A Hybrid Digital Video Watermarking Method Based on DCT and DWT

Haldun Sarnel, Kadir Ünal

Dokuz Eylül University, Department of Electrical and Electronic Engineering, İzmir, Turkey

E-mail: *haldun.sarnel@deu.edu.tr*, *kadirunal1985@yahoo.com*

Abstract

Digital video watermarking is one of the solutions for copyright protection of digital multimedia data and draws great attention of both researchers and digital video owners. In this work, a new hybrid digital video watermarking method which uses the discrete cosine transform (DCT) and discrete wavelet transform (DWT) is proposed. Each frame is processed with one of the DCT and DWT methods or a combination of them, half of a frame by the DCT method and the other half by the DWT method. Selection of the watermarking for each frame in the video is determined randomly. The proposed method has undergone several attacks in order to check its robustness and compare to the DCT and DWT methods. The proposed hybrid method including a watermarking diversity throughout the video improves robustness against attacks.

Keywords: Digital video watermarking, discrete cosine transform (DCT), discrete wavelet transform (DWT), copyright protection.

1. INTRODUCTION

Production and distribution of digital multimedia data have become much easier with the rapid growth of the internet and other digital technologies, and consequently they can reach a large number of people in a short time. On the other hand, this introduces copyright protection problem for the digital multimedia data with a legal owner. To prove the copyright of digital data, some methods must be applied on the original digital data whatever it is text, audio, image or video. Digital watermarking methods have been proposed for solving copyright protection problems (Cox et al. 2002). A secret message what is called a watermark is embedded to a digital data with a secret key. This process is called digital watermarking. The watermark can be a random number sequence, copyright messages, ownership identifiers, binary or gray level images, or other digital data formats. A digital watermarking method provides the copyright owner with proving his/her ownership by extracting and revealing the embedded watermark from the digital data in the case of an illegal usage of the digital data. The original digital data after watermarking can be distributed in a medium and may receive some manipulations and attacks, intentionally, or unintentionally, so as to degrade the watermark to disappear. Ideally, the watermark must remain intact, or safely extractable after the digital data has undergone some attacks to be able to prove the ownership.

Image and video watermarking methods can be classified in two groups, spatial domain and frequency domain methods. Spatial domain methods are based on modification of pixels values in embedding stage of the watermarking. These methods are very simple and have low computational cost, but are vulnerable to attacks and watermark can be easily distorted. Frequency domain methods are based on modification of frequency coefficients. The original content is transformed to frequency domain and coefficients in frequency domain are used to embed and recover the watermark. The most common transforms are the discrete cosine transform (Cox et al. 1997, Busch et al. 1999) and the discrete wavelet transform (Xia et al. 1997, Kundur and Hatzinakos 1998). The frequency domain methods are more robust and have more computational complexity. Hartung and Girod (1999) described a method of watermarking into both compressed and uncompressed MPEG-2 videos by modifying selected DCT coefficients. Langelaar and Lagendijk (2001) proposed a watermarking method that divides the video into groups of blocks and further divides the blocks into groups. A single watermark bit is embedded into selected DCT coefficients based on their energies within the groups. Chetan and Raghavendra (2010) proposed a DWT-based video watermarking scheme that embeds different parts of a single watermark into different scenes of a video after detecting scene changes.

In this work, a new hybrid digital video watermarking method which uses the discrete cosine transform (DCT) and discrete wavelet transform (DWT) is proposed. The watermark is divided to sub pieces and these pieces are embedded to frames. Each frame is processed with one of the DCT and DWT methods or a combination of them, half of a frame by the DCT method and the other half by the DWT method. The proposed hybrid method has undergone several attacks in order to check its robustness and compare to the DCT and DWT methods.

2. TRANSFORM –BASED DIGITAL VIDEO WATERMARKING

2.1 DCT-based Watermarking

The DCT coefficients $D(u,v)$ of a N x N block $p(x,y)$ in an image are computed as follows.

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

$$C(u) = \sqrt{\frac{1}{N}} \quad for\ u = 0, \qquad C(u) = \sqrt{\frac{2}{N}} \quad for\ u = 1,2,\ldots,N-1$$

The frequency bands of an 8x8 DCT block are shown in Figure 1. The top-left is the zero-frequency component. The lowest frequency band FL and the highest frequency band FH are not changed. A watermark bit is embedded in the highlighted medium frequency band FM by modifying its coefficients.

