A wavelet - Based Object Watermarking System for MPEG4 Video

A.Essaouabi abdessamad1977@yahoo.fr
Department of physics, LIMIARF Laboratory,
Mohammed V University
Rabat, Morocco

E.Ibnelhaj ibnelhaj@inpt.ac.ma
Institut National of Posts and Telecommunications,
Rabat, Morocco

F.Regragui regragui@fsr.ac.ma
Department of physics, LIMIARF Laboratory,
Mohammed V University
Rabat, Morocco

Abstract

Efficient storage, transmission and use of video information are key requirements in many multimedia applications currently being addressed by MPEG-4. To fulfill these requirements, a new approach for representing video information which relies on an object-based representation, has been adopted. Therefore, object-based watermarking schemes are needed for copyright protection. This paper presents a novel object based watermarking solution for MPEG4 video authentication using the shape adaptive-discrete wavelet transform (SA-DWT). In order to make the watermark robust and transparent, the watermark is embedded in the average of wavelet blocks using the visual model based on the human visual system. Wavelet coefficients n least significant bits (LSBs) are adjusted in concert with the average. Simulation results show that the proposed watermarking scheme is perceptually invisible and robust against many attacks such as lossy compression (e.g. MPEG1 and MPEG2, MPEG-4, H264).

Keywords: *Watermark*, *Visual model*, Robustness, Shape adaptive-discrete wavelet transform.

1. INTRODUCTION

With the emergence of new multimedia standards such as Mpeg-4, the notion of video-object or image object is more and more widespread [1][2]. Consequently, protecting the different objects of an image or a video appeared necessary. Therefore several object-based watermarking techniques, those consist at introducing an invisible signal known as a digital watermark into an image or video sequence, aim at solving this type of problem. Wu and al. [3] proposed a multiresolution object watermarking approach based on 2D and 3D shape adaptive wavelet transforms. The advantage of the multiresolution watermarking method is its robustness against image/video compression and computational saving. However, the main disadvantage is that original image/video object is required for watermark detection. Kim and al. [4] proposed an object-based video watermarking method using the shape adaptive-discrete cosine transforms

(SA-DCT). The SA-DCT method is superior to all other padding methods in terms of robustness against the image deformations. Yet, the watermark can be damaged by a wavelet-based image codec in the quantization stage. Therefore, this method limits their applications in the context of JPEG2000 and MPEG-4 due to the fact that the wavelet transform is playing an important role in JPEG2000 and MPEG-4. Piva and al. [5] propose an object watermarking system for MPEG-4 streams. Since this method applies the discrete wavelet transform (DWT) to the whole image and the watermark is embedded in all the wavelet coefficients belonging to the three detail bands at level 0, this may lead to loss of the watermark which is embedded in the region outside the object. Barni and Bartolini [6] proposed a method that consists in embedding a watermark in each video object of an MPEG-4 coded video bit-stream by imposing specific relationships between some predefined pairs of quantized DCT middle frequency coefficients in the luminance blocks of pseudo-randomly selected macroblocks. The quantized coefficients are recovered from the MPEG-4 bit-stream, they are modified to embed the watermark and then encoded again. The main drawback of this technique is that, since the code is directly embedded into the compressed MPEG-4 bit-stream, the copyright information is lost if the video file is converted to a different compression standard, like MPEG-2. In order to be robust against format conversions, the watermark has to be inserted before compression, i.e. frame by frame.

In order to satisfie the previous requirements we propose in this paper an object based watermarking solution for MPEG4 video authentication based on in place lifting SA-DWT . The watermark signal is embedded in the wavelet coefficients n LSBs before MPEG4 encoding and not embedded in the region outside object. Unlike most watermark schemes, watermark embedding is performed by modulating the average of the wavelet coefficients instead of the individual coefficients in the wavelet block. Visual model is employed to achieve the best tradeoff between transparent and robustness to signal processing. Watermark detection is accomplished without the original. Experimental results demonstrate that the proposed watermarking scheme is perceptually invisible and robust against unintentional and intentional attacks such as lossy video compression (e.g. MPEG2 and MPEG-4,MPEG1 and H264).

The rest of this paper is organized as follows:
In section 2, we briefly describe the SA-DWT, section 3 will describe the basic functionalities of watermarking embedding and extraction procedure, section 4 will give the simulations results, finally section 5 will give the conclusion.

