Real-time video watermarking based on extended m-sequences

Fuhao Zou, Zhengding Lu, Hefei Ling, Yanwei Yu
(College of Computer Science&Technology, Huazhong University of Science&Technology, Wuhan 430074, China)

ABSTRACT

This paper presents an oblivious real-time video watermarking scheme in which extended m-sequences is selected as watermark pattern. The good balance property of extended m-sequences is exploited to generate a stable reference point, which can be used to accurately compute modified amount between the watermarked and the original video in the case of original's absence, and furthermore, it helps to improve the robustness of watermarking scheme. To satisfy the real-time requirement, the processes of watermarking embedding and detection are directly operated in the variable length codeword (VLC) domain. To acquire better visual quality, the human visual system (HVS) is used to control the modification strength. Under the premise of assuring better visual quality and real-time requirements, our scheme is more robust than the schemes proposed by Lu and Frank et al..

1. INTRODUCTION

With the development of multimedia and Internet technique, the copy, distribution and transmission of digital works become much more convenient. It brings us advantages, at the same time, provides opportunities to invalid copy and pirates. Thus, how to effectively protect digital rights is a significant factor in making the digital technology progress. Presently digital watermarking is regarded as a critical technique for multimedia copyright protection. In existing watermarking research literature, image watermarking is far more than video's. However, video information is even more useful and should be protected with higher priority. As some special requirements such as blind detection, real-time, and constant bit-rate are required, so video watermarking scheme can not be simply extended from image watermarking scheme and should be specially developed for video sequence.

In last years, some video watermarking schemes[1,2,3,4] are proposed, and those watermarking schemes are operated either in spatial[1] or compressed[1,2,3,4] domain. Generally, most of video sequences are stored in compressed format, and therefore the scheme of compressed domain is more practical. In [1], Hartung and Girod propose a public video watermarking scheme of DCT coefficient domain which based on spread spectrum concept. They arranged a watermark sequence to be a two dimensional array with the same size as a video frame. Then, the watermark signal is 8×8 DCT transformed and added into the DCT coefficients of a video stream. Therefore, some preprocessing operations such as inverse entropy coding, inverse zigzag scanning and inverse quantization are needed before watermarking embedding. Furthermore, their scheme is fully decoded into spatial domain during watermarking detection, so their scheme is hard to satisfy real-time requirement. In [2,3], Langelaar et al. propose a DEW video watermarking of quantized DCT coefficient domain which embeds watermark bits by removing high-frequency coefficients. Although the authors claimed that it is not possible to remove the DEW watermark without causing perceptual degradation, the deadly shortcoming of their scheme is very sensitive to low-pass filtering attack. In [4], Lu et al. propose a real-time video watermarking based on mean-filtering concept. Because around 20%-30% of the macroblocks either have zero level values or generate new codewords that do not exist in the VLC table, watermark bits can not be really inserted. Therefore, the robustness of their scheme is not high.

From above review, a perfect video watermarking scheme should be satisfied the real-time and robustness requirement simultaneously. In our scheme, to satisfy the real-time requirement, the processes of the watermarking embedding and detection are directly performed in the VLC domain. Watermark pattern is extended m-sequences, whose good balance property is exploited to generate a stable reference point, and it will conduce to improve the robustness of watermarking scheme. To prevent from obvious degradation of visual quality, the human visual system (HVS) is used to control the modification strength. Under the premise of assuring better visual quality and real-time property, the capability of robustness is maximized. The rest of this paper organized as follows. In section 2 the watermarking scheme is described in detail. In section 3, our scheme is compared with the schemes proposed by Frank [1] and Lu [4] at the aspect of transparency, robustness and real-time. Finally, the conclusion is drawn in section 4.