In the DCT-based video watermarking, a frame is divided to 8x8 non-overlapping blocks. A watermark is embedded to video frames piece by piece. A watermark of size L bits is divided into m pieces. Since L/m bits has to be embedded in a frame, a subset of L/m block locations are selected randomly using a secret key. Always this subset is used to embed a watermark piece Wi into every frame. In recovering stage, the same secret key is utilized to locate those watermarked blocks. A watermark piece is selected randomly to embed it into the current frame. Every block in the subset is DCT transformed and its 22 medium frequency band coefficients are determined. Two uncorrelated pseudo random sequences, R0 and R1, of length 22 are generated to represent watermark data bits 0 and 1. A single bit is embedded to one of the blocks with coefficients D(u,v) in the subset by

$$D_W(u,v) = \begin{cases} D(u,v) + aR_k(u,v), & u,v \in F_M \\ D(u,v), & u,v \notin F_M \end{cases} \qquad (2)$$

where DW is coefficients of watermarked block and a is a strength constant. The balance between robustness and visibility of watermarking is adjusted by a. The watermarked block is then back-transformed to spatial domain by taking its inverse DCT.

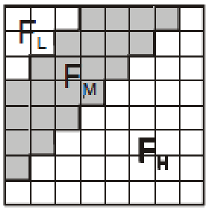

Figure 1: DCT frequency bands

LL1	LH1
HL1	HH1

Figure 2: Frequency bands of 1-level DWT

2.2 DCT-based Watermark Recovery

To recover the embedded watermark piece from each frame, first the watermarked blocks are determined and then the middle band frequency coefficients in their DCT are obtained. The pseudo random sequences R0 and R1 are computed and correlated with the middle band coefficients of each watermarked block. The higher of the two correlation values specifies the recovered possible watermark data bit. As a result, an array of higher correlation value C(n) and an array of recovered bits U(n) of the watermark piece are obtained for nth bit after all watermarked blocks are examined. U(n) is (binary) correlated with each of the real watermark pieces Wi(n) to identify it. If the maximum correlation value is smaller than a predefined threshold value Tp, U(n) is ignored assuming a failure of watermark detection in the current frame. When all frames are processed in this manner, a number of recovered instances Uij(n), indexed by j, of each watermark piece Wi(n) are obtained and recorded. Correlation values Cij(n) of all recovered bits are also formed from C(n) of every frame.

2.2.1 Reconstruction of Watermark

We use an algorithm for the reconstruction of watermark pieces Vi(n) from their recovered instances Uij(n) as follows. Set i=0 and count the numbers of 0 bits and 1 bits for every bit location n in Uij(n) discarding bits for which Cij(n)< T, a predefined bit correlation threshold value. Those are assumed as incorrectly recovered bits. Then, Vi(n) is assigned the bit with at least 70% majority of total count for bit n. If there is no such majority, the bit defaults to 0. This process is repeated after i is incremented until all watermark pieces are reconstructed. Finally, the entire watermark can be formed by concatenating all watermark pieces Vi(n) (i=0,1,2,…,m-1) in correct order.

2.3 DWT-based Watermarking

The basic idea of the DWT is to decompose a frame into a sub-image of different spatial domain and independent frequency districts. 1-level Haar DWT is used in this study for simplicity. Figure 2 shows the 1-level DWT of an image by frequency bands. The lower resolution approximation region LL of an image contains the most important pictorial information, hence, it is not used for watermark embedding. The frequency regions of LH, HL and HH represent the horizontal, vertical and diagonal details, respectively, of the image.

A watermark is embedded to video frames piece by piece as in DCT-based watermarking. In our study, the same watermark piece is embedded to both LH and HL bands. For each bit of the watermark, a different pseudo random sequence Rn (n=0,1,2,…,L-1 and R={ R1, R2,…}) with length equal to the length of these bands is generated. After a piece of the watermark is selected randomly, the sequences in R that represents the bits in that piece only are multiplied by a strength constant and added to both LH and HL band coefficients of the DWT of the current frame, in a way similar to that in (2). We embed only watermark bits 0 (black pixels in the binary watermark image) in the DWT-based method to improve imperceptibility of watermarking. Finally, the frame is transformed to spatial domain by taking its inverse DWT. In the recovering stage, every sequence Rn is correlated with LH band coefficients of the 1-level DWT of the current frame. Every bit n in the embedded watermark piece is recovered and stored in an array of U1(n) if C1(n) value of correlation of LH band coefficients with Rn exceeds a threshold value. Otherwise U1(n) defaults to 0. Then, U1(n) is divided into m pieces and each piece is correlated with each watermark piece Wi(n) to identify which watermark piece is embedded in the current frame. If the maximum correlation value is smaller than a predefined threshold value, U(n) is ignored assuming a failure of watermark detection in the LH band of the current frame. Otherwise, the piece giving the highest correlation value is overwritten on U1(n). Similarly, correlation of every Rn with HL band coefficients of current frame yield U2(n) and C2(n). When all frames are processed in this manner, a number of recovered instances Uij(n) of each watermark piece Wi(n) and the correlation values Cij(n) of all those recovered bits can be obtained and recorded. Note that Uij(n) and Cij(n) contains results obtained from both LH and HL band coefficients. The algorithm for the reconstruction of watermark pieces Vi(n) from their recovered instances Uij(n) is the same as explained in Section 2.2.1.