2. IN-PLACE LIFTING SHAPE ADAPTIVE-DISCRETE WAVELET TRANSFORM

Given an arbitrarily shaped object with shape mask information, with in place lifting SA-DWT[8][9], the number of transformed coefficients is equal to the number of pixels in the arbitrarily shaped segment image, and the spatial correlation across subbands is well preserved. Fig. 1 illustrates the result of one-level wavelet decomposition of an arbitrarily shaped object.

The in-place lifting DWT implementation has special implications for the SA-DWT[9], which can best be understood visually as shown in Fig. 1. As the SA-DWT is performed, the spatial domain shape mask remains intact with no requirement to derive a shape mask for each subband. How the subbands are arranged in this pseudo-spatial domain arrangement is shown in Fig. 2(a). Each subband can in fact be extracted from the interleaved subband arrangement using the lazy wavelet transform (LWT) [10]. After the one-level SA-DWT is performed, the LL1 subband can be extracted using a coordinate mapping from the interleaved subband coordinates (i,j) to the LL1 subband coordinates (i_{LL1},j_{LL1}) as follows:

$$(i_{LL1}, j_{LL1}) \leftarrow ([i/2], [j/2]) \tag{1}$$

Similarly, the mapping for the HL1 subband is $(i_{HL1}, j_{HL1}) \leftarrow ([i/2]+1,[j/2])$; for the LH1 subband $(i_{LH1}, j_{LH1})\leftarrow([i/2],[j/2]+ 1)$; and for the HH1 subband $(i_{HH1}, j_{HH1}) \leftarrow ([i/2]+1,[j/2]+1)$.

After the first level of the SA-DWT, the interleaved subband arrangement is made up of 2 × 2 basic blocks of coefficients. As shown in the left side of Fig. 2 (b), the top-left coefficient of each block is an LL1 subband coefficient, the top-right coefficient is an HL1 subband coefficient, and so on The second level SA-DWT is performed by first extracting the LL1 subband using the coordinate mapping (1) and then performing the one-level SA-DWT using the LL1 subband as the new input. The output is the four interleaved subbands, LL2, HL2, LH2, and HH2. This is then placed back into the original interleaved subband arrangement where the LL1 coefficients were extracted from. This creates a two-level interleaved subband arrangement. As shown in the middle of Fig.2(b), the two-level interleaved subband arrangement is made of a basic 4×4 coefficient block, with the top-left coefficient of each block being an LL2 coefficient. The coordinate mappings to extract the second and subsequent level subbands are simply derived by applying the one level coordinate mappings iteratively to the LL subband coordinate mapping from the previous level.

