MVA2009 IAPR Conference on Machine Vision Applications, May 20-22, 2009, Yokohama, JAPAN

VIDEO WATERMARKING ALGORITHM BASED ON PSEUDO 3D DCT AND QUANTIZATION INDEX MODULATION

Hui-Yu Huang
National Formosa University, Taiwan
hyhuang@nfu.edu.tw

Cheng-Han Yang, Wen-Hsing Hsu
National Tsing Hua University, Taiwan

ABSTRACT

In this paper, we propose an effective watermarking algorithm based on a pseudo 3D DCT and quantization index modulation for vide against the attacks. The watermark is principally inserted into the uncompressed domain by adjusting the correlation between DCT coefficients of the selected blocks, and the extraction of watermark is blind. This approach includes a pseudo 3D DCT, watermark embedding, and extraction. A pseudo 3D DCT obtained by taking twice DCT transformations will be firstly utilized to calculate the embedding messages. Using the quantization index modulation, we insert the watermark into the quantization regions from the successive frames and record the relative information to create a secret embedding key. This secret embedding key will further use to the extraction procedure. Experimental results demonstrate that our proposed method can gain a good performance in transparency and robustness against filtering, compression, and noise attacks.

1. INTRODUCTION

Owing to network technology rapidly advance, humans can arbitrarily and easily access or distribute any multimedia data from networks. Hence, the protection of intellectual property becomes more and more attentive and important for the society, Based on this scheme, many methods are developed [1–4]. Digital watermarking is a favorable method for copyright protection of the multimedia. It is a digital code embedded in the host data and typically contains information about origin, status, and/or destination of the data. Applications of watermarking technique include copyright protection, fingerprinting, authentication, copy control, tamper detection, and data hiding applications such as broadcast monitoring [4].

Many watermarking techniques have been proposed which was worked in the spatial domain [5] and frequency domain [1, 6, 7], etc. Lancini *et al.* [5] proposed a video watermarking technique in the spatial domain. In this approach, an important notion is mentioned: the compression algorithm will strongly decrease the chrominance quality. Kong *et al.* [8] proposed a video watermarking based on Singular Value Decomposition. The watermarks in this method are embedded in specifically selected singular values for the luminance channel of the video frames. Many of techniques based on the frequency domain contain the discrete cosine transform (DCT), discrete fourier transform (DFT), discrete wavelet transform (DWT), quantization index modulation (QIM) [9–11]. Li and Cox [10] proposed a watermarking system based on Watson's perceptual model to select the quantization step in the QIM method. The Watson's model can modify the quantization scale and provide a QIM algorithm that is invariant to valumetric scaling and further to improve fidelity. Thiemert *et al.* [12] designed a block based video watermark system which is robust against several image processing operations. The authors worked on quantized DCT blocks with size of 8×8 for luminance channel in MEPG1/2 compressed videos. This scheme enforces the relationships between block averages in groups of blocks to represent the embedded binary message, and chosen coefficients into each block to represent the message redundantly. In this paper, we propose an video watermark system based on the DCT domain to achieve various attacks and copyright protection. In order to avoid the distortion of the chrominance quality of video data, we mainly focus on the luminance component to perform our embedded system.

The rest of the paper is organized as follows. In Section 2, we describe our proposed method. The performance evaluation will be presented in Section 3. Section 4 presents the experimental results. Finally, Section 5 gives the brief conclusions.

2. PROPOSED METHOD

Our proposed system is based on DCT domain. Details of the whole method are described in the following.

2.1. Pseudo 3D DCT transformation

For video data, we take several successive frames as a group. Each frame within a group will be divided into a number of blocks which will be transformed into DCT domain by pseudo 3D DCT method. By means of pseudo 3D DCT method, our approach can reduce the computational complexity. First, we take four consecutive frames as a group, and every frame within a group is divided into some of blocks. Next, the DC value of each block located in the same position of successive frames for a group is transformed into the DCT domain again. After transforming the second DCT process, we will obtain a new DC value and several AC values. This procedure is called a pseudo 3D DCT. The pseudo 3D DCT diagram is shown in Fig. 1. And the sum of all absolute AC values with weights is expressed as

$$Sum(i,k) = \sum_{l} W_s(i,k,l) \times |AC(i,k,l)|, \qquad (1)$$

where $Sum(i,k)$, $W_s(i,k,l)$, and $AC(i,k,l)$ denote the sum of all AC values, the corresponding weight value, and the lth AC value corresponding to the kth blocks of successive frames within the ith group, respectively. Here, the initial wight value can be decided by user. Next, we arrange these sums and then achieve the embedding process.

