Fast compressed domain watermarking of MPEG multiplexed streams

D. Simitopoulos, S. A. Tsaftaris, N. V. Boulgouris, and M. G. Strintzis

Electrical and Computer Engineering Department
Aristotle University of Thessaloniki
Thessaloniki 54006, Greece

Informatics and Telematics Institute
1st Km Thermi-Panorama Road
Thermi-Thessaloniki 57001, Greece

***Abstract*--In this paper, a new technique for watermarking of MPEG compressed video streams is proposed. The watermarking scheme operates directly in the domain of MPEG multiplexed streams. Perceptual models are used during the embedding process in order to preserve the quality of the video. The watermark is embedded in the compressed domain and is detected without the use of the original video sequence. Experimental evaluation demonstrates that the proposed scheme is able to withstand a variety of attacks. The resulting watermarking system is very fast and reliable, and is suitable for copyright protection and real-time content authentication applications.**

I. Introduction

In parallel with the development and the introduction of Digital Versatile Disc (DVD) as the ultimate medium for the digital storage and distribution of audiovisual content, the MPEG-2 standard was established as the coding scheme for such content. These developments made the large-scale distribution and replication of multimedia very easy but at the same time also to a large extent uncontrollable. In order to protect multimedia content from unauthorized trading, many digital watermarking techniques have been introduced. However, very few of them deal with the very important issue of compressed domain watermarking for video [1, 2].

The important practical problem of watermarking MPEG multiplexed streams has not been addressed, up to date in the literature. Multiplexed streams contain at least two elementary streams, an audio and a video elementary stream. Thus, it is necessary to develop a watermarking scheme that operates with multiplexed streams as input.

In all watermarking systems the watermark is required to be imperceptible and robust against attacks such as compression, cropping, filtering, etc [3]. Apart from the above, video watermarking systems have additional requirements, such as fast embedding and detection, blind detection and file size preservation after the watermark is embedded. Drift error should also be avoided or compensated when watermark embedding is performed.

In this paper, a novel compressed domain watermarking scheme is presented which is suitable for MPEG multiplexed streams (MPEG-1 system and MPEG-2 program streams). Embedding and detection are performed without fully de-multiplexing the video stream. During the embedding process, the data that are going to be watermarked are extracted from the stream, watermarked and placed back into the stream. This approach leads to a fast implementation which is necessary for real-time applications and also when a large number of video-sequences have to be watermarked, as is the case in video libraries. The detection is also fast so that it can be incorporated to real-time content authentication systems. The resulting watermarking scheme is shown to withstand transcoding, as well as cropping and filtering.

II. Preprocessing Of Mpeg Multiplexed Streams

For several reasons, it is often preferable to watermark video in the compressed rather than the spatial domain. In fact, it is very often impractical (due to high storage capacity requirements) or indeed entirely not feasible to decompress and then recompress the entire video data. In addition, decoding and re-encoding an MPEG stream would also significantly increase the processing time, perhaps even prohibiting it from being used in real-time applications. For these reasons, the video watermark embedding and detection methods of the present paper are carried out entirely in the compressed domain.

MPEG multiplexed streams contain at least two elementary streams i.e. an audio and a video elementary stream. An obvious approach of watermarking MPEG multiplexed streams would be to de-multiplex the stream, then watermark the video data and finally multiplex the two elementary streams. The above process however is extremely costly computationally and very slow in implementation.

In order to achieve low complexity, a technique was developed that does not fully de-multiplex the stream before the watermark embedding, but instead deals with the multiplexed stream itself. Specifically, first the video elementary stream packets are detected in the multiplexed stream. For the video packets that contain I-frame data, the encoded video data are extracted from the video packets and variable length decoding is performed in order to obtain the quantized DCT coefficients. The headers of these packets are left intact. This procedure is schematically described in Fig. 1. The matrix of quantized

DCT coefficients is parsed to the watermark embedder. Then the watermarked coefficients are variable length coded. The video encoded data are partitioned so that they can fit into video packets that use the original headers. Audio packets and packets containing interframe data are not altered. Basically, the stream structure remains unaffected and only the video packets that contain coded I-frame data are altered.

