Digital Watermarking with Local Strength Adjustment for AVC-Compressed HDTV Bitstreams

Jakub Siast, Jakub Stankowski, Tomasz Grajek, Marek Domański
Chair of Multimedia Telecommunications and Microelectronics, Poznań University of Technology
60-965 Poznań, Polanka 3 Street, POLAND
{jsiast,jstankowski,tgrajek}@multimedia.edu.pl

***Abstract*— A computationally inexpensive watermark schema for MPEG-4 AVC/H.264 streams that use local mark strength adjustment minimizing visible artifacts is presented. The local mark strength adjustment mechanism and watermark embedding mechanism operate on residual data. This approach minimize the number of operations needed for video bitstream decoding and subsequent watermarked video data encoding. The experimental results included herein, described for real-world MPEG4 AVC/H.264 encoded HDTV video bitstreams prove its robustness to complex attacks, especially camcorder recording attack.**

I. Introduction

Along with development of digital TV systems and digital video cameras, quality of video material recorded from HDTV became better and better. Nowadays, it is possible to make a copy of video displayed on HDTV set with little quality degradation. Therefore, a method to protect movies with many-millions budget against illegal camcording and distribution in internet is highly desired. Watermarking is one of the methods that has been proposed to address this issue [1]. This article refers to watermarking for detection of illicit camcording of video content. Therefore, the watermarks should be robust against camcording, i.e. camcording attacks. This work is continuation of our previous research on this topic described in [2].

Generally watermark can be embedded either into uncompressed or compressed video. There are many approaches to embed watermark into uncompressed (raw) video e.g. [3,4]. Unfortunately, full decoding and re-encoding will be required to apply them into compressed video bitstream (see Fig. 1). In many applications it is not desired or even not possible to perform full decoding and again encoding, due to high computational complexity of such an approach. Consequently, watermarking of compressed video bitstreams has gained lot of attention e.g. [5-12]. Among compressed domain watermarking schemes, there are only several studies dealing with camcorder attacks e.g. [6-8] mostly for older video compression standards such as MPEG-2. Moreover, in the vast majority of papers all experiments are conducted on low resolution video materials and achieved results cannot be directly referred to HDTV video materials.

In this paper, we propose a fast watermarking method for MPEG-4 AVC/H.264 [13,14] bitstreams robust to very strong attacks, e.g. video camcorder recording. It was developed based on scaling matrices watermarking scheme described in our previous paper [2]. The computational complexity of scaling matrices watermarking was very low (see Fig 1.), but there was no possibility for local watermark adjustment. The proposed method is more complex because it requires partial decoding of bitstream to obtain prediction error transform coefficients (which are changed in the proposed watermark method). However, the strength of the watermark can be locally adjusted based on the values of residual transform coefficients. Additionally, this method provides an original approach to selection of the embedded coefficients. In contrary to scaling matrices watermarking, this method corresponds to transform domain watermarking (black dotted rectangle in Fig. 1).

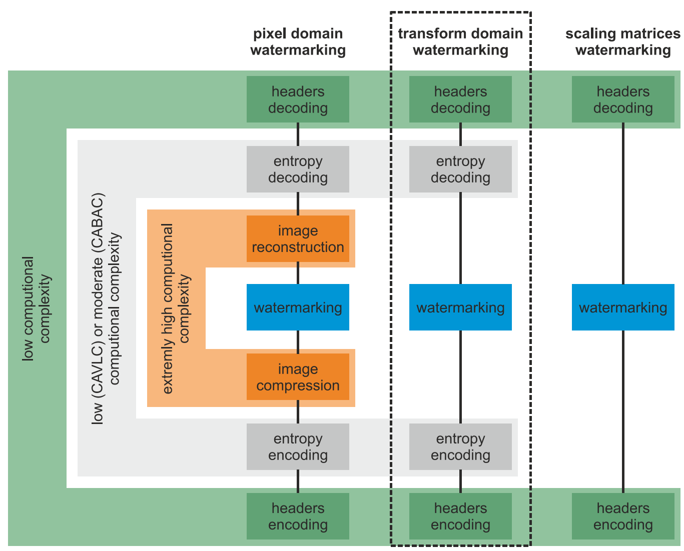

Figure 1. Three most common approaches to watermarking of compressed video sequences (pixel domain, transform domain and scaling matrices domain). For every approach the reqired operations and coresponding complexity on embedder side are shown.

