15th European Signal Processing Conference (EUSIPCO 2007), Poznan, Poland, September 3-7, 2007, copyright by EURASIP

3D-DCT VIDEO WATERMARKING USING QUANTIZATION–BASED METHODS

Patrizio Campisi and Alessandro Neri

Dip. Elettronica Applicata
Università degli Studi Roma TRE
Via della Vasca Navale 84, 00146 Roma, Italy
phone: +39.06.55177064, fax: +39.06.55177026
e-mail: (campisi, neri)@uniroma3.it

ABSTRACT

Video watermarking methods operating in the three-dimensional discrete cosine transform (3D-DCT) domain using quantization–based embedding techniques are here presented. Specifically, after having partitioned a video sequence into spatio-temporal units, they are projected onto the 3D-DCT domain. A set of transformed coefficients are taken according to a selection rule, ordered according to a predefined scan order, and eventually marked. Two embedding methods are here considered: Quantization Index Modulation watermarking (QIM) and Rational Dither Modulation (RDM). A perceptual criterion is also used to define the “local” quantization steps to be used in QIM. Finally, the inverse 3D inverse DCT (IDCT) is performed thus obtaining the marked video shot. Robustness of the proposed method against a variety of both intentional and unintentional video attacks has been observed through experimental results.

1. INTRODUCTION

In this paper we deal with the problem of copyright protection of digital video sequences by using digital watermarking. This issue is rising more and more concerns because of the development of the digital multimedia market and of the always decreasing price of storage devices and digital supports. Extensive surveys dealing with video watermarking applications, attacks and methodologies can be found in [1], [2].

The video temporal dimension is not exploited in the first developed video watermarking methods, where a video is considered as a succession of digital still images. Therefore existing schemes employed for still image watermarking have been employed by watermarking each frame independently of the others. A real time watermarking algorithm robust to video coding, D/A and A/D conversion, format conversion and limited editing, using a bidimensional spread spectrum approach operating in the spatial domain is presented in [3]. A spread spectrum approach to embed digital watermarks in both the uncompressed domain and the compressed MPEG-2 domain is presented in [4]. An attempt to exploit the static and dynamic temporal components of a video has been performed in [5] where a wavelet transform is applied along the temporal axis of the video to design a robust watermarking system. In order to make the embedding of the mark imperceptible, methods exploiting perceptual information have been presented in the recent past (see for example [6], [7]). In [7] a perceptual watermarking method, operating in the 2D-DCT domain, for embedding the mark into MPEG-2 compressed video streams, has been presented. However, in order to design video watermarking methods robust to both non-hostile and hostile video processing operations such as video format conversion, temporal desynchronization, video editing, and video collusion frame-by-frame watermarking strategy are not appropriate. In [8] collusion issues are discussed when frame-by-frame embedding strategies are enforced for video watermarking. In the recent past, to overcome the drawbacks related to frame-by-frame watermarking methods, approaches operating in three dimensional transform domains have been proposed. A watermarking approach operating in the 3D discrete Fourier transform domain has been proposed in [9]. A watermarking method operating in the wavelet domain and making use of BCH codes and 3D interliving is presented in [10]. In [11] a watermarking method operating in the 3D-DWT domain, by using pseudo random linear statistics of pseudo random connected regions in the DC subband of the 3D wavelet domain is described. In [12] a perceptual based video watermarking technique, operating in the 3D discrete wavelet transform (3D-DWT) domain, that jointly exploits both the spatial and the temporal dimension of a video sequence is described. Quantization based methods are used in [13] to mark 3D subbands in the 3D-DWT domain. In [14] a video watermarking technique using the 3D discrete cosine transform has been described.

Attempts to account for human visual attention in order to better localize video’s regions where to transparently embed the mark have been recently used to design saliency-based watermarking techniques. In [15] the payload is embedded in those regions that receive less attention by a viewer, so that the perception of the scene is not altered. Among the factors impacting the human visual attention there are: contrast, color, object size and shape, motion, and many more. Although several factors have been identified, their relative importance is not yet determined. However, in many studies, the temporal features are considered as characterizing the importance of a region in terms of human perception of quality [16]. For instance, foreground moving objects are prone to attract viewers’ attention making impairments in the background more difficult to detect. At the same time, when objects move faster with respect to the average motion of the scene, impairments in the object would have minor impact on the subject quality assessment. This would allow to increase the watermark strength in “low”-motion and “high”-motion areas, while reducing it for regions with “medium” temporal activity.

