Journal of Information Assurance and Security 2 (2007) 266-274

An Access Control Video Watermarking Method that is Robust to Geometric Distortions

Lino Coria[1], Panos Nasiopoulos[1], Rabab Ward[1] and Mark Pickering[2]

[1]Department of Electrical and Computer Engineering
University of British Columbia
2356 Main Mall, Vancouver, BC, Canada, V6T 1Z4
{linoc, panos, rababw}@ece.ubc.ca

[2]School of Information Technology and Electrical Engineering
University College, the University of New South Wales
Australian Defence Force Academy, Canberra ACT 2600
m.pickering@adfa.edu.au

Abstract: A new video watermarking algorithm for access control is introduced. This method is content-dependent and uses the dual tree complex wavelet transform (DT CWT) to create a watermark that is robust to geometric distortions and lossy compression. The watermark is a random array of 1's and -1's. A one-level DT CWT is applied to this watermark and the coefficients of this transformation are embedded into selected frequency components of the video sequence. The robustness of this method is tested against a joint attack, which involves rotation, scaling, cropping and H.264 video compression.

***Keywords*:** video watermarking, complex wavelets, geometric distortions, playback control**.**

1. Introduction

Piracy, the practice of selling, acquiring, copying or distributing copyrighted materials without permission has always been a concern for Hollywood studios and independent filmmakers. Theatrical camcorder piracy is one of the most common ways of illegally copying a movie [1]. This method consists of introducing a camcorder into a poorly supervised theater and creating a copy of the movie that is being shown. For a little more than a decade, watermarking techniques have been designed to, among other purposes, control access to digital content [2]. A watermark embedded in the video sequence provides information on whether video players are authorized to display the content or not. Compliant devices detect the watermark and obey the encoded usage restrictions. For this type of application, the decoding process must be blind as well as robust to geometric distortions (rotation, scaling), cropping and lossy compression.

Several watermarking methods have been design to resist geometric distortions. In [3], for example, two watermarks are employed. The first one is used to embed the message while the second one, a 1-bit watermark, is employed for geometric reference. This reference watermark is embedded in the spatial domain, which results in low robustness. A content-based image watermarking method is offered in [4], where robustness to geometric attacks is achieved by using feature points from the image. Although the scheme was shown to be successful to certain attacks, the method is computationally intensive and, therefore, not suitable for real-time video applications. In [5] the watermark is embedded in the video frame by applying the discrete wavelet transform (DWT) to it and replacing certain coefficients with the maximum or minimum value of its neighboring coefficients. While this scheme was proven to be robust to geometric attacks and compression, it provided no mechanism for controlling the amount of distortion introduced into the frame.

When dealing with signals that have more than one dimension, the Dual-Tree Complex Wavelet Transform (DT CWT) [6] is a particularly valuable solution since it adds perfect reconstruction to the list of desirable properties that regular complex wavelets have: shift invariance and directional selectivity. This makes the DT CWT ideal for designing watermarking algorithms that are capable of enduring slight geometric transformations. The use of this transform, however, is not straightforward since it is redundant and information might be lost when performing the inverse transformation. Therefore, only a handful of schemes have taken advantage of this tool. A watermark that consists of a pseudorandom sequence constructed with valid CWT transform coefficients is proposed in [7] and [8]. A four-level DT CWT is applied to the original content and the watermark is added to the coefficients from levels 2 and 3. Although the ideas portrayed in these efforts show some potential, the robustness of the schemes is never tested. Another watermarking method that uses the DT CWT is presented in [9]. In this method, the content is also subjected to a four-level DT CWT decomposition and the watermark is added to the two highest levels using the spread spectrum technique. However, the decoding process is not blind and, therefore, the applications for this scheme are very limited. More recently, a blind decoding watermarking scheme that overcomes geometric distortions by using the DT CWT was introduced in [10]. This method, designed specifically for access control of digital content, is robust to rotation, cropping and H.264 compression. However, the watermark needs to be extracted from a considerable number of frames in order to reach the correct decision, which is one of these

Received December 10, 2007.

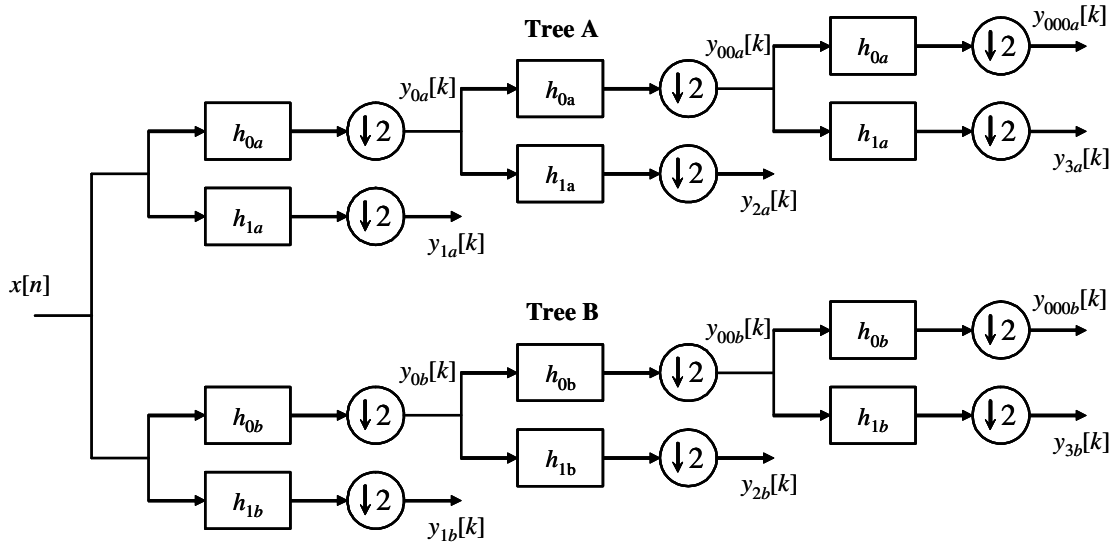

Figure 1. The two tree filter structure of the DT CWT.

two options: play or do not play the content.