II. Watermark Embedding

In the proposed method a watermark is embedded by direct modifications of the value of low frequency DCT coefficients of transformed residual signal. The increase of energy of selected residual DCT coefficient corresponds to increased energy of corresponding frequency in reconstructed picture. In proposed approach, energy of vertical or horizontal coefficients is changed according to bit value of embedded data. In order to embed invisible watermarks with sufficient data capacity enough to identify time and place of camcording watermark can be embedded into selected luma transform blocks, called Embedding Block (EB). Moreover, EBs can belong to selected frames called Embedding Frame (EF).

A. Frame selection

In order to preserve compatibility with embedding method described in [2], the watermark is embedded into P- and B- frames. Moreover modifying P- and B- frames allow to reduce the computational complexity and improve locality of the watermark in sequence of pictures. The complexity issue will be discussed later. Appropriate selection of EFs allow to control presence of watermark in selected periods of sequence. As in [2] the sequence of 8 watermarked images has been interleaved with the sequence of 8 non-watermarked images (Fig. 2.).

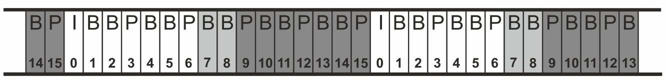

Figure 2. The location of watermarked and non-watermarked frames in watermarked video sequence. The frames marked in dark are watermarked (from [2]).

B. Embedding Block selection

The simplest approach is to watermark all blocks in the frame. Such a method will be very similar to scaling matrix watermarking [2]. However, direct access to transform coefficients allows for more advanced watermarking approach.

In MPEG-4 AVC/H.264 in P- and B- frame the both inter and intra prediction can be used. In case when intra prediction is not constrained to use another intra predicted block only (*constrained_intra_pred_flag* in PPS is equal to 0) the inter predicted block may be used as intra prediction source. This could cause drift and escalation of "desired error" introduced by watermarking. Moreover, the intra drift is the most disturbing kind of introduced distortion and occurs when watermarked blocks are used by intra predicted blocks.

In order to avoid the intra prediction issue the EB selection algorithm has been developed. The algorithm was designed to reduce visibility of watermarking artifacts. To limit the intra prediction issue every candidate block is investigated to check if any other block uses it as intra prediction reference. Such candidate blocks are excluded from EBs list. Therefore, embedder is able to reduce visibility of watermarking artifacts.

C. DCT coefficiens selection

As previously mentioned, the watermark is embedded by changing the energy of low frequency DCT coefficients. Increased energy of vertical coefficients corresponds to embedding information bit 1, and increased energy of horizontal coefficients corresponds to embedding information bit 0. For watermarking purposes, low-frequency AC coefficients have been chosen due to its camcorder attack robustness, as reported in the references [2]. Therefore, the so called 'embedding coefficients' (EC) have been selected. These are (0,1) and (1,0) coefficients for the 4x4 transform and (0,2), (0,3), (2,0) and (3,0) coefficients for the 8x8 transform (see Fig. 3). The selected 4x4 ECs and 8x8 ECs correspond to the same spatial frequencies.

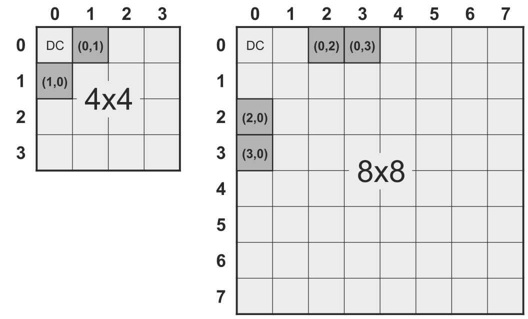

Figure 3. The modified transform coefficients, for 4x4 and 8x8 block size, respectively (from [2]).

D. Watermark strength adjustment

To utilize embedding block characteristic and its capacity for invisible watermark energy, embedding watermark strength adjustment mechanism was developed.

