Robust Video Watermarking Based On 3D-DWT Domain

Sadik. A.M .Al-Taweel
School of Computer Sciences
Universiti Sains Malaysia
11800, Penang, MALAYSIA
sadik@cs.usm.my

Putra Sumari
School of Computer Sciences
Universiti Sains Malaysia
11800, Penang, MALAYSIA
putras@cs.usm.my

***Abstract*—The digital watermarks have recently been recognized as a solution for protecting the copyright of the digital multimedia. In this paper a novel DWT-based video watermarking algorithm is proposed based on a three-level DWT using Haar filter which is robust against geometric distortions such as Downscaling, Cropping, and Rotation. It is also robust against Image processing attacks such as low pass filtering (LPF), Median filtering, and Weiner filtering. Furthermore, the algorithm is robust against Noise attacks such as Gaussian noise, Salt and Pepper attacks. The embedded data rate is high and robust. The experimental results show that the embedded watermark is robust and invisible. The watermark was successfully extracted from the video after various attacks.**

Keywords-component; Descret wavelet transform, video watermarking, geometric attacks.

I. INTRODUCTION

Digital Watermarking has existed since approximately the 13th century and the past watermarks were used on the papers to identify the mill, which made them [1]. 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. 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 into the digital data.

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 necessity 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 [2],[3].

The use of wavelets in image and video coding has increased significantly over the years, mainly due to the superior energy compaction property of wavelets compared with the traditional transforms like the DCT. Being invisible is the minimum requirement for a watermark. It is obvious that the author of any work (image, video, etc), would absolutely not bear any kind of degradation of his work even for the purpose of security. Studies conducted in different projects have shown that many authors require a watermarked image to have the same quality as the original one. Wavelet transform has a number of properties which will be described in the next section that improve both invisibility and robustness of the watermarking scheme.

II. THE WAVELET TRANSFORM

The base of the Discrete Wavelet Transform (DWT) goes back to 1976 when Croiser, Esteban, and Galand derived a technique to decompose discrete time signals. Crochiere, Weber, and Flanagan did similar work on the coding of speech signals in the same way. They named their analysis scheme Subband coding. In 1983, Burt renamed it pyramidal coding which is also known as multi-resolution analysis [4].

The fundamental idea of the DWT for a one dimensional signal is as follows. A signal is split into two parts, generally high frequencies and low frequencies. The edge components of the signal are largely restricted in the high frequency part. The low frequency part is split again into two parts of high and low frequency. This process is continued until the signal has been entirely decomposed or stopped by the application at hand. For compression and watermarking application, usually no more than five decomposition steps are computed. Furthermore, from the DWT coefficients, the original signal can be reconstructed. The reconstruction process is called the inverse DWT (IDWT). Mathematically, the DWT and IDWT can be stated as follows.

$$H(\omega) = \sum_{k} h_k . e^{-jk\omega}, \tag{1}$$

And

$$G(\omega) = \sum_{k} g_k . e^{-jk\omega} \tag{2}$$

Are a low-pass and a high-pass filter, respectively, which satisfy certain conditions for the reconstruction, stated later? A discrete signal, F (n) can be decomposed recursively as

978-1-4244-4547-9/09/$26.00 ©2009 IEEE TENCON 2009

$$f_{j-1}^{low}(k) = \sum_n h_{n-2k} f_j(n) \quad (3)$$

And

$$f_{j-1}^{high}(k) = \sum_n g_{n-2k} f_j(n) \quad (4)$$

For $j = J+1, J, ..., J_0$ where $f_{J+1}(k)$ = $F(f), k \in Z.J+1$ is the highest resolution level index and J_0 is the low resolution level index. The coefficients $f_{J_0}^{low}(k), f_{J_0}^{high}(k), f_{J_{0+1}}^{high}(k), ..., f_J^{high}(k)$ are called the DWT of the signal F (n), where $f_{J_0}^{low}(k)$ is the lowest resolution part of F (n) (the approximation) and the $f_j^{high}(k)$ are the details of F (n) at various bands of frequencies. Furthermore, the signal F (n) can be reconstructed from its DWT coefficients recursively.

