A Review Study on Image Digital Watermarking

Charles Way Hun Fung
CPGEI / UTFPR
Avenida Sete de Setembro, 3165
Curitiba-PR - CEP 80.230-910
E-mail: charleswhfung@gmail.com

Antônio Gortan
CPGEI / UTFPR
Avenida Sete de Setembro, 3165
Curitiba-PR - CEP 80.230-910
E-mail: gortan@dgmdesign.com.br

Walter Godoy Junior
CPGEI / UTFPR
Avenida Sete de Setembro, 3165
Curitiba-PR - CEP 80.230-910
E-mail: godoy@utfpr.edu.br

***Abstract*—There has been an increase in broadcasting media since the begin of this century, because many techniques had been developed to solve this problem. Watermarking is the greatest bet from many researchs 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.**

***Index Terms*—digital watermarking; wavelets; security; dwt; 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.

II. Classification

A watermarking system has requirements which must be met when implemented, however the application will dictate which features should be emphasized. In this section a classification of marks according with their requirements will be proposed.

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 occured [3]. With that in mind the following classification can be made:

1) Fragile: These marks can be destructed 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 e 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:

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 frame, in audio files 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 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].

This paper is financial supported by CAPES-Brazil.

2) Non-Blind: In this case, the receiver needs the original data, or some derived information form 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 influence 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. Applications

Before discussing watermarking algorithms let us review some common applications:

A. 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.

B. 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 overcame the problem.

C. 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].

D. 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.

IV. 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 basic model proposed is presented below:

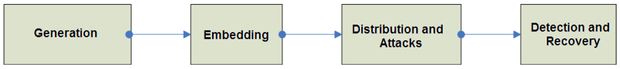

Fig. 1. Basic Model

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:

$$I_W = E(I, W) \tag{1}$$

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.

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:

- **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.

- **Singular Value Decomposition (SVD)**:

It 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:

$$A = USV^T \tag{2}$$

Where A is the original matrix, U and V are orthogonal matrices with dimensions MxM and NxN respectively, S is a diagonal matrix of the Eigenvalues of A and T indicates matrix transposition. [19] did the decomposition of the cover image and added the watermark using a scale coefficient α to get the following equation:

$$S + \alpha W = U_W S_W V_W^T \tag{3}$$

Multiplying matrices U, V^T and S_W result in the marked image A_W:

$$A_W = US_W V^T \tag{4}$$

This was possible due to the high stability of singular values (SV) of SVD. In another approach, the cover image was separated in blocks and the SVD applied to each block [20], in this case the dimension of watermark must be equal to the blocks size and a copy of the watermark is embeded in each block. This method improves watermark robustness and resistance against many kinds of attacks.

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:

- **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.

- **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.

- **Lifting Wavelet Transform (LWT)**:

Also called second generation wavelet transforms, its use has grown due to low memory consumption and easy implementation [26]. The following LWT scheme below is adapted from [7]:

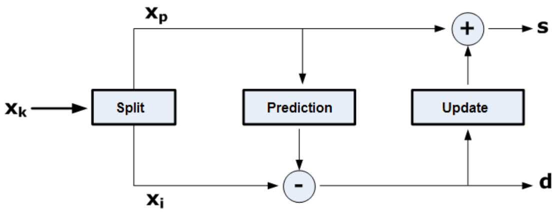

Fig. 2. Lifting scheme

In Figure 2, 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_p) 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:

$$h = x_i - P(x_p) \tag{5}$$

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

$$s = x_p - U(d) \tag{6}$$

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.

C. Distribution and Attacks:

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 was the original embedded mark in the cover image.

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.

V. 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 presented.

Many watermarking algorithms have been reviewed in the literature which show advantages in systems using wavelet transforms with SVD. 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.

In future works, the use of coding and cryptography watermarks will be approached. There is a large amount of literature on these topics showing that robustness increments can be gained through the addition of coding techniques.

Acknowledgments

The authors gratefully acknowledge CAPES-Brazil for financial support.

References

[1] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*. Morgan Kaufmann, 2008.

[2] M. El-Gayyar and J. von zur Gathen, “Watermarking techniques spatial domain,” University of Bonn Germany, Tech. Rep., 2006.

[3] J. Liu and X. He, “A review study on digital watermarking,” *First International Conference on Information and Communication Technologies*, pp. 337–341, 2005.

[4] M. Arnold, M. Schmucker, and S. D. Wolthusen, *Techniques and Applications of Digital Watermark and Content Protection*. Artech House, 2003.

[5] X. Wu, J. Hu, Z. Gu, and J. Huang, “A secure semi-fragile watermarking for image authentication based on integer wavelet transform with parameters,” *Australasian Information Security Workshop*, vol. 44, 2005.