It is well known that highly textured blocks has higher watermark embedding potential than plain ones [5]. In proposed watermarking technology no texture reconstruction process is done and no straight block pattern classification can be done to utilize this phenomena. From the other hand we observed, that nonzero coefficients in MPEG-4 AVC/H.264 residual data are strongly correlated with boundaries of objects or strong objects patterns. Therefore, value of residual coefficients is used as control data for watermark strength adjustment mechanism.

Visibility of watermark in MPEG-4 AVC/H.264 compressed domain depends not only on watermarked block change but also on a drift that it cause. Watermark intra drift artifacts are reduced as described in subsection B but developed watermark strength adjustment mechanism takes inter drift into consideration also. In proposed approach P frames has stronger constraints for watermark embedding, because they are a source for B frames and therefore source of inter drift artifacts.

In details, developed watermark strength adjustment mechanism is summarized in Table 1.

TABLE I. EB WATERMARK STRENGTH ADJUSTMENT

	4x4 block	8x8 block
B-frame	If(nonzero EC): EC' = EC*2 Else if(nonzero C): EC' = 3.5/qstep Else: EC' = 0	If(Nonzero EC): EC' = EC*2 Else: EC' = 0
P-frame	If(Nonzro EC): EC' = EC*2 Else: EC' = 0	

For blocks in P and B frames if embedding coefficient value is different from zero it is doubled. For B frames only if block has any nonzero coefficient (C) but embedding coefficient (EC) value is zero the nonzero value is inserted into EC position. The new EC value is computed as 3.5 divided by quantization step (dividing is a consequence of quantized representation for residual signal in a bitstream). In other situation embedding coefficients remains unchanged, effectively excluding EB from being watermarked. Proposed watermark strength adjustment method impose that energy of a watermark vary for sequences with different spatial frequency characteristic.

E. Embedding process

Watermark embedding is performed by changing the energy of residual coefficients (RC). Before watermark embedding RCs have to be extracted from the bitstream. In MPEG-4 AVC / H.264 partial entropy decoding of slice is not possible and whole slice has to be entropy decoded. After entropy decoding phase the embedder selects EBs from all available luma blocks. The EBs are modified by changing the selected RCs. In proposed approach only one bit per frame is embedded. Therefore all EBs in picture (slice) are modified in the same way (by changing coefficients corresponding to horizontal or vertical frequencies). After embedding process the slice is entropy encoded.

F. Embedding computional complexity

The most complex part of embedding process is entropy decoding and re-encoding. The MPEG-4 AVC / H.264 standard defines two different entropy encoders CABAC and CAVLC [14]. CABAC is based on binary arithmetic encoder enhanced by probability modeling system. CAVLC uses switched tables with predefined variable length codes. When compared to CAVLC, CABAC provides better adaptation to video characteristics and higher compression efficiency, however it has significant higher computational complexity and lower throughput. Therefore, the entropy decoding process complexity depends on entropy encoder used during generation of watermarked bitstream. The CABAC decoding complexity is rather significant but still it is only small part of full decoding burden.

The complexity of EBs selection and ECs modification is very low and can be neglected.

The entropy encoding complexity is similar to entropy decoding. However for CABAC encoded sequences the encoding complexity can be significantly reduced by changing the entropy encoder. The proposed approach is to use CAVLC instead CABAC to encode watermarked slices. In MPEG-4 AVC / H.264 switching between entropy encoders is allowed on slice level, and conformant bitstream can contain interleaved slices encoded using either CABAC or CAVLC.

G. Bitstream overhead

The bitstream increase caused by watermarking process is very small. During embedding process in P slices only existing coefficients are modified and no new coefficients are introduced. In B slices also new coefficients may be inserted in residual block containing any nonzero coefficient. It should be kept in mind that in MPEG-4 AVC / H.264 transform coefficients coding scheme the increase of coefficient absolute value causes less bitstream expansion than introduction of new coefficients [15].