For example, we take four frames as a group, and each of frames will be separated into the 8×8 size of blocks, and further the block will be firstly transformed to DCT domain by DCT method. Then we will pick the DC values of every block which locate the same position on the frames and

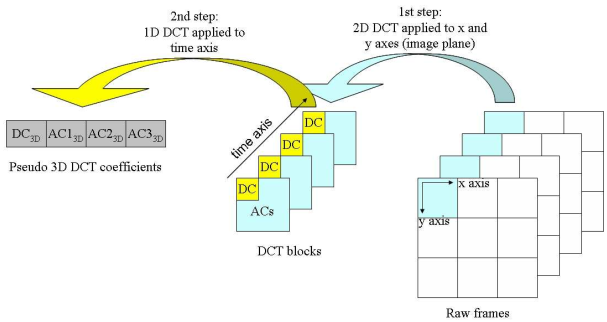


Fig. 1. A pseudo 3D DCT diagram.

transform these DC values to DCT domain again. After transforming process, we can obtain a new DC value and three AC values. According to Eq. (1), a sum of these AC values with corresponding weight values can be computed and obtained. By repeating above steps until all blocks of frames with the same group, we will acquire a sequence of sums of every block. Finally, the embedding information can be obtained to construct the embedding technique.

2.2. Watermark embedding

In order to embed the watermark bits, the quantization index modulation (QIM) method is employed. Based on the QIM algorithm, the embedding domain is divided into several regions. The interval of every region is the same value which equals to the threshold $T(i)$, and an index obtained by $Q(i,k)$ is assigned to each region. Every region represents a value of watermark. According to the $Q(i,k)$ and the embedded bit streams, we will further modify the values of $Sum(i,k)$ by means of the QIM method. The modification is expressed as:

$$Q(i,k) = \begin{cases} 2 \times p, & \text{if EW is 0, Matched,} \quad (a) \\ 2 \times p, & \text{if EW is 1, Unmatched,} \quad (b) \\ (2 \times p) + 1, & \text{if EW is 1, Matched,} \quad (c) \\ (2 \times p) + 1, & \text{if EW is 0, Unmatched.} \quad (d) \end{cases} \tag{2}$$

where p and EW denote a random non-negative integer and the embedded watermark bit, respectively. If the relationship of $Q(i,k)$ and the watermark bit conforms to Eq. (2) (a) or (c), the $Sum(i,k)$ is not modified. Otherwise, the $Sum(i,k)$ will be changed to fit this condition. In order to increase the robustness of our proposed system, $Sum(i,k)$ value will be changed to the center value corresponding to this section to gain the distortion tolerance. The insertion processing is illustrated as Fig. 2. After performing all blocks of frames within the group, we can derive the variation sequence $Diff(i)$ which is the difference denoted as $D(i,k)$ between the modified $Sum(i,k)$ and $Sum(i,k)$ for each block. It is reasonable that the blocks with the small variations will be selected to embed the watermark. Thus, the embedding positions can be determined according to the amount of $D(i,k)$. Since $Sum(i,k)$ is consisted of several AC values, the modification of $Sum(i,k)$ equals to the change of AC values. Note that the low frequency component is more robust and visually sensitive than the high frequency component. That is, if the low frequency component is modulated, it will cause the distortions more seriously, but it has higher ability to resist attacks than the high frequency component does. Therefore, we use the weights

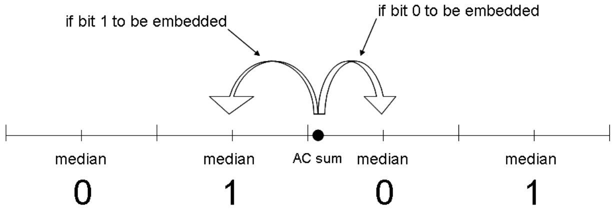


Fig. 2. Watermark insertion by QIM method.

