Digital Video Watermarking in the Discrete Cosine Transform Domain

Sadik Ali M. Al-Taweel, Putra Sumari, Saleh Ali K. Alomari and Anas J.A. Husain
School of Computer Sciences, University Sains Malaysia, 11800 Penang, Malaysia

Abstract: Problem statement: Nowadays, digital watermarks have recently become recognized as a solution for protecting copyright of the digital multimedia. For a watermark to be useful, it must be perceptually invisible and robust against any possible attack and image processing by those who seek to corsair the material, researchers have considered various approaches like JPEG compression, geometric distortions and noising. **Approach:** We proposed a framework based on the Hartung technique which depended on spread spectrum communication in discrete cosine transform (DCT). **Results:** For the experimental results, researchers had considered various approaches: JPEG compression, geometric distortions and noising. Results showed a good performance in the proposed method. **Conclusion:** The presented technique was applicable not only to MPEG-2 video, but also to other DCT coding videos like MPEG-1,H261 and H263. Future study could be on improving DCT and comparing it with existing methods. It could be on discrete wavelet transform that is relatively new and has useful properties for the image processing applications.

Key words: Data hiding, video, watermarking, geometric distortions, JPEG Compression, transform domain.

INTRODUCTION

Digital watermarking has recently become a popular area of research due to the proliferation of digital data (image, audio, or video) in the Internet and the need to find a way to protect the copyright of these materials. Recently, numerous digital watermarking algorithms are developed to help protect the copyright of digital images and to verify the multimedia data integrity. In spite of the existence of watermarking technique for all kinds of digital data, most of the literatures address the watermarking of the still images for copyright protection and only some are extended to the temporal domain for the video watermarking.

In order for a watermark to be useful, it must be perceptually invisible and robust against any possible attack and image processing by those who seek to corsair the material[1,2].

Digital watermark is a code that is embedded inside some innocent-looking cover data. Typically, this information is required to be robust against any intentional removal by malicious parties. In contrast to cryptography, where the existence but not the meaning of the information is known, watermarking aims to hide entirely the existence of the information. Watermarking has existed since approximately the 13th century and the past watermarks were used on the papers to identify the mill which made them[3].The oldest watermarked paper was found in archives dates back to 1292 and it originated from the town of Fabriano in Italy, which has played a major role in the evolution of paper industry. After their invention, watermarks quickly spread in Italy and then all over Europe. These are called physical watermarks because they were found on the physical media. Nowadays, physical watermarks are commonly used to authenticate important documents such as banknotes and passports.

Digital watermarks, on the other hand, are found with the advancement of the Internet and the ambiguity of digital data. Thus, it is natural to extend the idea of watermarking to the digital data. Several names have been used to describe and classify the watermarking techniques. These terms will be explained in the present section. Host or cover data is a piece of digital data in which the information is hidden, whereas payload refers to the hidden information.

Visible watermarks are visual patterns like logos, which are inserted into the digital data. Most watermarking systems involve marking imperceptible alteration on the cover data to convey the hidden information. This is called invisible watermarks. The watermarking scheme that allows extraction of embedded information using the original, un-watermarked data is known as non-blind or non-oblivious watermarking scheme, otherwise it is known as blind or oblivious.

Corresponding Author: Sadik Al-Taweel, School of Computer Sciences, University Sains Malaysia, 11800, Penang, Malaysia

Some watermarking schemes use key to enforce the security. Using a secret or public key, the watermarking techniques are usually referred to as the secret or public watermarking techniques respectively. Fingerprinting is the term that denotes the special application of watermarking. In this application, the embedded information is either a unique code specifying the author or originator of the cover data or a unique code out of a series of code specifying the receiver of the data.

Fragile watermarks are watermarks that have very limited robustness. They are used to detect modifications in the watermarked data rather than to extract un-erasable information[3-6].

