Available Online at www.ijecse.org **ISSN- 2277-1956**

Fragile Watermarking for Image Authentication: Survey

Pragya Jain [1] , Anand S. Rajawat [2]
[1, 2]Department of Computer Science and Engineering
[1, 2]Shri Vaishnav Institute of Technology and Science, Indore, India
[1] plathiya42@gmail.com
[2] rajawat_iet@yahoo.in

Abstract- **In this paper, we report results from comparative study on various related and relevant aspects of the digital watermarking such as image authentication techniques using fragile watermarking, fuzzy clustering and genetically inspired watermarking techniques for integrity verification. This is carried out with intent to develop an understanding of their working, contained challenges, possible attack scenarios, advantages and limitations.**

Keywords–Digital watermarking, fragile watermarking, Genetic algorithms, Fuzzy C-means clustering, PSNR

I. INTRODUCTION

Digitizing of multimedia data has enabled reliable, faster and efficient storage, transfer and processing of digital data [1]. The ease with which the multimedia data can be utilized, amplifies possibilities to perform illegal copying and redistribution of the multimedia data. Multimedia data in the form of image, audio and video, can be easily manipulated and reproduced in digital domain using present day multimedia manipulation tools. There is great demand of techniques for handling the problems associated with the multimedia data. In this context, it is important to develop systems for copyright protection, protection against duplication, and authentication of content [2]. Techniques of data hiding specifically watermarking can be applied to protect multimedia data against these types of manipulations and duplications. Digital watermarking is a method to attest the owner identification of the multimedia data and discourage the unauthorized copying, by embedding perceptually transparent pattern in digital data by specially designed algorithms. Generally, a watermark represents the owner and the user's information which could be owner's logo or some control information suitable for embedding in the cover multimedia. Watermarking techniques are judged on the basis of their performance on a small set of properties. These properties include robustness, transparency, watermarking capacity, blind detection and security. Watermarking schemes are developed according to the requirements of the application and all applications do not require each of these properties in their entirety i. e. watermarking requirements are application dependent and some most desirable properties for these applications are conflicting in nature. A huge trade-off among them is often involved.

Digital watermarking techniques are classified according to various criteria like robustness, perceptibility and embedding and retrieval methods. Robustness is an important criterion which means the ability of watermark to resist common image processing operations. Watermarking techniques based on robustness can be further divided into three main categories:

(1) Robust
(2) Fragile, and
(3) semi-fragile

Robust watermarking schemes are applied for proving ownership claims whereas fragile watermarking is applied to multimedia content authentication. These watermarking schemes have their own requirements in terms of robustness. Robust watermarks should be able to survive a wide range of friendly operations and malicious attacks, whereas fragile watermarks are intolerable to both malicious and content preserving operations. Fragile watermarking techniques are designed with a goal to identify and report every possible tampered region in the watermarked digital media. Semi-fragile watermarks are intermediate in robustness between the two and are also used for image authentication. Some critical applications like medical imagining and forensic image archiving also requires the fragile watermarks to be reversible. The different quantitative parameters such as PSNR, True and false positive may be used for the evaluation of the method of watermarking schemes.

IJECSE paper format font should be 10 in times new roman with single spacing. In recent years, the accessing of multimedia data or digital data has become very easy because of the fast development of the Internet. In other words, this development makes unauthorized distribution of multimedia data. For the protection of multimedia data, a solution known as watermarking is used. After the approximate 20 years' research, different kinds of watermarking algorithm based on different theory concepts were introduced [1-3]. A digital watermark encodes the owner's license information and embeds it into data. Watermarking may be used to identify the image of owners' license information and to track illegal copies.

The rest of the paper is organized as follows. Proposed embedding and extraction algorithms are explained in section II. Experimental results are presented in section III. Concluding remarks are given in section IV.

II. Literature Survey of Fragile Watermarking Techniques:

According to embedding and retrieval criteria, fragile watermarking techniques for image authentication can be divided into two categories: spatial domain and transform domain.

