Digital watermarking—A tutorial

Rajkumar Ramasamy and Vasuki Arumugam

©SHUTTERSTOCK.COM/DEEPADESIGNS

Digital watermarking is a technique in which secret information is embedded inside a host signal, which may be an audio, video, or image signal. *Watermarking* is also defined as a branch of information hiding that is used to hide proprietary information in digital media. Digital watermarking may be used for authentication, integrity, or information security. In this article, we present a comprehensive survey on watermarking techniques based on their characteristics and applications, including visible, invisible, robust, fragile, and semifragile watermarking methods. The parameters that are used for the analysis of watermarking algorithms are peak signal-to-noise ratio (PSNR), mean square error (MSE), embedding capacity, bit rate, and bit error rate (BER).

Digital Object Identifier 10.1109/MPOT.2015.2435802
Date of current version: 7 July 2022

Basic types of watermarks

In the past few years, a tremendous growth in digital data transmission has been achieved. Most of the information that is transmitted is copyright protected and secured with passwords. However, due to a lack of security, the transmitted data can be easily hacked and duplicated by unauthorized users. There are three different methods for the protection of our data: watermarking, steganography, and

Digital watermarking may be used for authentication, integrity, or information security.

cryptography. Watermarking is a technique in which we embed information into the original signal. In cryptography, the original information is changed into another form. Steganography hides the existence of data in the cover medium.

The basic types of watermarking techniques are

- reversible and irreversible
- visible and invisible
- fragile and robust.

In reversible watermarking, the original information can be completely recovered after the extraction of the secret message, but, in the irreversible technique, it cannot be completely recovered. A digital watermark is called *imperceptible* (*invisible*) if the original and marked signals are perceptually indistinguishable. However, if the marked signal is noticeable, then it is perceptible (visible) watermarking. In fragile watermarking, the secret information is not exactly detected after a slight modification of the original signal. Semifragile watermarking resists transformations but fails detection. A digital watermark is called *robust* if it resists the class of transformations, such as filtering, cropping, and so on.

The basic properties of digital watermarking include the following:

- The watermark should be highly invisible for secured communication.
- It should be statistically invisible so it cannot be detected or erased.
- The extraction of the data should be simple and accurate.
- Information should be robust to filtering, additive noise, compression, and other forms of image manipulation.

Watermarking principles

A watermarking system has two basic phases: embedding and extraction. In the embedding phase, an algorithm is developed by the sender to accept the host signal (original signal) and secret information (watermark signal), and it also hides the secret data in the host signal to produce a watermarked signal. Then, the watermarked signal (the original plus secret data) is received by a wired or wireless system. When the watermarked signal reaches the exact receiver, then the receiver enters into the extraction phase for the detection of the secret data from the watermarked signal. A general watermarking system is shown in Fig. 1.

Instead, if the watermarked signal is received or hacked by an unauthorized user, the secret information should not be visible to the unauthorized entity. This will be a perfect watermarking system for information security.

Requirements of watermarking

The major requirements of digital watermarking are transparency, robustness, and payload capacity (data capacity).

- *Transparency*: The secret information added in the original information should not degrade the original information. If distortions occurs in the image, then a hacker can easily identify that there is some information and can easily hack the secret data. This also reduces the commercial value of the image.
- *Robustness*: This is the most important requirement of a watermark. There are various attacks that can degrade the data, including cropping, compression, scaling, and so on. Therefore, the watermark should be invariant to these attacks.
- *Payload capacity*: The amount of secret information that can be hidden in an image is known as the *payload capacity*. This should be the maximum possible for the host signal with an assurance of the proper retrieval of the watermark (secret data) during the extraction process.

In watermarking algorithms, these are the three important factors, and they have tradeoffs with each other.

