View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Northumbria Research Link

Northumbria Research Link

Citation: Ait Saadi, Karima, Guessoum, Abderrezak, Bouridane, Ahmed and Khelifi, Fouad (2017) Content Fragile Watermarking for H.264/AVC Video Authentication. International Journal of Electronics, 104. pp. 673-691. ISSN 0020-7217

Published by: Taylor & Francis

URL: http://dx.doi.org/10.1080/00207217.2016.1242163 <http://dx.doi.org/10.1080/00207217.2016.1242163>

This version was downloaded from Northumbria Research Link: http://nrl.northumbria.ac.uk/27977/

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: http://nrl.northumbria.ac.uk/policies.html

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

www.northumbria.ac.uk/nrl

northumbria UNIVERSITY NEWCASTLE

Publisher: Taylor & Francis

Journal: *International Journal of Electronics*

DOI: 10.1080/00207217.2016.1242163

ACCEPTED MANUSCRIPT

Content Fragile Watermarking for H.264/AVC Video Authentication

K. Ait Sadi[1], A. Guessoum[2], A. Bouridane[3], F. Khelifi[3]

[1]*Centre de Développement des Technologies Avancées, Division Architecture des Systèmes, Alger, Algérie.*

[2]*Université Saad Dahlab, Institut d'Electronique, Blida, Algérie.*

[3]*School of Computing, Engineering and Information Sciences, Northumbria University, Northumbria, United Kingdom.*

Content Fragile Watermarking for H.264/AVC Video Authentication

Discrete Cosine transform (DCT) to generate the authentication data that are treated as a fragile watermark. This watermark is embedded in the motion vectors (*MV*s) The advances in multimedia technologies and digital processing tools have brought with them new challenges for the source and content authentication. To ensure the integrity of the H.264/AVC video stream, we introduce an approach based on a content fragile video watermarking method using an independent authentication of each Group of Pictures (GOPs) within the video. This technique uses robust visual features extracted from the video pertaining to the set of selected macroblocs (MBs) which hold the best partition mode in a tree-structured motion compensation process. An additional security degree is offered by the proposed method through using a more secured keyed function HMAC-

*Corresponding author. Email: aitsaadi@cdta.dz;ait_saadi@yahoo.com.

SHA-256 and randomly choosing candidates from already selected MBs. In here, the watermark detection and verification processes are blind, whereas the tampered frames detection is not since it needs the original frames within the tampered GOPs. The proposed scheme achieves an accurate authentication technique with a high fragility and fidelity whilst maintaining the original bitrate and the perceptual quality. Furthermore, its ability to detect the tampered frames in case of spatial, temporal and colour manipulations, is confirmed.

Keywords: Content protection; fragile watermarking; digital signature, video Authentication; H.264/AVC codec.

ACCEPTED MANUSCRIPT

1. Introduction

The rapid advances of technology and the widespread use of computers made it possible for digital data to be widely generated, distributed and stored electronically. Such advances, however, pose a serious problem of content protection, hence, the digital data can be easily copied from the original version and widely distributed. Various countermeasures have evolved to prevent and detect such unauthorised modification. Digital Watermarking is one such technique that can be useful to resolve ownership problem of original content (Upadhyay & Singh, 2011). Digital video watermarking is the best approach to hide ownership data in videos in order to preserve the authenticity of the originator or of the content when required.

The present paper covers a software implementation of an improved content-dependent authentication based watermarking scheme for the videos that are compressed by the current H.264/AVC standard (Wiegand, 2003). The watermark bits are embedded within frames which are defined by high motion activities resulting from the best 8×8 mode of the tree-structured hierarchical macroblock (MBs) partitions (Wiegand, 2003). The fragile watermark is generated by applying the Hashed Message Authentication Code-Secure Hash Algorithm-256 (HMAC-SHA-256) (J. Kim, Biryukov, Preneel, & Hong, 2006) on robust visual features that are produced within luminance and chrominance Discrete Cosine Transform (DCT). The authentication and verification processes are blind while the tampering detection is not. Indeed, to achieve the latter, the original frames within the tampered GOP are required. The algorithm is sensitive to spatial, temporal and colour manipulations, and is able to detect the frames tampering.

This paper covers a short overview of existing fragile watermarking schemes for content authentication. Section 2 covers the integrity verification of the compressed H.264/AVC video and followed by the proposed technique in section 3. Section 4 presents the experiments and the analysis of the results and section 5 covers some considerations for further improvements.

ACCEPTED MANUSCRIPT

2. Related works in fragile H.264/AVC video authentication

Most of the H.264/AVC video authentication methods are software based implementations, however, few present hardware implementation (Joshi, Mishra, & Patrikar, 2015; Liu, Chen, Gong, Ji, & Seo, 2015). Our work deals with content authentication software based implementation. The state-of-the-art authentication schemes for H.264/AVC video fall into two broad categories: (1) content-dependent also called digital signature-based authentication where the watermark can be fragile or semi-fragile (Farfoura, Horng, Guo, & Al-Haj, 2015) according to the integrity criteria, and (2) content-independent authentication in which the watermark is only fragile (Le, Nguyen, & Tran, 2014). The authentication based on fragile watermarking, also called hard authentication (Bovik, 2010; Feng, Siu, & Zhang, 2013; Zhu, Swanson, & Tewfik, 2004) is able to find the regions where the content has been modified. However, it is does not distinguish between malicious and un-malicious manipulations. In contrast, the authentication based on semi-fragile watermarking, called soft authentication (Feng et al., 2013), is robust against incidental modification such as compression, but fragile to other modifications. For the hard authentication, the following general requirements of the authentication system need to be met during implementation. These requirements are: (1) Sensitivity - meaning the approach must be able to detect any content modification or manipulation. For fragile authentication algorithms, not only content modification is sought but also the detection of any manipulation;

(2) Localization – meaning the system should be able to locate the altered regions within the frames; (3) Bitrate preserving – implying that the bitrate must be unchanged before and after watermarking; (4) Imperceptibility – entailing that the watermarking method should maintain the quality of the original video; (5) Security– meaning that the watermark of the authentication system should be resistant to any falsification attempts; and (6) Capacity (data payload) – referring to the amount of information that can be embedded within the frames of a video, i.e. the higher the capacity, the better the system.

Fragile authentication is compulsory in critical applications where the alterations can have drastic and costly effects. In our paper, we have targeted fragile watermarking techniques that have been developed for H.264/AVC video sequences. The watermark is embedded in the compressed domain (T. Kim, Park, & Hong, 2012; Kuo, Lo, & Lin, 2008; Qiu, Marziliano, He, & Sun, 2004; Wang & Hsu, 2008; Zhang & Ho, 2006), or directly in the compressed bitstream of the H.264/AVC encoder (Pröfrock, Richter, Schlauweg, & Müller, 2005; Razib, Shirmohammadi, & Zhao, 2007), or simply inserted as Supplemental Enhancement Information (SEI) in the H.264/AVC bitstream (Ramaswamy & Rao, 2006). In the compressed domain, Qiu et al. (Qiu et al., 2004) presented a hybrid watermarking scheme that used DCT coefficients for robust watermarking and motion vectors (*MV*) for fragile watermarking. The authentication scheme is content-independent, which means that the data is a binary sequence independent of the video content. The drawback of these kind of schemes is that the security cannot be guaranteed. The attackers can modify the unmarked coefficients to render the authentication non-operational, or to estimate the watermark from the watermarked video, and then embed it into other videos. Indeed, the authors (Qiu et al., 2004) mentioned that the robustness of their method is limited to transcoding attacks. Zhang et al. (Zhang & Ho, 2006) proposed a new complete fragile watermarking scheme that uses the tree-structured motion compensation, motion estimation (ME) and Lagrangian optimization of H.264/AVC.

