International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-1 Issue-5, June 2012

Fragile Watermarking Scheme for Images in Transform Domain

Nitish Agarwal

Abstract*— Watermarking had an explosive amount of work done in the past decade. The field has diverse application areas, and a large number of fields amalgamate to ensure different requirement in different circumstances is being met by watermarking. We will be focusing in this paper mainly on the techniques which come under authenticating an image as to see if the image which is being used as a proof is not tampered in any way. This comes as the subset named as fragile watermarking in the field of watermarking techniques. We propose a novel method which focuses mainly on JPEG format of images and takes the lossy compression into consideration. The method works for both grayscale as well as the color images.

***Index Terms*—Watermarking, Fragile Watermarking, Image Authentication, Transform Domain, JPEG, Hash, MD5, Digital Signature Algorithm, HMAC.**

I. INTRODUCTION

Watermarking is the field which is derived as a subset of Information Hiding field, but which has a completely different purpose as compared to the other part of information hiding i.e. Steganography. In Steganography we try to conceal the presence of the message in the medium which is carrying it. Also the main important material is the hidden message while the medium which is carrying the message is of little importance in the sense of value but it still has to conceal the presence of the message completely. Whereas in watermarking the main importance is given to the medium which is carrying the watermark. By the previous statement we mean that the quality of the medium which is carrying the watermark should not be reduced in the process of embedding the watermark as the medium here is of real value. Watermark is just for controlling the media in a broad sense. The digital media types are not restricted to images but also include audio, video and text documents etc. The watermarking field is classified into various types depending on the area of application i.e. the field of application. Watermarking has been classified into visible and invisible watermark type by [2]. But generally they are classified as fragile, semi-fragile and robust [16]. Robust Watermarking is most widely used for copyright protection. These method are robust in the sense that watermark cannot be removed from the digital image unless the digital media content is harmed in such a way that it becomes unusable.

Manuscript published on 30 June 2012.
***** Correspondence Author (s)
Nitish Agarwal, Electrical and Computer Department, University Of Waterloo, Waterloo, Canada.

They have to show robustness in terms of attacks on watermarks to remove them from the digital media. They can handle various types of intentional and unintentional attacks such as D/A and A/D conversion, lossy compression, cropping etc. Thus these methods are useful in various practical applications such as copyright protection, broadcast monitoring, copy prevention and control, fingerprinting etc. There have been a large amount of work done on this type of watermark and numerous algorithms have been proposed, some of the most popular are related to spread spectrum [5] and QIM [3]. The watermarking evolves basically based on improving the method once it has been broken i.e. filling up the security holes. Several methods to remove the watermark embedded into the digital media have been proposed. Some of the methods are really powerful like the attack by conspiracy, in which several copy of the watermarked data are used to create one copy of the media which does not have the watermark. There are many other attacks like inversion attack, uncorrelated noise attack, overmarking attack, etc [6].

In terms of the digital media under consideration our main focus here is on images. Thus from now on in this paper we will refer to the watermarked content as an image. Fragile Watermarking addresses completely different side of Watermarking application. Here the attacker has to take care of not to disturb the structure of the watermark. Here the main emphasis is on Image Authentication in terms of mainly data integrity check. Watermark is easily destroyed by small amount of alteration to the image. Thus the focus here is not to make the watermark robust but to make the watermark to easily detect any alterations done to the image. But one of the major drawbacks of this as we come across is that the simple operations such as lossy image compression are also in terms of violation of image authentication and thus image has to be stored with lossless compression or raw format which can take up a lot of space. The method proposed here handles this major concern with a simple method which works in the transform domain. Transform domain watermarking is also useful as it takes advantage of the imperceptibility in the embedding process. The idea is fairly simple, but has been adapted to be resistant to some of the known attacks on fragile watermarking systems. The fragile watermarking is although simple but does have some important practical applications in the areas such as image authentication which is required when some image is presented as proof of something in a court proceeding.

This also used in areas such as military application where any tampering to image content can be disastrous and also can be used in area of medical where modification to image content is not desirable characteristic and any change can make the image unusable.

Semi-fragile technique somehow comes as type of tradeoff technique which bridges the gap between the two main domains. They can be fairly robust to some types of changes but can be removed under fairly high amount of modification to the image. The paper flow goes as follows: In the following section we talk about various fragile methods currently present and what they fall short on. In the next section we propose the new method. In the last section we provide results which are marked on the basis of image quality. We conclude by providing some of the additional extensions which can be seen as an optional set for enhancing the security but are not completely necessary and heavily depend on the application.