Watermark Classification: Video watermarking techniques or algorithms are classified depending on the working domain, which is divided into spatial domain and transform domain techniques. In spatial domain technique, the watermark is embedded in the source video by selecting pixel positions and replacing bit. The spatial domain techniques are easy to perceive and it has low time complexity that does not exists in other domains. However, it has the lack in providing adequate robustness and imperceptibility requirements.

Lancini *et al*.[7] proposed exhaustive search scheme. In this scheme a transfer test image in every possible way to recover synchronization based on a training sequence robust resizing and cropping attacks Also, it entails inverting a large number of possible distortions and testing for a watermark after each one. However, the time complexity of this method increases exponentially and false positive probability become unacceptable.

Dumitru *et al*.[8] proposed spatial watermarking technique. In this scheme modifies a blocks of frames by a spatial watermark insertion and spatial mask of suitable size is used to hide data with less visual impairments robust filtering, resizing, cropping and rotation.

Chien *et al*.[9] proposed method and embeds watermarks into the feature blocks or non-feature blocks. The proposed method resists attacks such as frame reduction and frame shuttle. However to obtain synchronization will spend significant time to locating the feature blocks[10-13].

Discrete Cosine Transformation (DCT), Discrete Furrier Transform (DFT) and Discrete Wavelet Transformation (DWT) are generally used in transform methodologies. The host signal is transformed into a different domain and the watermark is embedded in selective coefficients. The ease and applicability of spatial transformed domain properties played a great role in making it preferable method versus other methods that lack these properties. For example, there are more advanced properties of the Human Visual System (HVS) to be applied and ensure better robustness and imperceptibility criteria when working in the frequency domain.

Spatial domain Watermarking: For spatial domain the watermark can be embedded in grayscale or color. So the watermark is performed by modifying values of pixel color of video frame. We can denote the picture to be watermarked by P and values of its pixel color samples by Pi, a watermarked version of picture P by P* and values of its pixel color samples by P*I. Let us have many elements of watermark W with values Wi as number of pixels in picture P. Watermark W hereby covering the whole picture P. Further, it is possible to increase the watermark strength by multiplying watermark element values using weight factor, a. Then the natural formula for embedding watermark W into picture P is:

$$P*i = Pi + aWi$$

The most common algorithm using spatial domain watermarking is:

Least significant bit modification: The easiest watermarking method in spatial domain is to immediately flip the Least Significant Bit (LSB) of selected pixels in a frame. A smaller object may be embedded several times in the this given method, particularly high channel capacity of using the entire cover for transmission even if the most of these objects are lost due to attacks, a single presented watermark would be deemed a success[14].

The insertion and scramble of the watermark is performed as follows: (move to the following line) Consider a grayscale image (I) with size M×N pixels and it is required to insert and an invisible watermark to create an image Iw of the same size as I by using a bi-level watermark image α, size (I×J) pixel blocks and watermarking procedure is performed on each block as follows.

Firstly the image I can be divided into (I×J) pixel blocks and watermarking procedure is made on each block as follows.

Let a be a bi-level image which will be used as the invisible watermark. Note that a is not necessary to be the same size as I. A matrix d is the same size as and it is generated by using an m-sequence and is used to scramble the watermark image a by exclusive o-ring a with d to form c(c = a XOR d). c is the new LSB plane that will be inserted into the image.

Repeating this procedure until the watermark bits are completed, after that, the resulting bits are combined together to give the watermarked image. d is considered as a secret key for watermark extraction, if d is chosen to be maximal length shift register sequence (m-sequence), then the feedback connection of the shift register and the primary conditions used to generate the m-sequence. For extracting watermark, the LSB plane is extracted from the image M×N blocks. During exclusive o-ring with d the watermark image a is recovered[15].

Frequency Domain Watermarking: The frequency methods are similar to spatial domain watermarking in the sense that the values of selected frequencies can be shifted. Since high frequencies will be lost by compression or scaling, the watermark signal is applied to lower frequencies, or adaptively applied to frequencies consisting important elements of the original image. For the inverse transformation, the watermarks applied to frequency domain will be spread over the entire spatial image, so these methods are not easy to defeat by cropping as the spatial techniques are easy to be defeated. On the other hand, the trade-off between invisibility and robustness, here, is greater[16].