$$f_j^{low}(n) = \sum_k h`_{n-2k} . f_{j-1}^{low}(k) + \sum_k g`_{n-2k} . f_{j-1}^{high}(k). \quad (5)$$

To ensure the above IDWT and DWT relationship, the following orthogonality condition on the filters $H(w)$ and $G(w)$ is needed:

$$| H(w) |^2 + | G(w) |^2 = 1. \quad (6)$$

An example of such $H(w)$ and $G(w)$ is given by:

$$H(w) = \frac{1}{2} + \frac{1}{2} e^{-jw} \quad (7)$$

And

$$G(w) = \frac{1}{2} - \frac{1}{2} e^{-jw} \quad (8)$$

Which is known as the Haar wavelet filter. Other common filters used in image processing are the family of Daubechies orthogonal (D-4, D-6, D-8, D-10, D-12) and bi-orthogonal (B-5/3, B-7/9) filters.

The Discrete Wavelet Transforms (DWT) provides us with one part of multiresolution approximation (MRA) and three parts of multiresolution representation (MRR) [5]. It is similar to hierarchical Subband system, where the Subbands are logarithmically spaced in frequency. The Subband LL1 (that is MRA) is further decomposed and critically subsampled; the Subbands labeled LH1, HL1, and HH1 of MRR represent the first scale wavelet coefficients to gain the next coarser scale of wavelet coefficients [6]. Some of the video watermarking techniques targeting geometric attacks are on raw videos [7], [8]. Numerous approaches video watermarking suggested handling geometric attacks and we can classify them into categories: invariant watermark,[9], [10] synchronization, [11] and autocorrelation [12].

Invariant watermarking embeds the watermark in a geometric-invariant transform, such as a log-polar wavelet transform, eliminating the need to identify and reverse the specific geometric distortions, such as rotation, and scaling. These kinds of techniques are very weak against slight geometric distortion, such as small-angle rotation and near-one scaling. Moreover, the computational cost is too high to obtain the invariant domain from the varied transform.

The synchronization is the exhaustive search which it entails inverting a large number of possible distortions and testing for a watermark after each one. Since the number of possible distortions increases, the positive probability and computational cost become unacceptable.

The autocorrelation technique is similar to the synchronization approach. It spreads lots of extra data, in addition to real watermark information to obtain synchronization for watermark detection by autocorrelation, which either further distorts the host media or sacrifices watermark payload.

Other published papers in this domain are Chan et al [13], the researchers presented a novel DWT-based video watermarking scheme with scrambled watermark and error correcting code. The scheme is robust against attacks such as frame dropping, frame averaging, and statistical analysis. Campisi et al [14] proposed perceptual mask, applied in the 3D DWT domain robust against MPEG2 and MPEG-4 compression, collusion and transcoding attacks.

In [15] propose a robust mpeg video watermarking in wavelet domain which they embedded in two bands (LL and HH) and chosen attacks JPEG compression, resizing, adding Gaussian noise, low pass filtering.

In [16] present a novel adaptive watermarking scheme based on error correction code and Human Visual System (HVS) in 3D-DWT domain. The proposed method is to resist signal processing attacks, Gaussian noise, and frame dropping.

In [17] propose a method based on 3D wavelet transforms. In this method the original video frames are divided into 3D-block according to HVS properties. The proposed method is robust against lossy compression, frame swapping, frame dropping, and median filtering.

III. THE PROPOSED ALGORITHM

The researchers proposed a new secret key based on a password. The key generator provides, KE. The embedded watermarking requires a password that should be known to the sender and receiver. In this method, the researchers used a password that contains sixteen characters (128-bits) as shown in Figure 1 where it was divided into two equal parts (K1, K2) each of which contains 64 bits which then converted each part to decimal values and stored each value in a variable m and n respectively. K1 and K2 are left shifted by these decimal values and swapped between these m and n mod 256 as shown in Figure 2.