In this paper we propose a novel video watermarking technique operating in the 3D-DCT domain, thus jointly exploiting both the spatial and the temporal dimension of a video sequence. Specifically, the video is projected onto the 3D-DCT domain, and the coefficients selection is performed by taking into account the energy compaction property of the DCT transform. Eventually, the selected coefficients set is marked by using two quantization-based approaches: the quantization index modulation (QIM) watermarking [17] and the rational dither modulation (RDM) [18]. The quantization-based approaches, which are gaining always increasing consensus with respect to other approaches such that spread spectrum based watermarking methods, consists in quantizing the selected coefficients by means of a quantizer chosen in a set on the basis of the coefficient to embed. We propose to perceptually design the quantization steps by using a qualitative Human Visual System (HVS) model. After a brief review of the 3D-DCT and of its inverse transform given in Section 2, the proposed watermarking methods are detailed in Section 3. Experimental results and conclusions are drawn in Sections 4.

2. 3D-DCT

Given a 3D real valued function $f[n_1,n_2,n_3]$, being n_1,n_2, and n_3 integers assuming values in the set $0 \le n_1 < N_1, 0 \le n_2 < N_2$, and $0 \le n_3 < N_3$, its forward 3D-DCT [20], $F[k_1,k_2,k_3]$, is given by:

$$F[k_1,k_2,k_3] = \alpha(k_1,k_2,k_3) \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \sum_{n_3=0}^{N_3-1} f[n_1,n_2,n_3] \cdot \\ \cdot \cos\frac{(2n_1+1)k_1\pi}{2N_1} \cos\frac{(2n_2+1)k_2\pi}{2N_2} \cos\frac{(2n_3+1)k_3\pi}{2N_3} \tag{1}$$

being

$$\alpha(k_1,k_2,k_3) = \sqrt{\frac{8}{N_1N_2N_3}} C(k_1)C(k_2)C(k_3) \tag{2}$$

and

$$C(k_i) = \begin{cases} \frac{1}{\sqrt{2}} & k_i = 0 \\ 1 & otherwise \end{cases} \tag{3}$$

with $0 \le k_i < N_i$, being $i = 1,2,3$.

The 3D IDCT [20], is given by:

$$f[n_1,n_2,n_3] = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \sum_{k_3=0}^{N_3-1} \alpha(k_1,k_2,k_3) F[k_1,k_2,k_3] \cdot \\ \cdot \cos\frac{(2n_1+1)k_1\pi}{2N_1} \cos\frac{(2n_2+1)k_2\pi}{2N_2} \cos\frac{(2n_3+1)k_3\pi}{2N_3} \tag{4}$$

where $\alpha(k_1,k_2,k_3)$ and $C(k_i)$ are given by (2) and (3) respectively.

It is worth pointing out that the 3D-DCT (3D-IDCT) is a separable transform, which implies that it can be implemented as a series of 1D-DCT (1D-IDCT). Therefore its computation can be performed by means of the fast algorithms which have been devised for computing the 1D-DCT (1D-IDCT). The 3D-DCT (3D-IDCT) computational complexity and computing time can be further reduced when fast 3D-DCT (3D-IDCT) algorithms are employed as discussed in [21].

3. THE PROPOSED WATERMARKING METHOD

The video sequence is first segmented into non-overlapping spatio-temporal plot units, called *shots*, that depict different actions. Each shot is characterized by no significant changes in its content which is determined by the background and the objects present in the scene. Although several video segmentation algorithms which detect the abrupt scene changes [23], [24], [25], have been proposed in literature, for the sake of simplicity, we have segmented the video sequence into shots of fixed length. Then each shot is projected onto the 3D-DCT domain and the coefficients where to embed the mark are selected and ordered according to a specific criterion. The selected coefficients are marked using either QIM or RDM, and eventually the 3D-IDCT is performed, thus obtaining the marked shot. Specifically, let us consider the m-th video sequence shot, $f_m[n_1,n_2,n_3]$, with $0 \le n_1 < N_1, 0 \le n_2 < N_2$, and $0 \le n_3 < N_3$, composed by N_3 frames of dimension $N_1 \times N_2$ pixels. Let us indicate with $F_m[k_1,k_2,k_3]$ the 3D-DCT of $f_m[n_1,n_2,n_3]$ obtained using (1).