II. FRAGILE WATERMARKING TECHNIQUES

Fragile Watermarking field was mainly proposed to for authentication of image content, so as to detect whether the image under consideration was modified or not. This is kind of data integrity check where the data is discarded if it is found that it has been tampered with. Similar is the case with images, once it is detected that image has been tampered with it is discarded. Here the main focus is to make the watermark sensitive to intentional changes to the image, so that they can be detected easily. Problem is with saving the integrity of the watermark against unintentional changes such as lossy compression, etc.

A large chunk of previously proposed method is working in the Spatial Domain. One of the earliest methods is more of the procedure for image authentication which proposed the authentication mechanism to be built into the camera itself from the image is generated [9]. This method did provide some of the basic structure but did not embed the data into the image, it appended the authentication data as a header or trailer to the image. For guaranteeing integrity of image as well as authenticity, some of the trusted cryptography methods can be used such as Hashing (MD5, SHA-1) and use of Digital Signatures [15]. One of the most worked upon method is Yeung and Mintzer scheme [20] it had a seemingly simple idea in terms of procedure, i.e. it used a simple watermark (normally a binary logo or binary image) which is embedded in the image and the embedding procedure resulted in a verification key. This verification key was used as a LUT for extraction and verification of the watermark which was embedded into the image. This method had a major weakness which is discussed in [8] with the assumption that same binary logo was used for multiple images. In a stronger form of attack given in [18], a single image and access to the verifier is enough to break the above system, this method learns mapping from the continuous sending of the images embedded with watermark with tampering of some pixels.

The next watermarking method discussed had less overhead in terms of processing as it did not implement any of the cryptography methods and thus was very fast in terms of implementation, it used a binary watermark which was embedded in to the image in the spatial domain with this XOR of the bitplane was also used [14]. One of the drawbacks is in terms of the watermark being a binary image, so if we can extract a some of the watermark we can reconstruct the whole watermark as it corresponds to a natural image which is in binary form. This is one of the major drawbacks which can be used to attack most of the system which used binary images as watermark but can be addressed by using a timestamp or pseudo random sequence instead of the binary image. Many methods proposed such as [7] have adopted this.

A new method proposed in [17] used public cryptography as major part of image watermarking scheme. This method also provided the localization property i.e. not only the modification to the watermark embedded image was detected but also the place where the modification done was also given. It used the private key to sign the watermark and embed it into the image so the authentication procedure was also in place in the algorithm. But the block wise independence in the above algorithm was exploited as shown in [10] which used a method to characterize pixel based on the watermark embedded, it assumed that several images were watermarked using the same key. This also has been addressed by a new method proposed in [1] but a lot of overhead has been added in place to thwart the above attack as the same procedure being applied multiple times at different hierarchical levels.

Another method works on a totally different criterion to embed the watermark in the image. It uses the Histogram shifting technique to make space to embed the watermark [11]. The location map is used to extract back the watermark, the experimental result show that a large amount of data can be embedded using this method being used iteratively. This method falls apart in case of image undergoing lossy compression, also the method is based on the assumption that shifting the histogram by a unit is normally imperceptible but this method may create problem in case of colored images with fixed pallette to choose from [13].

An elegant structure for privacy control has been specified in the paper [14]. They provide the basic structure of the Watermarking as well as the authentication procedure with or without involvement of the source parties. In transform domain the following method was proposed [19], this method uses a propretary LUT table to embed the data. The use of look up table is not very secure as can be seen from the paper [8]. Although the attack is proposed for Yeung method it can easily be extended to the method provided by this paper. Use of cryptographic algorithms is much more secure than using a LUT.

III. PROPOSED DATA EMBEDDING SCHEME

The method proposed here embeds the watermark in the Image which is undergoing JPEG compression. The embedding procedure is performed after the lossy quantization step so that the embedded watermark is not harmed, as after the quantization step Huffman encoding procedure is performed which is lossless.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The technique shown here can be seen as a two way authentication as well as verification procedure. The method provide a choice of using digital signature which uses two set of keys i.e. a private key and a public key or using a keyed hash function which eliminates the use of the a somewhat computationally intensive DSA with the following tradeoff a single key being used which has to be provided by the owner incase authenticity of the image is to be checked.