The watermarking data is encoded before it is embedded into the original frame of video to produce the watermarked video frame, so the method to encode watermarking data design depends on the Stream Cipher technique (RC4). A digital watermarking algorithm is proposed using the characteristics of a frame of video and the human visual system HVS for invisibility and robustness. The binary watermark is modulated before embedding into the perceptually significant wavelet coefficients. The watermark is embedded into the selected coefficients using different scale factors according to the level of decomposition.

A. *Watermark Modulation*

A binary sequence of watermark bits has to be embedded into the host video. A logo watermark is used as a perceptual meaningful watermark. The binary sequence {-1, 1} of watermark bits has to be embedded into the host video. This sequence spread with a factor to obtain the spread sequence. The purpose of spreading is to improve the robustness by adding redundancy to the embedded bits. Next, this redundant sequence is modulated by a binary pseudo-noise sequence P_i.

$$P_i \in \{-1, 1\};\ \iota=1, 2 \ldots M \tag{9}$$

Where M denotes length of the watermark bit sequence times the spreading factor. Again, since different adaptive scaling factors are used, the step of watermark amplification is postponed to the embedding step.

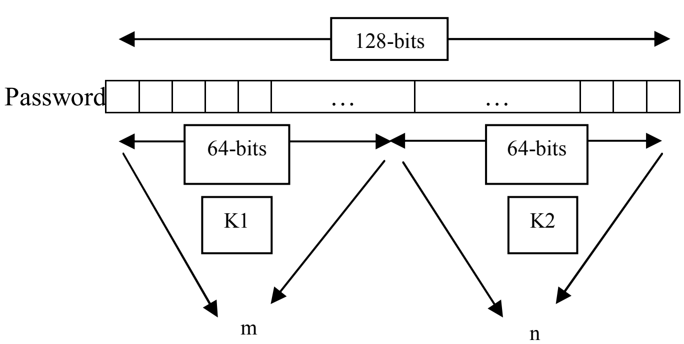


Figure 1. Key Generation

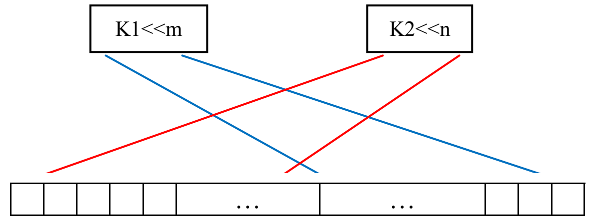


Figure 2. Shift Operation

B. *Wavelet-based Watermark Embedding Process*

The block diagram of the embedding process is shown in Figure 3 Splitting the video into frames I, B, P and the original frame is decomposed into a three-level DWT using Haar filter. The result of decomposition is ten frequency Subbands. Modulated watermark bits are embedded into the perceptual significant coefficients and the IDWT is applied to produce the watermarked frame. The original frame is replaced with the resultant watermarked frame.

1) *Selection of Perceptual Significant Coefficients*

The lowest band (approximation band) is the basic band of the wavelet-decomposed frame includes most of the energy from the original frame. Therefore, it has a crucial effect on the frame quality.

On the other hand, the basic band coefficients are not removed or modified by lossy compression and other common signal processing. Therefore, a portion of this band is selected and includes its highest coefficients. The highest wavelet coefficients on the first two levels (HH1,HH2) as well as those on the middle frequency bands on the lowest level (LH1,HL2) are excluded from the proposed algorithm because these coefficients can be easily eliminated and modified by lossy compression and other signal processing. The steps involved in the selection of coefficient of each Subband are:

- Calculating the energy of each selected Subband where the energy of a Subband Es is defined by [18].

$$E_s = \frac{1}{M.N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n)^2 \tag{10}$$

Where M, N denote the size of the Subband.