II.I *Literature Survey of Fragile Watermarking Techniques in Spatial Domain:*

In 2009, Chen et al., [7] proposed a spatial domain watermarking technique based on the idea of incorporating block-wise dependency information in watermarking procedure for thwarting VQ attack without compromising on localization capabilities of the scheme. The block-wise dependency relationship between the blocks of the image is established using fuzzy clustering criteria; a fuzzy C-means algorithm is used for this purpose. This method allows one piece of data to belong to two or more clusters unlike other traditional hard clustering schemes like k-means algorithm that assign data points to a specific cluster. The scheme consists of authentication data embedding procedure and tamper detection procedure. The original 8 bit gray scale image of size M x M pixels is divided into non-overlapping 2 x 2 blocks which are arranged in a specific order. Two LSB of each pixel within each block is then set to zero. The FCM clustering of these image blocks is then performed to fuzzy cluster given image blocks into C clusters. Outcomes of FCM are a set of cluster centers V and a fuzzy membership matrix U. The membership matrix is then used to obtain a feature sequence F, which is XORed with a random sequence created by pseudorandom number generator (PRNG), seeded with the secret key SK. This generates authentication data which is embedded into two least significant bits of each image block to produce watermarked image. The set of cluster centers V and the secret key SK is kept secure for use in tamper detection and localization phase. After the embedding process, the watermarked image enters the watermark channel where it may be subject to malicious or non-malicious attacks. During tamper detection and localization, by verifying the authentication data embedded in each image block, it is possible to determine whether an image block has been tampered with or not.

In 2011, Bhattacharya et. al. [10] proposed a new approach which makes use of both fragile and robust watermarking techniques. The embedded fragile watermark is used to assess the degradation undergone by the transmitted images. Robust image features, on the other hand, are used to construct the reference watermark from the received image, for assessing the amount of degradation of the fragile watermark. In the same year, Hernandez and Torres-Huitzil [11] presented a chaos theory based fragile watermarking scheme for image authentication in mobile devices. In such scheme a fragile signal that is sensitive to manipulations is embedded in the image so as to detect the image tampering inconsistency.

In 2011, Yan et. al. [12] presented a blind watermarking approach to protecting vector geo-spatial data from illegal use. The presented method is rarely affected by data format change, random noise, similarity transformation of the data, and data editing.

In 2012, Chen et. al. [13] proposed A Watermarking Technique based on the Frequency Domain. A modified algorithm is presented to improve the defect of the JPEG quantification in order to reduce the bit error rate (BER) of the retrieved watermark. Addition, two parameters are regarded as the controlling factors. They are used to adjust the value of the DCT coefficient in order to trade-off the qualities between the watermarked images and retrieve watermark. Moreover, the proposed algorithm is design as a blind mechanism. Thus, the original image and watermark are not needed for extracting watermark.

II.II *Literature Survey of Fragile Watermarking Techniques in Transform Domain:*

Frequency domain techniques have proved to be more effective than spatial domain techniques in achieving high robustness against attacks and can embed more bits of watermark. Commonly used frequency domain transforms are Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT), and Discrete Fourier transform (DFT). DWT has been given special attention in digital image watermarking due to its excellent spatial localization and multi-resolution characteristics, which are similar to the theoretical models of the human visual system. Transform domain techniques have been applied for copyright protection and image authentication. A brief description of some techniques for image authentication is covered below.

In 2008, Wang H. et al., [9] proposed a chaotic watermarking scheme for authentication of JPEG images. The quantized DCT coefficients after entropy decoding are mapped to the initial values of the chaotic system, and then the generated watermark information by chaotic iteration is embedded into JPEG compressed domain. Re-quantization operation does not invalidate tamper detection due to direct modification of DCT coefficient after quantization. Extraction is also performed in the compression domain. Extraction is fast and complexity of method is claimed to be low.

In 2012, Kannammal et. al. [14] proposed a digital watermarking framework in which the Electrocardiograph (ECG) and Patients demographic text ID act as double watermarks. By this method the medical information of the patient is protected and mismatching of diagnostic information is prevented.

Transform domain techniques are in greater use now a days in place of spatial domain techniques as much is known about the properties of these transforms to achieve better watermark characteristics.

II.III ***Literature Survey of Fragile Watermarking Techniques using GAs:***

Genetic Algorithms (GAs) are promising evolutionary method for solving difficult engineering, scientific and industrial problems. GA's are a part of a relatively new movement in computer science that explores biologically inspired approaches to computation. Genetic algorithms (GAs) provide a randomized, parallel, and global search approach to find the optimum solution of problems, based on the mechanics of natural selection and natural genetics [3]. In recent years, the optimization capabilities of genetic algorithms have been explored in many different areas such as music generation, digital watermarking, genetic synthesis, VLSI technology, strategy planning, and machine learning for achieving better results in many combinatorial optimization problems from these domains.