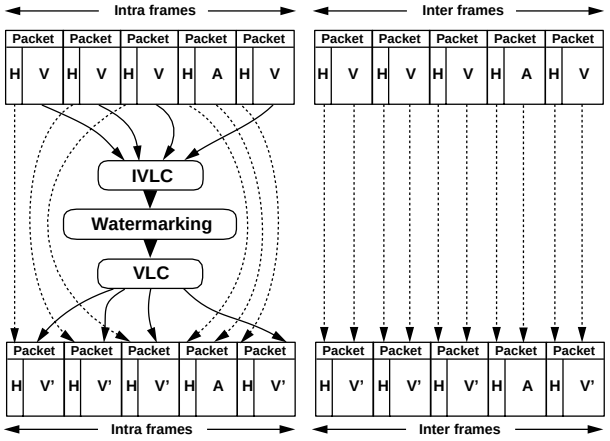

Fig. 1 Operations of the proposed scheme performed on a MPEG multiplexed stream (V: encoded video data, A: encoded audio data, H: elementary stream packet header, Packet: elementary stream packet, V': watermarked encoded video data).

III. Compressed Domain Watermark Embedding and Detection

A. Imperceptible watermark embedding in the quantized DCT domain

The values of the embedding watermark sequence W are either –1 or 1. This sequence is produced from an integer random number generator by setting the watermark coefficient to 1 when the generator outputs a positive number and by setting the watermark coefficient to -1 when the generator output is negative. The result is a zero mean, unit variance process. The random number generator is seeded with the result of a hash process. This process consists of the MD5 algorithm [4] that produces a 128 bit key from a meaningful message (owner ID) and another hash function that accepts the 128 bit key as input and produces a 32 bit integer suitable for seeding.

The proposed watermark embedding scheme (see Fig. 2) alters only the quantized AC coefficients $X_{\kappa,\lambda}(m,n)$ (where κ is the index of the current macroblock, λ is the index of the block within the current macroblock and m, n are indices indicating the position of the current coefficient in an 8x8 DCT block) of luminance blocks of I-frames and leaves the chrominance blocks unaffected. In order to make the watermark as imperceptible as possible, perceptual analysis [5] and block classification techniques [6] are combined as in [7]. These are applied in the DCT domain in order to select which coefficients are best for watermarking. For each selected coefficient in the DCT domain, the product of the embedding watermark coefficient $W_{\kappa,\lambda}(m,n)$ with the corresponding parameters that result from the perceptual analysis (embedding mask $M_{\kappa,\lambda}(m,n)$ and block classification (classification mask $C_{\kappa,\lambda}(m,n)$ is added to the corresponding quantized coefficient:

$$X'_{Q\kappa,\lambda}(m,n) = X_{Q\kappa,\lambda}(m,n) + \\ + C_{\kappa,\lambda}(m,n) M_{Q\kappa,\lambda}(m,n) W_{\kappa,\lambda}(m,n) \qquad (1)$$

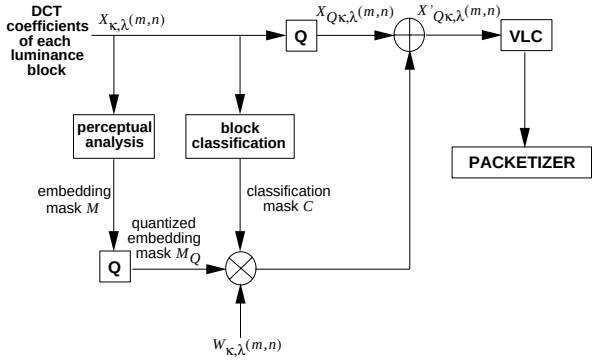

Fig. 2 Watermark embedding scheme.

If the watermark were embedded in the DCT coefficients, the quantization process would possibly eliminate it. This case applies for a large number of DCT coefficients, making the detection process unreliable. Thus, in order to avoid reduced detection performance due to MPEG quantization, the watermark is embedded in the quantized DCT coefficients, since the MPEG coding algorithm performs no other lossy operation after quantization. Therefore any information embedded as in Fig. 3 does not run the risk of being eliminated by the subsequent processing. Thus, the watermark exists intact in the quantized coefficients when the detection process is carried out and the quantized DCT coefficients $X_{Q\kappa,\lambda}(m,n)$ are watermarked as in (1).