- The number of watermark bits Nob in each of the detailed Subband is proportional to the energy of the Subband.

$$No_b = \frac{b}{t} * n \tag{11}$$

Where b denotes the band energy, t denotes the total energy and n denote number of bits.

- Coefficients of each Subband are sorted. Watermark bits are embedded in the highest coefficients.

2) Selection of Adaptive Scalars

The proposed method uses an appropriate value of scale factor α for each Subband according to DWT decomposition characteristics. The mean of the wavelet coefficients is reduced by a half with each descending level and the values of wavelet coefficients for the approximation Subband (LL3) are relatively larger than other Subbands in the same decomposition level. This adaptive scale factor improves the performance of both robustness and invisibility.

To the selected coefficients, the watermark is embedded using equation (12).

$$V_i' = V_i + \alpha\, a_i \qquad (12)$$

Where V_i the selected wavelet coefficients and a_i is the modulated watermark vector.

C. Watermark Extraction Process

The video is split into frames I, B, P, the original frame and the possibly altered watermarked frame are decomposed to three levels DWT. Same password is used for generating secret key K_D and the same coefficients used in the embedded process are selected and the original frame coefficients are subtracted from the watermarked frame coefficients to remove the major components of the frame itself. The result is then demodulated with the pseudo-noise signal Pi that is the same as the one used for embedding. Demodulation is followed by a summation over a window of length equal to the chip-rate, and threshold which yields the watermark bits ai. The result of the decoded watermark bits recovers the original logo.

IV. Results and Performance Evaluation

We evaluated the performance of the proposed 3D-DWT video watermarking algorithm with respect to two metrics, imperceptibility and robustness. The metrics were evaluated using a colored video clip having a size 450,150, and 149 frames, with size of frame 240x352 pixels as shown in the table (1). For the watermark we used a grayscale image having a size of 64x64 pixels. Both the video clip and watermark were processed as described in details in the previous section.

A. Imperceptibility Results

The perceived quality of the host image should not be distorted by the presence of the watermark. As a measure of the quality of a watermarked image, the peak signal to noise ratio (PSNR) is typically used. In the proposed method, the watermark was embedded in the I-frame according to 3D-DWT. The average PSNR for all watermarked frame was 47.66dB. With this PSNR value no quality degradation in the watermarked video was perceived.

The proposed method is compared with the existing methods. Table (2) shows the quality of the watermarked video for the mentioned methods.

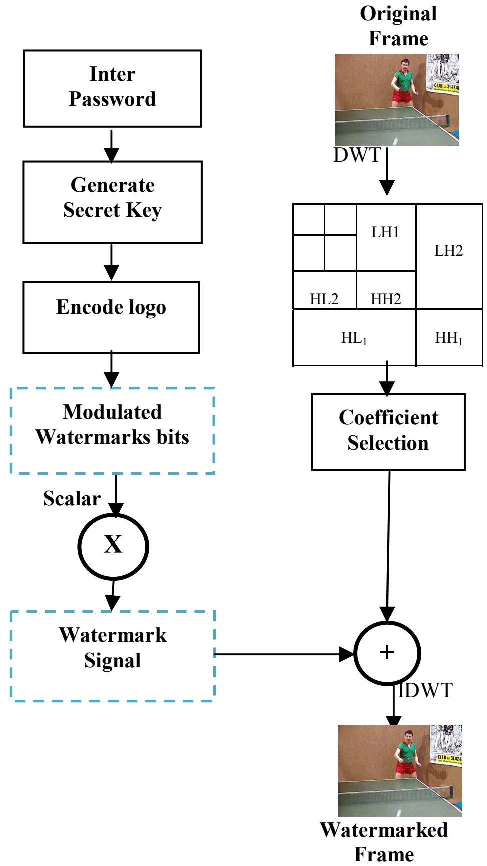


Figure 3. Watermark embedding process

B. Robustness Results

Robustness is a measure of the invulnerability of the watermark against attempts to remove it or degrade it by different types of digital signal processing attacks.