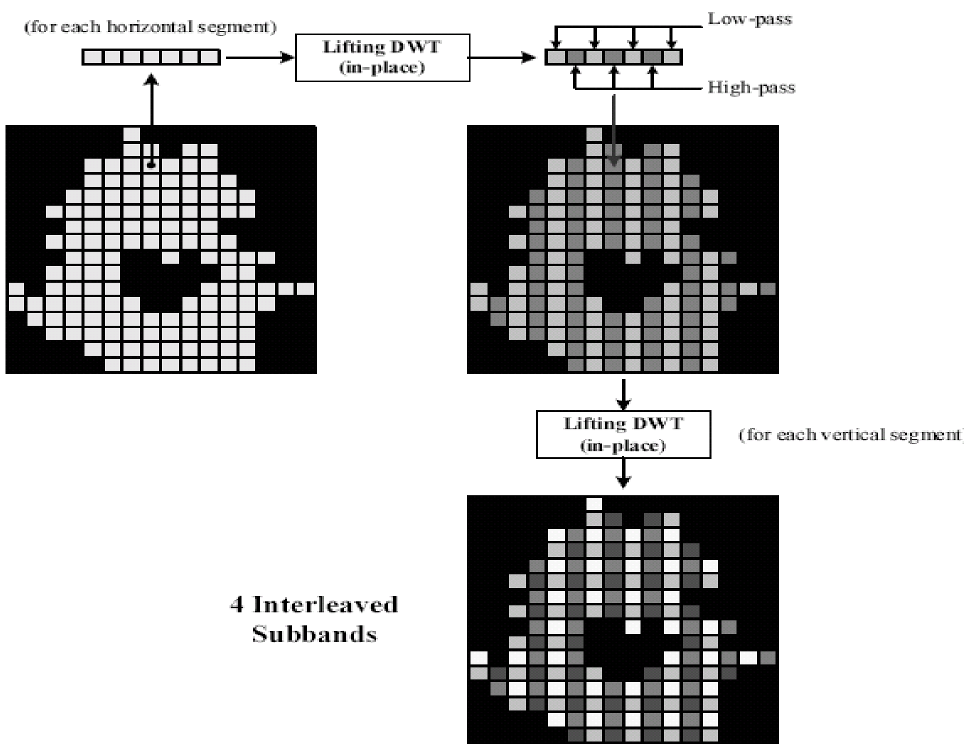


FIGURE 1: One-Level Two-Dimensional SA-DWT Using In-Place Lifting DWT Implementation

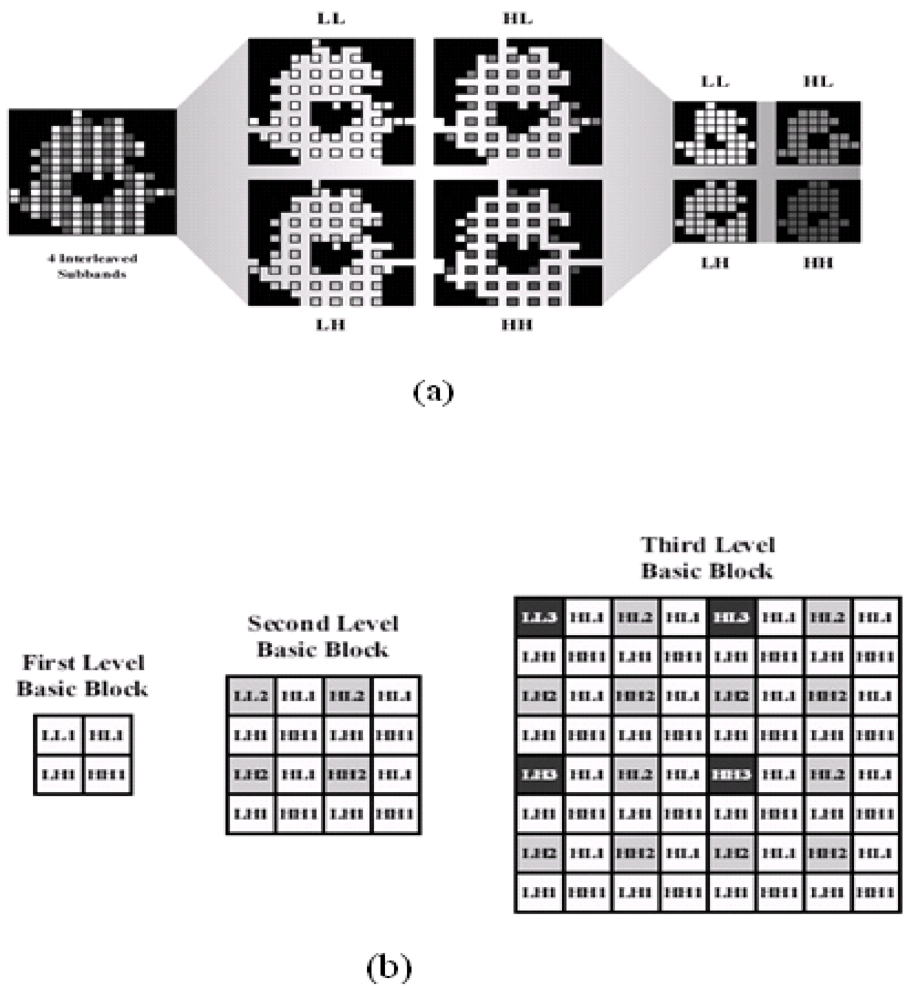


FIGURE 2: (a) Interleaved Subband Abstraction (b) Basic Group of Coefficients for Each Level of In-Place DWT

3. PROPOSED WATERMARKING SCHEME

A content-based watermarking system for content integrity protection is illustrated in Fig. 3.