ACCEPTED MANUSCRIPT

The watermark embedding is based on the best mode decision strategy for the detection of any spatial or temporal alterations. Their algorithm is able to perform hard authentication in which the tampering contents is detected by the sensitive mode change. However, it is ineffective in locating the attacked areas. This weakness point is also witnessed in the approach presented by Wang et al. in (Wang & Hsu, 2008), where the authors introduced a complete fragile watermarking for H.264/AVC based on fixed watermark independent video content. In their results, the analysis of the recompression and GOP removal attacks ensured the authentication. A video authentication technique for H.264/AVC has been proposed by Kim et al. (T. Kim et al., 2012) where watermark bits are inserted into the *MV* of the inter coded MBs or in the mode number for intra coded MBs achieving a high payload, while ensuring low video quality degradation along with the same required size (in terms of bit rate) of the compressed watermarked bitstream. Kuo et al. (Kuo et al., 2008) presented a fragile video watermarking in H.264 via *MV* in which they extracted random features from 4x4 DCT transform of the previous video frame and embedded them in *MV* at the current video frame. Using the H.264/AVC Rate-Distortion cost function, the best embedding locations are found by a statistical analysis of the *MV*. Moreover, their algorithm achieved the fragility in the case of transcoding while keeping a high visual quality video, however, the spatial and temporal tampering attacks have not been included.

Pröfock et al. (Pröfrock et al., 2005) proposed a blind, fragile and effaceable watermarking approach to ensure that the integrity of the H.264 encoded video bitstream. This watermark is embedded in selected skipped MBs of H.264. They used the hard authentication process and managed to reach low video quality degradations and a low data rate, however, the achieved watermark payload is low. For the same purpose, the method in Razib et al. (Razib et al., 2007) uses MPEG-21 generic Bitstream Syntax Description (gBSD) to perform the adaptation, the encryption, and the authentication of a compressed H.264 video. Using MPEG-21 gBSD requires the preservation and the knowledge of the bitstream as well as the content

ACCEPTED MANUSCRIPT

structure that is stored in the form of metadata. However, the metadata preservation, the decryption processing, the authentication and the adaptation lead to an additional overhead, and more importantly they might be inefficiently slow for the conventional applications such as video conferencing. The authors in (Ramaswamy & Rao, 2006) presented a digital signature-based authentication where the signature, generated from the video content and treated as a watermark, is embedded as SEI in H.264/AVC bitstream. Their algorithm can detect the cause of the authentication failure and can localise the tampered frames, however, it increases the video bitrate that requires an extra bandwidth. A low complexity content-based hard authentication scheme is proposed by Horng in (Horng et al., 2014), in which the scheme concept is to extract fragile features from intra and inter prediction modes of intra and inter MBs. These constitute the authentication code which is embedded and extracted in a GOP within the related elements of the Network Application Layer (NAL) units of the compressed H.264 bitstream. In their approach, a generated content-based key is used to ensure the security of the authentication scheme, and the selection of the last nonzero quantized ac residuals are used to achieve the fragility. The hard authentication process is used. The algorithm achieved low video quality degradations as well as a low data rate, however, the preservation and the knowledge of the bitstream for the features extracted and embedding location is required. New features of H.264/AVC are studied in other algorithms (such as context-adaptive entropy coding, intra prediction mode and the reference index) for watermarking (Xu D & R, 2011; Xu, Wang, & Shi, 2014), but these algorithms are fragile to some common attacks. To provide complete fragile watermarking technique for H.264/AVC, more research is needed. All the above mentioned papers show clearly that the reported algorithms cannot achieve all the requirements of the H.264 video authentication system. Some of the existing techniques should be cleverly combined to improve the performance.

The proposed approach targets the enhancement of our authentication scheme presented in (Ait Sadi, Bouridane, & Guessoum, 2009), and in a further improvement in our work in (Ait

ACCEPTED MANUSCRIPT

Saadi et al., 2010) by fulfilling the maximum number of the mentioned-above requirements, more particularly the bitrate preservation criteria, in which Ramaswamy's technique has shown limitations (Ramaswamy & Rao, 2006). It is worth noting that we have already tackled in (Ait Sadi et al., 2009; Ait Saadi et al., 2010) the impairment of bitrate preservation reported in (Ramaswamy & Rao, 2006) by modifying the embedding process. Indeed, we have shown that the stream size after the entropic coding is preserved and no extra bandwidth for the transmission of the video is required. The embedded watermark process is achieved in the last significant bits (LSB) of the selected *MV*s associated with areas of the highest motion activity.

Compared to our previous work (Ait Sadi et al., 2009), the present work proposes an enhancement of the authentication performance in terms of security and sensitivity by bringing the following modifications: (1) for the watermark generation (as partly looked at in (Ait Saadi et al., 2010)) and (2) the security of the embedding which is reinforced by performing it on two levels. The watermark generation is based on the features of the luminance and chrominance components in order to identify spatial and colour manipulations. In relation to the security improvement, a more secured keyed function HMAC-SHA-256 is adopted at the first level to generate the watermark instead of the hash function MD5 (Ait Sadi et al., 2009) and unkeyed SHA-256 (Ait Saadi et al., 2010). The added second level of security depends on a pseudo random sequence that is used to select the embedding position of the MBs. The selection of the embedding areas, which is achieved by using two fixed thresholds in (Ait Sadi et al., 2009; Ait Saadi et al., 2010), is performed quite differently in this work, where the threshold is dynamically selected based on the highest activity motion. In addition, the keyed watermark bits are substituted into the two last LSBs of the *MV* components instead of one bit as in (Ait Sadi et al., 2009) in order to increase the embedding payload of the 256 bits, which is delivered at the output of HMAC-SHA256 by contrast to the 128 bits delivered by MD5 or unkeyed SHA256 in (Ait Saadi et al., 2010).

ACCEPTED MANUSCRIPT

3. The proposed video authentication system

The embedding process should be performed in a way that it does not increase the video stream‘s file-size, while maintaining the perceptual visual quality. Otherwise, the watermarking becomes a burden for the available bandwidth and the method cannot fulfil the bitrate maintaining requirement. Figure 1 shows the schematic block diagram of the proposed embedding process.

3.1 Content fragile watermark generation

Robust visual features, to which the human eye is sensitive, should be used to ensure the efficiency of authentication. It was demonstrated that robust visual features in the DCT domain (Chang, 2008; Farfoura et al., 2015; Horng et al., 2014; M. E. Farfoura, 2013 ; Q. Sun, He, & Tian, 2006; Weng & Preneel, 2007) are mostly DC coefficients which represent the mean values of every block and carry most of the energy in the block. The robust visual features used for a fragile watermark generation consist of a set of coefficients extracted from INTRA and INTER prediction MBs including INTRA 16×16, INTRA 4×4, and INTER 4×4 prediction MBs for luma component and INTRA 8×8 prediction for chroma components (Wiegand, 2003). The feature data for the watermark generation consists of the quantized DC and the first two lower AC coefficients that are part of the low frequency coefficients in a zig-zag scan order of every block within INTRA 4×4 and INTER 4×4 MBs. The choice of these coefficients is based on: (1) the DC coefficient which is a measure of the average energy over all the 4x4 pixels and on (2) the largest energy contained within the first several low frequency coefficients. The high-frequency coefficients are almost close to zero and thus they are ignored during the quantization of DCT coefficients. In addition to the above and according to (Fridrich, 1999; Upadhyay & Singh, 2011), the DC and the two first AC coefficients are more stable than the other coefficients in the image manipulation. All the non-zero quantized Hadamard transform coefficients from both INTRA 16×16 luma and 8×8 chroma components form the feature data

ACCEPTED MANUSCRIPT

for a MB. These features are then collected in a buffer for each coded MB within each frame until reaching an instantaneous decoder refresh (IDR) mentioning the end of the GOP within H.264/AVC bitstream. The presence of an IDR indicates that there is no further subsequent pictures in the bitstream requiring references in the decoding order to the frames just before the I-frame (Wiegand, 2003). In our experiments, the IDR frames are I-frames which are not referenced by any frames outside the current GOP (Closed GOP). GOP consists of the I-frame and all the other P-frames, which are temporally enclosed between of IDR frames. Therefore, a coded video subsequence begins with an IDR frame and ends when a new IDR frame is received, signalling the availability of a new subsequence to be coded or the end of transmission. At the end of each GOP, the features present in the buffer are hashed using a secure keyed function HMAC-SHA256 (J. Kim et al., 2006) which is created based on the SHA-256 hash function and it is used as a Hash-based Message-Authentication Code (HMAC). The HMAC process initially mingles the features present within the buffer with the chosen secret key *K*. Then it hashes the resulting sequence with the SHA-256, and mixes it again with the secret key *K*, and then it applies the SHA-256 processing. The resulting 256 bits length hash sequence is used as a fragile watermark embedded in the *MV*s of the H.264/AVC.