We introduce in this paper a new watermarking method that is based on the DT CWT and proven to be robust to both lossy compression and geometric attacks. The decoding process from our approach is blind and the amount of frames required to retrieve the watermark is small when compared to the scheme presented in [10]. In our method, the watermark is a random set of 1's and -1's. A one-level DT CWT is applied to this watermark and the coefficients of this transformation become the data embedded into the video sequence, thus avoiding the inconveniences of a redundant transformation. Every frame of the original video sequence is transformed with a four-level DT CWT and, after inspecting the content, the watermark coefficients are weighted and added to the coefficients of levels 3 and 4. Our algorithm is tested against lossy compression and geometric distortions.

2. Proposed Method

The dual-tree complex wavelet transform (DT CWT) [11] has the desirable properties of perfect reconstruction, shift invariance, good directional selectivity, limited redundancy and efficient order-N computation [6]. This transform is a variation of the original DWT implementation but the main difference is that the DT CWT uses two filter trees instead of one as shown in Fig. 1. If the level or scale of the filter outputs are denoted by s then the set of high-pass complex wavelet coefficients produced by the DT CWT at scale s are given by

$$y_s[k] = y_{sa}[k] + j\, y_{sb}[k] \tag{1}$$

where $y_{sa}[k]$ denotes the output coefficients of Tree A at scale s and $y_{sb}[k]$ denotes the output coefficients of Tree B at scale s. The complex coefficients can also be written in polar form as follows

$$y_s[k] = m_s[k]\, e^{j\theta_s[k]} \tag{2}$$

where the magnitude of each coefficient is given by

$$m_s[k] = \sqrt{y_{sa}^2[k] + y_{sb}^2[k]} \tag{3}$$

and the phase is given by

$$\theta_s[k] = \tan^{-1}\left(\frac{y_{sb}[k]}{y_{sa}[k]}\right) \tag{4}$$

The number of complex coefficients produced at scale s is given by $N/2^s$ where N is the number of samples in the input signal. Fig. 2 illustrates the shift invariant properties of the DT CWT. Fig. 2(a) shows a number of input signals that contain an edge that is shifting by one sample to the right relative to the previous input. Fig. 2(b) shows the level 4 DWT output coefficients and Fig. 2(c) shows the magnitude of the level 4 DT CWT coefficients for each of these input signals. Fig. 2(b) clearly shows the variation in the DWT coefficients as the input signal shifts to the right. In contrast to the DWT coefficients, the magnitude of the complex coefficients of the DT CWT shown in Fig. 2(c) remain relatively unaffected by small shifts in the input signal. This property is particularly useful when detecting watermarks that have undergone some geometric distortion. The more invariant a transform is to small shifts in the input signal the the greater the probability that a watermark embedded in the signal will be detectable after undergoing geometric distortion.

Fig. 3(a) shows the two dimensional impulse responses of the reconstruction filters in the 2D DT CWT. Each level of the transform produces complex coefficients that correspond to the output of six directional filters. The set of high-pass complex wavelet coefficients at scale s and direction d can be written as

$$y_{s,d}[u,v] = m_{s,d}[u,v]\, e^{j\theta_{s,d}[u,v]} \tag{5}$$

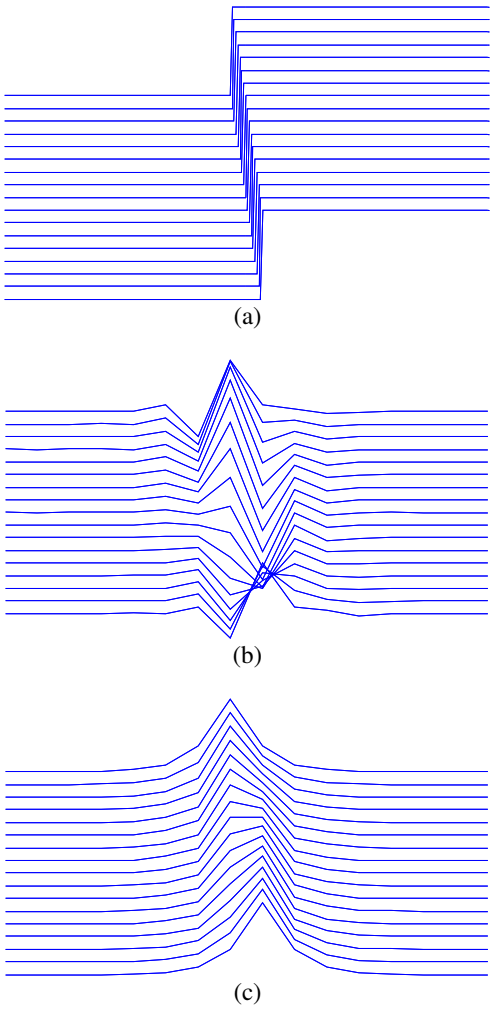

Figure 2. (a) Input signals that contain an edge that is shifting by one sample to the right. (b) Level 4 DWT coefficients. (c) Level 4 DT CWT coefficients.