2. THE PROPOSED VIDEO WATERMARKING IN VLC DOMAIN

2.1. The scheme of watermarking embedding

To avoid the time-consuming operations (i.e. DCT, inverse DCT and motion compensating), the watermarking scheme for MPEG video sequence should tightly cooperate with MPEG compressed standard. Because our scheme directly performed in VLC domain, only inverse entropy coding and inverse zigzag scanning operations are needed before watermarking embedding and detection. In VLC domain, every codeword corresponds to a run-level pair denoted as (r,l). For a given run-level pair (r,l), run r represents the number of DCT coefficients with magnitude zero preceding the current run-level pair and level l denotes the quantization value of the current DCT coefficient Our scheme embeds watermark bits through slightly modifying the level of run-level pair of DC coefficients in luminance space of 'I' frame. Why the luminance component rather than chrominance component is selected as embedding space is that our scheme is to avoid synchronization problem due to chroma-format varying. The reasons that DC coefficients are selected as embedding objects are as follows. On the one hand the operations of modifying DC coefficients will prevent the run r of other DCT coefficients from changing; on the other hand watermark embeds in perceptual significant component is of great benefit to the robustness of watermarking scheme. In literature [5], Huang et al. apply quantitative analysis method to illuminate the rationality of DC coefficients being selected as embedding objects.

The extended m-sequences are selected as watermark pattern in our scheme. The advantages of extended m-sequences mainly have two aspects: firstly the even periods of extended m-sequences are convenient for the realization of hardware; secondly extended m-sequences have a good balance property. Let the extended m-sequences $W=\{w(i)\,|\,1\le i\le 2^n \wedge i\in(-1,1)\wedge n\in Z\}$ denotes the original watermark, then its balance property is appeared as:

$$\sum_{i=1}^{2^n} w(i)=0 \tag{1}$$

This property will be exploited to generate a stable reference point, whose stability will be discussed detailedly in watermarking detection section.

Taking 8×8 DCT block as basic unit, the luminance component of 'I' frame is partitioned into 2^n disjointed regions according to the length of watermark pattern, namely each region will be watermarked a bit. Suppose (r_{ij},l_{ij}) denotes run-level pair corresponding to the $j-th$ DC coefficient in the $i-th$ region, then, the sum of absolute value of level of DC coefficient in the $i-th$ region can be represented as:

$$S_i=\sum_{i=1}^{L_{region}}\left|l_{ij}\right| \tag{2}$$

where L_{region} denotes the number of 8×8 DCT blocks in the $i-th$ region. The mean value of sequence $S=\{S_i,1\le i\le 2^n\}$ is:

$$u=\frac{1}{2^n}\sum_{i=1}^{2^n}S_i \tag{3}$$

In our scheme, using the $i-th$ watermark bit's embedding process as example, watermarking embedding is realized by adjusting the sum of absolute value of level of DC coefficient in the $i-th$ region from S_i to S_i^h as follows:

$$S_i^h=\begin{cases}u+T, if\ w(i)=1\\ u-T, if\ w(i)=-1\end{cases}, or\ S_i^h=u+w(i)gT \tag{4}$$

T can be determined as follows:

$$\begin{aligned}&Min_T=\min\{\left|S_1-u\right|,...,\left|S_{2^n}-u\right|\}\\&Max_T=\max\{\left|S_1-u\right|,...,\left|S_{2^n}-u\right|\}\\&T=\lambda g Min_T+(1-\lambda)g Max_T, 0\le\lambda\le 1\end{aligned} \tag{5}$$

where λ is scale factor. When $\lambda=1$, our scheme reaches the minimum distortion and the worst robustness; when $\lambda=0$, our scheme reaches the maximum robustness and the worst visual quality. According to the relation between S_i and S_i^h, each run-level pair corresponding to DC coefficient is respectively modulated, and the modulated strategies can be divided into positive modulation (PM) and negative modulation (NM). In the case of PM or NM, the level value l_{ij} of run-level pair (r_{ij},l_{ij}) can be modulated as:

$$l_{ij}^h=\begin{cases}l_{ij}+sign(l_{ij})g\left|S_i^h-S_i\right|g\dfrac{JND_{ij}}{\sum_{j=1}^{L_{region}}JND_{ij}}, PM\\[2ex] l_{ij}-sign(l_{ij})g\left|S_i^h-S_i\right|g\dfrac{JND_{ij}}{\sum_{j=1}^{L_{region}}JND_{ij}}, NM\end{cases} \tag{6}$$

Where $sign(g)$ is sign function, JND_{ij} is the maximum permission modified magnitude of l_{ij}, which is computed according to Watson visual model [6]. Because the modified objects are all DC coefficients, which means that the change of l_{ij} will not affected the value of r_{ij}, the modified run-level pair of (r_{ij},l_{ij}) is (r_{ij},l_{ij}^h).