3.2 Watermark embedding

The watermark embedding is performed on *MV* components within P-frames characterized by high motion activities and belonging to selected MBs. To secure the watermark insertion, the same key *K* used in HMAC-SHA-256 function is employed to generate a pseudo random sequence which is chosen as the embedding position of the MBs. Two constraining aspects have been considered to carefully select the *MV* to be embedded. First, the neighbouring blocks of skipped MBs are discarded because, in the decoder, the *MV* of skipped MBs are derived from motion vector prediction only. Therefore, a motion vector error in the skipped MB, that cannot be compensated, is introduced due to the insertion in the neighbouring blocks. On the other

ACCEPTED MANUSCRIPT

hand, modifying these blocks may increase the video bitrate. The second limitation we found is that we need to avoid insertion in zero motion and low motion blocks. Indeed, any alteration in these blocks will be sensitive to human visual system and results in video bitrate increase. Hence, the *MVs* within the best 8×8 mode (including four sub-block modes chosen from 4×4, 4×8, 8×4 and 8×8) are selected. These selected blocks correspond to the areas with a high motion. In Figure 2, the degradation of the subjective visual quality introduced by the embedding process does not take into account the two restrictions mentioned above.

The frames with high motion activities are picked according to the intensity of the motion activity within each frame. For a given P-frame, the spatial activity matrix is calculated as (X. Sun, Divakaran, & Manjunath, 2001):

$$C = \{MV(i,j)\} \tag{1}$$

and

$$MV(i,j) = \sqrt{(MV_x(i,j))^2 + (MV_y(i,j))^2} \tag{2}$$

where *(i,j)* indicates the block indices within MBs.

For each P-frame, the average of the activity matrix C^{avg} is given as:

$$C^{avg} = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} C(i,j) \tag{3}$$

where *M* and *N* are the width and height of the MBs (in our case, *M=N=16*).

The motion activity of the frame which is defined as standard deviation of motion vector magnitude, is calculated as:

$$\sigma_{Fi} = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (C(i,j) - C^{avg}(i,j))^2} \tag{4}$$

The frames with high motion activity are only chosen if the standard deviation satisfies the following condition:

$$\sigma \geq T \tag{5}$$

ACCEPTED MANUSCRIPT

where *T* is are thresholds which are obtained experimentally.

By applying the proposed selection, it is difficult for the human eye to detect distortions introduced during the watermark embedding process. In the proposed scheme, one unit of length of each component in *MVx* and *MVy* corresponds to ¼ pel. Before being watermarked, each component of the selected *MV is* quantized to the nearest full-pel position. By doing so, the maximum distortion of a *MV* is only ½ pel for each component. To comply with such a scheme, the quantization process is applied as:

$$Q(MVx) = \begin{cases} (2 + MVx)\,\&\,(0xFFFC) & MVx \geq 0 \\ -((2 - MVx)\,\&\,(0XFFFC)) & MVx < 0 \end{cases} \qquad (6)$$

$$Q(MVy) = \begin{cases} (2 + MVy)\,\&\,(0xFFFC) & \mathrm{MVy} \geq 0 \\ -((2 - \mathrm{MVy})\,\&\,(0\mathrm{XFFFC})) & \mathrm{MVy} < 0 \end{cases}$$

where the operator ‘&’ (AND) is used to remove the last LSB bits of both above components.

The watermark bits are then hidden in the last two original LSBs of *MVx* and *MVy* as:

$$\overline{MVx} = \begin{cases} Q(MVx) - \alpha & \text{if } Q(MVx) \geq 0 \text{ and } \alpha = 2 \\ Q(MVx) + \alpha & \text{otherwise} \end{cases} \qquad (7)$$

$$\overline{MVy} = \begin{cases} \mathrm{Q(MVy)} - & \text{if } \quad = 2 \text{ and } \mathrm{Q(MVy)} \geq 0 \\ Q(MVy) + & \text{otherwise} \end{cases}$$

where *α* represents the watermark bits with the values of -1, 0, 1 and 2 corresponding to a pair of bits 11, 00, 01, 10, respectively.

In order to minimize the distortion caused by the inserted watermark, the embedding process must ensure the following synchronization condition:

$$Q(\overline{MVx}) = Q(MVx) \quad and \quad Q(\overline{MVy}) = Q(MVy) \qquad (8)$$

The output *MV* within the P-frame is composed of the blocks in which the motion vector has been watermarked $\overline{MV} = \{\overline{MVx}, \overline{MVy}\}$ and the un-watermarked blocks, leading to the following *MVs* at the P-frame:

ACCEPTED MANUSCRIPT

$$\overline{MV}=\begin{cases}\{\overline{MVx},\overline{MVy}\} & if\ MV\in selected\ MB\\ \{MVx,MVy\} & otherwise\end{cases} \tag{9}$$

Figure 3 outlines the flowchart of the content fragile watermark generation and the embedding process.

3.3 Watermark extraction and Digital signature verification

The extraction process is performed at the H.264/AVC decoder, and the watermark is extracted in a blind fashion implying that the original video sequence is not needed at the extraction. This is a similar to the embedding process. This process includes the computation of the frames with the highest activity motion, the selection of the embedding position of the MBs based on the key *K*, the entropy decoding of *MV* and the application of the two conditions which are performed in the insertion process. The watermark bits are then extracted from the last two bits of LSB of $\overline{MVx}$ and $\overline{MVy}$ components for each $\overline{MV}$.

At the verification stage, the visual features used are the same as in the embedding. These features are extracted before the inverse quantization and the transform operations in H.264/AVC decoder. At the end of each GOP, these features are encrypted using the HMAC-SHA-256 function based on the same key *K* to produce the hash values which are compared with the extracted watermark bits. Any alteration to the features or to the hash values induces in a mismatch as consequence of the secret key *K* which is used in the hashing. Therefore, the content is authenticated only if the original and computed hash values correspond. In case of an authentication failure, the tampered frames detection is performed offline. To determine the tampered frames location within the tampered GOP, the receiver calculates the hash digests of all the frames in the GOP at the decoder level, and it also requests the sender to compute and send the hash values of every frame. Afterward, the receiver compares the hash values frame by

ACCEPTED MANUSCRIPT

frame between the altered and original GOPs. If the hash values of the encoder and decoder of any frame does not match, the decoder indicates the corresponding tampered frames.

4. Experiments and results

Our approach is incorporated in the H.264 JM-10.1 reference software using constrained baseline profile (Suehring). This profile, common to the main and high profiles, remains the core of the H.264 extensions developed by the JVT (Joint Video Team) group. Thus, the input video sequence is configured to produce an I-frame every 15 frames, therefore, generating a GOP with IPPPP… structure. The results obtained are independent of the version of the Joint Model (JM) reference software, since each version of H.264/AVC is an update of the previous one (Ohm & Sullivan, 2013). Consequently, the extension to current version of H.264/AVC software (JM 18.5) only requires the change of the configuration file in the baseline profile (Profile IDC=66) according to the parameters illustrated in Table 1.

To show the effectiveness of the proposed scheme, our experiments are performed on video sequences in QCIF format (YUV 4:2:0) which are commonly used in video processing to assess the fragility of the authentication against spatial, temporal, and colour tampering. The tempering includes frames reordering, frames replacing, spatial cropping, rotation and colour changing. The simulations comprise Akio, Miss America and Claire sequences (Group A) which are characterized by a low spatial detail and a low amount of motion. They also include Foreman, Flowers, Carophone and Table sequences (Group B) that exhibit a medium spatial detail and high amount of motion, or a high spatial detail and a medium amount of motion. The performance of the proposed scheme is evaluated in terms of (1) the maximum payload; (2) the perceptual quality of a watermarked video expressed in terms of Peak Signal-to-Noise Ratio (PSNR), the Video Quality (VQM) and Structural SIMilarity index (SSIM) metrics; (3) file-size preservation; (4) tamper localization capability; and (5) the fragility against incidental distortions.

ACCEPTED MANUSCRIPT

4.1 Effect on Payload

The amount of payload bits or watermark capacity with respect to the video degradation is an important metric in the fragile watermarking. In our proposed algorithm, the watermark capacity per frame is mainly determined by (a) the motion activity and (b) the distortion introduced by the embedding. For the video sequences with significant motion activities (the motion activity of frame $\sigma \geq 3{,}870$), many MBs are allocated with the 8×8 mode. Therefore, a significant amount of *MV*s are produced. However, if the video contains modest motion activities, the number of suitable *MV* decreases and less watermark bits can be embedded into the video sequences. Table 2 illustrates the comparative results between the embedding performed by changing only the last LSB bit of *MV* (Ait Sadi et al., 2009; Ait Saadi et al., 2010).