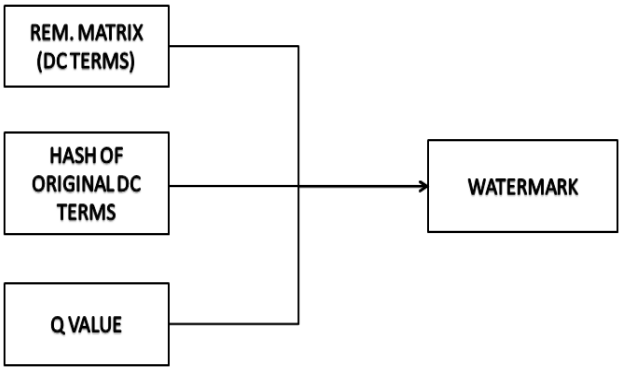

Fig. 1. Watermark Structure

The watermark structure is formed from the remainder obtained from the DC component in the DCT coefficients. These components are compressed using Huffman coding to reduce the size of data to be embedded. Also the original DC components are hashed and stored. This does not require much bits as the resultant after hashing is just 128 bit. The quality factor value is concatenated with the hash value. The quality factor value is required at the time of retrieval of watermark. These two data set are embedded into the image after the quantization step in the DCT coefficients in the LSB, leaving the first value of each DCT coefficient set as this value is not be disturbed as is being used for retrieval purpose. As each of the DCT coefficients is related to the 64 pixels in the block thus the imperceptibility criterion is satisfied as changing the LSB of the coefficients has no perceptible change in the image. The binary data set which is embedded as watermark is not a binary image and is thus secure from the point of view of attack that used the structure of the binary image to find the watermark.

The embedding procedure is not blockwise and all the dataset are connected as each element in the dataset correspond to each block DC component thus the attack previously specified [10] regarding the independent block embedding attack does not work here.

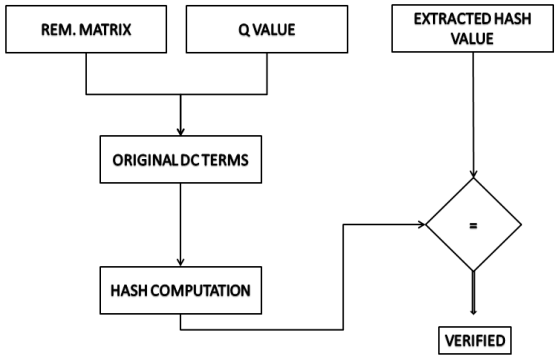

Fig. 2. Extraction Procedure

During the extraction procedure the dataset one is extracted back from the image given. This is decoded to get back the remainder matrix. The values obtained after decoding are used with the quality factor which corresponds to the matrix used for quantization to get back the original DC coefficients. As the quantization matrix are standard for JPEG compression thus these matrix are readily available for original coefficient computation. These coefficients are hashed using technique md5 and compared to the second dataset hash coefficients. If they match the image is verified otherwise we can say that the image has been modified. Due to hashing property a one bit change changes approximately 50\% of the bits in the resulting hash values (Avalanche Effect), this shows that a small amount of modification will totally change the hash value and thus can be detected easily. The method discussed above has been mainly tested on grayscale images but could be extended on color images without any modification, the embedding step would be performed same as previously after the lossy quantization step.

IV. EXPERIMENTAL RESULTS

The algorithm was tested for various images and it produces appreciable results in terms of quality of final image seen. This is in terms of the perceptibility criterion. This algorithm was tested against other algorithm on the basis of the PSNR values. Different quality factor was taken into consideration for lenna image and the PSNR value calculated was compared to the values obtained by other algorithm.

The image shown in Fig.3 corresponds to JPEG compression with quality factor of 50. The PSNR value obtained for this image was 43.078dB. As can be seen from the two images the watermark embedding is not at all visible and perceptibility criterion is being satisfied. For the lenna image the comparison result with scheme [12] is given in the Table 1.

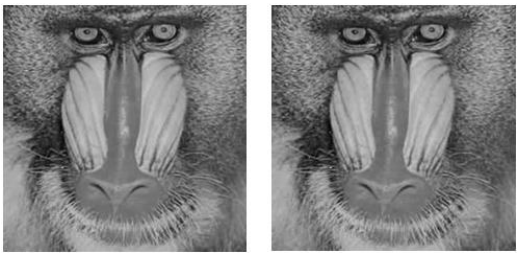

Fig. 3. [L]Original Image [R] Watermarked Image (Q=50)

Table 1. Comparisons of PSNR values

Quality Factor	Our Method (PSNR, dB)	Prev. Method (PSNR, dB)
50	43.02	40.8
75	48.83	48.1
85	53.02	51.5

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

90	59.49	54.0

As can be seen our method embeds much more data (depending on image size and resolution) as compared to the other method and it also maintains the PSNR values to fairly respectable levels. Table 2 is provided for the values obtained for different quality factor for various standard images:

Table 2. PSNR (dB) values for different Q-factor

Image	Q = 40	Q = 45	Q = 50	Q = 75	Q = 80
Lena	40.98	42.29	43.02	48.83	50.91
Baboon	41.11	42.31	43.07	49.14	51.89
Airfield	40.93	42.39	42.94	48.84	51.08
Bridge	41.05	42.23	43.05	48.81	50.96
House	40.87	42.20	43.06	48.92	50.95

V. CONCLUSION

It can be seen the above technique does well in terms of imperceptibility as expected from embedding in the Transform Domain. The procedure was adapted to plug in the loop holes of the previous algorithms.