The CABAC entropy encoding has higher compression efficiency than CAVLC [15]. Therefore, slice conversion from CABAC to CAVLC entropy encoder causes noticeable bitstream increase. In proposed approach only some part of slices are marked and transcoded to CAVLC (see Fig. 2). Moreover, the embedder modifies P and B slices only which are much smaller than I slices. Therefore, the average bitstream expansion is limited.

The total bitstream expansion has been investigated experimentally. For case when CABAC encoded slices are converted to CAVLC during watermarking process the obtained average bitstream expansion was about 3.8%.

III. WATERMARK DETECTION

In order to maintain compatibility between scaling matrices watermarking described in [2], the watermark detection process is performed in the same way as for scaling matrices watermarking method. The detector uses both original and investigated video sequence (in pixel domain) to detect the watermark. Both sequences must be temporally aligned but spatial alignment is not critical. The detector calculates the spatial frequency characteristics for all pictures in original and investigated sequence. The differences in frequency characteristic caused by watermarking process are calculated.

In the watermarking process only some pictures are watermarked, therefore watermarked sequence contains interleaved groups of watermarked and original pictures. The presence of original pictures allows to eliminate attacks influence (different changes in frequency characteristics) in detection process. More details about watermark detection algorithm can be find in [2].

IV. RESULTS

The watermarking and the detection techniques were implemented and tested for various attacks robustness and visibility. Four real-world HDTV sequences (Blu-ray quality) were selected for tests: *Soap opera 1*, *Soap opera 2*, *Historical novel*, *Computer-animated film*. Proposed sequence set represents material with quite different spatial frequency characteristic. This is important because of watermark strength adjustment method as described in section II D. First two sequences can be described as first-plan-in-focus, historical novel is all-in-focus sequence, and the last one is animated movie with low noise and weak textures.

All sequences used in the tests were FullHD resolution (1920x1080), 25 frames per second and 10 minutes long. Sequences were encoded by MPEG-4 AVC/H.264 encoder conformant to High profile (bitrate - 15Mbit/s). The prepared bitstreams have been watermarked. Two different random bit streams (A and B in Table 2) have been used as the watermark payload. Watermarked sequences have been displayed on an LCD HDTV set.

Watermark visibility tests have been performed on group of non-expertise students. Viewers assessed overall quality of the watermarked as well as original (not watermarked) video material. Analysis of the subjective test results shown that the distortion introduced by watermarking process was rated as unnoticeable.

In order to rate robustness of the proposed watermarking method four different attack scenarios have been tested:

- Scenario 1: HD camcorder recording,
- Scenario 2: HD camcorder recording and downscaling to SD resolution,
- Scenario 3: Camcorder recording and frame rate reduction by frame dropping (from 25 to 12,5 fps).

All scenarios start with HD camcorder recording of watermarked sequence displayed on a LCD HDTV set and further processing of captured sequence. Camcorder and LCD HDTV set were configured to obtain as good quality of recorded sequence as possible. Therefore, automatic mechanism in both devices were disabled (e.g. auto white balance, autofocus, auto exposure) and day light strength in a test room was limited. The watermarked and attacked sequence has been decoded and passed to the input of the watermark detector. The percentage of properly decoded bits from the watermark payload is shown in Table 2.

TABLE II. EXPERIMENTAL RESULTS

Sequence type	**Payload**	**Correctly detected bits [%] in scenario:**		
		1.	***2.***	***3.***
Soap opera 1	A	86.8	82.6	85.5
	B	87.4	82.6	86.1
Soap opera 2	A	82.6	75.6	82.2
	B	86.9	79.3	85.3
Historical novel	A	90.9	87.9	88.0
	B	97.2	92.9	95.4
Computer-animated film	A	79.0	73.1	77.0
	B	76.9	70.0	75.6
Average		86.0	80.5	84.4

Robustness of proposed watermarking method depends mostly on sequence spatial frequency characteristic. Computer-animated film as a one with the lowest low spatial frequency characteristic and the lowest embedded watermark energy has the lowest correctly detected bits. Historical novel with the highest energy of low spatial frequencies has the highest detection ratio. Comparing different scenarios using camcorder recording without further processing (scenario 1) and scenarios with further processing (scenario 2 and 3) one can see that differences in correctly detected bits are low with maximum of 7%. That makes proposed method robust to captured sequence post processing. Average correctly detected bits ratio above 80% and minimum of 70% shows that with proper correction codes watermark payload can be correctly decoded. Based on experimental results we may say that proposed watermarking method is robust to very strong and complex attacks.