3.1 Host coefficients selection

The host coefficients where to embed the mark must be chosen in such a way to guarantee both the robustness and the imperceptibility of the mark. Therefore, the low frequency coefficients, which usually present the higher energy, are left unaltered to guarantee transparency. To achieve robustness, similar rule applies to the high frequency coefficients, which are usually drastically quantized by video encoders due to their low energetic content. Following this criterion, the watermark is embedded in the DCT coefficients that correspond to the middle spatial and temporal frequencies. Specifically, the coefficients choice is performed by selecting those coefficients, in the 3D-DCT volume, whose indices (k_1,k_2,k_3) belong to the region R limited by the two hyperboloids $(k_1+1)(k_2+1)(k_3+1) = C_l$ and $(k_1+1)(k_2+1)(k_3+1) = C_u$, being $C_l < C_u$ two properly chosen constants.

Once the embedding region has been selected, its coefficients are ordered in a vector $\mathbf{V}_R = [v_1, v_2, \cdots, v_L]$, according to a predefined scan order, which we assume known at the receiving side. In our approach the coefficients scan is made by picking the coefficients, belonging to the region R, first along the direction k_1, then along the direction k_2, and finally along the direction k_3. This scan order is determined by the observation that in most cases the coefficients are spread over the smaller values of k_3.

3.2 Perceptual QIM watermark embedding

The mark embedding, based on the use of the QIM approach, relies on the quantization of the selected host coefficients by means of a quantizer chosen among a set of quantizers according to the data to be embedded. However, to minimize the distortion perception produced by the watermarking, the quantizers can be designed taking into account the spatio-temporal properties of the HVS. Specifically, the HVS has low contrast sensitivity for high horizontal (k_1), vertical (k_2), and temporal (k_3) frequencies, and even lower sensitivity for high joint frequencies ($k_1-k_2, k_1-k_3, k_2-k_3, k_1-k_2-k_3$). Therefore the quantization step should increase with an increase of (k_1), (k_2), and (k_3), and increase even more for

$(k_1-k_2, k_1-k_3, k_2-k_3, k_1-k_2-k_3)$ [19]. These qualitative considerations on the HVS allow us properly choosing the quantization step as a function of the coefficient index (k_1,k_2,k_3) as follows:

$$\Delta(\vec{k}) = \Delta(k_1,k_2,k_3) = \lfloor \frac{\Delta}{4}(1+k_1^p+k_2^p+k_3^p)\rfloor. \tag{5}$$

Then, given the coefficients to be marked, ordered in the vector $\mathbf{V}_R = [v_1, v_2, \cdots, v_L]$, the corresponding quantization steps, defined according to (5) are collected in the vector $\mathbf{\Delta}_R = [\Delta_1, \Delta_2, \cdots, \Delta_L]$.

Given a binary representation $\mathbf{b} = \{b_i, 1 \le i \le L\}$ of the watermark, and the host coefficients $\mathbf{V}_R = \{v_i, 1 \le i \le L\}$ the perceptual QIM system we define consists of quantizing v_i by letting b_i choose between one of two uniform quantizers $\mathcal{U}_0$ and $\mathcal{U}_1$ defined as:

$$\begin{aligned} \mathcal{U}_0 &= \{u = l\Delta_i + d \mid l \in Z\} \\ \mathcal{U}_1 &= \{u = l\Delta_i + \Delta_i/2 + d \mid l \in Z\} \end{aligned} \tag{6}$$

being $\Delta_i \in \mathbf{\Delta_R}$ the quantization step and $d \in [0,1)$ a key used to improve security. It is straightforward to note that this method can be expressed in terms of additive watermarking. In fact, by posing

$$q_i = \mathcal{Q}_{\Delta_i}\left[v_i - \Delta_i\left(\frac{b_i}{2}+d\right)\right] - \left[v_i - \Delta_i\left(\frac{b_i}{2}+d\right)\right], \tag{7}$$

being $\mathcal{Q}_{\Delta_i}$ a uniform quantizer of step Δ_i, the watermarked data can be expressed as

$$v_i^w = v_i + q_i. \tag{8}$$

The so obtained marked coefficients, are first grouped in the vector $\mathbf{V}_R^w = [v_1^w, v_2^w, \cdots, v_L^w]$, then inserted back in the selected region R in the DCT volume according to the defined scan order, thus obtaining the marked 3D-DCT volume $F_m^w[k_1,k_2,k_3]$. Eventually, the 3D-IDCT is performed according to (4) thus obtaining the marked m-th video sequence shot $f_m^w[n_1,n_2,n_3]$.