3. THE PROPOSED HYBRID VIDEO WATERMARKING METHOD

3.1 Hybrid Watermarking

In the hybrid method, there are three ways to embed a piece of watermark in the current frame, using DCT, DWT, and a combination of the DCT-DWT transforms. In embedding stage, one of the three different watermarking algorithms and a watermark piece for the current frame are selected randomly. Depending on the selection, the entire frame is watermarked using the algorithms and parameters given in Section 2.1 (for DCT-based) or in Section 2.3 (for DWT-based) in the hybrid method. If the random selection comes out to be the combined method, the frame is divided into two equal halves by a horizontal line. The watermark piece is embedded to the upper half by the DCT-based algorithm and to the lower half by the DWT-based algorithm. In this case, the concepts are the same except that the maximum number of 8x8 DCT blocks and size of frequency regions of DWT in a frame will be halved. The hybrid method is expected to perform better by exploiting the advantages of the both transform based methods.

3.2 Watermark Recovery for Hybrid Method

The watermarked video is used as input for the DCT based, DWT based and combined DCT-DWT based watermark recovery modules, individually. Some attacks drop several frames from watermarked frame or change the order of frames in video. This makes any record of watermarking history of frames void, hence, all three modules must try to recover watermark data blindly from each frame. The recovering stages of the DCT based and DWT based watermarked frames are the same as given in Sections 2.2 and 2.3. If watermark data is to be recovered using The combined DCT-DWT recovery module applies to the upper half of each frame the DCT-based recovery algorithm and to the lower half the DWT-based algorithm. In the upper half frame, the same subset of DCT blocks watermarked by the combined method is located to recover the watermark bits. The outputs of two full frame independent recovery algorithms (DCT-based and DWT-based) and two half frame recovery algorithms for each frame are recorded separately. When all frames are processed 4 individual Uij(n), one from each recovery algorithm, are merged into one. The same merging applies to Cij(n) too. The algorithm for the reconstruction of watermark pieces Vi(n) from their hybrid-recovered instances Uij(n) is the same as explained in Section 2.2.1.

Some attacks can distort the watermark data embedded using one specific frequency domain transform. In this case, the advantage of the hybrid method appears. If one of the transform based methods cannot resist to some attacks, the other transform based method may hopefully resist and recover the lost watermark data from either the same frame or other frames.

4. EXPERIMENTAL RESULTS

4.1 Details of Tests

A digital video with 104 frames of size 640x480 and two different watermark data with dimensions of 20x20 and 24x24 pixels, respectively, are used in the tests (Figure 3).

(b) (c)

Figure 3: (a) A frame from the test video, (b) and (c) two watermarks used in the tests.

The test video is also watermarked with the DCT-based and DWT-based methods, individually, in order to compare their performances to that of the proposed hybrid method. The tests are performed under 4 different scenarios. The watermark 1 is divided to 2 and 4 pieces in the scenarios 1 and 2, respectively, and watermark 2 is divided to 2 and 6 pieces in the scenarios 3 and 4, respectively.

4.2 Results of Simulated Attacks

Noise addition attack adds to the watermarked video two types of noise, salt & pepper' noise and Gaussian noise depending on selection. Because a video contains a large amount of redundancies between frames, frame dropping is an attractive attack to destroy watermark data. In the tests, the frames of the watermarked videos were dropped up to 70% of the total number of frames in the test video. The frame averaging attack collects a number of successive frames and averages them out to generate an output frame. This is repeated at every frame of watermarked video producing an averaged video. The compression attack is simulated by compressing the watermarked video using a codec for 'wmv3' video format in MATLAB environment. Median filtering attack smooths images without blurring edges significantly. Intensity adjustment attack, maps the values in intensity image to new values such that normalized intensity values less than 0.01 and higher than 0.99 are saturated at 0 and 255, respectively. Contrast enhancement attack applies histogram equalization to every frame. The normalized correlation value between a reconstructed watermark and an original watermark is computed and given in Table 1 as a measure of watermark detection robustness of the tested methods under a given scenario and an attack. Simulation results show that the hybrid method improves watermark robustness against the tested attacks.