In our method the video watermarking application robustness is measured against three types of attacks; geometric distortion, image processing attacks, and noise attacks.

Geometric distortion includes frame dropping, scaledown, cropping, and rotation. , Image processing includes low pass filtering, median filtering, and wiener filtering. Noise attacks include Gaussian noise, salt & pepper noise. For all types of attacks we measured the similarity between the original and extracted watermarks using the correlation factor a it may take values between 0 and 1.

TABLE I. Video Clips Used in Testing

Video test sequence	Size	Format	Frames	Resolution
Susie on the phone	CIF	4.2.0	450	352x240
Flower garden	CIF	4.2.0	150	352x240
Football	CIF	4.2.0	150	352x240
Mobile and calendar	CIF	4.2.0	450	352x240
Tempete	CIF	4.2.0	149	352x240
Table Tennis	CIF	4.2.0	150	352x240

TABLE II. THE WATERMARKED SIGNAL STRENGTH OF SELECTED

Method	PSNR (in dB)
Anqiang [16]	36
Xu Da-Wen[17]	39.6
Our method	47.66

1) Robustness performance results against Scaling, Cropping and Rotation.

Table III shows that the watermarked frame is 50% scaled down using bilinear interpolation method, cropping version of the watermarked frame in which only the central quarter of the frame remains and the rotated watermarked frame by –17° using bilinear interpolation.

2) Robustness performance results against LPF, Median filtering and Wiener filtering

Table IV shows a 3×3 averaging filter with coefficients of 1/9 is used for the low pass filtering, median filtering and 5x5 Weiner filtering attack.

3) Robustness performance results against Gaussian Noise, Salt& Pepper Noise

Table V Shows Noise Attack, Gaussian Noise with Mean 0 and Variance 0.005 Is Used On the Watermarked Frame and Salt and Pepper Noise Is Added To the Watermarked Frame with 0.02 Noise Densities

TABLE III. ROBUSTNESS PERFORMANCE RESULTS AGAINST SCALING, CROPPING AND ROTATION

Attack	NC	Watermark
Scaling	0.520803	
Cropping	0.420051	
Rotation	0.700949	

TABLE IV. ROBUSTNESS PERFORMANCE RESULTS AGAINST LPF, MEDIAN FILTERING AND WIENER FILTERING

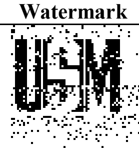
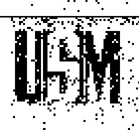
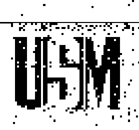
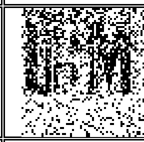
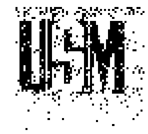
Attack	NC	Watermark
Low pas Filtering	0.664079	
Median Filtering	0.83005	
Wiener Filtering	0.897899	

TABLE V. ROBUSTNESS PERFORMANCE RESULTS AGAINST GAUSSIAN NOISE, SALT & PEPPER NOISE

Attack	NC	Watermark
Gaussian Noise	0.47856	
Salt & pepper Noise	0.81263068	

CONCLUSIONS

In this paper, a novel DWT-based video watermarking algorithm is proposed based on a three-level DWT using Haar filter which is robust against geometric distortions such as Downscaling, Cropping, and Rotation. It is also robust against Image processing attacks such as low pass filtering (LPF), Median filtering, and Weiner filtering. Furthermore, the algorithm is robust against Noise attacks such as Gaussian noise, Salt and Pepper attacks. Simulation results demonstrated the effectiveness of our proposed method.

V. ACKNOWLEDGMENT

The authors would like to thank the Universiti Sains Malaysia (USM) for supporting this study.