3.3 QIM Watermark extraction

At the receiving side, the m-th shot of the received marked video sequence, $\widehat{f}_m^w[n_1,n_2,n_3]$, is 3D-DCT transformed. The region R, where the mark has been embedded, is then selected according to the aforementioned criterion. The coefficients belonging to the region R are then extracted and ordered in the vector $\widehat{\mathbf{V}}_R^w = [\hat{v}_1^w, \hat{v}_2^w, \cdots, \hat{v}_L^w]$. Given the generic element $\hat{v}_i^w$ belonging to $\widehat{\mathbf{V}}_R^w$, and the quantization steps vector $\mathbf{\Delta}_R = [\Delta_1, \Delta_2, \cdots, \Delta_L]$, the decoded value $\hat{b}_i$ is obtained by first extracting y_i according to the following rule:

$$y_i = \mathcal{Q}_{\Delta_i}[\hat{v}_i^w - d\Delta_i] - [\hat{v}_i^w - d\Delta_i], \tag{9}$$

and then by using hard-decision decoding. Because of the adopted embedding rule (7), the values y_i should be close to zero when the bit b_i to be embedded is equal to 0 and close to $\pm\Delta_i/2$ when $b_i = 1$. Therefore, the decoded value $\hat{b}_i$ can be obtained according to the following decision rule:

$$\hat{b}_i = \begin{cases} 0 & |y_i| < \Delta_i/4 \\ 1 & |y_i| \ge \Delta_i/4. \end{cases} \tag{10}$$

3.4 RDM watermark embedding

However, as well known QIM based watermarking schemes are largely vulnerable to amplitude scalings. Therefore, in [18] a variation of the QIM embedding approach robust to amplitude scalings while still retaining the simplicity of the QIM method has been presented. Specifically, given the i-th information symbol, $b_i \in \{-1,1\}$, and the host coefficients vector $\mathbf{V}_R = [v_1, v_2, \cdots, v_L]$, the embedding rule is expressed as follows:

$$v_i^w = g(\mathbf{V}_{i-1})\mathcal{Q}_{b_i}\left(\frac{v_i}{g(\mathbf{V}_{i-1})}\right) \tag{11}$$

where $\mathcal{Q}_{b_i}$ represents the quantizers induced by the lattices:

$$\begin{aligned} \mathcal{U}_{\text{-}1} &= \{u = 2l\Delta - \Delta/2 \mid l \in Z\} \\ \mathcal{U}_1 &= \{u = 2l\Delta + \Delta/2 \mid l \in Z\}. \end{aligned} \tag{12}$$

The vector $\mathbf{V}_{i-1}$ contains the past W samples of vector $\mathbf{V}$ taken at instant $i-1$ that is $\mathbf{V}_{i-1} = \{v_{i-1}, v_{i-2}, \cdots, v_{i-W}\}$. The function $g(\cdot)$ is chosen within the functions such that for any $\rho > 0$, $g(\rho\mathbf{V}_i) = \rho g(\mathbf{V}_i)$. Some functions satisfying this constraint are the l_p vector-norm:

$$g(\mathbf{V}_i) = \left(\frac{1}{W}\sum_{m=k-W}^{k-1} |\, v_m \,|^p\right)^{1/p}. \tag{13}$$

The so obtained marked coefficients, are first collected in the vector $\mathbf{V}_R^w = [v_1^w, v_2^w, \cdots, v_L^w]$, then inserted back in the DCT volume according to the defined scan order. Eventually, the 3D-IDCT is performed thus obtaining the marked m-th video sequence shot $f_m^w[n_1,n_2,n_3]$.

3.5 RDM watermark extraction

At the receiving side, the region R where the mark has been embedded is selected after having 3D-DCT transformed the m-th shot of the received marked video sequence $\widehat{f}_m^w[n_1,n_2,n_3]$. Let us indicate with $\widehat{\mathbf{V}}_R^w = [\hat{v}_1^w, \hat{v}_2^w, \cdots, \hat{v}_L^w]$ the coefficients belonging to the marked region R.