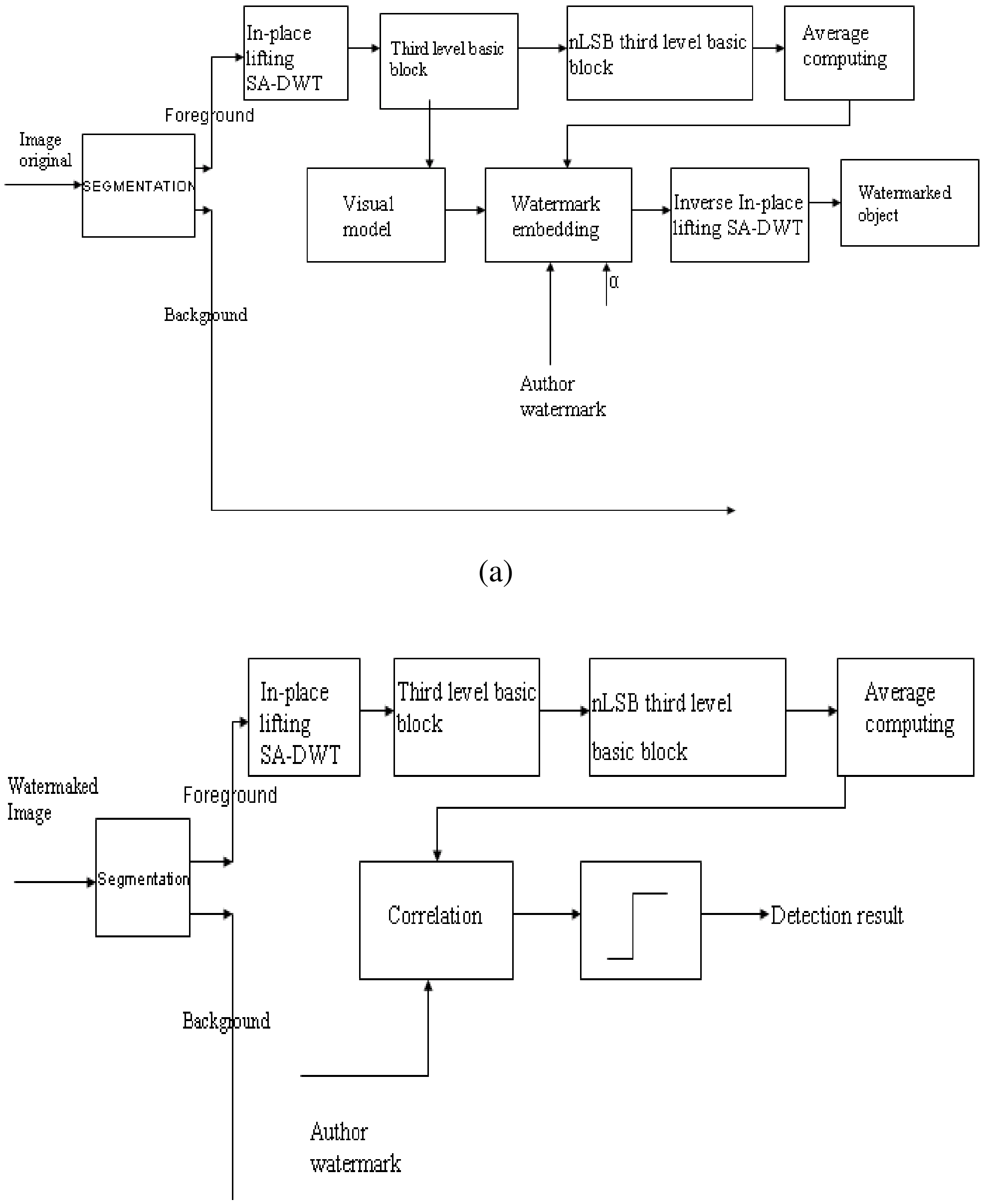


FIGURE 3 : Block Diagrams for The Proposed Watermarking Scheme. (a) Watermark Embedding (b) Watermark Detection.

3.1 Watermark Embedding

Fig.3 (a) shows the watermarking embedding procedure. Firstly, the MPEG-4 coded video bit stream is decoded obtaining a sequence of frames. Each frame is segmented into foreground (objects) and background and we apply the three levels SA-DWT to foreground object frame by frame. Then, we apply the algorithm scheme at each third level basic block (see fig 2(b)). NxN is the size of the matrix wavelet block and $I_i(k)$ is the ith wavelet coefficient in the kth wavelet block where $i \in [1, N\times N]$.