Watermarking domains

The watermarking techniques can be implemented in two different domains: spatial and frequency. In spatial domain methods, the watermark data are directly embedded into the pixels of the original image. The original pixel values are directly changed to produce the watermarked pixel. The basic spatial

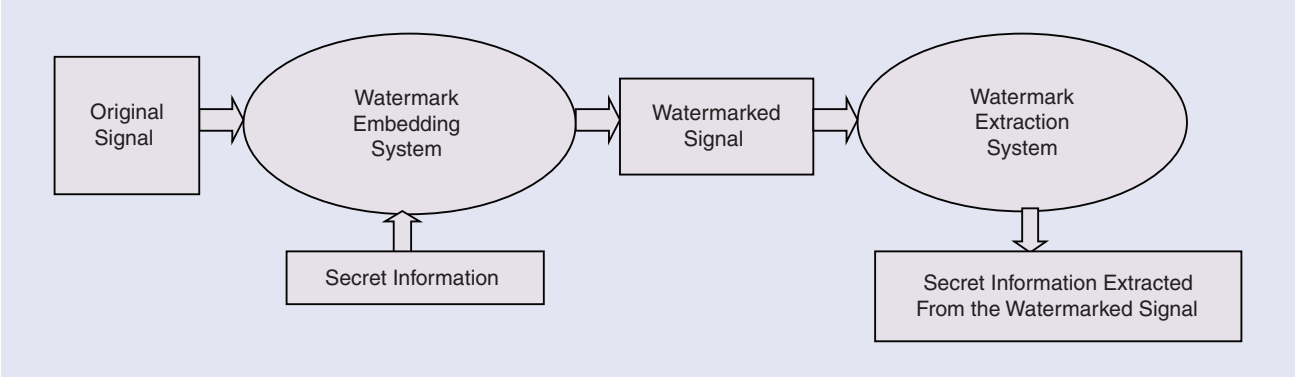

FIG1 A general watermarking system.

domain technique is the least significant bit (LSB) embedding method. A general spatial domain watermarking method is shown in Fig. 2.

The transform domain method is used to produce a high-quality watermarked image by first transforming the original image into the frequency domain by using image transformations. Some of the image transforms are the discrete Fourier (DFT), discrete cosine (DCT), and discrete wavelet transform (DWT) approaches. By using these, the watermarks are not added to the intensities of the image but to the values of its transform coefficient in the embedding phase. A general transform domain watermarking method is shown in Fig. 3. Then, by taking the inverse transformation for the watermarked coefficients, the secret data can be extracted from the watermarked image.

Spatial domain watermarking techniques

Some of the spatial domain techniques are the

- additive watermarking approach
- LSB coding method
- patchwork-based technique
- correlation-based technique
- spread spectrum-based technique
- histogram-based technique.

Additive watermarking

This is the most simple and straightforward method of spatial domain watermarking. Here, the secret data may be a digital data (ones and zeros) or integer values. To ensure the watermark, a secret key is generated between the sender and receiver. The simple example shown in Fig. 4 describes the additive watermarking process. Consider that the image matrix shown is to be embedded with a secret message of the sequence (2, 1, 0, –2, 1, –1, 2, 0, 1). The secret values are just added or subtracted with the input matrix. This method is very simple, and a high distortion occurs in the original image if the secret sequence has larger values.

LSB coding method

This was one of the earliest techniques used for the watermarking process. Each pixel value in a gray-scale image can be represent by 8 b. In this approach, the LSBs of the image pixel values are replaced with the secret data sequence. This method can embed only the digital data (ones and zeros), as shown in Fig. 5. The change in the LSB will have a drastic effect on the gray level. In LSB coding, the secret data can be exactly retrieved, but the robustness of the watermark is very low.

The major requirements of digital watermarking are transparency, robustness, and payload capacity (data capacity).

Patchwork-based technique

This is another simple spatial domain technique that is used to implement the secret message in an invisible manner. This is a statistical

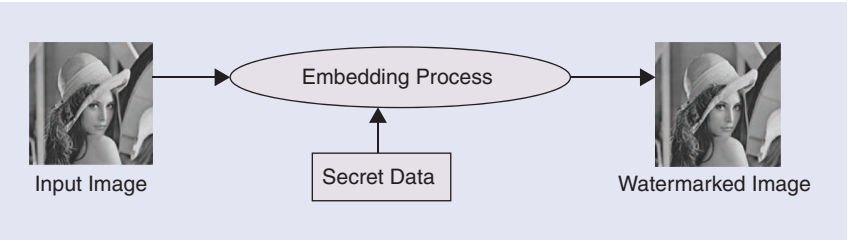

FIG2 The general spatial domain watermarking method.