The proposed scheme shows an improvement in terms of average payload per GOP compared to (Ait Sadi et al., 2009; (Ait Sadi et al., 2009; Ait Saadi et al., 2010) while keeping the subjective visual quality. In addition, the size of the watermarked video file is preserved. The average payload per GOP, given in Table 2, is nearly three times higher for most the sequences belonging to Group B compared to (Ait Sadi et al., 2009). Thus, the motion activity is significant due to the fact that the foreground and the background are both in a moving scene. The number of best 8×8 modes increases resulting in large values of *MV* in each frame that can be watermarked. However, for the video sequences of Group A, despite the fact that the average payload per GOP has practically doubled, the number of *MVs* belonging to the best 8x8 mode is not sufficient to embed the 256 bits of fragile watermark. To achieve the full embedding of the 256 bits in each GOP the sequences that pertain to Group A, we have increased the number of frames within a given GOP to 25 frames as shown in Table 3.

ACCEPTED MANUSCRIPT

4.2 Imperceptibility test

The H.264 reference software JM-10.1 provides users with PSNR which is an objective measure of the visual quality of the compressed video. This metric is then read from the log file built during the execution by the H.264/AVC codec. The PSNR remains almost unchanged as seen in Table 2 for the sequences of group A which includes video clips with low motion and many homogeneous regions (the sequences are mostly static with limited motion in some frames). However, a slight decrease has been noticed in the PSNR values (0.025 to 0.315 dB) for video sequences of Group B with more motions and textured areas. For an objective evaluation, the simulation results of the frame by frame YPSNR performance on the luma of the two video sequences Table and Foreman is illustrated in Figure 4. One can see clearly from this figure that the average PSNR of the luma (Y) samples in all frames of the original and watermarked video sequences is less than 0.268 (dB) for all of the intra frames concluding that the visual quality is preserved.

The software tools VQM and SSIM (MSU, Graphics & Media Lab (Video Group)) are used to measure the temporal and structural perceptual qualities. A value of VQM that is close to 0 means the presence of a smaller distortion. In our experiments, the evaluation has been carried out only on Group B sequences. As shown in Table 4, VQM values are all under 0.4 in the range of 0.292 to 0.398, resulting in an insignificant discrepancy between the original and the watermarked video sequences. It is clear that our approach, by maintain the low values of VQM, provides a good imperceptibility. The SSIM index is generally used in the evaluation of the watermarked videos against the original in terms of similarities or discrepancies by calculating luminance, contrast and structural similarities. It combines these similarities to return a unique value. As reported in the literature it ranges from 0 to 1. A SSIM value close to 1 indicates a high similarity of two videos and 0 is a total discrepancy. Table 4 clearly shows that the obtained values range between the minimum of 0.968 and the maximum of 0.998 with most

ACCEPTED MANUSCRIPT

the remaining values above 0.97. Consequently, there is no perceptual alteration on the video after the embedding process as all values are very close to 1.

4.3 Sensitivity test

The localization capability of the proposed fragile watermarking is assessed on Foreman sequence (Figure 5(a)). in the first experiment, we apply spatial alterations including attacks by DC, cropping and rotation of frames within the watermarked sequence. Figure 5(b) shows the watermarked Foreman sequence, whilst considering the embedding conditions. Figures 5(c) to 5(d) show the results obtained from different tests of spatial alterations. We started by DC attack which is performed by tampering the original coefficients of the 4×4 block while maintaining a constant mean value of the block. The mean value in DCT domain corresponds to the DC component. Figure 5(c) illustrates the noted distortion when the DC attack is carried out on the 52 th MB of the 10th frame within the first GOP. These results show clearly that the change only occurs at the AC values, and that it leads to an imperceptible quality degradation and a digital signature modification. One can see that all embedded signatures are similar to the one extracted from GOPs of the sequence with the exception of the first GOP signature which is different. The modified area is therefore detected and localized in the first GOP. The cropping manipulation, shown in Figure 5(d), consists in deleting the word "SIEMENS" displayed on the background of the 8th frame within the first GOP. The method has also the ability to detect the illegal manipulations such as rotation of the 9th frame (Figure 5(e)). The digital signature within the first GOP (manipulated) differ from the original.

In the second experiment, the sensitivity to temporal attacks is investigated through re-encoding, transcoding, reordering and frame dropping. For the compression attack, the watermarked Table sequence is compressed using MPEG2 standard. As a result, the watermarks within GOPs are completely removed due to the change of the GOP frames' structure and to the alteration of the MVs. To simulate a transcoding attack, the watermarked Table video is

ACCEPTED MANUSCRIPT

recompressed with the same encoding parameters as the original encoder used to watermark the video, apart from the value of the quantization parameter which has been changed to 32. Similarly, the watermarks are damaged as in the previous experiment. Therefore, the proposed algorithm is able to efficiently detect the transcoding attack while keeping the video perception unaltered (Figure 6). Frame dropping and reordering are intentional attacks which take advantage of the temporal redundancy of video sequences to destroy efficiently the embedded watermark, without causing any visual degradation in the video sequence. In our method, every GOP contains an independent watermark to authenticate itself. The reodring attack is reproduced by swapping the 9th and the 10th frmas around within the first GOP of Foreman sequence (Figure 7(a)). The dropping attack is performed by removing the 56th and the 58th frames in the fourth GOP of the same sequence (Figure 7(b)).

The sensitivity of temporal attacks comes from the fact that, in the proposed algorithm, (1) the temporal features are taken into account by including the set of coefficients extracted from INTER 4×4 prediction MBs, and (2) the embedding location is based on the *MV*s of the best 8×8 mode MBs. As a result, any frame manipulation applied to the content or the position of the frame(s) within a GOP leads to a change of the extracted features from INTER 4×4 prediction MBs within the original frame(s), and it also leads to a new mode allocation. Effectively, by applying a small change in the content, the original best 8×8 mode can change to large partition modes and the locations of suitable *MV*s change in the video. Consequently, resulting best mode affects the correct extraction of the embedded watermark within it. This modification is propagated to the following watermark bits which eventually destroys the watermark sequence. Figure 8 illustrates the Normalize Correlation (*NC*) values resulting from temporal attacks where *NC* measures the correlation between the embedded and the original watermarks within each GOP. NC varies from 0 indicating a complete destruction of the embedded watermarks, to 1 meaning that the watermark is unchanged. Our results indicate that

ACCEPTED MANUSCRIPT

the *NC* values are close to zero within the temporal attacked GOPs indicating that the watermark video are completely destroyed.

Figure 9 illustrates the benefit of including the chrominance components as features in the watermark generation. The attack is achieved by changing the colour of the wall behind Foreman in the 7th frame of the sequence, thus resulting in a different signature of the altered GOP.

It is worth noting that the efficiency of the adopted authentication operation for the aforementioned attacks emanates from the hash viable properties. In fact, since the adopted HMAC-SHA-256 hash is a one way process and the probability of yielding the same hash value from two different sets of inputs is close to zero, the hash values produced from the original and the attacked videos are quite different and thus detectable.

4.4 Computational complexity of the watermarking

The computational complexity is related to the processing time required by the different stages in our architecture. For the video encoding part, it encompasses the time spent by the H.264 encoder to generate the watermark of each GOP and the one needed to embed it within the *MV* components. This computational time is evaluated as the difference between the run times with and without watermarking. The simulation has been performed on a personnel computer equipped with 3.2 GHz dual-core CPU and 2GB RAM. Table 5 shows the corresponding average run times. This evaluation has been carried out only on Group B sequences. It can be seen from Table 5 that the encoding time when including the watermarking is increasing by 3.3-11.4 seconds compared to the original video. This slight increase arises from not only the watermark generation but also the identification of the best 8×8 mode and the *MV* components where to insert the watermark.

ACCEPTED MANUSCRIPT

For the video decoding part on the other hand, Table 5 shows that the authentication and the detection of the failure in GOPs is done seamlessly since the additional run time in this case is insignificant. It is worth noting that the additional run time is the time pertaining to the GOP processing. As mentioned earlier, a time penalty is experienced when localizing the tampered frame within the unauthenticated GOP.