The rest of the watermarking embedding procedure is presented in and resumed in the following. The n LSBs of $I_i(k)$ is defined as :

$$\hat{I}_i(k)=\mathrm{mod}(I_i(k),2^n) \tag{2}$$

The average of the wavelet block is defined as follows:

$$Average\ (k) = \frac{\sum_{i=1}^{N\times N} \hat{I}_i(k)}{N \times N} \tag{3}$$

In the proposed watermarking, we choose the blocks with an average value different from zero. If a few of I_i (k) are changed by Ω due to some distortions, the average of the wavelet block will only have a small change. Assuming that $I'_i(k)$ is the ith wavelet coefficient in the kth wavelet block after the watermark embedding, $\hat{I}_{i'}(k)$ is the n LSBs of $I'_i(k)$ and Average′(k) is the average of $\hat{I}_i'(k)$ in the kth wavelet block accordingly. The watermark W, consisting of a binary pseudo random sequence, $W(k) \in \{-1, 1\}$, is embedded by adjusting the average of wavelet blocks in this way :

$$Average'(k)\in\begin{cases}[0,2^{n-1}), if & W(k)=-1\\ [2^{n-1},2^n), if & W(k)=1\end{cases} \tag{4}$$

To adapt the watermark sequence to the local properties of the wavelet block, we use the model based on HVS in the watermark system. The model is similar to that proposed in [7], but it is developed independently. The visual model takes into account the brightness sensitivity and texture sensitivity of the wavelet block to noise. The visual model function Vm(k) is defined as:

$$Vm(k)=brightness\ (k)x\ texture(k)^{\beta} \tag{5}$$

where
$$texture(k)=\frac{\sum_{i=1}^{N\times N}\left[brightness(k)-I_i(k)\right]^2}{N\times N}$$

$$brightness(k)=\frac{\sum_{i=1}^{N\times N} I_i(k)}{N\times N}$$

β is a parameter used to control the degree of texture sensitivity. This visual model function indicates that the human eye is less sensitive to noise in the highly bright and the highly textured areas of the image. Hence, the wavelet blocks are divided into two parts depending on the value of Vm(k): high activity wavelet block and low activity wavelet block. For simplicity, the threshold Tc is set to the average of Vm(k). The following function can be applied to distinguish high or low activity wavelet block:

$$T(k)=sign(Vm(k)-Tc) \tag{6}$$

Considering the tradeoff between robustness and transparency, the proposed watermark embedding algorithm can be formulated as follows:

$$I'_i(k)=I_i(k)+\alpha W(k)F_i(k)[2^{n-2-S(k)}+T(k)x2^{n-3}] \tag{7}$$

where α is a scaling factor used to control the strength of the inserted watermark. The flag function is defined as follows:

$$F_i(k)=sign((2^{n-1}-I_i(k))xW(k)) \tag{8}$$

where

$$sign(x)=\begin{cases} 1 & if \quad x\geq 0 \\ -1 & if \quad x<0 \end{cases}$$

The strength function is defined as follows:

$$S(k)=sign(X(k)) \tag{9}$$

Where

$$X(k)=(2^{n-1}-Average(k))x\,W(k)$$

Details concerning the flag function and the strength function are described in table 1.

W(k)	$2^{n-1}-\hat{I}_i(k)$	2^{n-1}-Average(k)	$F_i(k)$	S(k)
-1	>0	>0	-1	-1
-1	≤0	≤0	1	1
1	>0	>0	1	1
1	≤0	≤0	-1	-1

TABLE 1: The Detailed Results of $F_i(K)$ and S(K)

In light of the above, the *n* LSBs of wavelet coefficients have been adjusted by using equation (7). Naturally, their average has been updated depending on the requirement of *W(k)* as show in equation (4). In other word, the watermark has been embedded.