REFERENCES

[1] Karzenbeisser S., Perircolas F., “Information Hiding Techniques for Steganography and Digital Watermarking”, Artech House, 2000.

[2] Voloshynovskiy S., Pereira S., Herrigel A., Baumgartner N., & Pun T. Generalized “watermarking attack based on watermark estimation and perceptual remodulation,” SPIE 3971, Security and Watermarking of Multimedia Content II, San Jose, CA,2000.

[3] Voloshynovskiy S., Pereira S., Pun T., Eggers J J., & Su J. K. “Attacks on digital watermarks classification, estimation based attacks, and benchmarks,” IEEE Communications Magazine,39(8),2001,pp. 118–126.

[4] Polikar R. “The Wavelet Tutorial”, 2002.

[5] S.G.Mallat, “A theory for multiresolution decomposition: The Wavelet Representation,” IEEE Trans. Pattern Anal. R. Mach. Intell., voI.1 I, no.7, pp.674-693, 1989.

[6] Hisashi Inoue, Akio Miyazaki, Takashi Araki, and Takashi Katsura, “A Digital Watermark Method Using the Wavelet Transform for Video Data,” IEICE Transform for Video Data,” IEICE Trans. Fundamentals, vol. E83-A, no.1, January 2000.

[7] Y Li,. X Gao,. Ji.H.,: “A 3D wavelet based spatial-temporal approach for video watermarking”. in Proc. 5th Int. Conf. Computational Intelligence Multimedia Applications, Sep. 27–30, 2003, pp. 260–265

[8] Liu N .H.,. Chen J,. Huang X. Huang,. Shi .Y. Q,. :” A robust DWT based video watermarking algorithm”. in Proc. IEEE Int. Symp. Circuits Systems, vol. 3, May 26–29, 2002, pp. 631–634.

[9] H. Inoue, A. Miyazaki, T. Araki, and T. Katsura, “A digital watermark method using the wavelet transform for video data,” in Proc. IEEE Int. Symp. Circuits Systems, vol. 4, May 30, 1999, pp. 247–250.

[10] M. Ejima and A. Miyazaki, “A wavelet-based watermarking for digital images and video,” in Proc. Int. Conf. Image Processing, vol. 3, Sep. 10–13, 2000, pp. 678–681.

[11] C. V. Ambroze, M. A. Tomlinson, and J. G. Wade, “Adding robustness to geometrical attacks to a wavelet based blind video watermarking system,” in Proc. IEEE Int. Conf. Multimedia Expo, vol. 1, Aug. 26–29, 2002, pp. 557–560.

[12] C. V. Serdean, M. A. Ambroze, M. Tomlinson, and J. G. Wade, “DWT based high capacity blind video watermarking, invariant to geometrical attacks,” in Proc. IEE Vision, Image, Signal Processing, vol. 150, 2003, pp. 51–58.

[13] Pik-Wah Chan and Michael R. Lyu1 ,”A DWT-based digital video watermarking scheme with error correction code”. Fifth International Conference on Information and Communications Security ICICS 2003

[14] Patrizio Campisi and Alessandro Neri.” perceptual video watermarking in the 3D-DWT domain using a multiplicative approach”, IWDW 2005, LNCS 3710, pp. 432–443, 2005.

[15] Elbasi, E ,. Eskicioglu .A.M, ”robust mpeg video watermarking in wavelet domain”, Trakya Univ J Sci, 8(2): 87-93, 2007.

[16] Lv Anqiang, Li Jing, “A Novel Scheme for Robust Video Watermark in the 3D-DWT Domain”, First International Symposium on Data, Privacy and E-Commerce, 2007.

[17] Xu Da-Wen, “A Blind Video Watermarking Algorithm Based on 3D Wavelet Transform”, 2007 International Conference on Computational Intelligence and Security.

[18] Meerwald P., “Digital Image Watermarking in the Wavelet Transform Domain”, M. Sc. Thesis, Salzburg University, Germany, 2001.