4.5 Comparison with previous works

Table 6 summarizes the comparison of the proposed scheme with other fragile authentication techniques using either content-dependent or content-independent watermarking for H.264/AVC codec. First, when comparing the content-dependent watermarking techniques, the scheme in (Ramaswamy & Rao, 2006) requires higher memory resources since it has introduced a larger file-size whereas the proposed technique does not entail any increase in file-size. More importantly, the proposed method achieves excellent performance in terms of video quality. Indeed, as seen in Table 6, the PSNR degradation is evaluated at only 0.315 dB which is perceptually insignificant. Furthermore, in terms of the tampering detection, the proposed scheme exhibits a high sensitivity to spatial, temporal and colour manipulations compared to Kuo et al's (Kuo et al., 2008) method which did not include the effect of temporal and colour alterations. Second, when comparing the current approach with the content-independent watermarking algorithms, it can be noted from Table 6 that even though most of algorithms are sensitive to some spatial manipulations such as transcoding or compression, they either provide no information necessary to characterize the attacks or they have neglected to include some attacks such as spatial tampering (Qiu et al., 2004), temporal and colour tampering (Kuo et al., 2008). By contrast, our approach is able to localize where the authentication failure at GOP level is, while keeping the video quality and the file-size of the compressed bitstream unchanged. The only impairment is the additional time consumption induced by the detection of the failure within the frames of the identified tampered GOP.

ACCEPTED MANUSCRIPT

While the work in (Horng et al., 2014) achieves better performance compared to our proposed approach both in terms of YPSNR and VQM, it implied inconsiderable perceptual distortion because the embedding is performed only on I-frames. On the other hand, our proposed MV-based embedding ensures a better perceptual quality of the video sequences.

Furthermore, the SSIM values of our proposed method span approximately the same interval as in (Horng et al., 2014) and their reported average values are very close to the advantage of a higher payload capacity with our scheme. Indeed, in spite of its ability in preserving the coding efficiency, the work in (Horng et al., 2014) experiences less payload capacity, more particularly with the video sequences of the Group B.

In terms of computational complexity, the method adopted in reference (Horng et al., 2014) allows to attain a more viable performance. This comes from the fact that the different steps of the embedding and detection processes are performed on the bitstream level using the syntactic elements of the NAL. Accordingly, when applying our approach in a similar way, it would show its advantage fullness and bring a valuable complexity reduction.

5. Conclusion

In our paper, an efficient content fragile watermarking scheme of H.264/AVC video for authentication and frames tampering detection is explained. This approach individually authenticates each GOP within the video. An extraction of robust visual features in the H.264/AVC encoder transform's domain are performed in order to create a unique fragile watermark that is afterward inserted into a selected the MB with high intensity motion. The results of this simulation have demonstrated that almost all the criteria of the content authentication are satisfied. This approach not only verifies the authenticity and the integrity of the video, but also maintains the same bitrate in the encoder output. For the video characterized by high motion activities and textured area, the average PSNR value of the watermarked video

ACCEPTED MANUSCRIPT

content decreases by relatively small value of 0.3 dB, while preserving the file-size of the watermarked H.264 stream. Moreover, the scheme has a good sensitivity and provides the detection of spatial, temporal and colour frames tampered within the video. However, the algorithm takes more time to detect the tampered frames. Our future research will focus on reducing the computational complexity by enhancing the run-time of the tampered frames within the unauthenticated GOP to render the algorithm more suitable to real-time applications.

ACCEPTED MANUSCRIPT

References

Ait Sadi, K. A, Bouridane, A., & Gessoum, A. (2010). *H. 264/AVC video authentication based video content.* Paper presented at the I/V Communications and Mobile Network (ISVC), 2010 5th International Symposium on.

Ait Sadi, K., Bouridane, A., & Guessoum, A. (2009). *Combined fragile watermark and digital signature for H. 264/AVC video authentication.* Paper presented at the Signal Processing Conference, 2009 17th European.

Bovik, A. C. (2010). *Handbook of image and video processing*: Academic press.

Chang, L. (2008). *Comparison of Transformed-based Visual features for Automatic Lip Reading* University of Sheffield.

Farfoura, M. E., Horng, S.-J., Guo, J.-M., & Al-Haj, A. (2015). Low complexity semi-fragile watermarking scheme for H. 264/AVC authentication. *Multimedia Tools and Applications*, 1-29.

Feng, D., Siu, W.-C., & Zhang, H. J. (2013). *Multimedia information retrieval and management: Technological fundamentals and applications*: Springer Science & Business Media.

Fridrich, J. (1999). *Robust bit extraction from images.* Paper presented at the Multimedia Computing and Systems, 1999. IEEE International Conference on.

Horng, S.-J., Farfoura, M. E., Fan, P., Wang, X., Li, T., & Guo, J.-M. (2014). A low cost fragile watermarking scheme in H. 264/AVC compressed domain. *Multimedia Tools and Applications, 72*(3), 2469-2495.

Joshi, A. M., Mishra, V., & Patrikar, R. (2015). Design of real-time video watermarking based on Integer DCT for H. 264 encoder. *International Journal of Electronics, 102*(1), 141-155.

Kim, J., Biryukov, A., Preneel, B., & Hong, S. (2006). On the security of HMAC and NMAC based on HAVAL, MD4, MD5, SHA-0 and SHA-1 *Security and Cryptography for Networks* (pp. 242-256): Springer.

Kim, T., Park, K., & Hong, Y. (2012). Video watermarking technique for H. 264/AVC. *Optical Engineering, 51*(4), 047402-047401-047402-047412.

Kuo, T.-Y., Lo, Y.-C., & Lin, C.-I. (2008). *Fragile video watermarking technique by motion field embedding with rate-distortion minimization.* Paper presented at the Intelligent Information Hiding and Multimedia Signal Processing, 2008. IIHMSP'08 International Conference on.

Le, B., Nguyen, H., & Tran, D. (2014). A robust fingerprint watermark-based authentication scheme in H. 264/AVC video. *Vietnam Journal of Computer Science, 1*(3), 193-206.

Liu, S., Chen, D. B.-W., Gong, L., Ji, W., & Seo, S. (2015). A Real-Time Video Watermarking Algorithm for Authentication of Small-Business Wireless Surveillance Networks. *International Journal of Distributed Sensor Networks, 501*, 789536.

M. E. Farfoura, S. J. H., P. Fan, J. Guo, and X. Wang. (2013). Low complexity semi-fragile watermarking scheme for H.264/AVC authentication *Multimedia Tools and Applications, May, 2015.* .

MSU. (Graphics & Media Lab (Video Group)). MSU Quality Measurement Tool: Metrics information *http://compression.ru/video/quality_measure/info_en.html*.

Ohm, J.-R., & Sullivan, G. J. (2013). High efficiency video coding: the next frontier in video compression [Standards in a Nutshell]. *Signal Processing Magazine, IEEE, 30*(1), 152-158.

Pröfrock, D., Richter, H., Schlauweg, M., & Müller, E. (2005). *H. 264/AVC video authentication using skipped macroblocks for an erasable watermark.* Paper presented at the Visual Communications and Image Processing 2005.

Qiu, G., Marziliano, P. A. T. S. H., He, D. J., & Sun, Q. B. (2004). *A hybrid watermarking scheme for H.264/AVC video*. Paper presented at the In Proc. 17th Int. Conf. Pattern Recogn., U.K.

Ramaswamy, N., & Rao, K. (2006). *Video authentication for H. 264/AVC using digital signature standard and secure hash algorithm.* Paper presented at the Proceedings of the 2006 international workshop on Network and operating systems support for digital audio and video.

Razib, I., Shirmohammadi, S., & Zhao, J. (2007). *Compressed Domain Authentication of Live Video.* Paper presented at the Signal Processing and Communications, 2007. ICSPC 2007. IEEE International Conference on.

Suehring, K. JVT JM reference software home page.

Sun, Q., He, D., & Tian, Q. (2006). A secure and robust authentication scheme for video transcoding. *Circuits and Systems for Video Technology, IEEE Transactions on, 16*(10), 1232-1244.

Sun, X., Divakaran, A., & Manjunath, B. (2001). A motion activity descriptor and its extraction in compressed domain *Advances in Multimedia Information Processing—PCM 2001* (pp. 450-457): Springer.

Upadhyay, S., & Singh, S. K. (2011). Video Authentication- An Overview. *International Journal of Computer Science & Engineering Survey (IJCSES), 2*(4), 75-96.