The decoding is performed using a minimum Euclidean distance rule:

$$\hat{v}_i = \underset{-1,1}{\operatorname{argmin}}\left|\left(\frac{v_i}{g(\widehat{\mathbf{V}}_{i-1}^w)}\right) - \mathcal{Q}_{b_i}\left(\frac{v_i}{g(\widehat{V}_{i-1}^w)}\right)\right|^2. \tag{14}$$

4. EXPERIMENTAL RESULTS AND CONCLUSIONS

Experimentations have been performed on uncompressed videos composed by $N_3 = 384$ frames each of size $(N_1 \times N_2) = (352 \times 288)$ pixels (CIF format). The employed test sequences comprise news programs, computer generated videos, musical videos, as well as sport video clips. Each video has been partitioned into shots having fixed length equal to 20 frames. The region R limited by the two hyperboloids, where to embed the watermark, is specified by considering $C_l = 300$ and $C_u = 400$, thus selecting a middle-frequency region in the 3D-DCT domain. A watermark composed by 600 bits, drawn from a uniform distribution,

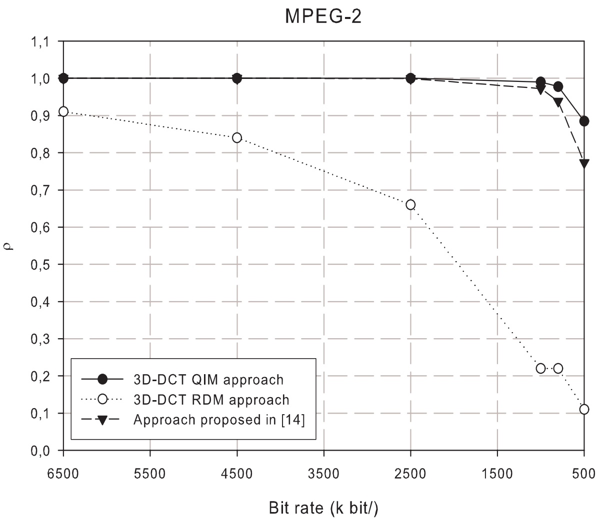

Figure 1: Performance of both the proposed watermarking methods (3D-DCT QIM and 3D-DCT RDM) in comparison with the method proposed in [14] when the video is MPEG-2 compressed.

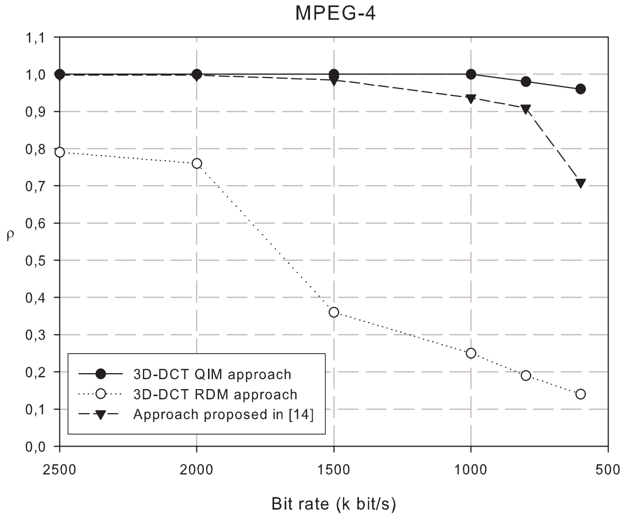

Figure 2: Performance of both the 3D-DCT QIM and 3D-DCT RDM proposed methods in comparison with the method proposed in [14] when the video is MPEG-4 compressed.

is embedded in the selected region by using the approach described in Section 3.

When the perceptual QIM embedding approach is used, the embedding is performed by means of (7)-(8) using a value for Δ in (5) equal to 185. The values of both the constants C_l, C_u and the parameter Δ have been empirically chosen, on a wide range of videos, in order to trade off between the imperceptibility of the mark and its robustness against most intentional attacks.

The imperceptibility of the mark has been tested by using the Video Quality Metric (VQM) Software [26] used to measure the perceptual effects on the video of impairment such as blurring, block distortion, unnatural motion, noise and error blocks. Specifically, the general model metric, denoted as VQM$_G$, which can assume values comprised in the interval $[0,1]$, has been used. Higher values of the metric correspond to higher distorted scenes. For the values of the parameters C_l, C_u, Δ used in the experimentations a value VQM$_G = 0.2$ has been observed (see Table 1). This value guarantees the imperceptibility of the mark as can be verified by direct observation of the marked video sequence. From Table 1 it is evident how the video quality degrades when both the embedding region became closer to the origin of the 3D-DCT domain, thus involving lower frequency coefficients, and when the quantization step increases.