for $d = 1, 2, ..., 6$. The variables u and v specify the location of the complex coefficients in each sub-band. Using matrix notation the coefficients in a subband are denoted by $\mathbf{Y}_{s,d}$ and the number of complex coefficients in each subband is $N/2^s$ x $M/2^s$ where N and M are the dimensions of the video frame in pixels. Fig. 3(b) shows a wavelet-type output structure for the six directional sub-bands at each level of the DT CWT.

Since the application that is being considered for our watermarking scheme is access control, there are only two possible outcomes needed from the decoder, that is, the watermark is either detected or not. This means that there is no information associated with the watermark other than its presence or absence. If a watermark is detected by a compliant DVD player, for instance, the copy will be considered of illegal origin and the movie will not be played.

2.1 Creating the watermark

In our method, the watermark is inserted on every frame of the video sequence. The watermark is a 2D array that is sixty-four times smaller than the video frame where it will be embedded (i.e. its height and width are one eighth of the frame's height and width, respectively). The watermark $\mathbf{\Omega}$ is a pseudorandom matrix of 1's and -1's that is created using a key $K = K_O + K_F$. K_O is provided by the user and is constant for the whole video sequence while K_F changes every β frames.

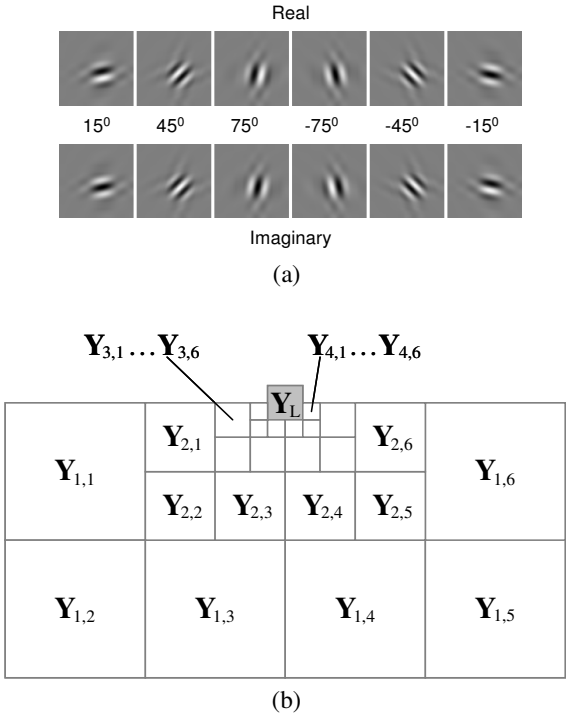

Figure 3. (a) Two dimensional impulse responses of the reconstruction filters in a 2D DTCWT; (b) Structure of the DT CWT coefficients for a four level decomposition.

Usually, watermarking algorithms rely on the addition of a pseudorandom sequence to the host content coefficients in some frequency domain. This approach, however, cannot be used directly when working in the DT CWT domain. The DT CWT is a redundant transform and some components of an arbitrary pseudorandom sequence in the DT CWT domain may be lost during the inverse transform. The loss of information corresponds to the sequence that lies in the null space of the inverse DT CWT [8]. One way to reduce this information loss is by embedding in the host content the DT CWT coefficients of the watermark instead of the actual watermark. Having this in mind, the one-level DT CWT transform is applied to the watermark $\mathbf{\Omega}$. This results in a low-pass component, $\mathbf{W}_L$, and six subbands that contain the details, $\mathbf{W}_1$ ... $\mathbf{W}_6$. The coefficients from these subbands form the data to be embedded in the frame. This process is illustrated in Fig. 4.

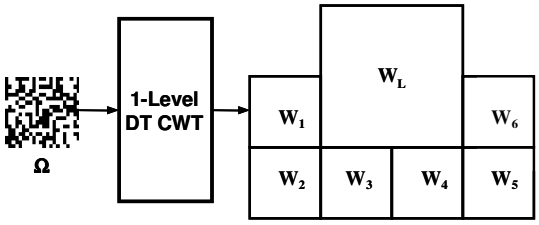

Figure 4. The one-level DT CWT transform is applied to the watermark.

2.2 Embedding the watermark

A four-level DT CWT is applied to every video frame. The watermark is then embedded in the level 3 and level 4 coefficients of these frames. The strength of the watermark is decided by using information from level 2 coefficients. The embedding algorithm is described in detail next.