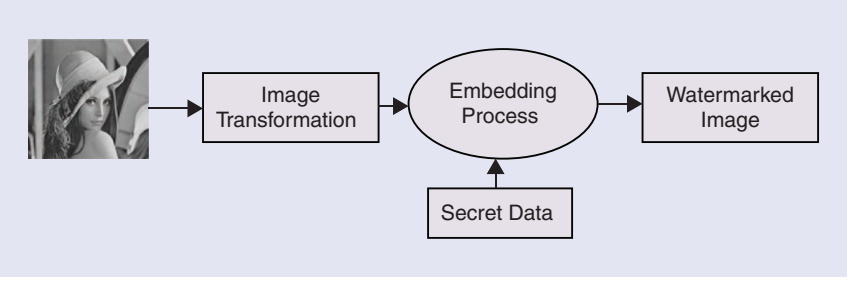

FIG3 The general transform domain watermarking method.

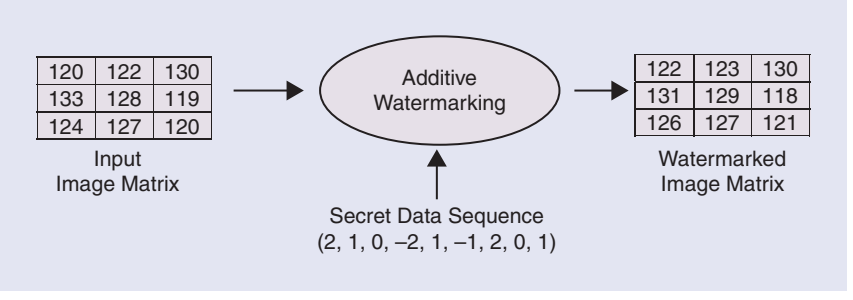

FIG4 Additive watermarking.

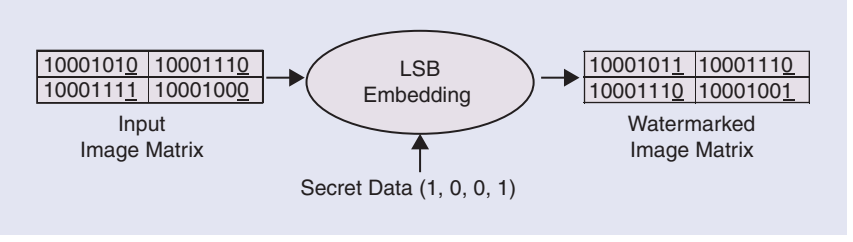

FIG5 The LSB embedding model.

The relation between the DCT and DFT is that the DCT uses the real parts of the DFT coefficients.

method in which the watermark is embedded by changing the distribution of some image pixel pairs. Here, the image pixels are randomly selected and divided into two patches (pairs); the watermark is embedded in one patch by increasing the brightness and in the other by decreasing the brightness (pixel values). The difference between the two pixels is calculated. If the randomly chosen pixel pair has a difference greater than the threshold, then a bit is embedded; otherwise, another pair is chosen. Here, the detection process is done by the linear correlation method by correlating the image with an image pattern consisting of one and –1 values.

Correlation-based technique

In this method, the watermark image W(x, y) pixels are directly added with the original cover image I(x, y) by multiplying it with a gain factor (g). The watermarked image $I_w(x, y)$ can be represent as

$$I_w(x, y) = I(x, y) + g\, W(x, y)$$

Here, the robustness of the image can be strengthened by increasing the gain value, but the watermark will be more perceptible, which leads to a decrease in the quality of the watermarked image.

Spread spectrum-based technique

In spread spectrum communication, narrow-band signals are transmitted over a wide bandwidth, so the amount of energy of any signal is very low for a wide range of frequencies. Similarly, in spread spectrum watermarking, the watermark signals with low energy levels are spread over the range of frequencies. The location of the watermark cannot be easily identified. The frequency regions should be selected in such a way that the watermark does degrade the original image quality. In this method, the watermark is inserted in the cover signal, as given by $V_{iw} = V_i + \alpha W_i$ where V_i is the input; W_i is the watermark signal; V_{iw} is the watermarked signal; and α is the scaling factor, which defines the level of watermark that can be added with the input.

Histogram-based techniques

A histogram is the graphical representation of an image; i.e., the distribution of the gray-scale values of the pixels is plotted (the gray-scale values versus number of pixels). In this method, a histogram of the cover image is plotted. Depending upon the watermarking algorithm, the histogram (pixel values) of the cover image is varied for the purpose of inserting the watermark into the cover image. The basic histogram-based watermarking technique performs the histogram-shifting (creating an empty space in selected pixels) operation. In the shifted areas of the histogram, the watermark values are embedded.