to modulate the $Sum(i,k)$ which is defined as:

$$\begin{aligned} Sum'(i,k) &= \sum_{l} W_s(i,k,l) \times |AC(i,k,l)| \\ &\quad + W_e(i,k,l) \times D(i,k), \end{aligned} \tag{3}$$

where the $D(i,k)$ represents the difference between $Sum'(i,k)$ and $Sum(i,k)$. $AC(i,k,l)$ and $W_e(i,k,l)$ denote the AC value and the weights corresponding to the kth block within the ith group in the frames, respectively. $W_e(i,k,l)$ can be adjusted by user and the sum of $W_e(i,k,l)$ must equal to one. After determining the embedding position, we change the original value in the position into the center of the corresponding section by using Eq. (3). By repeating above procedures until all watermark bits are inserted, the embedding process will be achieved. Finally, all embedding positions, the secret seed S, weights $W_s(i,k,l)$, and the threshold $T(i)$ will be recorded as the secret embedding key. This embedding key will provide the important information to exactly extract the embedded watermark.

2.3. Watermark extraction

The extraction process is the inverse of the embedding process. First of all, the raw video sequence is separated into several groups of frames and each frame is divided into blocks. When we determine the embedding blocks, the selected blocks were transformed by using pseudo 3D DCT, we can further obtain $Sum_e(i,k)$ by using Eq. (1). Then we divide $Sum_e(i,k)$ based on the relative threshold $T(i)$ and the secret embedding key to calculate the quotient $Q_e(i,k)$. According to $Q_e(i,k)$, The embedded bit can be detected and given by

$$\text{EW} = \begin{cases} 0, & \text{if } Q_e(i,k) = 2 \times p, \\ 1, & \text{if } Q_e(i,k) = 2 \times p + 1, \end{cases} \tag{4}$$

If $Q_e(i,k)$ is odd value, the embedded bit is 1. If $Q_e(i,k)$ is even value, the embedded bit is 0. By repeating above steps, we can exactly determine the embedded bits gradually until all watermark bits are extracted. Finally, using the secret seed S recorded in the secret embedding key, the embedded watermark can be effectively detected.

3. PERFORMANCE MEASUREMENT

There are two important factors to measure the performance of watermark system:transparency and robustness. For transparency, we use the peak-signal-to-noise ratio (PSNR) to present this characteristic expressed as

$$PSNR = 10 \times \log \frac{S_{max}^2}{MSE}, \tag{5}$$

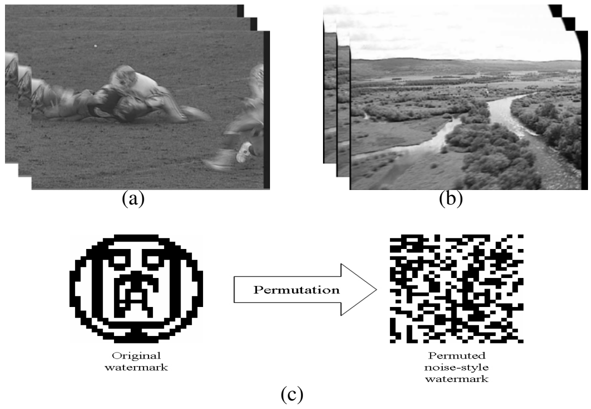


Fig. 3. (a) and (b) Original frames. (c) Watermark image.

$$MSE = \frac{1}{h \times w} \sum_{y}^{h} \sum_{x}^{w} |S_1(x,y) - S_2(x,y)|^2, \qquad (6)$$

where S_1 and S_2 denote the corrupted and original images, respectively. Variables h and w denote the height and width of the image. For a gray level image, S_{max} represents 255 gray value. For robustness, we use the NC value to represent this characteristic. The NC value which measures the similarity between the original watermark $W(i,j)$ and the extracted watermark $\hat{W}(i,j)$ is given by

$$NC = \frac{\sum_{i=0}^{w} \sum_{j=0}^{h} W(i,j) \times \hat{W}(i,j)}{\sum_{i=0}^{w} \sum_{j=0}^{h} [W(i,j)]^2}, \qquad (7)$$

where h and w denote the height and width of the watermark.