We consider one frame at a time. A perceptual mask for each of the six level 3 subbands is created. In order to obtain these masks we apply a low-pass filter to every level 2 subband $\mathbf{Y}_{2,1}$ ... $\mathbf{Y}_{2,6}$ and then down-sample the resulting arrays by a factor of 2. The elements of these arrays are divided by a step value Δ and then rounded to the next higher integer value (this operation is represented by the symbol $\lceil x \rceil$).The resulting arrays $\mathbf{P}_{3,1}$ ... $\mathbf{P}_{3,6}$ have the same dimensions as the level 3 subbands. This process is described in (6).

$$\mathbf{P}_{3,d} = \left\lceil \frac{\left[(\downarrow 2)(\mathbf{Y}_{2,d} * \mathbf{h})\right]}{\Delta} \right\rceil \tag{6}$$

for d = 1, 2, ..., 6 and the low pass filter is given by

$$\mathbf{h} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix} \tag{7}$$

The masks for the level 4 subbands are created as follows. The low-pass filter **h** is applied to every level 2 subband $\mathbf{Y}_{2,1}$ ... $\mathbf{Y}_{2,6}$ and then the resulting arrays are down-sampled by a factor of 2. The same process of low-pass filtering and down-sampling is applied again to these arrays. The elements of the resulting arrays are divided by the step value Δ and then rounded to the next higher integer value. These new arrays become the masks $\mathbf{P}_{4,1}$ ... $\mathbf{P}_{4,6}$ for the level 4 subbands and share the same dimensions as these subbands. This is process is described in (8).

$$\mathbf{P}_{4,d} = \left\lceil \frac{\left[(\downarrow 2)\left[(\downarrow 2)(\mathbf{F}_{2,d} * \mathbf{h})\right] * \mathbf{h}\right]}{\Delta} \right\rceil \tag{8}$$

for d = 1, ..., 6.

These perceptual masks provide an estimate of the strength that can be used to embed the watermark in every coefficient from levels 3 and 4. The elements in these masks will be used as weights during the embedding process. Since the watermark is not embedded in the level 2 coefficients, the masks can be retrieved at the decoder.

The watermarked frame is then created by adding the magnitude of the watermark coefficients to the magnitudes of the level 3 and level 4 coefficients of the current frame. For level 4 coefficients this process is defined by the following equation:

$$\hat{y}_{4,d}[u,v] = \left(m_{4,d}[u,v] + \alpha\, p_{4,d}[u,v] \left| w_d[u,v] \right|\right) e^{j\theta_{3,d}[u,v]} \tag{9}$$

for d = 1, 2, ..., 6. In equation (9), $\hat{y}_{4,d}[u,v]$ denotes the coefficients at position $[u,v]$ of the watermarked level 4 subbands $\hat{\mathbf{Y}}_{4,d}$. Similarly, $p_{4,\mathrm{d}}[u,v]$ denotes the elements of the perceptual mask arrays $\mathbf{P}_{4,d}$ and $\left| w_d[u,v] \right|$ denotes the magnitude of the complex coefficients of the level 4 watermark subband $\mathbf{W}_d$. The parameter α is always positive and is used to control the fidelity impact of the watermark.

For level 3 coefficients the watermark coefficients are embedded using the process defined by the following equation:

$$\hat{y}_{3,d}[u,v] = \left(m_{3,d}[u,v] + \alpha\, p_{3,d}[u,v] \left| z_d[u,v] \right|\right) e^{j\theta_{3,d}[u,v]} \tag{10}$$

for d = 1, 2, ..., 6. In equation (10), $\hat{y}_{3,d}[u,v]$ denotes the coefficients of the watermarked level 3 subbands $\hat{\mathbf{Y}}_{3,d}$, $p_{3,\mathrm{d}}[u,v]$ denotes the elements of the perceptual mask array $\mathbf{P}_{3,d}$ and $\left| z_d[u,v] \right|$ denotes the magnitude of the complex coefficients of the level 3 watermark subband $\mathbf{Z}_d$, which is created as follows:

$$\mathbf{Z}_d = \begin{bmatrix} \mathbf{W}_d & \mathbf{W}_d \\ \mathbf{W}_d & \mathbf{W}_d \end{bmatrix} \tag{11}$$

The original level 3 and level 4 coefficients are then replaced by the watermarked coefficients and the inverse DT CWT is calculated to produce the watermarked video frame.

2.3 Decoding the watermark

The decoding process is blind, that is, the watermark is decoded without relying on any information from the original video file. For every frame of the watermarked video sequence, the 4-level DT CWT is applied. The masks for levels 3 and 4 are obtained via (2) and (3), respectively. The level 3 subbands and perceptual mask arrays are then partitioned as follows:

$$\hat{\mathbf{Y}}_{3,d} = \begin{bmatrix} \hat{\mathbf{Y}}_{\mathrm{A},d} & \hat{\mathbf{Y}}_{\mathrm{B},d} \\ \hat{\mathbf{Y}}_{\mathrm{C},d} & \hat{\mathbf{Y}}_{\mathrm{D},d} \end{bmatrix} \tag{12}$$

$$\mathbf{P}_{3,d} = \begin{bmatrix} \mathbf{P}_{\mathrm{A},d} & \mathbf{P}_{\mathrm{B},d} \\ \mathbf{P}_{\mathrm{C},d} & \mathbf{P}_{\mathrm{D},d} \end{bmatrix} \tag{13}$$

An estimate of the magnitude of the 1-level DT CWT coefficients of the watermark can then be obtained using the partitioned level 3 watermarked coefficients and the level 4 watermarked coefficients as follows:

$$|w'_d[u,v]| = \frac{|\hat{y}_{A,d}[u,v]|}{p_{A,d}[u,v]} + \frac{|\hat{y}_{B,d}[u,v]|}{p_{B,d}[u,v]} + \frac{|\hat{y}_{C,d}[u,v]|}{p_{C,d}[u,v]} + \frac{|\hat{y}_{D,d}[u,v]|}{p_{D,d}[u,v]} + \frac{|\hat{y}_{4,d}[u,v]|}{p_{4,d}[u,v]} \quad (14)$$

for $d = 1, 2, ..., 6$.