Discrete Fourier transformation: the Discrete Fourier Transformation (DFT) is considered in the field of watermarking because it controls the frequency of the host signal. It enables the schemes further to embed the watermark with the magnitude of its coefficients. Given a two-dimensional signal f(x, y), the DFT is defined as[16,17]:

$$F(u,v) = \frac{1}{MN} \sum_{X=0}^{M-1} \sum_{Y=0}^{N-1} f(x,y) e^{(-j2\pi(ux/M+vy/N))} \qquad (1)$$

For u = 0, 1, 2…,M-1, v = 0, 1, 2,..,N-1 and $j = \sqrt{-1}$

The inverse DFT (IDFT) is given by:

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) e^{(j2\pi(ux/M+vy/N))} \qquad (2)$$

where, (M, N) are the dimensions of the image.

The DFT is useful for watermarking purposes because it helps in selecting the adequate parts of the image for embedding, in order to achieve the highest invisibility and robustness.

Discrete cosine transform: The DCT domain permits a host signal (image or video) to be broken into different frequency bands, making it a lot easier to embed watermark information into the middle frequency bands where these bands avoid the most visual important parts of the host signal without over-exposing themselves to removing through noise attacks and compression (high frequencies). Therefore the middle frequency bands are chosen.

A void the most visual important parts of the host signal (low frequencies) without over-exposing themselves to removal through noise attacks and compression (high frequencies). The original signal is divided into 8×8 blocks of pixels and the 2-D DCT is applied independently to each block. The two dimensional DCT pair is given by[13]:

$$C(0,0) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \qquad (3)$$

$$C(u,v) = \frac{1}{2N^3} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)$$
$$\left[\cos(2x+1)u\pi\right] \left[\cos(2y+1)v\pi\right] \qquad (4)$$

For u, v = 1, 2… N-1 and the inverse DCT is given by:

$$f(x,y) = \frac{1}{N} C(0,0) + \frac{1}{2N^3} \sum_{u=1}^{N-1} \sum_{v=1}^{N-1} C(u,v)$$
$$\left[\cos(2x+1)u\pi\right] \left[\cos(2y+1)v\pi\right]$$

Several algorithms have considered DCT in the watermarking process. Some of them added the DCT coefficient of image to the coefficients of the watermark[18,19] or select some of the image DCT coefficients for embedding[20,21]. Other publications in the DCT domain are[22-25].

The wavelet transform: The base of the Discrete Wavelet Transform (DWT) was in[26] derived a technique to decompose the discrete time signals. And also they did a similar research on the coding of speech signals. They named their analysis scheme as sub-band coding. In 1983, Burt named it pyramidal coding and it is also known as multi resolution analysis.

MATERIALS AND METHODS

The algorithm was designed by Hartung and Girod[27]. In their research, they presented an algorithm for embedding the digital watermarks into compressed and uncompressed video sequences. The basic principle is borrowed from the spread spectrum communications. In the spread spectrum communication, a narrow band

signal is transmitted over a much larger bandwidth such that the signal energy that present in any single frequency is undetectable. Similarly, the watermark bits are spread by a large factor called chip-rate so that it is imperceptible.

Watermark generation: Let N be the total number of pixels in the video signal. A chip-rate refers to the amount of information bits being spread (Fig. 1). Let cr be the chip-rate. Then, a total of $\frac{N}{cr}$ information bits are embedded in the video signal.

Let $a_j \in \{-1, 1\}$ be the sequence of information bits that has to be embedded into the video stream.