C_l	C_u	Δ	VQM_G
200	300	285	0.30
300	400	185	0.20
700	800	135	0.15

Table 1: VQM$_G$ for the QIM watermarked video.

When considering the RDM watermark embedding the same region R used in the QIM apporach has been considered ($C_l = 300$ and $C_u = 400$). Moreover the quantization step $\Delta = 0.38$ together with the chosen value for (C_l and C_u). guarantee a VQM_G equal to 0.2, thus obtaining the same perceptual transparency obtained by using the QIM approach.

The robustness of the proposed methods has been tested using a variety of unintentional and intentional video attacks such as MPEG-2 and MPEG-4 compression, collusion, transcoding, and frame dropping. As a figure of merit, the ratio ρ between the number of correctly decoded watermark bits and their total number is computed. For the sake of comparison, the method described in [14], operating in the 3D-DCT domain, has been implemented and the experimental results have been here reported. In Figs.1, 2 the performance of both the proposed methods and the one presented in [14] in terms of the ratio ρ versus the bit rate of the marked video sequence compressed using both MPEG-4 and MPEG-2 coder have been reported. From Figs.1, 2 we observe that both the presented approaches are more robust both against MPEG-2 and MPEG-4 compression than the one proposed in [14]. Type I and type II collusion video attacks have also been considered. As for the first attack a subsequence obtained by picking one frame belonging to the marked video sequence every ten is generated. Then, an estimation of the mark is obtained by averaging the subsequence's frames. Type II collusion is performed by averaging two consecutive frames in a subsequence of ten consecutive frames drawn from the marked video. Then the watermark is estimated by applying the decoding procedure to the so obtained nine frames sequence. The experimental results in Table 2 point out that the proposed method is robust against the collusion attacks. This is explained by the consideration that the watermark is "spread" along the temporal axis since the embedding is performed in the 3D-DCT domain. Moreover, this approach exhibits good performance also with respect to transcoding attacks and frame dropping. The methods are sensitive to temporal desynchronization attacks. Robustness to the gain attack is obtained by using the RDM embedding based approach.

ATTACK	3D-DCT QIM	3D-DCT RDM	Approach in [14]
Type I Collusion	0.89	0.87	0.79
Type II Collusion	0.90	0.88	0.82
MPEG-2 (2500 Kb/s) → MPEG-4 (600 Kb/s)	0.92	0.83	0.12
MPEG-2 (1000 Kb/s) → MPEG-4 (600 Kb/s)	0.90	0.8	0.10
Drop 1 frame every 5	0.54	0.46	0.16

Table 2: Performance of the proposed video watermarking schemes (3D-DCT QIM and 3D-DCT RDM) for different video attacks.

REFERENCES

[1] G. Doërr, J.-L. Dugelay "A guided tour to video watermarking," *Signal Processing: Image Communication* 18, pp.263-282, 2003.

[2] D. Kundur, K. Su, and D. Hatzinakos, "Digital Video Watermarking: Techniques, Technology and Trends," in **Intelligent Watermarking Techniques,** Chapter 10, P. J.-S. Pan, H.-C. Huang and L. Jain, eds., World Scientific Publishing Company, pp. 265-314, 2004.

[3] T. Kalker, G. Depovere, J. Haitsma, and M. Maes, "A video watermarking system for broadcast monitoring," *Proc. SPIE* Vol. 3657, pp.103-112, January 1999.

[4] I. Cox, J. Kilian, F. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Transactions on Image Processing* Vol. 6, No. 12, pp.1673-1687, Dec. 1997.

[5] M. Swanson, B. Zhu, and A. T. Tewfik, "Multiresolution scene-based video watermarking using perceptual models," *IEEE Journal on Sel. Areas in Comm.* Vol. 16, No. 4, pp.540-550, 1998.

[6] C. De Vleeschouwer, J.-F. Delaigle, B. Macq, " Invisibility and application functionalities in perceptual watermarking an overview," *Proc. of the IEEE* Vol. 90, No.1, pp. 64-77, Jan. 2002.