The estimated magnitudes of the watermark coefficients are then combined with their corresponding phases and the inverse DT CWT is performed to obtain the decoded watermark. Since the low-pass component $\mathbf{W}_L$ was not encoded in the watermarked video sequence, $\mathbf{W}'_L$ is considered to be an array of zeros when performing this inverse DT CWT.

The resulting decoded watermark is then correlated with the original watermark for each frame and the results are added until a certain amount of frames is reached (a hundred or more is recommended). When a watermark is decoded from every frame, the cumulative correlation strength will be a relatively high positive number when compared to the correlation value obtained after an un-watermarked video sequence has undergone the same decoding process. By looking at the resulting correlation strength, a decision can be made at the decoder on whether the video sequence contains an embedded watermark or not.

3. Experimental Results

To test the proposed watermarking method we employed ten QCIF (176 × 144) video sequences. Each consists of 300 frames. Five of them were the standard video files Container, Hall Monitor, Mother and Daughter, News, and Suzie. The other five were formed using different short frame sequences from these standard video files. The combined sequences had a change of scene every 30 frames. We refer to the combined sequences as Mix 1, Mix 2, ..., Mix 5. Five different watermarked video sequences were created for each of the 10 sequences, by using a different key K_O each time. For every 15 frames different values of K_F were used. The watermarks were embedded in the luminance components. Tests were performed on each of the 10 × 5 = 50 sequences. For each sequence, 300 frames were used to determine the strength of the watermark. For our tests, α was set to 15 so that the average PSNR of the watermarked frames was 41 dB. Computer experiments showed that setting Δ to 25 yielded the best results.

In order to study the performance of our method we compared our results against two algorithms that employ the regular discrete wavelet transform (DWT). These two methods also have the advantages over others in that they are blind and computationally less demanding. Thus the original content is not needed to retrieve the watermark and the frames do not need to be geometrically restored before detecting the watermark. The first method we compare with is basically the same algorithm as proposed in this chapter except the DWT replaces the DT CWT. We will refer to this method as DWT1. The second method is the one presented in [5], which is also based on the DWT. In this method, which we denote as DWT2, video frames are watermarked by replacing the values of some coefficients of the frames by the highest or lowest values of their neighboring coefficients (depending if a 0 or a 1 is being embedded). The average PSNR of the sequences watermarked with these methods was also set to 41 dB.

The watermark decoder measures the correlation between the transform coefficients of every frame and the coefficients of the watermark. The measured correlation for a particular frame will be very small but, over several frames (300 for these tests), the accumulation of the correlations will provide an indication as to whether or not the video sequence is watermarked. When the normalized value of the correlation strength over 300 frames is between ±10, the decoder deduces that no watermark has been embedded. Otherwise the decoder decides that a watermark is present. Fig. 5 illustrates the difference between the strength values obtained for watermarked and non-watermarked sequences.

In order to compare the performance of the different methods, the correlation strength values were normalized: for every method, the highest correlation (after 300 frames) was set to 100 and the other values were proportionally modified. For sequences which did not suffer any distortion attack, the results obtained using the three methods are shown in Table 1. The normalized correlation strength values are displayed for both watermarked (W) and non-watermarked (NW) sequences. To fit the information in one table, the 5 tests performed for every video sequence were averaged.

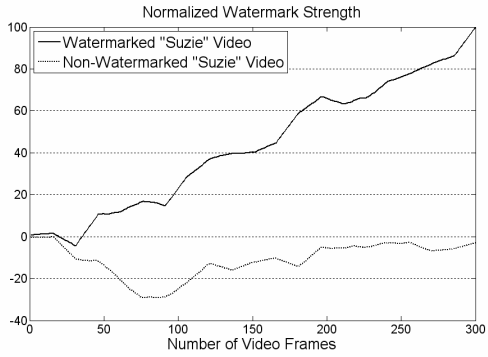

Figure 5. After 300 frames, the decoder detects a high watermark strength value for the watermarked Suzie sequence while the strength decoded for the non-watermarked Suzie video is close to zero.

Table 1. Comparison of normalized correlation values obtained by the three watermarking methods: DT CWT, DWT1 and DWT2.

VIDEO	DT CWT		DWT1		DWT2	
	NW	W	NW	W	NW	W
Container	7	98	-5	100	1	99
Hall	-1	91	4	79	2	99
Mother	0	93	-3	84	-5	99
News	0	100	-1	74	0	100
Suzie	0	87	1	100	-1	100
Mix 1	3	95	1	91	-2	100
Mix 2	2	95	-3	80	1	100
Mix 3	0	93	-6	88	-2	99

Table 2. Comparison of normalized correlation values for three watermarking methods: DT CWT, DWT1 and DWT2. Watermarked and Non-Watermarked sequences are subjected to scaling (by 5%, 10% and 15%) and cropping.