4.3 Imperceptibility and Capacity

Imperceptibility indicates how invisible the watermark is. This requirement has a trade off relation with two other requirements, robustness and capacity. The imperceptibility of the watermarked data is measured using peak signal-to-noise ratio (PSNR) between original and watermarked data. Imperceptibility values of all methods computed from the watermarked video frames were extremely high (over 80 dB for every frame) so that existence of any watermark in video could not be visually detected. Capacity is the amount of the data that can be embedded in a digital data. Increasing the size of watermark data embedded in a video decreases the visual quality of the video. We can compute the capacity of the DCT based method only, because in the DWT based method, number of bits to be embedded in a frame is theoretically infinite. For a given video, capacity of the DCT based method is equal to the number of blocks with a given size (usually 8) in a frame. The DCT based capacity is 4800 bits/frame for full frame watermarking and 2400 bits/frame for half-frame (i.e., for combined DCT-DWT based) watermarking.

Table 1: Performances of the DCT-based, DWT-based and hybrid methods for several attacks. Given are the normalized correlation values between recovered and original watermarks.

	DCT				DWT				Hybrid Method			
	Scenario				Scenario				Scenario			
Attack name	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
salt & pepper noise density= 0.02	0.98	0.98	0.97	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
salt & pepper noise density=0.04	0.69	0.74	0.96	0.65	0.99	0.95	1.00	1.00	0.89	0.95	1.00	1.00
Gaussian noise (mild)	0.97	0.88	0.98	0.87	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
frame dropping (70%)	0.99	0.98	0.97	0.87	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
frame averaging over 2 frames	0.98	0.98	0.97	0.78	1.00	1.00	0.90	0.79	1.00	0.99	1.00	1.00
frame averaging over 6 frames	0.92	0.81	0.93	0.70	0.86	0.86	0.86	0.58	1.00	0.99	1.00	0.95
video compression (quality=90)	0.87	0.70	0.86	0.72	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00
3x3 median filter	0.98	0.93	0.90	0.84	0.00	0.00	0.01	0.20	0.97	0.91	0.97	0.84
intensity adjustment	0.99	0.99	0.99	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
contrast enhancement	0.99	0.99	0.98	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

5. CONCLUSIONS

In this study, a hybrid digital video watermarking method is proposed. This hybrid method contains two frequency domain watermarking methods because of their advantages: DCT based and DWT based methods. Thus, the advantages of the both methods are utilized in the proposed method. The watermark is divided to sub-pieces and these pieces are embedded to frames. The hybrid method has better robustness compared to the individual methods it merges without significant reductions in the capacity and imperceptibility requirements. The size of a piece of watermark to be embedded into a frame must not be too small. Otherwise recovery of such small pieces from an attacked video may fail.

REFERENCES

Busch, C., Funk, W. and Wolthusen, S. (1999) Digital Watermarking: From Concepts to Real-Time Video Applications. IEEE Computer Graphics and Applications, 19, 25-35.

Chetan K.R, and Raghavendra K. (2010) DWT Based Blind Digital Video Watermarking Scheme for Video Authentication. International Journal of Computer Applications, 4, 19-26.

Cox, I., Kilian, J., Leighton, F. and Shamoon, T. (1997) Secure Spread Spectrum Watermarking for Multimedia. IEEE Transactions on Image Processing, 6, 1673-1687.

Cox, I., Miller, M. and Bloom, J. (2002) Digital Watermarking, Academic Press, USA

Hartung, F. and Girod, B. (1998). Watermarking of uncompressed and compressed video. Proceedings Signal Processing, 66, 283–301

Kundur, D. and Hatzinakos, D. (1998) Digital watermarking using multiresolution wavelet decomposition. Int. Conf. on Acoustics, Speech and Signal Processing, 2969-2972.

Langelaar, G., and Lagendijk, R. (2001) Optimal differential energy watermarking of dct encoded images and video. IEEE Transactions on image Processing, 148–158

Xia, X., Boncelet, C., and Arce, G. (1997) A Multiresolution Watermark for Digital Images. Proc. IEEE Int. Conf. on Image Processing, vol. I, 548-551.