This sequence is spread by a large factor, which is the chip-rate cr, to obtain the spread sequence b_i:

$$b_i = a_j, \text{ where } (j-1).cr+1 \leq i < (j-1).cr+cr \qquad (6)$$

The spread sequence b_i is then modulated by a pseudo noise sequence:

$$p_i, p_i \in \{-1,1\}, i = 1, 2, \ldots, N \qquad (7)$$

Such a sequence can be generated by the feedback shift register or by any other random number generator. By using the shift registers, at each clock time, the register shifts all the contents to the right (Fig. 2). The sequence p_i is generated according to the recursive formula:

$$p_i = c_1 p_i\text{-}1 + c_2 p_i\text{-}2 + \ldots + c_n p_i\text{-}n \qquad (8)$$

The modulated signal is then scaled with the scalar α.

$$W_i = \alpha.b_i. \ p_i, \ i = 1,2,\ldots,N \qquad (9)$$

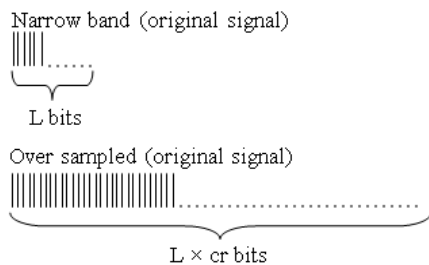where, W_i is the spread spectrum watermark.

Watermark embedding: The watermark W_i is directly added to the digital video signal V_i, to produce the watermarked video signal, V_i':

$$V_i' = V_i + W_i \qquad (10)$$

Figure 3 visualizes the principle of watermark embedding process.

Watermark extraction: The watermark could be extracted with or without using the original, un-watermarked signal. Watermark extraction without using the original signal can be done by first, high pass filtering the watermarked video sequence $\widetilde{V}$ to remove the major components of the video signal and produce a filtered watermarked video signal $\overline{\widetilde{V}}$.

The second step is demodulation, where the filtered watermarked video signal is multiplied by the same pseudo noise signal P_i that was used for embedding. This is followed by summation of window of length with the chip rate, yielding the sum S_j for the Jth information bit:

$$s_j = \underbrace{\sum_{i=(j-1).cr+1}^{(j-1).cr+cr} P_i.\tilde{v}_i = \sum_{i=(j-1).cr+1}^{(j-1).cr+cr} P_i.\overline{v}_i}_{\Sigma_1}$$

$$+ \underbrace{\sum_{i=(j-1).cr+1}^{(j-1).cr+cr} \overline{P_i.P_i.\alpha_i.b_i}}_{\Sigma_2} \qquad (11)$$

where, the two terms Σ_1 and Σ_2 represent the contribution to the correlation sum from the filtered video signal and the filtered watermark signal respectively.
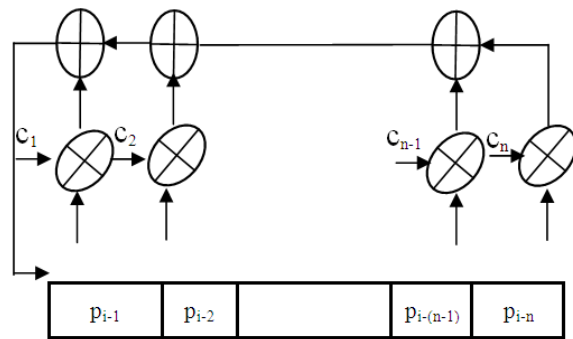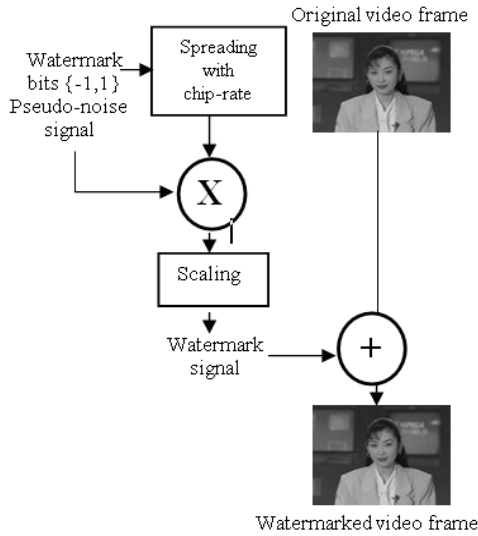