2.2. The watermarking detection scheme

Generally, the watermarked video sequence probably suffered from some attacks, which could be classified two categories: one is unintentional attack such as common signal processing, the other is intentional attack whose goal is to destroy or tamper watermark. To be convenient for analysis, we assume that watermarked video sequence don't suffer from any attacks. Like watermarking embedding, the luminance component of 'I' frame is partitioned into 2^n disjointed regions. And then, we can compute the sum of absolute value of level of DC coefficient for each region and obtain a sum sequence $S^e=\{S_i^e\,|\,1\le i\le 2^n\wedge S_i^e=S_i^h\}$, the mean of sequence S^e can be computed as follows:

$$\begin{aligned} u^e &= \frac{1}{2^n}\sum_{i=1}^{2^n} S_i^e = \frac{1}{2^n}\sum_{i=1}^{2^n} S_i^h = \frac{1}{2^n}\sum_{i=1}^{2^n}(u + w(i)\cdot T) \\ &= \frac{1}{2^n}\sum_{i=1}^{2^n} u + \frac{1}{2^n}\sum_{i=1}^{2^n} w(i)\cdot T = u + 0 = u \end{aligned} \tag{7}$$

From equation (7), it apparently appears that mean value point doesn't change before and after watermarking embedding. Because the mean values are all reference points during watermarking embedding and detection process, we can believe that the reference point is considerably stable. According to the relation between S_i^e and u^e, we can extract watermark $W^e = \{w^e(i), 1 \le i \le 2^n\}$, the value of $w^e(i)$ is determined as:

$$w^e(i) = bipolar(S_i^e - u^e) \tag{8}$$

where $bipolar(\cdot)$ can be defined as:

$$bipolar(t) = \begin{cases} 1, if\ t \ge 0 \\ -1, if\ t < 0 \end{cases} \tag{9}$$

Finally we compute the correlation coefficient (also called detector response) between the detected watermark W^e and original watermark W:

$$\rho(W, W^e) = \frac{\sum_i w(i)\cdot w^e(i)}{\sqrt{\sum_i w^2(i)}\sqrt{\sum_i (w^e(i))^2}} \tag{10}$$

Judging from the relation between $\rho(W, W^e)$ and detection threshold $Threshold_e$, we can identify whether the watermark exists in detected video sequence, the detailed process can be represented as:

$$\begin{cases} existing\ watermark & \rho(W, W^e) \ge Threshold_e \\ no\ watermark & \rho(W, W^e) < Threshold_e \end{cases} \tag{11}$$

Due to the size of video sequence doesn't permit to change during the watermarking embedding process; the specific method is adopted to maintain constant bit-rate. Computing the difference of the video size between before and after watermarking embedding, if the size is increased, we insert corresponding number bits between slice, else we delete some high-frequency coefficient.

3. EXPERIMENTAL RESULTS

To evaluate the performance of our scheme, the schemes respectively proposed by Frank [1] and Lu [4] are selected to compare. Test video sequence is a famous 'flower-garden' video sequence, the size of frame is 352×240, the number of frames is 150 ('I'-26, 'P'-25, 'B'-99), and bit-rate is 8Mbit/s. The performances of scheme vary due to different parameters, and therefore parameter selection is very important. To objectively evaluate the performances of three schemes, we embed one watermark pattern in each 'I' frame and render the value of detector response to a same value by adjusting parameters using three schemes respectively. In our simulation experiment, the detector response is assigned 0.95, and then we evaluate performance at the aspect of visual quality, robustness, and real-time.