Genetic algorithms (GAs) were invented by John Holland in the 1960s during his studies of cellular automata. Holland's original goal was to formally study and mathematically analyze the phenomenon of adaptation as it occurs in nature and to develop ways in which the mechanisms of natural adaptation might be imported into computer systems for finding approximate solutions to complex problems. The optimized GA can be used for the clustering, which can further be used for the watermarking of the images [4].

In 2007 Chen et al. [32], proposed use of GAs in image authentication procedure to improve the image quality of the protected image. Correlations between important DCT coefficients and user defined thresholds constitute the image authentication message.GA was employed to find near optimal position for embedding authentication data.

In 2011, Lai [15] has proposed a robust digital image watermarking scheme based on singular value decomposition (SVD) and a tiny genetic algorithm (Tiny-GA). With the proposed scheme, the embedded watermark can successfully survive after attacked by image-processing operations. The proposed approach were using human visual systems and comparing the results with current state of the art SVD-based image watermarking techniques.

In 2011, Lipinski and Stolarek [16] presented a genetic-based enhancement of digital image watermarking in the Discrete Wavelet Transform domain is presented. The proposed method is based on adaptive synthesis of a mother wavelet used for image decomposition. Effectiveness of the proposed method is demonstrated by comparing watermarking results using synthesized wavelets and the most commonly used Daubechies wavelets. Experiments demonstrate that mother wavelet selection is an important part of a watermark embedding process and can influence watermarking robustness, separability, and fidelity.

III. Desirable Properties of Fragile Watermarking Schemes:

Every watermarking technique is designed by keeping a particular application in mind. The features and their relative importance that watermarking technique should possess are also application dependent. Giving paramount attention to this, we now present desirable features of fragile marking systems [6, 7]:

1. Tamper detection

IJECSE
International Journal of Electronics and Computer Science Engineering
WWW.IJECSE.ORG

2. Perceptual Transparency
3. Detection should not require the original image
4. Detector should be able to locate and characterize alterations made to a marked image
5. The watermarks generated by different marking keys should be "orthogonal" during watermark detection
6. The marking key spaces should be large
7. The marking key should be difficult to deduce from the detection side information
8. The insertion of a mark by unauthorized parties should be difficult

A good fragile watermarking scheme should demonstrate all desirable properties enlisted above. The initial research in this field was confined to improving perceptual transparency, tamper detection rate and designing blind systems. The use of cryptographic techniques in watermark embedding and extraction is adopted by many authors. Special attacks like collage attack, VQ attacks [8], transplantation attacks were carefully designed to test the effectiveness of fragile watermarking techniques. In recent years, the issues of accurate tamper localization and recovery have been given equal importance as tamper detection and many techniques in these areas have been proposed.

IV. Attacks on Fragile Watermarks:

Fragile watermarks are embedded in the cover media to detect any occurrence of tampering in it. If the alterations are so performed on the watermarked image that they do not disturbed the embedded watermark, then the altered image can still be authenticated defeating the purpose of watermark embedding. Many block wise independent techniques are to be vulnerable to counterfeiting attacks [8]. Some counterfeiting attacks common to fragile watermarks are briefly defined in this section:

(a) Cut and paste attack: In cut-and-paste attack, the attacker modifies the content of a watermarked image by cutting regions from the same or another watermarked image and pasting them together to form a new image.

(b) Birthday attack/collage attack: Birthday attacks constitute a more sophisticated and powerful means of subverting digital signatures. The attacker searches for collisions i.e. pairs of blocks that hash to the same value, thus having the same signature. A hash function that produces a bit string of length l, the probability of finding at least two blocks that hash to the same output is greater than 0.5 whenever roughly 2l/2 watermarked blocks are available. The idea of the attack is to forge a new watermarked image (a collage) from a number of authenticated images watermarked with the same key and watermark/logo by combining portions of various authenticated images while maintaining their relative positions in the forged version. In general, the only protection against birthday attacks is to increase the hash size. The attack is also termed as collage or VQ attack. In vector quantization attack, a VQ codebook generated from a set of watermarked images. Since each block is authenticated by itself, the counterfeit image appears authentic to the watermarking scheme.