3.2 Watermark Extraction and Detection

The watermark sequence can be extracted without the original object. From the process of the watermark embedding, we can obtain the watermarked objects by applying the function of equation (3). Thus, for a given watermarked object, the watermark can be extracted as:

$$W'(k)=\begin{cases} -1, & if \quad Average(k)\in\left[0,2^{n-1}\right) \\ 1, & if \quad Average(k)\in\left[2^{n-1},2^{n}\right) \end{cases} \tag{10}$$

In order to detect the watermark *W'* extracted from the watermarked object, we firstly evaluate the detector response (or similarity of *W'* and *W*) as :

$$\rho(W',W)=\frac{\sum_{k=1}^{L}W'(k)\times W(k)}{\sum_{k=1}^{L}\|W'(k)\|^2}=\frac{\sum_{k=1}^{L}W'(k)\times W(k)}{L} \tag{11}$$

where, L is the length of the watermark signal. The Threshold $T\rho$ is set so as to minimize the sum p of the probability of error detection and the probability of false alarm. If $\rho \geq T\rho$, we considered the watermark is present, otherwise absent.

4. EXPERIMENTS RESULTS

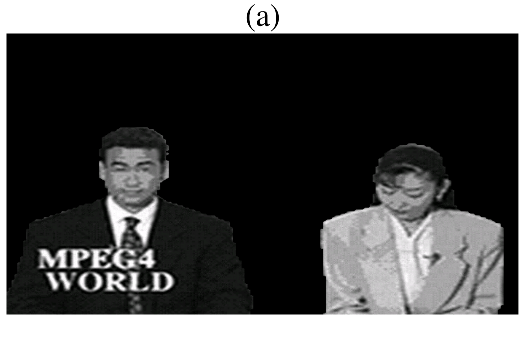
We tested our scheme on a number of video ("akiyo", "news", "sean") as show in Fig.4, we only report in detail results for 'Akiyo'. In our experiments, the parameters considered are : the threshold $T\rho = 0.1$, $\beta = 0.318$, $n = 5$, N=8 ,wavelet-level = 3, wavelettype ='haar', L=1600 and scaling factor $\alpha \in [0.1, 0.5]$.

In order to test the performance of the proposed watermarking scheme, 200 watermarks were randomly generated.
The PSNR result between the original object and the watermarked object is 39.26 dB, As shown in Fig. 5, the watermark is perceptual invisible and the object with watermark appears visually identical to the object without watermark.

In Fig. 6 the absolute difference between the original object and the watermarked one, it is evident that there is no watermark embedded in the region outside the object. Fig. 7 shows the response of the watermark detector to 200 randomly generated watermarks of which only one matches the watermark present. The response to the correct watermark (i.e. number 100) is much higher than the responses to incorrect watermarks.

Added experiment results are described in details in the following and added experiment for other video are listed in table 2.

(a)



(b)

(c)

FIGURE 4: (a)Original Video (Object) Akiyo, (b)Original Video(Object) News, (c)Original Video(Object) Scene.

FIGURE 5: Watermarked Object Akiyo (PSNR=39.26 dB)

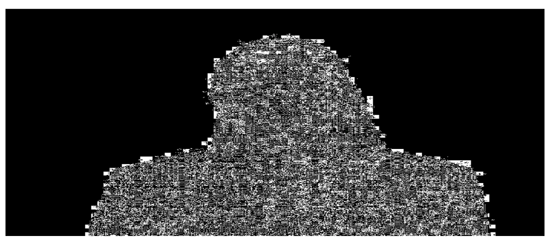


FIGURE 6: Absolute Difference Between the Original Object and The Watermarked

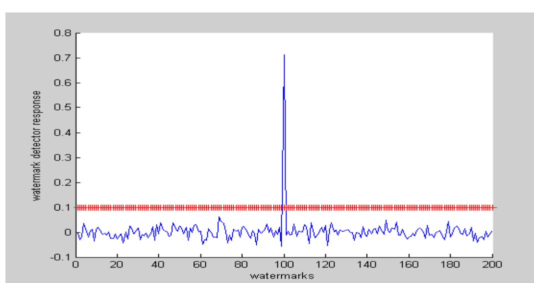


FIGURE 7: Detector Response of the Watermarked Object Akiyo for 200 Randomly Generated Watermark

4.1 MPEG-4 Compression

The watermarks are embedded in video objects, frame by frame, using MPEG-4 video object watermarking scheme, as shown in Fig. 3(a). The MPEG-4 video stream is next decompressed and two different objects are obtained as shown in Figs 8,9and 10, where the watermark detection process is applied, as shown in Fig. 3(b). The watermark detector responses of the decoded foreground objects of akiyo sequence are 0.4735 (foreground) and 0.8834(background), as shown in Fig. 11 and 13. The responses are well above the threshold $T\rho$ and indicate that our proposed watermarking scheme is robust to MPEG-4 compression.