4. EXPERIMENTAL RESULTS

In the experiments, we used the 720×480 raw video sequence with 80 frames and made four frames as a group, and each frame is divided into the number of 8×8 blocks. The watermark with the size of 36×20 is prepermuted into a binary pattern by pseudorandom generator. The threshold is obtained by computed the median of the quantization region. Figure 3 shows the testing data and watermark. Figure 4 shows the PSNR values of 80 frames comparing with our proposed method, Kong *et al*.'s [8] method and Thiemert *et al*.'s [12] method. From these results, it is clearly obvious that the transparency of our proposed method is superior to Thiemert *et al*'.'s and Kong *et al*.'s results. For robustness, the compared results for the MPEG-1 compression are shown in Fig. 5. Figures 6-9 illustrate the results of different attacks. According to above experiments, it is obvious that our proposed method is more robust than Kong *et al*.'s method and Thiemert *et al*.'s method.

5. CONCLUSIONS

In this paper, we have proposed an effective video watermarking algorithm based on 3D pseudo DCT and QIM method to achieve the copyright protection in DCT domain. We use twice DCT method to obtain the embedded information and the QIM method to decide the embedded bits of watermark as previously discussed. Experimental results demonstrate that our proposed approach is feasible and can obtain the good performance in transparency and robustness.

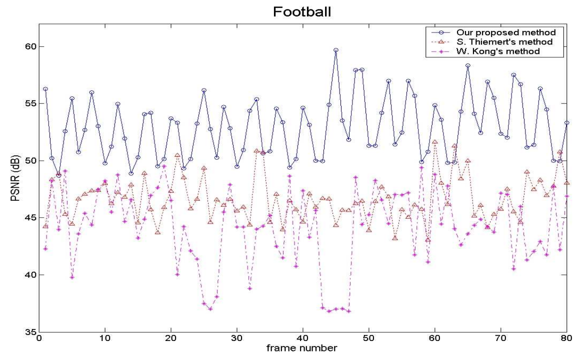


Fig. 4. The PSNR values compared with our proposed method, Kong *et al.*'s method, and Thiemert *et al.*'s method.

6. ACKNOWLEDGEMENTS

This work was supported in part by the National Science Council of Republic of China under Grant No. NSC 95-2221-E-007-187-MY3 and NSC 97-2221-E-150-065.

References

[1] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1108–1126, July 1999.

[2] I. J. Cox, F. J. Kilian, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. on Image Processing*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.

[3] C. T. Hsu and J. L. Wu, "Hidden digital watermarks in images," *IEEE Trans. on Image Processing*, vol. 8, no. 1, pp. 58–68, Jan. 1999.

[4] C. I. Podilchuk and E. J. Delp, "Digital watermarking: algorithms and applications," *IEEE Signal Processing Magazine*, vol. 18, no. 4, pp. 33–46, July 2001.

[5] R. Lancini, F. Mapelli, and S. Tubaro, "A robust video watermarking technique in the spatial domain," in *Proc. of the 8th IEEE Int. Symposium on Video/Image Processing and Multimedia Communications*, June 2002, vol. 17, pp. 251–256.

[6] A. S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," *IEEE Trans. on Image Processing*, vol. 1, no. 2, pp. 244–250, April 1992.

[7] N. M. Charkari and M. A. Z. Chahooki, "A robust high capacity watermrking based on DCT and spread spectrum," in *Proc. of IEEE Int. Symposium on Signal Processing and Information Technology*, 2007, pp. 194–197.

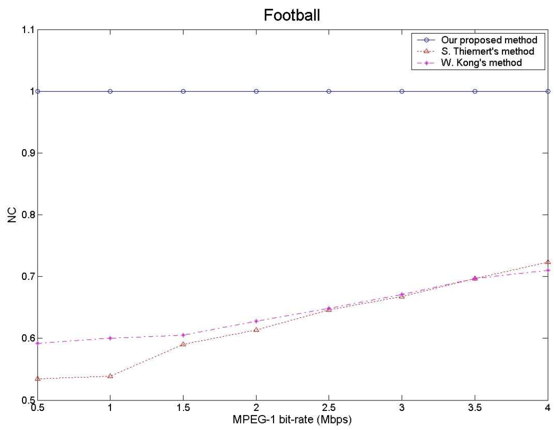
[8] W. Kong, B. Yang, D. Wu, and X. Niu, "SVD based blind vido watermarking algorithm," in *Proc. of the first Int. Conf. on Innovative Computing, Information and Control*, 2006, pp. 265–268.