VIDEO	DT CWT						DWT1						DWT2					
	5%		10%		15%		5%		10%		15%		5%		10%		15%	
	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W
Container	9	69	5	41	3	26	0	41	0	9	0	-2	18	54	12	22	-14	-6
Hall	-4	60	-4	39	-7	22	1	34	1	7	0	0	19	36	18	18	-1	4
Mother	0	67	-1	42	-2	26	-1	25	0	5	0	1	8	50	1	0	-7	-12
News	-1	72	0	46	0	30	0	23	-1	6	0	1	-4	54	3	1	-4	-13
Suzie	1	66	2	43	1	28	0	31	0	1	1	-1	6	48	-19	5	-11	-5
Mix 1	2	67	0	42	-3	25	-1	31	-1	4	-1	-2	-1	49	-7	6	-15	-9
Mix 2	2	67	1	43	-2	25	0	27	0	5	1	0	-8	53	0	8	-7	-2
Mix 3	1	67	1	42	0	26	0	33	0	5	1	-1	12	41	1	8	-5	-6
Mix 4	-1	65	-2	40	-3	25	0	34	1	10	0	1	27	44	25	13	-1	-8
Mix 5	2	67	2	44	2	30	0	31	1	6	1	0	16	55	-4	11	-9	-6
Mean	**1**	**67**	**0**	**42**	**-1**	**26**	**0**	**31**	**0**	**6**	**0**	**0**	**9**	**48**	**3**	**9**	**-7**	**-6**

Mix 4	0	94	3	92	-1	100
Mix 5	0	92	1	84	2	100
Mean	**1**	**94**	**-1**	**87**	**0**	**100**

The three watermarking schemes show similar results in sequences which did not undergo any distortion attack. There is a significant difference between the normalized correlation values obtained from watermarked sequences and the values obtained from non-watermarked ones. For each method, when the decoder attempts to extract a watermark from a non-watermarked video sequence the resulting correlation values are very low (an average of 1, -1 and 0 for DT CWT, DWT1 and DWT2, respectively) since the information in the video frames is not correlated with the actual watermark. In contrast, correlation values obtained from watermarked sequences are constantly high (average values of 94, 87 and 100 for DT CWT, DWT1 and DWT2, respectively).

We then tested the robustness of our method to common distortions. In the first experiment, watermarks were decoded after the video sequences had undergone scaling and cropping distortion. For the second test, the video sequences were rotated by a few degrees and the watermark was later decoded. We also tested the effects of additive Gaussian noise and lossy compression. Finally, all these distortions: scaling, rotation, cropping, additive noise and lossy compression were put together as a joint attack.

3.1 Frame scaling and cropping

We examined the robustness of all three methods when the watermarked sequences were subjected to scaling and cropping. Every video sequence was scaled up by 5%, 10% and 15% using bicubic interpolation. The frames were then cropped back to their original size (176 × 144). A visual example of this process can be seen in Fig. 6. Results are summarized in Table 2.

3.2 Frame rotation

Robustness to frame rotation was then tested. Each frame was rotated counterclockwise by 3^o, 6^o and 9^o. Bilinear interpolation was employed and the resulting images were cropped to fit the QCIF format. An example of this attack can be seen in Fig. 7. Table 3 shows comparisons of the performance by the three watermarking methods to this particular type of distortion.

Once again, it can be observed that, for the case of DT CWT, the higher the degree of rotation, the lower the strength of the watermark correlation. Even though the watermark correlation strength is low when frames are rotated by 9^o, this value (18 on average) is high enough to be disassociated with a non-watermarked sequence. Regarding the other two methods, DWT1 and DWT2, it is evident that they are extremely fragile to rotation. None of these schemes was able to withstand a rotation attack higher than 3^o. The watermarks embedded with these schemes went completely undetected.

Table 3. Normalized correlation values for sequences subjected to rotation (by 3°, 6° and 9°) and cropping.

VIDEO	DT CWT						DWT1						DWT2					
	3°		6°		9°		3°		6°		9°		3°		6°		9°	
	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W
Container	9	71	3	35	2	19	0	20	0	2	0	0	23	53	10	-10	4	-2
Hall	-2	61	-2	33	-4	15	0	18	0	1	0	-2	8	17	4	1	-4	14
Mother	0	62	-2	33	-2	17	-1	19	-1	-1	-1	-2	16	60	-13	1	-9	-2
News	-1	68	1	39	2	21	0	14	0	-5	0	-2	3	49	-20	2	-8	-5
Suzie	-1	56	-1	30	-2	16	0	16	0	0	0	2	-4	48	-5	0	8	-6
Mix 1	3	65	3	37	4	22	0	21	-1	-1	-1	2	23	43	1	1	11	-3
Mix 2	2	65	1	35	-1	18	0	14	0	-4	0	-5	-3	46	-24	-4	-18	4
Mix 3	0	63	2	35	3	21	0	21	0	-1	0	-2	23	38	-7	-7	-2	-5
Mix 4	-1	63	-6	28	-5	13	0	16	0	-1	0	-2	3	39	-2	1	-14	4
Mix 5	1	63	1	34	-5	13	1	16	0	2	0	3	1	61	7	3	14	-1
Mean	**1**	**64**	**0**	**34**	**-1**	**18**	**0**	**17**	**0**	**-1**	**0**	**-1**	**9**	**45**	**-5**	**-1**	**-2**	**0**

Figure 6. A watermarked frame of the sequence Suzie is scaled and then cropped: (a) 5%, (b) 10% and (c) 15% scaling.