Fig. 1: Spreading using chip_rate

Fig. 2: Random sequence generator

Fig. 3: Watermark embedding

Fig. 4: Watermark embedding process in DCT domain

Let us assume that the sum Σ_1 is zero. This means that the video signal has been filtered out in V and that $Pi.\ \alpha i.\ bi \approx Pi.\ \alpha i.\ bi \quad Pi.\ \alpha i.bi \approx \overline{Pi.\ \alpha i.\ bi}$ which means that the high pass filtering has a negligible influence on the white pseudo noise watermark signal. Under these assumptions, we have:

$$s_j = \Sigma_1 + \Sigma_2 \approx \sum_{i=(j-1).cr+1}^{(j-1).cr+cr} Pi^2.\alpha_i.b_i = P_i^2.\alpha_i.a_j.cr \quad (12)$$

Thus, s_j, is a calculation of the embedded information bit in the video format:

$$Sign(s_j) = sign(\underbrace{a_j.P_i^2.cr.\alpha_i}_{>0}) = sign(a_j) = a_j \quad (13)$$

The recovery of the watermark bits is more robust, if the original, un-watermarked signal is available. The signal can be subtracted from the watermarked video signal before demodulation instead of the filtering operation because the subtraction removes the interference between the video signal and the embedded watermark.

Adjustment using DCT: In order to transform the spatial watermark to the frequency or spectral watermarking, the watermark bits, wi are transformed by using the DCT and are added to the frame of the video DCT coefficients. Figure 4 shows the watermark embedding process.
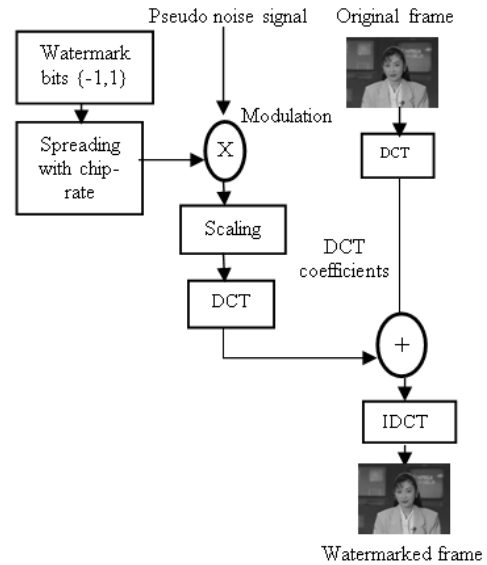
RESULTS AND DISCUSSION

To evaluate the algorithm, the video clips with 352×240 are used as shown in the Table 1 and they are spread with chip-rate = 8. We use the PSNR (peak signal to noise ratio) to estimate the performance of the invisibility and the detection ratio of the watermarks to estimate the performance of robustness. Figure 5 shows the watermarked frame with PSNR = 36.742 dB and Fig. 6 shows the watermark image 64×64. Typical correlation result of the detected watermark after the compression attack and detection is shown in Fig. 7. A 3×3 averaging filter with coefficients of 1/9 is used in the LPF. The LPF attack causes the decoded watermark to be noisy with the correlation of 0.135. Although the correlation is relatively small, the detection score remains acceptable. The result of cropping 50% of the watermarked frame is shown in Fig. 8.

The watermark frame of the video is rotated-17° by using the bilinear interpolation. The detection score is 0.52 (Fig. 9). Figure 10 shows the detection score after 50% scale down attack with correlation = 0.15. As for noise attack, Gaussian noise with mean 0 and variance of 0.005 are used on the watermarked frame. Figure 11 shows the detection score. The correlation value is 0.5. Salt and pepper noise with the density of 0.02 is also added to the watermarked signal (Fig. 12). In most cases, although the correlation between the original and extracted watermark is relatively small, the watermark is distinguishable in the random watermark set.

