploiting the zero-tree in the rearranged DCT coefficients. This technique is reasonable to apply in a real time system as it can directly extract the inserted watermark from the watermarked image devoid of the original image.

A robust digital image watermarking using hybrid DWT-DCT-SVD technique [22] is proposed by S. Murty and P.R. Kumar. They apply DCT to an image results in three frequency sub-bands: low-frequency, mid-frequency and high frequency sub-bands. They calculated the DCT coefficients; for the transformed output image by using following equation:

$$C(u, v) = \alpha(u) \alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \left( \frac{(2x+1)\pi u}{2N} \right) \cos \left( \frac{(2y+1)\pi v}{2N} \right)$$

Where

$$\begin{align*}
\alpha(u) &= \begin{cases} 
1 & u = 0 \\
\sqrt{2} & 1, u = 1, 2, \ldots, N - 1 
\end{cases} \\
\alpha(v) &= \begin{cases} 
1 & v = 0 \\
\sqrt{2} & 1, v = 1, 2, \ldots, N - 1 
\end{cases}
\end{align*}$$

In the above equation, is the input image having N x M pixels, is the intensity of the pixel in row and column of the image, and is the DCT coefficient in row and column of the DCT matrix.

(ii) DISCRETE WAVELET TRANSFORM (DWT)

The wavelet transform decompose the image in four channels (LL, HL, LH and HH) with the same bandwidth thus creating a multi-resolution perspective. The advantage of wavelet transforms is to allow for dual analyses taking into account both frequency and spatial domains.

Wavelets are being widely studied due to their application in image compression, owing to which compression resistant watermarks may be achieved through their use. Another interesting feature of the DWT is the possibility to select among different types of filter banks, tuning for the desired bandwidth. The most commonly used filters are: Haar, Daubechies, Coiflets, Biorthogonal, Gaussian.

When the DWT is applied to an image, the resolution is reduced by a $2^k$, where K is the number of times the transform was applied.

These algorithms are called the “Wavelet based Watermarking” [8]. The watermark is inserted by substituting the coefficients of the cover image for the watermark’s data. This process improves mark robustness, but depends on the frequency. The low frequency (LL) channel houses image contents in which a coefficients change, however small, will damage the cover image, which in turn challenges the fidelity propriety. However when this region of the spectrum is watermarked, a robust mark against compressions like JPEG and JPEG2000 is attained. Furthermore, when the middle and high frequency channels are marked, some benefits against noise interference and several types of filtering show up. Therefore these algorithms tend to be adapted for human visual system (HSV) to avoid small modification in the cover image being perceptible.

Taskovski et al. [21] implemented two watermarks using binary marks in LL2 and HH2 respectively, resulting in a mark which is robust against manipulations like compression and weak against cropping and rescaling. Similarly, [22] created a watermark adapted to JPEG2000 using two algorithms to modify the wavelet coefficients of the LH2 band of the cover image, introducing only minimal differences between the watermarked image and the original. The decision, which algorithm to use, is based on which one produces the smallest change.
To create a watermark which is resistant against noise and some kinds of processing [23] proposed an algorithm that makes three watermarks: pseudo-random, luminance and texture. The first mark is embedded in LL1 band and the others are inserted by segmenting the cover image in blocks and ordering according to the sum of coefficients and standard deviation. This algorithm is robust against cropping, noise and several compression levels.

In order to increase its recovery capacity, error correcting codes can be applied to the watermark; however, its storage capacity will be reduced due to the additional redundancy. A performance comparison of the Hamming, BCD, and Reed Solomon codes is presented in [12]. For small error rates, the codes are effective in error elimination when compared to no coding; on the other hand for higher rates, no benefit has been observed.

Mixing spatial and transform analysis, we have a robust watermark with different features. An algorithm that applies the SVD in all bands of the first level of DWT is proposed [24], making this a watermarking process in all frequencies. Bao [25] made a watermark of the singular values (SV) of each band of the cover image, in order to achieve the least possible distortion according to the human visual system. This watermark is resistant against JPEG encoding, but is fragile against filter manipulation and random noises. An algorithm with greater robustness against cropping, Gaussian noise and compression is proposed in [24]. Initially, the DWT is applied to HL1 or HH1. In the selected band, HH2 or HL2 must be selected and divided into 4x4 blocks. Finally, SVD is applied to each block, and the watermark is embedded into the S matrix.

(iii) LIFTING WAVELET TRANSFORM (LWT)

Also called second generation wavelet transforms, its use has grown due to low memory consumption and easy implementation [16]. The following LWT scheme below I adapted from [7]:

![Lifting scheme](image)

Figure 6: Lifting scheme

In Figure 6, there are three basic operations: split, predict and update. In split stage the input \( x_k \) is separated into odd \( (x_i) \) and even \( (x_o) \) samples, so that each of these variables contains half the number of samples of \( x_k \). In the prediction stage, even samples are used to predict the odd samples. The details coefficients or high frequency \( (h) \) are calculated as prediction errors of the odd samples through the use of the prediction operator \( P: \)

\[
h = x_i - P(x_o)
\]  

(6)

To create the low frequency samples \( s \), the even samples are updated through the update operator \( U: \)

\[
s = x_o - U(d)
\]  

(7)

Some approaches use the LWT in spatial domain operations, as in [27], which embedded a watermark in band LL3, changing the least significant bits of the wavelet coefficients. On the other hand, [26] used a combination of SVD and LWT to apply two levels of wavelet to the cover image and select among one of bands: HH2, HL2 and LH2. In that approach, SVD is applied separately to the watermark and the selected band. The resulting S matrices must be combined into an S matrix, which will be used to create the watermarked image. This process is not blind, however it is robust against many types of manipulations like: noises, rotation, JPEG compression and quantization. It also exhibits very good performance concerning PSNR and normalized correlation values.

VII. ATTACKS ON WATERMARK

The transmission media can cause some loss in the signal implying in a damaged content. These attacks may be intentional or accidental [3]. Intentional attacks use all available resources to destroy or modify the watermark making it impossible to extract it, the methods usually used are: signal processing techniques, cryptanalysis, steganalysis. On the other hand, accidental attacks are inevitable, because every image processing or transmission noise may introduce distortions.

Hartung et al. [28] classified these attacks in classes:

1) **Simple Attacks**: These attacks change the data of the cover image without attempting to target the watermark location. Example: Noise addition, cropping, conversion to analog and wavelet-based compression.

2) **Disabling Attacks**: The goal of these attacks is to attempt to break the correlation between the watermark and the cover image, making extraction impossible. Example: Geometric distortions, rotation, cropping and insertion of pixels.

3) **Ambiguity Attacks**: These attacks confuse the receptor embedding a fake watermark, making it impossible to discover which the original embedded mark in the cover image was.

4) **Removal Attacks**: In this type of attack a study of the watermark is carried out, estimating the watermark content and attempting to separate it from the host image. Example: Certain non-linear filter operations and attacks tailored to a specific watermark algorithm.

VIII. CONCLUSION

In this paper, we have reviewed some recent algorithms, proposed a classification based on their intrinsic features, embedding methods and detection forms. Also a basic four steps model for the watermark process was studied.

Many watermarking algorithms have been reviewed in the literature which show advantages in systems using wavelet transforms. These marks are robust against several different attacks. Another highlight is the replacement of DWT by LWT which improves computational performance and has an easier hardware implementation, but at the cost of extraction quality.

In future works, the use of a secret key insertion for making watermarked image encrypted, so that only user having this secret key can access to the watermark file. Hence, data will be double secured. Also in general model need of original image at the time of extraction, should be avoided, while implementing it.
IX. REFERENCES