FIGURE 8: A Frame of The Video Sequence ‘Akiyo’ The Video Object 3 ‘Background’.

FIGURE 9: A Frame of The Video Sequence ‘Akiyo’ The Video Object 3 ‘Foreground’.

FIGURE 10 : A Frame of The Video Sequence ‘Akiyo’ The Video Object 3.

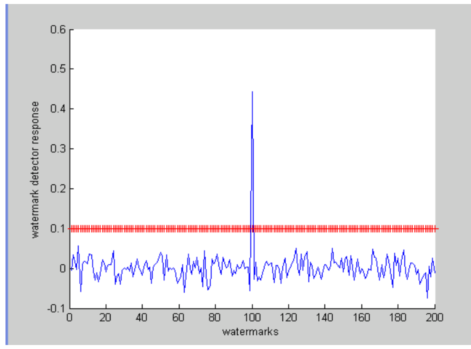


FIGURE 11: Watermark Detection Response Relating To The Video Object 3(Foreground) After MPEG-4 Compression.

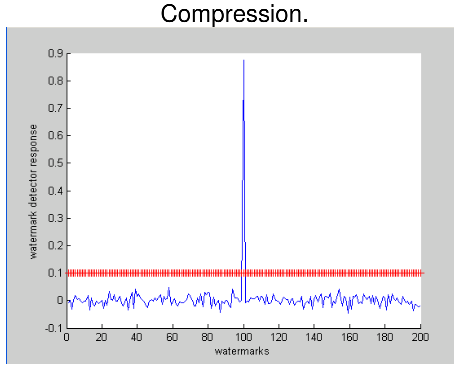


FIGURE 12 : Watermark Detection Response Relating To The Video Object 3(Background) After MPEG-4 Compression.

4.2 Format Conversion from MPEG-4 to MPEG-2

The watermarked MPEG-4 video bitstream is decompressed and frames are obtained. These frames are compressed MPEG-2 coded video bitstream using AVS video converter 6.2 and Easy video converter V.4.2. Next, the MPEG-2 coded video bitstream is decompressed, and each frame is separated so different objects are obtained, where the watermark detection process is applied. As shown in Fig.13 and 14, both watermarks embedded in the two objects are easily detected which indicates that the proposed scheme is robust to conversion from MPEG-4 to MPEG-2.

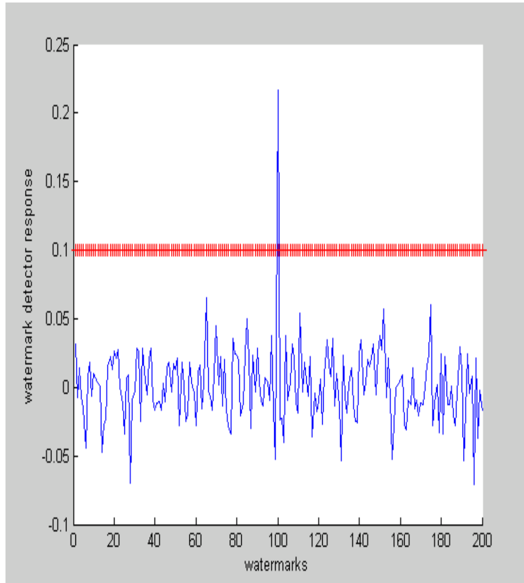


FIGURE 13 : Watermark Detection Response Relating To The Video Object 0 (Foreground) After Format Conversion From MPEG-4 To MPEG-2.