Wang, C.-C., & Hsu, Y.-C. (2008). *Fragile watermarking for H. 264 video stream authentication.* Paper presented at the Intelligent Systems Design and Applications, 2008. ISDA'08. Eighth International Conference on.

Weng, L., & Preneel, B. (2007). *On Encryption and Authentication of the DC DCT Coefficient.* Paper presented at the SIGMAP.

Wiegand, T. (2003). Draft ITU-T recommendation and final draft international standard of joint video specification. *ITU-T rec. H. 264| ISO/IEC 14496-10 AVC.*

Xu D, & R, W. (2011). Watermarking in H.264/AVC compressed domain using Exp-Golomb code words mapping. *Opt Eng, 50*(9), 267-279

Xu, D., Wang, R., & Shi, Y. Q. (2014). Data Hiding in Encrypted H.264/AVC Video Streams by Codeword Substitution. *IEEE Transactions on Information Forensics and Security Vol. 9*(No. 4), 596-606.

Zhang, J., & Ho, A. T. (2006). *Efficient video authentication for H. 264/AVC.* Paper presented at the Innovative Computing, Information and Control, 2006. ICICIC'06. First International Conference on.

Zhu, B. B., Swanson, M. D., & Tewfik, A. H. (2004). When seeing isn't believing [multimedia authentication technologies]. *Signal Processing Magazine, IEEE, 21*(2), 40-49.

ACCEPTED MANUSCRIPT

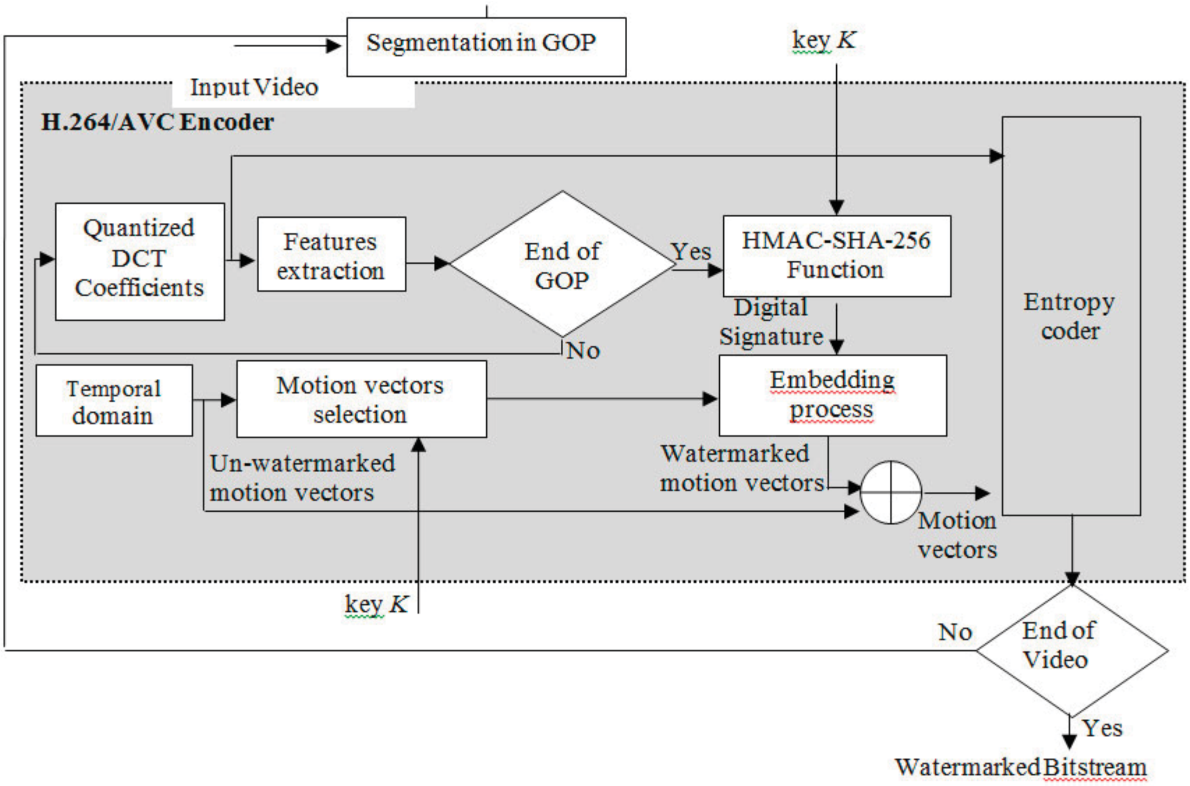


Figure 1. Schematic block diagram of the proposed embedding process.

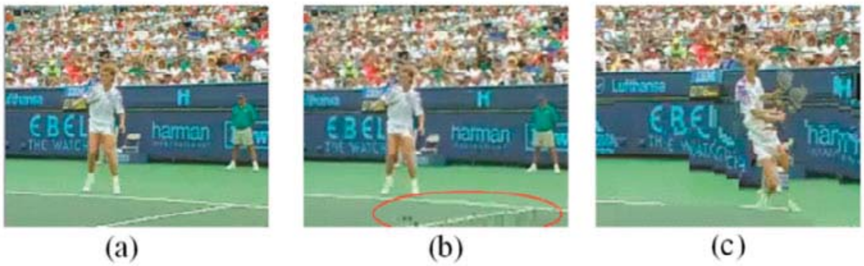


Figure 2. Visual quality results without respecting the two first conditions of insertion (a) original frame from Table sequence (b) Insertion in *MV* belonging to 16×16 partition mode (c) Insertion in *MV* belonging to 8x8 partition mode without observing the two restrictions.