Fig. 5: Watermarked frame

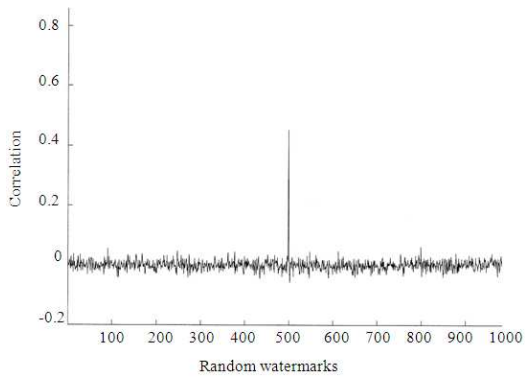Fig. 6: Logo used as a watermark

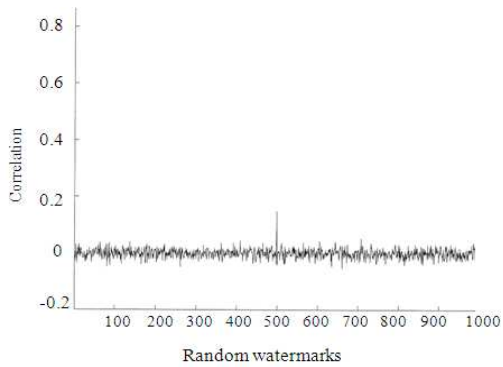Fig. 7: Random watermark detection results under Compression attack

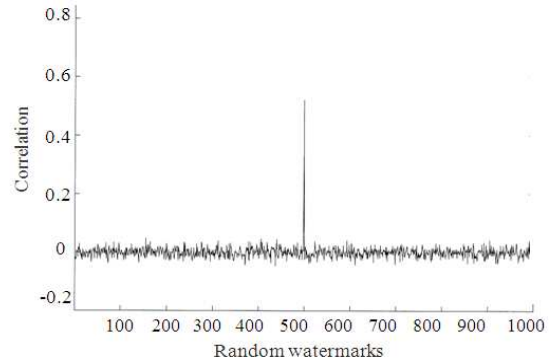Fig. 8: Random watermark detection results under cropping attack

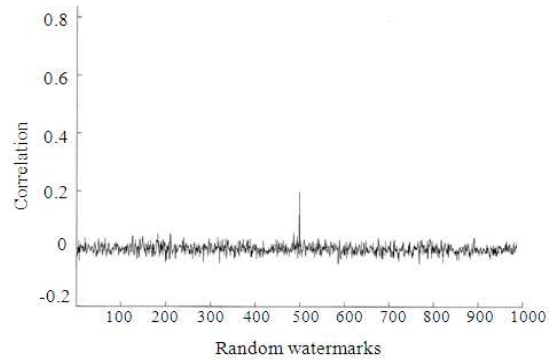Fig. 9: Random watermark detection results under rotation attack

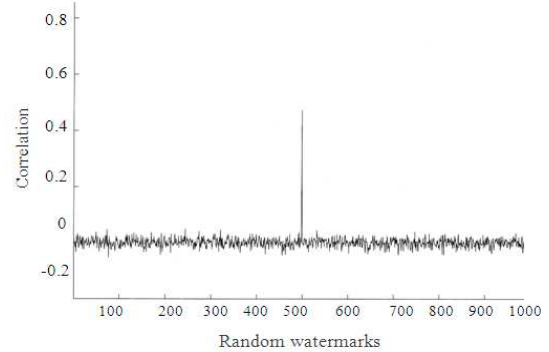Fig 10: Random watermark detection results under scaling down attack

Fig. 11: Random watermark detection results under Gaussian noise attack

Table 1: Video clips used in testing

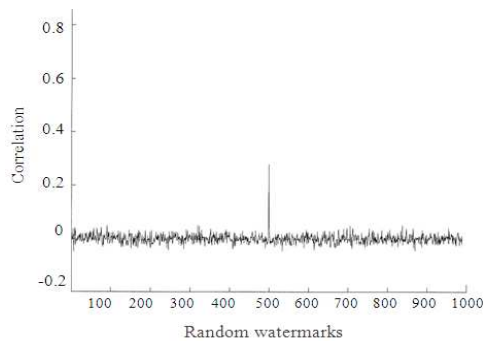| Video test sequence	Size	Format	Frames	Resolution
Susie on the phone	CIF	4.2.0	450	352×240
Flower garden	CIF	4.2.0	450	352×240
Mobile and calendar	CIF	4.2.0	450	352×240
Table tennis	CIF	4.2.0	450	352×240
Akiyo	CIF	4.2.0	450	352×240