Figure 7. A watermarked frame of the sequence Suzie is rotated and then cropped: (a) 3°, (b) 6° and (c) 9° rotation.

3.3 Additive noise

We tested the robustness of the schemes to additive noise. Gaussian noise with a standard deviation of 5 was added to the frames. All three watermarking methods proved to be robust to additive noise. The average detected watermark strength values for DT CWT, DWT1 and DWT2 are, respectively, 90, 87, and 96. Table 4 shows the results.

3.4 Compression

In order to test the robustness of the proposed scheme to compression we encoded the video sequences using H.264/AVC. Every 15th frame was set to be an I-frame and the rest were chosen to be P-frames. The quantization parameter QP for both I and P frames was set to 15, which results in a compression ratio of around 40:1. In this instance, the three watermarking methods demonstrate to be robust to compression. The average detected watermark strength is 85, 77 and 93 for DT CWT, DWT1 and DWT2, respectively. Results are summarized in Table 4.

Table 4. Normalized correlation values for sequences subjected to additive noise with a standard deviation of 5 and also to H.264 compression with a QF of 15.

VIDEO	DT CWT				DWT1				DWT2			
	Noise		H.264		Noise		H.264		Noise		H.264	
	NW	W	NW	W	NW	W	NW	W	NW	W	NW	W
Container	7	93	7	89	-5	99	0	88	1	97	-20	90
Hall	-1	87	-1	83	4	79	0	70	2	96	15	93
Mother	0	90	0	84	-3	84	-1	73	-5	97	-10	96
News	0	97	0	88	-1	74	0	64	0	96	-9	90
Suzie	0	82	0	79	1	100	0	88	-1	97	5	96
Mix 1	3	91	3	86	1	90	0	80	-2	97	3	93
Mix 2	2	91	2	86	-3	81	0	70	1	97	-13	95
Mix 3	-1	89	0	84	-6	88	0	77	-2	94	-26	91
Mix 4	0	89	0	84	3	91	0	81	-1	96	13	91
Mix 5	0	88	0	83	1	83	0	75	2	97	6	94
Mean	**1**	**90**	**1**	**85**	**-1**	**87**	**0**	**77**	**0**	**96**	**-4**	**93**

3.5 Joint attack

The final experiment involved all the previous attacks together. For this joint attack, we scaled the video frames by 5% and rotated them by 5°. The frames were later cropped to fit their original size (176 × 144). Gaussian noise with a standard deviation of 5 was added and H.264/AVC was used to compress the video sequences (same compression ratio as before). An example of a video frame that has gone through this joint attack can be seen in Fig. 8. The performance of the three methods is shown in Table 5.

(a)

(b)

Figure 8. (a) A watermarked frame of the sequence Suzie. (b) The same frame is subjected to a joint attack (rotating by 5°, scaling up by 5%, cropping, adding noise and using H.264 compression).

Results for the DT CWT indicate that the method can successfully survive a joint attack. After 300 frames, the average watermark correlation strength value was 37. Considering the poor performance of DWT1 and DWT2 to the previous distortions, it is not surprising that these methods are not able to withstand a joint attack. Although these approaches are very robust to noise and lossy compression, they can only endure mild geometric distortions.

Table 5. Normalized correlation values for sequences subjected to a joint attack (additive noise with a standard deviation of 5, rotating by 5°, scaling up by 5%, cropping back to QCIF and H.264 compression with a QF of 15).

VIDEO	JOINT ATTACK					
	DT CWT		DWT1		DWT2	
	NW	W	NW	W	NW	W
Container	7	39	0	3	16	-1
Hall	-3	36	1	1	9	-16
Mother	-3	34	0	0	-3	-14
News	1	42	1	0	-3	-2
Suzie	-1	32	0	-1	-7	-10
Mix 1	-1	35	-1	3	15	-10
Mix 2	0	37	0	-3	-11	-8
Mix 3	1	37	0	1	-3	-9
Mix 4	-2	35	1	2	8	-8
Mix 5	2	39	1	0	3	-8
Mean	**0**	**37**	**0**	**0**	**3**	**-9**

4 Conclusion

A new video watermarking algorithm for playback control that takes advantage of the properties of the dual-tree complex wavelet transform (DT CWT) is introduced. This transform maintains the advantages but avoids the shortcomings of regular wavelets. The DT CWT provides important features such as perfect reconstruction, approximate shift invariance and good directional selectivity. Our method relies on these characteristics to create a watermark that is robust to geometric distortions.

The robustness of our method was tested against several attacks, which included lossy compression, additive noise, rotation, scaling, cropping and a joint attack, which involved a combination of all the previous distortions. Our method successfully detected the presence of the watermarks in all the corrupted video sequences. The joint attack was employed to simulate a video sequence that has been recorded from a movie screen with a handheld camcorder and then stored in a digital format.

In order to compare the performance of our scheme and evaluate the advantages of using DT CWT as a watermarking tool, we subjected two DWT-based watermarking algorithms to the same fidelity standards and the same attacks. Although these methods were robust to compression and noise, they did not survive any significant geometric distortions.

Our proposed method is simple to implement; this is important when considering the added cost and complexity to DVD players. Furthermore, it is robust to geometric distortions and to lossy compression. All these characteristics make our algorithm suitable for access control of digital video.

5 Acknowledgment

The authors would like to thank Dr. Nick Kingsbury for providing the software to perform the DT CWT transform operations.