ACCEPTED MANUSCRIPT

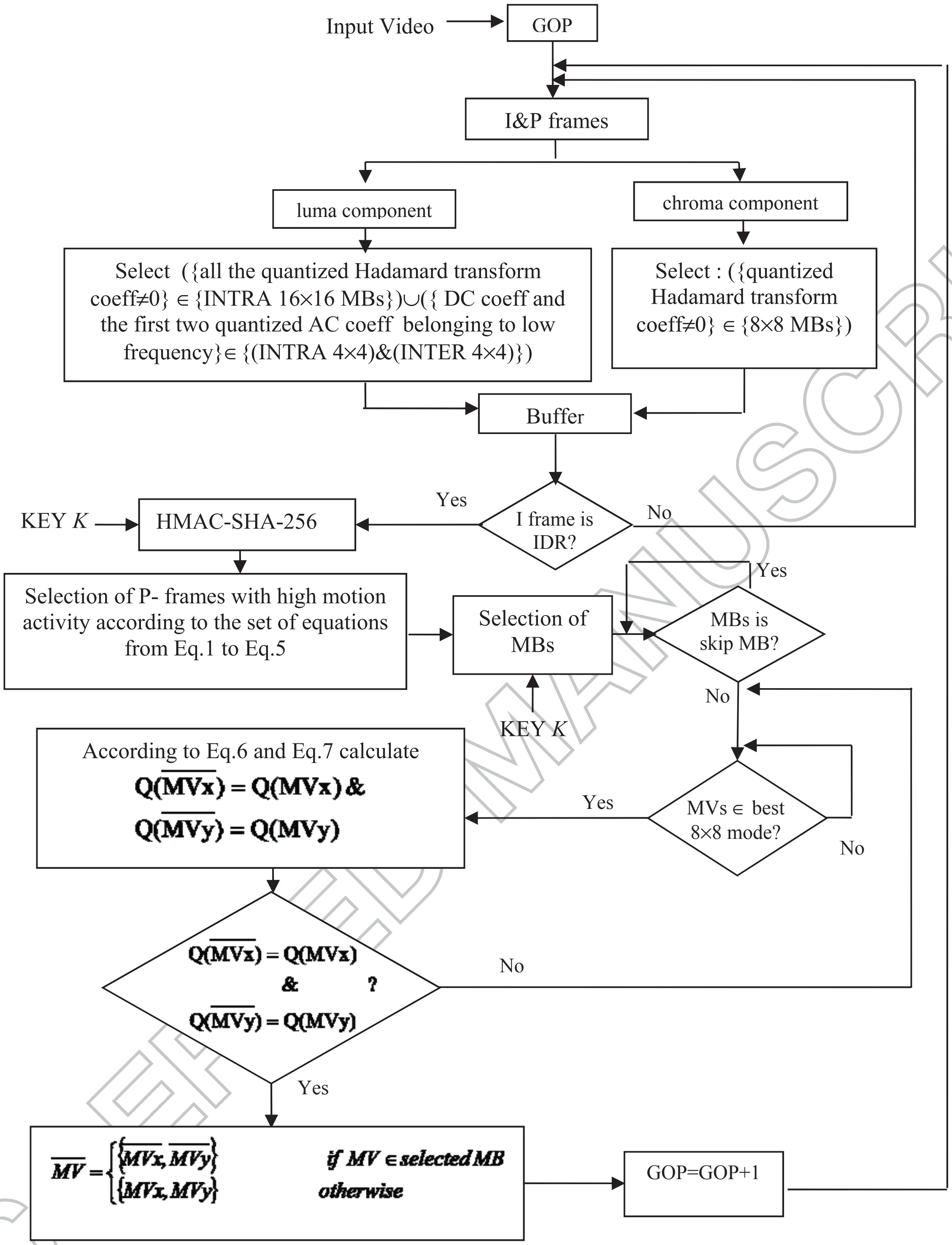

Input Video
GOP
I&P frames
luma component
chroma component
Select ({all the quantized Hadamard transform coeff≠0} ∈ {INTRA 16×16 MBs})∪({ DC coeff and the first two quantized AC coeff belonging to low frequency}∈ {(INTRA 4×4)&(INTER 4×4)})
Select : ({quantized Hadamard transform coeff≠0} ∈ {8×8 MBs})
Buffer
Yes
I frame is IDR?
No
KEY K
HMAC-SHA-256
Selection of P- frames with high motion activity according to the set of equations from Eq.1 to Eq.5
Selection of MBs
KEY K
Yes
MBs is skip MB?
No
According to Eq.6 and Eq.7 calculate
$Q(\overline{MVx}) = Q(MVx)$ &
$Q(\overline{MVy}) = Q(MVy)$
Yes
MVs ∈ best 8×8 mode?
No
$Q(\overline{MVx}) = Q(MVx)$ & $Q(\overline{MVy}) = Q(MVy)$?
No
Yes
$\overline{MV} = \begin{cases} \{\overline{MVx}, \overline{MVy}\} & \text{if } MV \in selected\,MB \\ \{MVx, MVy\} & otherwise \end{cases}$
GOP=GOP+1

Figure 3. Flowchart of the proposed embedding process.

ACCEPTED MANUSCRIPT

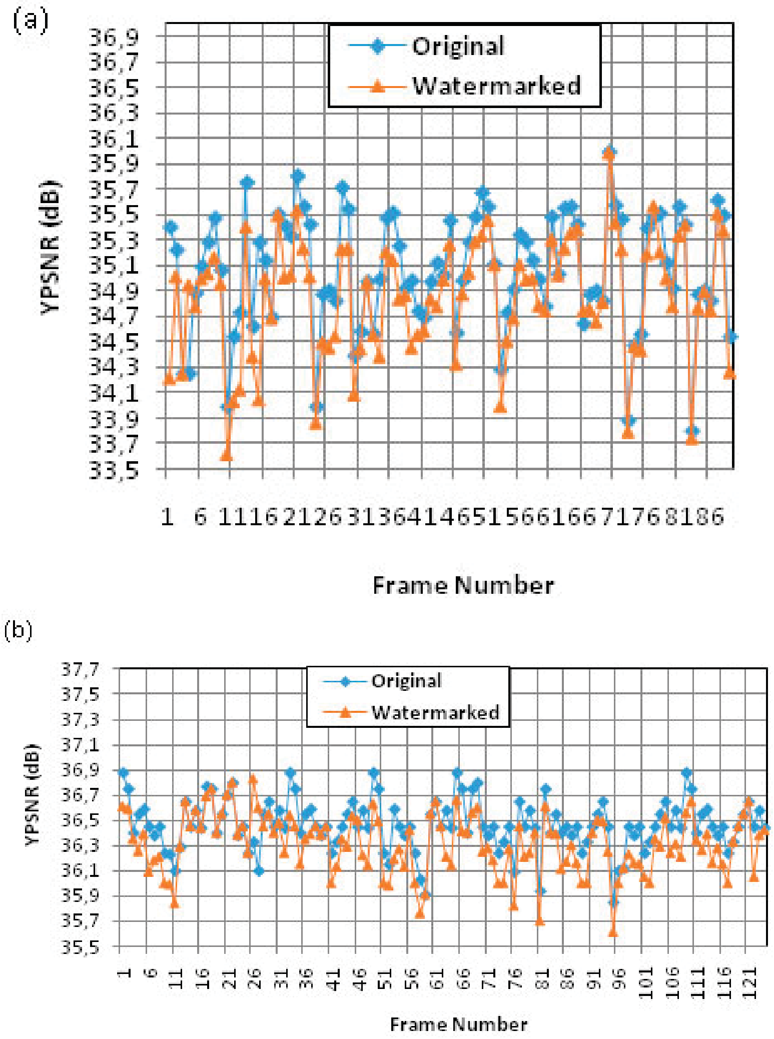


Figure 4. Frame by-frame YPSNR of the original and watermarked sequence (a) Table, (b) Foreman with QP=28.

ACCEPTED MANUSCRIPT

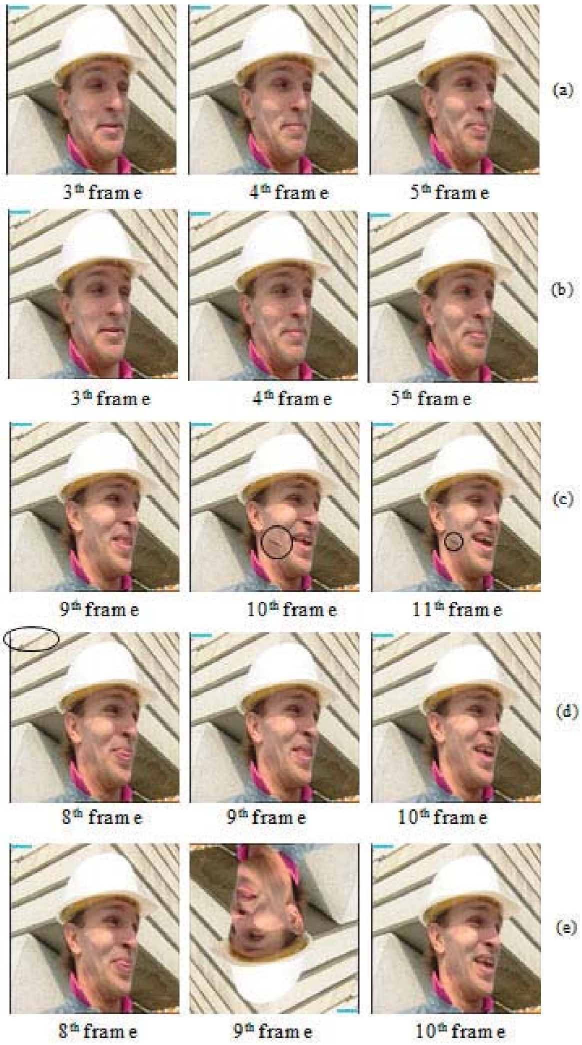


Figure 5. Foreman frames within the 1st GOP: (a) Original frames, (b) watermarked frames, (c) DC attack, (d) cropping attack and (e) rotation attack.

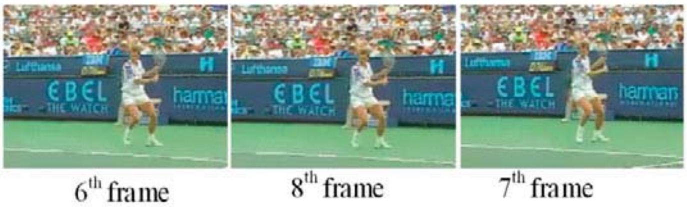
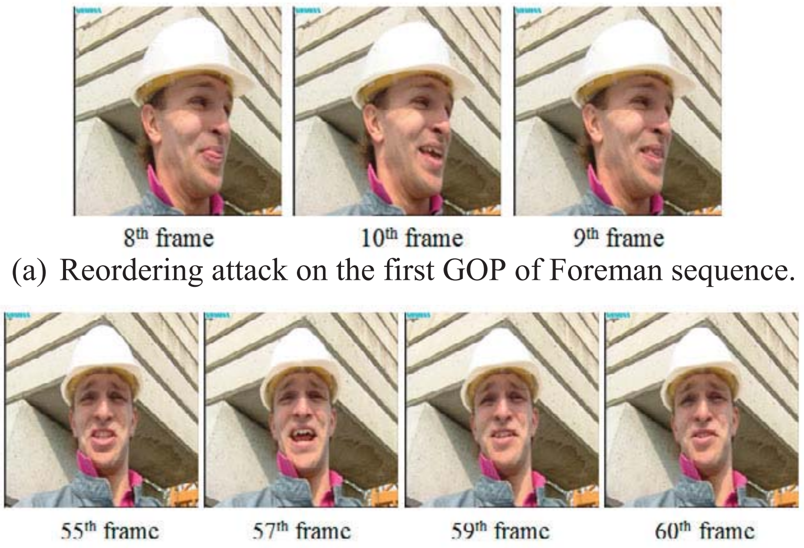


Figure 6. Table Frames within the 1st GOP after transcoding attack.