An additional security factor if wanted can be added is the use of pseudo-random generator to select the pixel to embed the sequence or an adaptive technique to select pixels to embed the sequence which would lead to even better quality final images. Inclusion of pseudo-random generator will increase the overhead as the number of pixels which already have data embedded into them have to be avoided to prevent overwriting, thus we have a greater memory requirement for this. These extensions can be chosen based on the application.

ACKNOWLEDGMENT

I would like to thank my friend Sunil Ganatra for proof reading and pointing out any changes. Also would like to thank my professor for his valuable comments.

REFERENCES

[1] M. U. Celik, Gaurav Sharma, E. Saber, and A. M. Tekalp. Hierarchical watermarking for secure image authentication with localization. *IEEE Transactions on Image Processing*, June 2002.

[2] R. Chandramouli, N. Memon, and M. Rabbani. Digital watermarking. *Encyclopedia of Imaging Science and Technology*, 2002.

[3] B. Chen and G. W. Wornell. Quantization index modulation methods: A class of provably good methods for digital watermarking and information embedding, May 2001.

[4] D. Coppersmith, F. Mintzer, C. Tresser, Chai Wah Wu, and M. M. Yeung. Fragile imperceptible digital watermark with privacy control. *SPIE conference on Security and Watermarking of Multimedia Contents*, January 1999.

[5] I. J. Cox, J. Killian, F. T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia, December1997.

[6] F.Minitzer, G.W. Braudaway, and A. E. Bell. Opportunities for watermarking standards. *Communications of ACM*, July 1998.

[7] J. Fridrich, M. Goljan, and Arnold C. Baldoza. New fragile authentication watermark for images. *Proc. IEEE International Conference Image Processing*, September 2000.

[8] J. Fridrich, M. Goljan, and N. Memon. Further attacks on yeung-mintzer watermarking scheme. *Proc. SPIE Electronic Imaging*, January 2000.

[9] G. L. Friedman. The trustworthy digital camera: Restoring credibility to the photographic image. *IEEE Transactions on Consumer Electronics*, November 1993.

[10] M. Holliman and N. Memon. Counterfeiting attacks on oblivious block-wise independent invisible watermarking scheme. *IEEE Trans. Image Processing*, pages 432–441, March 2000.

[11] J. Hwang, J. Kim, and J. Choi. A reversible watermarking based on histogram shifting. *Springer-Verlag IWDW*, pages 348–361, 2006.

[12] R. Du J. Fridrich, M. Goljan. Invertible authentication watermark for jpeg images. *Proc. SPIE Photonics West, Security and Watermarking of Multimedia Contents III*, pages 197–208, Janurary 2001.

[13] N. F. Johnson and S. Jajodia. Exploring steganography: Seeing the unseen. *IEEE Computer*, pages 26–34, 1998.

[14] H. Lu, R. Shen, and Fu-Lai Chung. Fragile watermarking scheme for image authentication. *Electronics Letters*, June 2003.

[15] C. Paar and et al. Understanding Cryptography. *Springer*, 2010.

[16] Christine I. Podilchuk and Edward J. Delpi. Digital watermarking: Algorithms and applications. *IEEE Signal Processing Magazine*, 2001.

[17] P. W. Wong. A public key watermarking for image verification and authentication. *Proc. IEEE International Conference Image Processing*, pages 425–429, October 1998.

[18] J. Wu, B. B. Zhu, S. Li, and F. Lin. Efficient oracle attacks on yeung-mintzer and variant authentication schemes. *Proc. Of International Conf. on Multimedia and Expro*. Taiwan, pages 301–306, 2004.

[19] M. Wu and B. Liu. Watermarking for image authentication. *IEEE Proc. of ICIP*, October 1998.

[20] Minerva M. Yeung and F. Mintzer. An invisible watermarking technique for image verification. *Proc. IEEE Int. Conf. Image Processing*, October 1997.

Nitish Agarwal received his bachelors degree from Manipal Institute Of Technology and is currently doing his masters from University Of Waterloo. His research interest includes timing in embedded system, watermarking and steganography, ad-hoc network security and data clustering using artificial neural network and fuzzy logic.