Other sophisticated attacks have also been designed like transplantation attacks, which require the block wise dependency to be nondeterministic.

V. Methodology Image Authentication Framework:

Since ancient times, measures have been taken to satisfy the need for authentication of important documents and valuable art works, such as the signing and seal stamping of these items [5]. The importance of authentication and IPR protection has became more apparent and sensitive in this digital information era. Given the power of digital multimedia processing tools for perfect duplication and reproduction along with the internet technology for fast transfer of digital contents, forgery and impersonation have become major concerns of digital age. Therefore, image authentication and integrity verification have become an active research area in recent years.

Various types of watermarking schemes have been proposed for different applications according to their needs. Watermarks for copyright protection are robust in the sense that they are designed to survive various kinds of manipulation to some extent, provided that the visual acceptability and commercial value of the altered images is retained [4]. On the contrary, the schemes for authentication and verification of content integrity are usually fragile in the sense that, when attacked, the embedded watermark should be entirely or locally destroyed, depending on whether the attack is a global or local tampering, so that alarms can be raised when wrong watermark is extracted [5]. Such watermarking schemes work as a convenient tool for authentication of visual information, tamper detection, and verification of image integrity. Fragile watermarking is usually compared with pure cryptographic techniques for authentication. Watermarking as opposed to pure cryptographic tools enables us to localize tampered or damaged areas or even authenticate with a degree without having to store any additional information about the image. This is

typically achieved at the expense of introducing a slight amount of distortion into the image. Because the auxiliary authentication information is embedded in the image itself rather than appended, it gives us more flexible and convenient tool for investigating the image integrity.

In real life scenarios like medical, forensic, broadcasting, and military, content verification and identity authentication are much more of a concern; more emphasis is focused on the capability of the watermarking schemes to detect forgeries and masquerade. For example, the staff in a military headquarter always has to be sure about the authenticity and content integrity of the digital images before planning any action. For all such cases fragile watermarking schemes have been used successfully.

In a fragile marking system, a signal (watermark) is embedded within an image such that subsequent alterations to the watermarked image can be detected with high probability. The framework for embedding and detecting a fragile mark is similar to that of any watermarking system [6]. In embedding phase (Fig 1.1), the authentication watermark is generated using a secret marking key. The generated watermark is embedded in to an original image by an owner (or an independent third party authority). The original image is kept secret or may not even be available in some applications such as digital camera. The marked image is which is perceptually identical to the original image under normal observation may be transmitted, presented, or distributed.

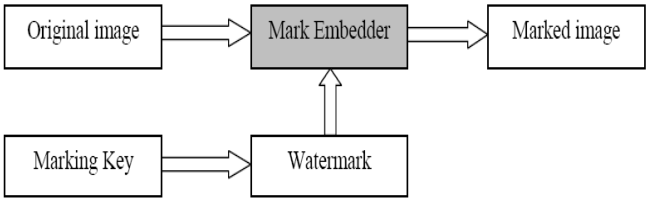

Fig V.I- An image authentication framework: Watermark Embedding

Upon arrival of an image the user uses the detector to evaluate the authenticity of the received image (see Figure 1.2). The detection process also requires knowledge of "side information" like the marking key, the watermark, the original image, or other information. The detector is usually based on statistical detection theory whereby a test statistic is generated and from that test statistic the image is determined to be authentic. If it is not authentic then it would be desirable for the detector to determine where the image has been modified.

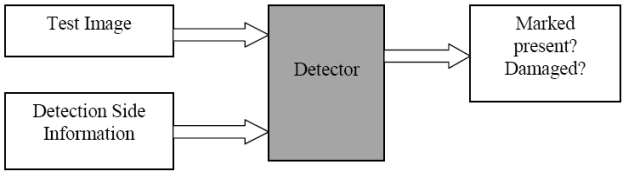

Fig V.II - An image authentication framework: Watermark Detection

The side information used by the detector is very important in the overall use of a fragile- watermark. Techniques that require that the detector have the original image are known as informed techniques while techniques that do not require the detector to refer to the original image are known as blind. To be effective a fragile watermarking system should be blind. In many applications the original image may never be available since it might have been watermarked immediately upon creation as in the case of digital camera.