Fig. 12: Random watermark detection results under salt and pepper attack

CONCLUSION

The need for digital watermarking on electronic distribution of copyright material is becoming more prevalent. In this study, we have improved the existing method which is the robust watermarking algorithm. Various applications of watermarks were introduced and necessary requirements of such watermarks were also introduced. An overview of the existing watermarking techniques and the attacks were given. We have also demonstrated that the discrete cosine transform (DCT) resembles the human visual system.

Although there are many digital watermarking techniques developed in recent years, the capability of the traditional watermarking techniques is yet to develop completely. Some video watermarking techniques are sensitive to geometric distortions, such as rotation, scaling and cropping. In this research, we proposed a framework for a robust digital watermarking for MPEG-2 video against the global geometric attacks such as cropping, scaling and rotation. Our future work will be on improving the DCT and comparing it with the existing methods. It will be on the discrete wavelet transform that is relatively new and has useful properties for the image processing applications.

ACKNOWLEDGMENT

REFERENCES

1. Voloshynovskiy, S., S. Pereira, A. Herrigel, N. Baumgartner and T. Pun, 2000. Generalized watermarking attack based on watermark estimation and perceptual remodulation. Proceedings of the Security and Watermarking of Multimedia Content II Conference, Jan. 24-26, San Jose, CA., pp: 548. http://cat.inist.fr/?aModele=afficheN&cpsidt=1379132

2. Voloshynovskiy, S., S. Pereira, T. Pun, J.J. Eggers and J.K. Su, 2001. Attacks on digital watermarks: Classification, estimation based attacks and benchmarks. IEEE. Commun. Mag., 39: 118-126. DOI: 10.1109/35.940053

3. Karzenbeisser, S. and F. Perircolas, 2000. Information Hiding Techniques for Steganography and Digital Watermarking. Artech House, ISBN: 1580530354, pp: 240.

4. Meerwald, P., 2001. Digital image watermarking in the wavelet transform domain. MSc Thesis, Salzburg University, Germany. http://www.cosy.sbg.ac.at/~pmeerw/Watermarking/MasterThesis/

5. Qiao, L. and K. Nahrstedt, 1999. Watermarking schemes and protocols for protecting rightful ownership and customer rights. National Science Foundation Career Grant. http://cairo.cs.uiuc.edu/publications/papers/jvcir.ps

6. Delaigle, J., 2000. Protection of intellectual property of images by perceptual watermarking. PhD. Thesis, Catholique University, Belgique.

7. Lancini, R., F. Mapelli and S. Tubaro, 2002. A robust video watermarking technique in the spatial domain. Proceeding of the 4th International Symposium on Video/Image Processing and Multimedia Communications, (VIPMC'02), IEEE Xplore Press, USA., pp: 251-256. DOI: 10.1109/VIPROM.2002.1026664

8. Burdescu, D.D., L. Stanescu, A. Ion and C.M. Mihaescu, 2007. A spatial watermarking algorithm for video images. Commun. Comput. Inform. Sci., 1: 402-407. DOI: 10.1007/978-3-540-73986-9_35

9. Ko, C.C. and B.Z. Yang, 2007. An integrated technique for video watermarking. Proceeding of the International Conference on Computer and Information Science, July 11-13, IEEE Xplore Press, Melbourne, Qld., pp: 37-42. DOI: 10.1109/ICIS.2007.59

10. Hernandez, J. and F. Gonzalez, 1998. Shedding more lights on image watermarks. Lecture Notes Comput. Sci., 1525: 191-207. http://portal.acm.org/citation.cfm?id=731549