ACCEPTED MANUSCRIPT



(a) Reordering attack on the first GOP of Foreman sequence.

(b) Foreman frames after dropping the 56th and 58h frames within the fourth GOP

Figure 7. Reordering and dropping attacks on the first and Fourth GOPs of Foreman sequence.

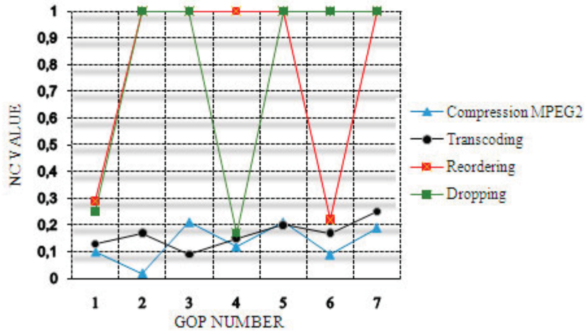


Figure 8. Sensitivity to temporal attacks applied on Table sequence video.

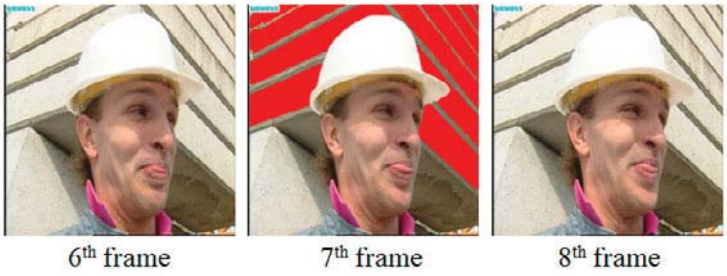


Figure 9. Colour attack applied on the 7th frame of Foreman sequence.

ACCEPTED MANUSCRIPT

Table 1. Configuration parameters of the encoder

Profile	**Baseline (Profile IDC=66)**
Number of frames	150 for all test sequences except for Table which includes 88 frames
Frame rate	30 fps
Source Bit Depth luma	8
Source Bit Depth chroma	8
Motion estimation	Full Search
RD optimization	0
Entropy coding	CAVLC
Search range	16
Quantization parameter	28
Intra period	15, only the first frame is intra

Table 2. Simulation results of payload and PSNR for sequences belonging to the two groups with $\sigma \geq 3{,}870$ and comparison with the work in (Ait Sadi et al., 2009).

		Average Payload per GOP(bits)		**PSNR (dB)**					**The file-size of the video stream (MBytes)**	
	Video sequence	**proposed approach**	**Work in (Ait Sadi et al., 2009)**	**Original video**	**Proposed video Watermarking**	**Difference between original and proposed watermarked video**	**Work in (Ait Sadi et al., 2009)**	**Difference between proposed work and (Ait Sadi et al., 2009)**	**Original video**	**Proposed video Watermarking**
Group A	Miss America	162	88	40.056	40.056	0	40.04	0.016	5.43	5.43
	Claire	238	135	39.681	39.681	0	39.67	0.011	5.43	5.43
	Akiyo	248	133	38.205	38.205	0	38.17	0.035	5.43	5.43
	Bridge-close	518	264	34.847	34.847	0	34.85	0.03	5.43	5.43
Group B	Carphone	2960	1495	37.340	37.315	0.025	37.29	0.025	5.43	5.43
	Coastguard	5334	2806	34.181	34.026	0.155	34.02	0.006	5.43	5.43
	Flower	8578	4304	34.336	34.308	0.028	34.31	0.02	5.43	5.43
	Foreman	7688	2569	36.687	36.45	0.237	35.75	0.3	5.43	5.43
	Suzie	2760	1349	37.357	37.141	0.216	37.12	0.021	5.43	5.43
	Table	9992	5182	35.23	34.915	0.315	34.91	0.05	3.22	3.22

Table 3. Simulation results of payload and PSNR for sequences belonging to group A with 25 Frames within GOP ($\sigma \geq 3{,}870$).

Video sequence Of Group A	**Average Payload per GOP(bits)**	**PSNR (dB)**	**The file-size of the video stream (MByte)**
Miss America	265	40.056	5.43
Claire	283	39.681	5.43
Akiyo	301	38.205	5.43
Bridge-close	555	34.847	5.43

ACCEPTED MANUSCRIPT

Table 4. Simulation results of VQM and SSIM.

	Videos	VQM	SSIM		Videos	VQM	SSIM
Group A	Miss America	0.192	0.998	Group B	Carphone	0.292	0.987
	Claire	0.206	0.978		Coastguard	0.234	0.998
	Akiyo	0.187	0.995		Flower	0.290	0.995
	Bridje close	0.245	0.981		Foreman	0.351	0.976
					Suzie	0.289	0.983
					Table	0.398	0.968

Table 5. Run time video coding for sequences belonging to Group B

Video Sequences Group B	Run-time video coding (sec)			Run-time video decoding (sec)		
	Without watermarking	With watermarking	Difference	Without watermarking	With watermarking	Difference
Carphone	21.60	24.86	3.26	15.1	15.31	0.21
Coastguard	22.21	26.31	4.10	16.2	16.44	0.22
Flower	23.39	30.33	6.94	16.47	16.73	0.26
Foreman	30.48	35.54	5.06	17.6	17.78	0.18
Table	32.80	44.20	11.4	17.81	18.05	0.24

ACCEPTED MANUSCRIPT

Table 6. Comparison of the proposed scheme with other hard authentication techniques for H.264/AVC codec.

Parameters	(Qiu et al., 2004)	(Zhang & Ho, 2006)	(Wang & Hsu, 2008)	(T. Kim et al., 2012)	(Kuo et al., 2008)	(Ramaswamy & Rao, 2006)	(Horng et al., 2014)	Proposed method
Type of watermarking	content-independent				Content-dependent (Features extracted in transform Domain)			
Features extraction					luma	luma	luma	luma & chroma
Type of data of every block					AC Coeffs.	DC & AC coeffs. of luma	nonzero quantized ac residuals	DC & AC coeffs. of luma & chroma
Embedding space	Motion Estimation	AC coeffs. (I frame)	AC coeffs. (P & B frames)	MV for inter-coded MBs or on the mode number for intra-coded MBs	Motion vectors	SEI (Supplemental Enhancement Information) of GOPs	I4-blocks of I-frames	Motion vectors
Payload	1 bit /MBs	Low (150 bits)	Low (400 bits)	High	Not mentioned	Signatures for all video (y*160) bits y : total number of GOP in video	High	High : Signature for all video (y*256) bits y : total number of GOP in video
Increase in file-size	None	None	+ 1%	+1% ~2%	+ 4%	Large	None	None
Detect spatial manipulations	no	yes	Only GOP remove & re-encoding	Frames removal & recompression	transcoding	yes	yes	yes
Detect temporal manipulations	no	no	no	no	Not mentioned	no	yes	yes
Detect color manipulations	no	no	no	no	Not mentioned	no	no	yes
Detect tampered location	no	Yes (Frame level)	Yes (Frame level)	no	Yes (Frame level)	Yes (Frame level)	yes	Yes (Frame level)
PSNR	Small degradation	Maintained	- 0.12dB	-0.06 dB	- 0.27 dB	Maintained	--0.005	- 0.3 dB
Computational complexity	Low	Low	Low	High	Not mentioned	Low	Low	High (in tampered frames locations)

ACCEPTED MANUSCRIPT