V. CONCLUSIONS

The main advantage of the proposed method is local watermark strength adjustment performed without full video content decoding. Therefore, the watermark can be embedded with moderate computational burden and without significant bitstream overhead. Additionally, distortion introduced by the watermarking process in perceptually unnoticeable for the viewer. The experimental results show that on average over 80% of embedded data bits are correctly decoded. This number of correctly decoded bits are sufficient for the modern forward error correction (FEC) techniques to restore transmitted information. Proposed method provides sufficient robustness against strong attacks including camcorder recording.

REFERENCES

[1] G. Doërr, J.-L. Dugelay: “A guide tour of video watermarking”, Signal Processing: Image Communication vol. 18, no. 4, pp.263-282, Elsevier 2003.

[2] J. Stankowski, T. Grajek, M. Domański: “Fast Watermarking of MPEG-4 AVC/H.264 Encoded HDTV Video Bitstreams”, 29th Picture Coding Symposium, PCS 2012, pp. 265-268, Kraków, Poland, 2012.

[3] A. Koz and A. A. Alatan: “Oblivious spatio-temporal watermarking of digital video by exploiting the human visual system”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 3, pp. 326–337, Mar. 2008.

[4] F. Hartung, B. Girod: “Watermarking of Uncompressed and Compressed Video”, Signal Processing, vol. 66, no. 3, pp. 283-301, Elsevier, 1998.

[5] M. Noorkami, R.M. Mersereau: “Digital Video Watermarking in P-Frames With Controlled Video Bit-Rate Increase”, IEEE Transactions on Information Forensics and Security, vol. 3, no 3, pp. 441-455, 2008.

[6] M. Celik, J. Talstra, A.N. Lemma, S. Katzenbeisser: “Camcorder Capture Robust Low-Complexity Watermarking of MPEG-2 Bit-Streams”, IEEE International Conference on Image Processing, ICIP 2007, pp. 489-492, San Antonio, USA (2007).

[7] D. Choi, H. Do, H. Choi, T. Kim: “A blind MPEG-2 video watermarking robust to camcorder recording”, Signal Processing, vol. 90, no. 4, pp. 1327-1332, Elsevier, 2010.

[8] H. Do, D. Choi, H. Choi, T. Kim: “Digital Video Watermarking Based on Histogram and Temporal Modulation and Robust to Camcorder Recording”, IEEE International Symposium on Signal Processing and Information Technology, pp. 330-335, Sarajevo, 2008.

[9] Y. He, G. Yang, N. Zhu: “A real-time dual watermarking algorithm of H.264/AVC video stream for Video-on-Demand service”, International Journal of Electronics and Communications, vol. 66, no. 4, pp. 305–312, Elsevier, 2012.

[10] J. Zhang, A.T..S. Ho, G. Qiu, P. Marziliano: “Robust Video Watermarking of H.264/AVC”, IEEE Transactions on Circuits and Systems—II: Express Briefs, vol. 54, no. 2, pp. 205-209, 2007.

[11] C.-V. Nguyen, D.B.H. Tay, G. Deng: “A Fast Watermarking System for H.264/AVC Video”, IEEE Asia Pacific Conference on Circuits and Systems, APCCAS 2006, pp. 81-84, Singapore, 2006.

[12] P.-C. Su, M.-L. Li, I.-F. Chen: “A Content-Adaptive Digital Watermarking Scheme in H.264/AVC Compressed Videos”, International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP '08, pp. 849-852, Harbin, China, 2008.

[13] ISO/IEC 14496-10: Inform. Tech. – Coding of Audio-Visual Objects – Part 10: Advanced Video Coding, 2012.

[14] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 7, pp. 560 –576, 2003.

[15] D. Marpe, H. Schwarz, T. Wiegand, “Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard”, IEEE Transactions on Circuits and Systems for Video Technology vol. 13, no. 7, pp. 620–636, 2003.