VI .CONCLUSION

Image authentication is a problem that can be efficiently solved with either fragile or semi-fragile watermarks. A framework for fragile watermarking scheme is presented with its desirable properties. Many fragile watermarking techniques in spatial domain and transform domain have been discussed in details. Evolutionary algorithms like GAs are getting thrust in their use in watermarking techniques. They are capable of solving complicated optimization problems with great simplicity, utilizing principle of evolution and survival of fittest. In watermarking, general optimization criteria are robustness and/or image quality. On the basis of above literature survey a genetically based scheme for image authentication and tamper proofing would be used for fragile

IJECSE
International Journal of Electronics and Computer Science Engineering
WWW.IJECSE.ORG

watermarking. The proposed approach may detect effective alterations for the watermarking of images and would provide better results than the existing method of fragile watermarking schemes.

VII. REFERENCE

[1] I. Cox, M. Millar and J. Bloom (2002), "Digital watermarking", Morgan-Kaufmann, San Francisco, CA.

[2] C. Rey, and J. L. Dugelay (2002), "A survey of watermarking algorithms for image authentication", EURASIP Journal on Applied Signal Processing, Vol. 6, pp. 613-621.

[3] Juergen Seitz (2005), "Digital Watermarking For Digital Media", University of Cooperative Education Heidenheim, Germany.

[4] S. Radharani, M. L. Valarmathi (2010), "A study of watermarking scheme for image authentication", International Journal of Computer Applications, Vol. 2, No.4, pp. 24-32.

[5] C.-T. Li, and F.-M. Yang (2003), "One-dimensional neighborhood forming strategy for fragile watermarking", Journal of Electronic Imaging, Vol.12, No. 2, pp. 284-291.

[6] E. T. Lin, and E. J. Delp (1999), "A review of fragile image watermarks", In Proceedings of the ACM multimedia and security workshop, pp. 25–29.

[7] W.C. Chen, and M.S. Wang (2009), "A Fuzzy c-Means Clustering based Fragile Watermarking Scheme for Image Authentication", Expert Systems with Applications, Volume 36, Issue 2, Part 1, pp. 1300-1307.

[8] M. Holliman, and N. Memon (2000), "Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemes", IEEE Trans. Image Processing, vol. 9, no. 3, pp. 432-441.

[9] Hongxia Wang, Ke Ding, Changxing Liao (2008), "Chaotic Watermarking Scheme for Authentication of JPEG Images", International Symposium on Biometrics and Security Technologies, pp. 1-4.

[10] Ankan Bhattacharya, Sarbani Palit, Nivedita Chatterjee, and Gourav Roy (2011), "Blind assessment of image quality employing fragile watermarking", 7th International Sym. on Image and Signal Processing and Analysis (ISPA 2011) Dubrovnik, Croatia, pp. 431-436.

[11] Cynthia Palma Hern andez, Cesar Torres-Huitzil (2011)," A fragile watermarking scheme for image authentication in mobile devices", 8th International Conference on Electrical Engineering Computing Science and Automatic Control (CCE), pp. 1-6.

[12] Haowen Yan, Jonathan Li, Hong Wen (2011), "A key points-based blind watermarking approach for vector geo-spatial data", Elsevier Journal of Computers, Environment and Urban Systems, Volume 35, Issue 6, pp. 485–492.

[13] Huang-Chi Chen, Yu-Wen Chang, Rey-Chue Hwang (2012), "A Watermarking Technique based on the Frequency Domain", Journal of Multimedia, Vol. 7, No. 1, pp. 82-89.

[14] A. Kannammal, K. Pavithra, S. Subha Rani (2012), "Double Watermarking of Dicom Medical Images using Wavelet Decomposition Technique", European Journal of Scientific Research, Vol. 70, No. 1, pp. 46-55.

[15] Chih-Chin Lai (2011), "A digital watermarking scheme based on singular value decomposition and tiny genetic algorithm", Elsevier Journal of Digital Signal Processing, Vol. 21, pp. 522–527.

[16] Piotr Lipiński, and Jan Stolarek (2011)," Digital Watermarking Enhancement Using Wavelet Filter Parameterization", Adaptive and Natural Computing Algorithms, Lecture Notes in Computer Science Adaptive and Natural Computing Algorithms, Vol. 6593/2011, pp. 330-339.