11. Nikolaidis, N. and I. Pitas, 1999. Robust image watermarking in the spatial domain. Sign. Process., 66: 385-403. http://portal.acm.org/citation.cfm?id=284792

12. Hartung, F., J. Su and B. Girod, 1999. Spread spectrum watermarking: Malicious attacks and counterattacks. Sign. Process., 3657: 147-158. http://cat.inist.fr/?aModele=afficheN&cpsidt=1828354

13. Tzeng, J., W. Hwang and I. Chern, 2002. Enhancing image watermarking methods with/without reference images by optimization on second-order statistics. IEEE Trans. Image Proc., 11: 771-782. DOI: 10.1109/TIP.2002.800895

14. Hsu, C.T. and J.L. Wu, 1998. DCT-based watermarking for video. IEEE Trans. Consum. Elect., 44: 206-216. DOI: 10.1109/30.663749

15. Rizk, M.R.M., A.E. Youssef and S.E. El-Khamy, 2006. Adaptive watermarking techniques based on multi-scale morphological image segmentation. Proceeding of the IEEE Mediterranean Electrotechnical Conference, May 16-19, IEEE Xplore Press, Malaga, pp: 520-823. DOI: 10.1109/MELCON.2006.1653224

16. Young, I., J. Gerbrands and L. Vliet, 1995. Fundamentals of image processing. Delft University of Technology, ISBN: 9075691017.

17. Gonzalez, R. and P. Wintz, 1977. Digital Image Processing. Addison-Wesley Publishing Co., ISBN: 0471018880, pp: 750.

18. George, M., J. Chouinard and N. Georganas, 1999. Spread spectrum spatial and spectral watermarking for images and video. http://www.mcrlab.uottawa.ca/papers/cwitt99paper.pdf

19. George, M., J. Chouinard and N. Georganas, 1999. Digital watermarking of images and video using direct sequence spread spectrum techniques. Proceeding of the IEEE Canadian Conference on Electrical and Computer Engineering, May 9-12, IEEE Xplore Press, Edmonton, Canada, pp: 116-121. DOI: 10.1109/CCECE.1999.807181

20. Cox, I., J. Kilian, T. Leighton and T. Shamoon, 1997. Secure spread spectrum watermarking for multimedia. IEEE Trans. Image Process., 6: 1673-1687. DOI: 10.1109/83.650120

21. Koch, E. and J. Zhao, 1995. Towards Robust and Hidden Image Copyright Labeling. Proceeding of the IEEE Workshop on Nonlinear Signal and Image Processing, June 20-22, IEEE Xplore Press, USA., pp: 1-4. http://www.igd.fraunhofer.de/igd-a8/syscop/Publications/Zhao95a.pdf

22. Ali Murshid, S., A. Taweel and P. Sumari, 2008. Blind compressed video watermarking in DCT domain robust against global geometric attacks. Proceeding of the International Cryptology Workshop and Conference.

23. Hyun, P. and H. Sunh, 2006. Adaptive video watermarking utilizing video characteristics in 3D-DCT domain. Lecture Notes Comput. Sci., 4283: 397-406. http://cat.inist.fr/?aModele=afficheN&cpsidt=19105154

24. Gerado, P., H. Ayman, G. Mohamed, Y. Takashi and L. Jianming, 2006. Robust watermarking in the DCT domain using dual detection. Proceeding of the IEEE International Symposium on Industrial Electronics, July 9-13, IEEE Xplore Press, Montreal, pp: 579-584. DOI: 10.1109/ISIE.2006.295523

25. Satyen Biswas, Sunil R. Das, Life Fellow and Emil M. Petriu, 2005. An adaptive compressed MPEG-2 video watermarking scheme. IEEE Trans. Instrumentat. Measure., 54.

26. Polikar, R., 2002. The wavelet tutorial. http://engineering.rowan.edu/~polikar/wavelets/wttutorial.html

27. Hartung, F. and B. Girod, 1998. Watermarking of uncompressed and compressed video. Sig. Process., 66: 283-301. http://portal.acm.org/citation.cfm?id=284776