[9] H. Y. Huang, C. H. Fan, and W. H. Hsu, "An effective watermark embedding algorithm for high jpeg compression," in *Proc. of Machine Vision Applications*, May. 2007, pp. 256–259.

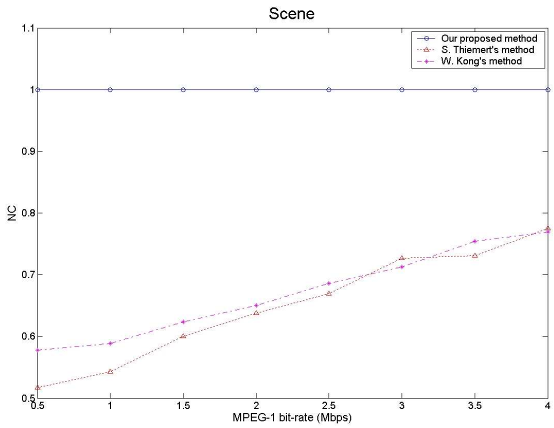
[10] Q Li and I. J. Cox, "Using perceptual models to improve fidelity and provide resistance to valumetric scaling for quantization index modulation watermarking," *IEEE Trans. on Infor. Forensics and Security*, vol. 2, no. 2, pp. 127–139, June 2007.

[11] B. Chen and G. W. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embeeding," *IEEE Trans. on Information Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[12] S. Thiemert, T. Vogel, J. Dittmann, and M. Steinebach, "A high-capacity block based vido watermark," in *Proc. of the 30th Euromicro Conference*, 2004, pp. 457–460.



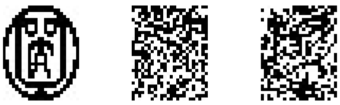
(a)



(b)

Fig. 5. NC values for the different MPEG-1 compression compared with our proposed method, Thiemert *et al.*'s method, and Kong *et al*'s. method.

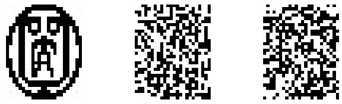
(a)



(b) NC=1 (c) NC=0.5903 (d) NC=0.6361

Fig. 6. (a)Watermarked frame with Wiener filtering. Extraction watermark results. (b)-(d)The proposed method, Thiemert *et al.*'s method, and Kong *et al*'s method, respectively.

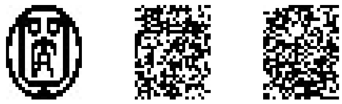
(a)



(b) NC=1 (c) NC=0.5889 (d) NC=0.6042

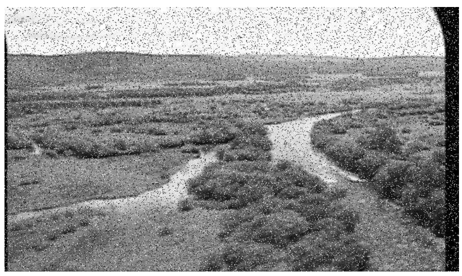
Fig. 7. (a)Watermarked frame with Wiener filtering. Extraction watermark results. (b)-(d)The proposed method, Thiemert *et al.*'s method, and Kong *et al*'s. method, respectively.

(a)



(b) NC=1 (c) NC=0.6181 (d) NC=0.5903

Fig. 8. (a)Watermarked frame with pepper and salt noise. Extraction watermark results. (b)-(d)The proposed method, Thiemert *et al.*'s method, and Kong *et al*'s method, respectively.



(a)

(b) NC=0.9986 (c) NC=0.6278 (d) NC=0.5833

Fig. 9. (a)Watermarked frame with pepper and salt noise. Extraction watermark results. (b)-(d)The proposed method, Thiemert *et al.*'s method, and Kong *et al*'s. method, respectively.