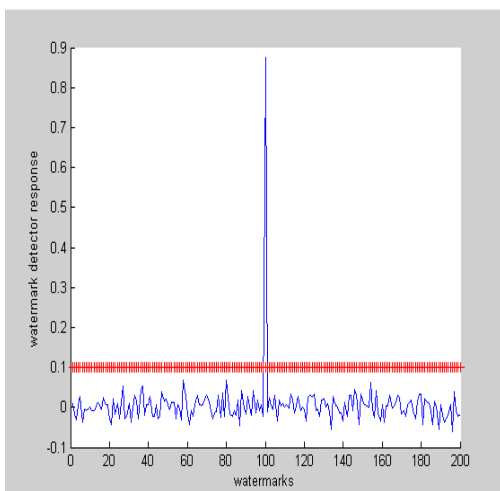


FIGURE 14: Watermark Detection Response Relating To The Video Object 1 (Background) After Format Conversion From MPEG-4 To MPEG-2.

Detector responses	akiyo	News	Sean
No attack(foreground)	0.7101	0.7212	0.7131
MPEG-4 Compression (foreground)	0.4506	0.5535	0.5944
MPEG-4 Compression (background)	0.8723	0.8077	0.8606
MPEG-4 to MPEG-2 (foreground)	0.2376	0.3253	0.3962
MPEG-4 to MPEG-2 (background)	0.8491	0.8652	0.8245
MPEG-1(foreground)	0.4231	0.4179	0.5481
MPEG-1(background)	0.8296	0.8019	0.8022
H264(foreground)	0.4673	0.4798	0.4022
H264(background)	0.8024	0.8266	0.8070

TABLE 2: Watermark Detector Responses After Attacks

5. CONSLUSION

In this article, a novel blind object watermarking scheme for MPEG4 video streams using the SA-DWT has been proposed. To make the watermark robust and transparent, we embed it in the average of the wavelet blocks using visual model. The visual model takes into account sensitivity to brightness and texture. Efficiency of the method is revealed on the basis of the following results:

(1) The average has a smaller change than that of individual coefficient. Thus, unlike most watermarking schemes, the watermark is not embedded by just an individual wavelet coefficient but by modulating the average of the wavelet blocks.

(2) Visual model allowed to achieve the best tradeoff between transparency and robustness.

(3) Watermark detection is accomplished without the original.

(4) Many parameters can be used as private key to that they are unknown to public.

6. REFERENCES

1. MPEG Requirements Group. "*MPEG-4 requirements*". Doc. ISO/IEC JTC1/SC29/WG11 N1595, Sevilla MPEG meeting, February 1997

2. F. Hartung and M. Kutter. "*Multimedia watermarking techniques*". In Proceedings of the IEEE, vol. 87, no. 7,pp. 1079–1107, july 1999.

3. X. Wu, W. Zhu, Z. Xiong and Y. Zhang. "Object-*based multiresolution watermarking of images and video*". In Proceedings of ISCAS'2000, Geneva, Switzerland, pp. 212–215, May 23–31, 2000

4. G.Y. Kim, J. Lee and C.S. Won. "*Object-based video watermarking*". In Proceedings of ICCE'99, pp. 100–101, June 22–24, 1999.

5. A. Piva, R. Caldelli and A.D. Rosa. "*A DWT-based object watermarking system for MPEG-4 video streams*". In Proceedings of ICIP'2000,Vancouver, Canada, September 2000, vol. III, pp. 5–8, 2000.

6. M. Barni, F. Bartolini, V. Capelini, and N. Checacci. "*Object Watermarking for MPEG-4 video streams copyright protection*". In Proceedings of IST/SPIE's : Security and Watermarking of Multimedia Content II, SPIE Proceedings, San Jose, CA, vol 3971, pp. 465-476, 2000

7. K. Xiangwei, L. Yu, L. Huajian and Y. Deli. "*Object watermarks for digital images and video*". Image and Vision Computing, Vol.22, No.8, pp.583-595, Aug.2004.

8. M .Karl, L.Rastislav and, N. P. Konstantinos. " SPIHT-based Coding of the Shape and Texture of Arbitrarily-Shaped Visual Objects". Circuits and Systems for Video Technology, IEEE Transactions on Volume 16, Issue 10, pp:1196 – 1208 Oct. 2006

9. S. Li and W. Li. "*Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding*". In Proceedings of IEEE Trans. Circuits Syst. Video Technol., vol. 10, pp. 725–743, Aug. 2000.

10. W. Sweldens. "*The lifting scheme: A new philosophy in biorthogonal wavelet constructions*". In Wavelet Applications in Signal and Image Processing III, A. F. Laine and M. Unser, Eds. Proc. SPIE 2569, pp. 68–79, 1995.