[7] R. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," *Proc. of the IEEE* Vol. 87, No.7, pp. 1108-1126, July 1999.

[8] G. Doërr, J.-L. Dugelay "Security pitfalls of frame-by-frame approaches to video watermarking," *IEEE Transactions on Signal Processing* Vol. 52, No. 10, Oct. 2004, pp. 2955 - 2964.

[9] F. Deguillaume, G. Csurka, J. O'Ruanaidh, T. Pun, "Robust 3D DFT video watermarking," *Proc. SPIE, Security and Watermarking of Multimedia Content II*, Vol. 3971, pp. 346-357, 2000.

[10] H. Liu, N. Chen, J. Huang, X. Haung, Y. Q. Shi, "A robust DWT-based video watermarking algorithm," *IEEE International Symposium on Circuits and Systems,* pp.631-634, 2002.

[11] M. Kucukgoz1, Ö. Harmanci, M. K. Mihçak and R. Venkatesan, "Robust Video Watermarking via Optimization Algorithm for Quantization of Pseudo-Random Semi-Global Statistics" *Proc. SPIE, Security, Steganography, and Watermarking of Mult. Cont. VII*, Vol. 5681, 2005.

[12] P. Campisi, A. Neri, "Video watermarking in the 3D-DWT domain using perceptual masking," *IEEE Int. Conference on Image Processing 2005* Sept. 2005, Genoa, Italy.

[13] P. Campisi, "Video watermarking in the 3D-DWT domain using quantization-based methods," *IEEE International Workshop on Multimedia Signal Processing 2005 (MMSP 2005)*, October 30-November 2, 2005, Shanghai, China.

[14] J.H. Lim, D.J. Kim, H.T. Kim, C.S. Won "Digital video watermarking using 3D-DCT and Intra-Cubic Correlation," *Proc. SPIE, Security and Watermarking Contents III* Vol. 4314 pp.54-72, 2001.

[15] A. Neri, P. Campisi, M. Carli, E. Drelie Gelasca, "Watermarking hiding in video sequences," *First International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM 2005*, January 2005, Scottsdale, AZ, USA.

[16] O. Le Meur, D. Thoreau, P. Le Callet, D. Barba, "A Spatio-Temporal Model of the Selective Human Visual Attention," *IEEE Int. Conference on Image Processing 2005* Sept. 2005, Genoa, Italy.

[17] B. Chen, G. W. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory,* Vol. 47, No. 4, pp. 1423-1443, May 2001.

[18] F. Perez-Gonzalez, C. Mosquera, M. Barni, and A. Abrardo, "Rational dither modulation: a high-rate data-hiding method invariant to gain attacks," *IEEE Transactions on Signal Processing,* Vol. 53, No. 10, pp. 3960-3975, Oct. 2005.

[19] N. Bozinovic and J. Konrad, "Scan order and quantization for 3D-DCT coding," *Proc. of the SPIE, Symp. on Image and Video Comm. and Proc.,* July 2003.

[20] K. Rao, P. Yip, **Discrete cosine transform: algorithms, advantages, applications.** Acad. Press, 1990.

[21] S. Boussakta and H. O. Alshibami, "Fast Algorithm for the 3D DCT-II" *IEEE Transactions on Signal Processing,* Vol. 52, No. 4, pp. 992-1001, April 2004.

[22] R.K.W. Chan, M.C. Lee, "3D-DCT quantization as a compression technique for video sequences", *Proc. of the International Conference on Virtual Systems and Multimedia*, pp. 188-196, September 1997.

[23] P. Campisi, A. Neri, "Synthetic summaries of video sequences using a multiresolution based key frame selection technique in a perceptually uniform color space", *Proc. of IEEE Inter. Conf. on Image Proc. 2000*, Vol. 2, pp. 299-302, Sept. 2000.

[24] N. Vasconcelos, and A. Lippman, "Statistical models of video structure for content analysis and characterization", *IEEE Transactions on Image Processing,* Vol. 9, No.1, pp. 3-19, Jan. 2000.

[25] Liu Lijie, Fan Guoliang, "Combined key-frame extraction and object-based video segmentation," *IEEE Trans. on Circuits and Systems for Video Technology,* Vol. 15, No. 7, pp.869-884 July 2005.

[26] Video Quality Metric (VQM) Software Download, v. 2.0, *http://www.its.bldrdoc.gov/n3/video/.*