Transform domain watermarking techniques

Some of the transform domain techniques are based on the

- DFT
- DCT
- DWT
- Karhunen–Loeve transform
- singular value decomposition (SVD)
- radon transform.

DFT-based technique

The Fourier transform is a complex-valued transformation in which an image is represented with magnitude and phase values. Here, the magnitude values of the transformation contain only a little information about the image, but the phase coefficients contain all of the important information about it. In communications, phase modulation is highly noise resistive when compared to amplitude modulation. Therefore, if the watermark is added with the phase coefficients, the quality of the cover image is decreased, but the robustness of the watermark is high. When the watermark is embedded by modifying the magnitude coefficients, it will be less robust and may be affected by image compression methods.

DCT-based technique

The relation between the DCT and DFT is that the DCT uses the real parts of the DFT coefficients. The important information about the image is stores in the low-frequency components. The DCT divides the image components into the different frequency bands. The watermark can be embedded in the low- to middle-frequency bands of the DCT, which carries the information about the image. Watermarking using the DCT can be classified as global DCT (the whole image is considered) and block-based DCT watermarking (the image is separated into blocks).

DWT-based technique

This transform is based on finite duration signals called *wavelets*. In wavelet transform, the image is decomposed in three spatial directions, such as horizontal, vertical, and diagonal. The wavelet transform uses scaling functions related to a low-pass filter and a wavelet function associated with a high-pass filter. The input image is decomposed into four different subbands: the low–low (LL), high–low (HL) (horizontal), low–high (LH) (vertical), and high–high (HH) (diagonal) pass subbands. The LL subband has the DWT coefficients of larger magnitude, which are the most significant. The watermark can be added in any one of the subbands. The DWT has an advantage when compared with the other transforms; i.e., it provides the frequency spread of the watermark and spatial localization.

SVD technique

In this approach, the transformed images are represented in terms of linear algebraic methods instead of normal frequency domains. SVD-based watermarking algorithms have high robustness over various attacks. The SVD of image matrix I(x, y) can

be represented as $I = U \cdot S \cdot V^T$, where U and V are orthogonal column matrices, and S is a diagonal matrix that contains nonnegative values and has a size equal to the input image. The singular values correspond to the brightness of the image. In this method, the image is applied to have an SVD transformation. The watermark W is multiplied with the scaling factor a and added with the diagonal S matrix; i.e., $D = S + a \cdot W$. D matrix is applied for the transformation with SVD again, such as $D = U_w \cdot S_w \cdot V_w^T$. The watermarked image is calculated as $I_w = U \cdot S_w \cdot V^T$.

Parameters of calculation

The parameters that are used for an analysis of the watermarking algorithms are the PSNR, MSE, embedding capacity, bit rate, and BER. The squared difference between the original and watermarked images is known as the *MSE*

TABLE 1. A comparison of various watermarking methods.

TYPES	CLASSIFICATIONS	BASIC CONCEPTS	ADVANTAGES	DISADVANTAGES
Spatial domain watermarking	Additive watermarking	The secret data (integers) are directly added into the pixels.	• Simple method	• Low payload capacity • Low robustness
	LSB coding method	The LSB bit of each pixel is modified based on the secret data (one or zero).	• Implementation is simple.	• Low payload capacity • Secret data can be easily hacked.
	Patchwork-based technique	Image pixels are classified into two groups (patches). Then, one part of the secret data is added in one patch, and the other part is subtracted in another patch.	• Due to the random selection of patches, the security level of the data is increased.	• The watermark can be identified if the patches are known. • Low robustness
	Correlation-based technique	The secret data are multiplied with a gain factor; then, they are added directly to the image pixels	• The gain factor decides the robustness of watermark.	• If the robustness increases, the perceptibility also increases.
	Spread spectrum-based technique	The watermark signal is spread over a wide range of frequencies. Then, it is added with host image regions of the same frequencies.	• High embedding capacity • Identification of the watermark is difficult.	• Low robustness • The quality of the reconstructed image decreases.
	Histogram-based techniques	A histogram of the cover image is varied (creating space) for the purpose of inserting a watermark signal into the image.	• High embedding capacity • Highly robust	• High perceptibility • The quality of the reconstructed image decreases.
	DFT-based technique	The Fourier transform output can be represented by combining the magnitude and phase values. The watermark signal can be added in either the magnitude or phase components.	• High payload capacity • The robustness will be high if the watermark is added in the phase components.	• High perceptibility • Easily affected by compression methods
Frequency domain watermarking	DCT-based technique	The output of the DCT can be divided into frequency bands (low, middle, and high). The watermark can be added in the low- or middle-frequency bands.	• Increased payload capacity • Highly imperceptible	• Low robustness
	DWT-based technique	In the DWT approach, the image is decomposed into four subbands: LL, LH, HL, and HH. The significant values for the image reconstruction are stored in the LL band. Therefore, the watermark can be added in the other three subbands.	• Carries high payload • Highly robust	• Perceptibility is high.
	SVD-based technique	SVD transformation represents the images in terms of linear algebraic forms. The secret data are multiplied with a scaling factor and added to the SVD equations.	• High payload capacity	• SVD has low energy compaction. • Low robustness