References

[1] Motion Picture Association of America, 2007. Available: http://www.mpaa.org/piracy.asp

[2] P. B. Schneck, "Persistent Access Control to Prevent Piracy of Digital Information," Proceedings of the IEEE, vol. 87, pp. 1239-1249, July, 1999.

[3] C. V. Serdean, M. A. Ambroze, M. Tomlinson and J. G. Wade, "DWT-based high-capacity blind video watermarking, invariant to geometrical attacks," IEE Proceedings -- Vision, Image & Signal Processing, vol. 150, pp. 51-58, Feb 2003.

[4] P. Bas, J. M. Chassery and B. Macq, "Geometrically invariant watermarking using feature points," IEEE Transactions on Image Processing, vol. 11, pp. 1014, 2002.

[5] P. W. Chan, M. R. Lyu and R. T. Chin, "A novel scheme for hybrid digital video watermarking: approach, evaluation and experimentation," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 15, pp. 1638-1649, Dec. 2005.

[6] N. Kingsbury, "Image processing with complex wavelets," Philosophical Transactions. Mathematical, Physical, and Engineering Sciences, vol. 357, pp. 2543, 1999.

[7] P. Loo and N. Kingsburry, "Digital watermarking using complex wavelets," in International Conference on Image Processing, ICIP, 2000, pp. 29-32.

[8] P. Loo and N. Kingsbury, "Digital watermarking with complex wavelets," in IEE Seminar on Secure Images and Image Authentication, 2000, pp. 10/1-10/7.

[9] N. Terzija and W. Geisselhardt, "Digital image watermarking using complex wavelet transform," in MM&Sec '04: Proceedings of the 2004 Workshop on Multimedia and Security, 2004, pp. 193-198.

[10] M. Pickering, L. E. Coria and P. Nasiopoulos, "A novel blind video watermarking scheme for access control using complex wavelets," in International Conference on Consumer Electronics, 2007, pp. 1-2.

[11] N. Kingsbury, "The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters," Proc. 8th IEEE DSP Workshop, pp. Paper no. 86, 1998.

Author Biographies

Lino Coria was born in Morelia, Mexico. He obtained his Bachelor degree in electronics engineering from Instituto Tecnológico de Morelia in Morelia, Mich., Mexico in 1996, his Master's degree in electrical engineering from McMaster University in Hamilton, ON, Canada in 1998, and his PhD degree in electrical and computer engineering from the University of British Columbia in Vancouver, BC, Canada in 2008.

He is currently a Professor at Instituto Tecnológico y de Estudios Superiores de Occidente (ITESO) Department of Electronics, Systems and Informatics in Tlaquepaque, Jalisco, Mexico. His interests include image and video processing and his research has mainly focused on digital watermarking.

Dr. Coria was awarded two scholarships for his graduate studies (1996-1998 and 2003-2007) by the Mexican Council for Science and Technology (CONACYT). He is a member of IEEE and ACM.

Panos Nasiopoulos holds a Bachelor degree in physics from the Aristotle University of Thessaloniki, Greece, and a Bachelor, Master and PhD in electrical and computer engineering from the University of British Columbia, Canada.

He is an Associate Professor with the University of British Columbia (UBC) department of Electrical and Computer Engineering, the holder of the Professorship in Digital Multimedia, and the current Director of the Master of Software Systems Program at UBC. Before joining UBC, he was the President of Daikin Comtec US and the Executive Vice President of Sonic Solutions. He was voted as one of the most influential DVD executives in the world. He is recognized as a leading authority on DVD and multimedia and has published numerous papers on the subjects of digital video compression and communications. He has organized and chaired numerous conferences and seminars and he is a featured speaker at multimedia/DVD conferences worldwide.

Dr. Nasiopoulos has been an active member of IEEE, ACM, the Standards Council of Canada (ISO/ITU and MPEG) and the In-Flight-Entertainment committee.

Rabab Kreidieh Ward was born in Beirut, Lebanon. She obtained her Bachelor degree in electrical engineering from the University of Cairo in 1966 and her Master's and PhD, degrees from University of California, Berkeley in 1969 and 1972 respectively.

She has made significant research contributions in digital signal processing and its applications to cable TV, high definition TV, video compression and medical images, including mammography, microscopy and cell images. Her research ideas have been transferred to industry. She was the chair of IEEE International Conference on Image Processing 2000 and IEEE ISSPIT, and was the vice chair of IEEE ISCAS 2004.

Dr. Ward is the recipient of the "Society Award" of the IEEE Signal Processing Society, the R.A. McLachlan Memorial Award, of the Association of Professional Engineers and Geoscientists of British Columbia, and the UBC Killam Research Prize. She is a fellow of the Royal Society of Canada, the IEEE, the Canadian Academy of Engineers and the Engineering Institute of Canada.

Mark R. Pickering was born in Biloela, Australia in 1966. He received a Bachelor of Engineering from Capricornia Institute of Advanced Education, Rockhampton, Australia in 1988 and Master of Engineering and PhD degrees in electrical engineering from the University of New South Wales, Sydney, Australia in 1991 and 1995 respectively.

He is currently a Senior Lecturer in the School of Information Technology and Electrical Engineering at

the University College, University of New South Wales, Australian Defence Force Academy, Canberra, Australia.

Dr Pickering's research interests include video and audio coding, medical imaging, data compression, information security, data networks and error-resilient data transmission.