[6] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet transform,” *IEEE Transaction on Image Processing*, vol. 1, pp. 205–220, 1992.

[7] I. Daubechies and W. Sweldens, “Factoring wavelet transforms into lifting steps,” *SIAM Journal on Mathematical Analysis*, 1997.

[8] V. M. Potdar, S. Han, and E. Chang, “A survey of digital image watermarking techniques,” *3rd IEEE International Conference on Industrial Informatics*, pp. 709–716, 2005.

[9] A. Tefas, A. Nikolaidis, N. Nikolaidis, V. Solachidis, S. Tsekeridou, and I. Pitas, “Performance analysis of watermarking schemes based on skewtent chaotic sequences,” *NSIP’01*, 2001.

[10] S.Teskeridou, V. Solochidis, N. Nikolaidis, A. Tefas, and I. Pitas, “Bernoulli shift generated chaoticwatermarks: Theoretic investigation,” *SCIA2001*, 2001.

[11] W. Yan, Z. Shi-qiang, and W. Yan-chun, “Wavelet digital watermark based on chaotic sequence,” *ICICIC’08*, 2008.

[12] K. L. W. G. Natasa Terzija, Markus Repges, “Digital image watermarking using discrete wavelet transform: Performance comparison of error correction codes,” *Visualization, Imaging and Image Processing*, 2002.

[13] L. Haiyan, Z. Xuefeng, and W. Ying, “Analysis of the performance of error correcting coding in audio watermarking,” *3rd IEEE Conference on Industrial Electronics and Applications*, pp. 843–848, 2008.

[14] P. Cika, “Watermarking scheme based on discrete wavelet transform and error-correction codes,” *16th International Conference on Systems, Signals and Image Processing*, pp. 1–4, 2009.

[15] A. Bastug and B. Sankur, “Improving the payload of watermarking channels via ldpc coding,” *Signal Processing Letters*, vol. 11, pp. 90–92, 2004.

[16] C. Nafornita, A. Isar, and M. Kovaci, “Increasing watermarking robustness using turbo codes,” *International Symposium on Intelligent Signal Processing*, pp. 113–118, 2009.

[17] N. Pantuwong and N. Chotikakamthorn, “Line watermark embedding method for affine transformed images,” *ISSPA 2007*, pp. 1–4, 2007.

[18] S. Riaz, M. Y. Javed, and M. A. Anjum, “Invisible watermarking schemes in spatial and frequency domains,” *International Conference on Emerging Technologies*, 2008.

[19] R. Liu and T. Tan, “An svd-based watermarking scheme for protecting rightful ownership,” *IEEE TRANSACTIONS ON MULTIMEDIA*, vol. 4, pp. 121–128, 2002.

[20] R. A. Ghazy, N. A. El-Fishawy, M. M. Hadhoud, M. I. Dessouky, and F. E. A. E.-S. Samie, “An efficient block-by-block svd-based image watermarking scheme,” *National Radio Science Conference*, pp. 1–9, 2007.

[21] D. Taskovski, S. Bogdanova, and M. Bogdanov, “Digital watermarking in wavelet domain,” *FIRST IEEE BALKAN CONFERENCE ON SIGNAL PROCESSING, COMMUNICATIONS, CIRCUITS, AND SYSTEMS*, 2000.

[22] G. Hai-ying, L. Guo-qiang, L. Xu, and X. Yin, “A robust watermark algorithm for jpeg2000 images,” *Fifth International Conference on Information Assurance and Security*, 2009.

[23] D. R. Sans, “Identificao de propriedade em imagens com marcas d’gua no domnio da transformada wavelet,” Master’s thesis, Universidade Federal do Paran - UFPR, 2008, in portuguese.

[24] E. Ganic and A. M. Eskicioglu, “Robust dwt-svd domain image watermarking: Embedding data in all frequencies,” *Proceedings of the 2004 workshop on Multimedia and security*, pp. 166 – 174, 2004.

[25] P. Bao and X. Ma, “Image adaptive watermarking using wavelet domain singular value decomposition,” *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, vol. 15, pp. 96–102, 2005.

[26] K. Loukhaoukha and J. Y. Chouinard, “Hybrid watermarking algorithm based on svd and lifting wavelet transform for ownership verification,” *11th Canadian Workshop on Information Theory*, pp. 177–182, 2009.

[27] F. B. C. Mendes, “Uma proposta de assinatura digital para imagens por meio de marca d’gua,” Master’s thesis, Universidade Federal de Uberlandia, 2008, in portuguese.

[28] F. Hartung, J. K. Su, , and B. Girod, “Spread spectrum watermarking: Malicious attacks and counterattacks,” pp. 147–158, 1999.