We have tried to give complete basic information about digital watermarking, which will help new researchers to get the maximum knowledge in this domain.

$$\text{MSE} = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{i=1}^{M} (x(i, j) - w(i, j))^2,$$

where $x(i, j)$ is the original image, and $w(i, j)$ is the watermarked image.

The PSNR can be calculated as

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}} \right) (\text{dB}).$$

The *embedding capacity* of an image can be defined as the total number of secret message bits that can be added into an image without the degradation of its original quality. The *bit rate* of an image can be defined as the ratio of the number of bits added in an image to the number of bits present in the image

$$\text{bitrate} = \frac{\text{embedding capacity}}{\text{total number of pixels}}$$

The BER for a watermarking system can be defined as the ratio of the number of bits correctly retrieved by the receiver to the number of bits transmitted. Table 1 shows a comparison of various watermarking methods with their advantages and disadvantages.

Conclusion

In this article, we have presented various aspects of digital watermarking, including the principles, requirements, various domains with basic algorithms, and comparison parameters. Also, we have tried to give complete basic information about digital watermarking, which will help new researchers to get the maximum knowledge in this domain.

Read more about it

- P. K. Sharma Rajni, "Analysis of image watermarking using least significant bit algorithm," *Int. J. Inf. Sci. Techn. (IJIST)*, vol. 2, no. 4, pp. 95–101, Jul. 2012, doi: 10.5121/ijist.2012.2409.
- M. Kaur, S. Jindal, and S. Behal, "A study of digital image watermarking," *J. Res. Eng. Appl. Sci.*, vol. 2, no. 2, pp. 126–136, Feb. 2012.
- C. I. Podilchuk and E. J. Delp, "Digital watermarking: Algorithms and applications," *IEEE Signal Process. Mag.*, vol. 18, no. 4, pp. 33–46, Jul. 2001, doi: 10.1109/79.939835.
- M. Wu and B. Liu, "Data hiding in image and video .I. Fundamental issues and solutions," *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 685–695, Jun. 2003, doi: 10.1109/TIP.2003.810588.
- M. Wu, H. Yu, and B. Liu, "Data hiding in image and video .II. Designs and applications," *IEEE Trans. Image Process.*, vol. 12, no. 6, pp. 696–705, Jun. 2003, doi: 10.1109/TIP.2003.810589.

About the authors

Rajkumar Ramasamy (rajkumar ramasami@gmail.com) earned his B.E. degree in electronics and communication engineering, his master's degree in communication systems, and his Ph.D. degree in the domain of image processing. He is currently a professor in Department of Electronics and Communication Engineering, Siddartha Institute of Science and Technology, Puttur, Andhra Pradesh, 517583, India. His research interests include the fields of image processing, data mining, and wireless communications.

Vasuki Arumugam (vasuki.a.ece@kct.ac.in) earned her B.E. degree in electronics and communication engineering, her master's degree in applied electronics, and her Ph.D. degree in image compression from Anna University Chennai in 2010. She is currently a professor in the Department of Mechatronics Engineering, Kumaraguru College of Technology, Coimbatore, Tamilnadu, 641049, India. Her research interests include the field of data/image compression. She has published more than 30 papers in journals and at conferences.