3.1. The evaluation of visual quality

Whether the watermarked video evidently degenerates can be judged via the value of PSNR between watermarked and original video, generally the lager the PSNR is, the better the visual quality is. The PSNR values corresponding to three schemes are plotted in Fig. 1, the axis-X represents the index of 'I' frame, and the axis-Y represents the value of PSNR. From Fig. 1, it is evident that the PSNR values corresponding to our scheme are all larger than the other two, which means that modification magnitude of our scheme is smaller than the other two in the case of obtaining the same detector response. Furthermore, the plot-lines corresponding to the other two schemes are flat, which reflects the other two schemes don't adjust the watermarking embedding strength according to the content of video sequence.

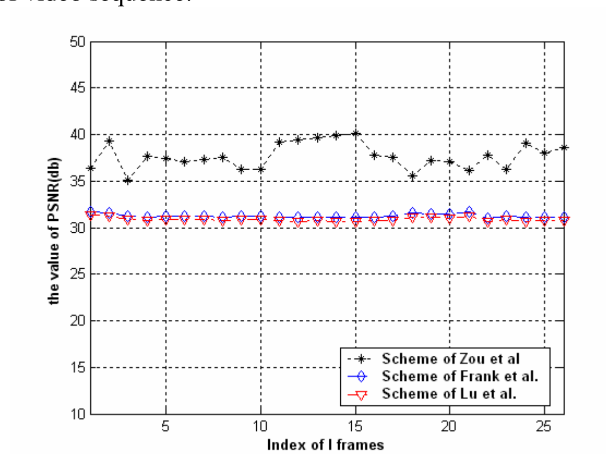


Fig. 1 the PSNR value of watermarked video sequences corresponding to three schemes

3.2. The estimate of robustness

To estimate the robustness of three schemes, some common attacks such as MPEG compression (i.e. re-encoding the watermarked video sequence from bit-rate 8Mbit/s to 6Mbit/s, 4Mbit/s, and 2Mbit/s respectively), low-pass filtering, median filtering, Gaussian noise, Gaussian filtering, sharpening, and random bend. Comparing the values of detector response from attacked video sequences, we can judge the robustness of three schemes. Generally, the larger the detector response is, the more robust the scheme is. In Fig. 2 and Fig. 3, the axis-X represents the index of 'I' frame, and the axis-Y represents the value of detector response. Fig. 2 is the result of MPEG compression attacks, and the other attack results are respectively plotted in Fig. 3. From Fig. 2 and Fig. 3, we can find that the detector response values of our scheme are larger than the other two, and it means that our scheme is more robust than the other two.

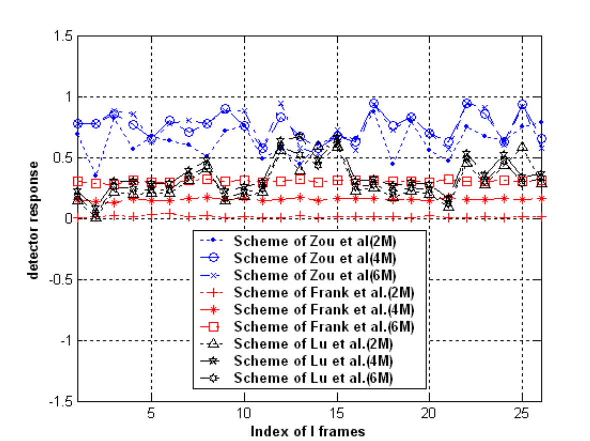
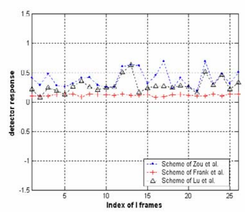
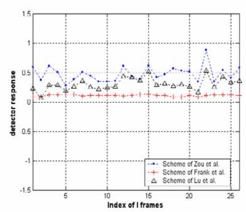


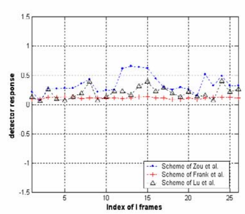
Fig. 2 the detector response corresponding to bit-rate 6Mbit/s, 4Mbit/s, 2Mbit/s re-encoding of three schemes



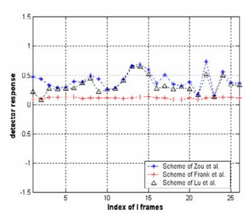
(a) low-pass filtering



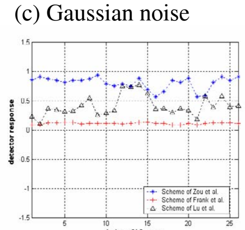
(b) median filtering



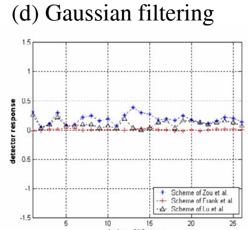
(c) Gaussian noise



(d) Gaussian filtering



(e)Sharpening



(f) random bend

Fig. 3 detector response corresponding to different attacks

3.3. The evaluation of real-time

To demonstrate our scheme can satisfy the real-time requirement, we compare each time cost of eight operations to 'flower-garden' video sequence in Fig. 4. In Fig. 4, the axis-X represents the serial number of operation, and the axis-Y represents the time cost. The eight operations are as follow: (1) fully decode MPEG video sequence from bit domain to spatial domain; (2) the embedding process of Zou; (3) the detection process of Zou; (4) the embedding process of Lu; (5) the detection process of Lu; (6) the embedding process of Frank; (7) the detection process of Frank;(8) only decode to VLC domain. From Fig. 4, it appears that time cost of the $\boldsymbol{2-th}$ operation and the $\boldsymbol{3-th}$ operation are very close to the $\boldsymbol{8-th}$ operation, namely our scheme can satisfy the real-time requirement.

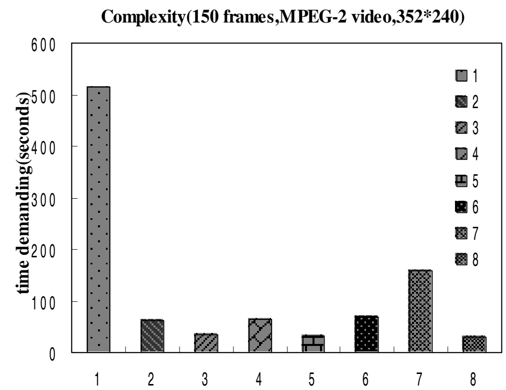


Fig. 4 the comparison of time cost corresponding to eight operations.

4. CONCLUSION

This paper proposes a real-time watermarking scheme, in which the extended m-sequences are selected as watermark pattern. Under the premise of assuring better visual quality and real-time property, our scheme is more robust than the scheme proposed by Lu and Frank et al. Through the theory analysis and experimental result comparison, our scheme is much better.

5. ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China under Grant No. 60502024, the Natural Science Foundation of Hubei Province under Grant No. 2005ABA267, the Electronic Development Fund of Ministry of Information Industry of China and the Innovation Fund for Technology Based Firms of Ministry of Science and Technology of China under Grant No. 04C26214201284.

6. REFERENCES

[1] Hartung Frank,Girod Bernd. Watermarking of uncompressed and compressed video. Signal Processing,1998,66(3):283~301.

[2] Langelaar Gerrit C., Lagendijk Reginald L., Biemond Jan. Real-time labeling of MPEG-2 compressed video. Journal of Visual Communication and Image Representation,1998,9(4):256~270.

[3] Langelaar Gerrit C., Lagendijk Reginald L.. Optimal differential energy watermarking of DCT encoded images and video. IEEE Transactions on Image Processing,2001,10(1):148~158.

[4] Chun-Shien Lu, Jan-Ru Chen,Liao H.-Y.M., Kuo-Chih Fan. Real-time MPEG2 video watermarking in the VLC domain. In: Proc. Of 16th International Conference on Pattern Recognition. Los Alamitos, CA, USA: IEEE Comput. Soc, 2002. 552~555.

[5] Huang Jiwu, Shi Yun Q., Shi Yi. Embedding image watermarks in DC components. IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(6):974-979.

[6] Watson A.B.. DCT quantization matrices visually optimized for individual images. In: Proc. of conference of Human Vision, Visual Processing, and Digital Display IV. San Jose, CA, USA:SPIE,1993.202~216.