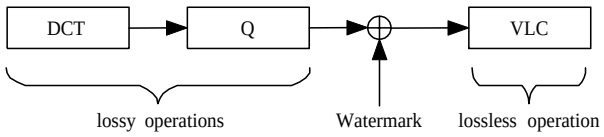

Fig. 3 MPEG encoding operations.

B. Watermark detection

The detection of the watermark is performed *without* using the original data. The original meaningful message that produces the watermark sequence W is needed in order to check if the specified watermark sequence exists in a copy of the watermarked video. Then, a correlation based detection approach is taken, similar to that analyzed in [3].

Variable length decoding is first performed to obtain the quantized DCT coefficients. The block classification and perceptual analysis procedures are performed as in embedding in order to define the set $\{X\}$ of the N coefficients that are expected to be watermarked with the sequence W. Each coefficient in the set $\{X\}$ is multiplied

with the corresponding watermark coefficient $W_{\kappa,\lambda}(m,n)$ resulting to the data set $\{X_W\}$.
The statistical characteristics (mean and variance) of the data set $\{X_W\}$ are calculated as follows

$$mean = E\{X_W\} = \frac{1}{N}\sum_{l=0}^{N-1} X_W(l) \quad (2)$$

$$\text{var}\,iance = E\left\{\left(X_W - mean\right)^2\right\} =$$

$$= \frac{1}{N}\sum_{l=0}^{N-1}\left(X_W(l) - mean\right)^2 \quad (3)$$

Finally, the statistical correlation metric c for each frame is calculated as

$$c = \frac{mean\sqrt{N}}{\text{var}\,iance} \quad (4)$$

The correlation metric c is compared to the threshold T_c, which is an adaptive threshold calculated for each frame of the video sequence as in [3]. In [3] the authors set the threshold T_c to the half of the mean value of the random variable c. If the correlation metric c exceeds the threshold T_c, the examined frame is considered watermarked.

IV. Detector Implementation

The proposed correlation based DCT domain detection described in Section III.B can be implemented using two types of detectors.
The first detector (*detector-A*) assumes that the sequence under examination is the original watermarked sequence or has the same GOP structure with the original watermarked sequence but is encoded at a different bit-rate using one of the techniques proposed in [8]. Therefore, this detector simply detects the watermark only in I-frames during their decoding by applying the procedure as it is described in Section III.B. The detection is very fast due to the fact that it actually introduces negligible additional computational load to the decoding operation. This enables the proposed system to be used not only for copyright protection but also to be incorporated in compliant to this scheme real-time decoders/players that accommodate immediate content authentication.
The second detector (*detector-B*) assumes that the GOP structure may have changed due to transcoding and frames that were previously coded as I-frames may now be coded as B- or P-frames. This detector decodes and partially re-encodes each frame in order to detect the watermark using the procedure described in Section III.B. The decoding operation performed by this detector may also consist of the decoding of non-MPEG compressed or uncompressed video streams.
In the case that transcoding and I-frame skipping are performed on an MPEG video sequence, then *detector-B* will try to detect the watermark in frames that were previously coded as B- and P-frames. If the motion of the objects in the scene is not intense or slow camera zoom or pan has occurred, then the watermark will be detected in B- and P-frames. On the contrary, if this assumption is reversed, the watermark may not be detected in any of the video frames but, in this case, the quality of the transcoded video would be highly decreased due to frame skipping (jerkiness in scenes will be created or visible motion blur will appear if interpolation is used) and it is very unlikely that an attacker will use such an attack.

V. Experimental Evaluation

The video sequence used for the experiments was the MPEG-2 video *sportnews*, which contains two fast motion scenes and is part of a TV broadcast. This is a MPEG-2 program stream i.e. multiplexed stream that contains video and audio. It was produced using an OPTIBASE hardware MPEG-1/2 encoder from a PAL VHS source. The use of such test video sequences in place of most commonly used sequences like *table tennis* or *foreman* was that the latter are short video-only sequences that are not multiplexed with audio streams, as is the case in practice. Therefore, in order to test the overall performance of the system all experiments were made using the multiplexed sequence *sportnews*, although the system also supports video-only MPEG-1/2 streams. In general, the embedding and detection scheme supports constant and variable bitrate main profile MPEG streams. Our software implementation makes use of the MPEG Software Simulation Group (MSSG) software codec.

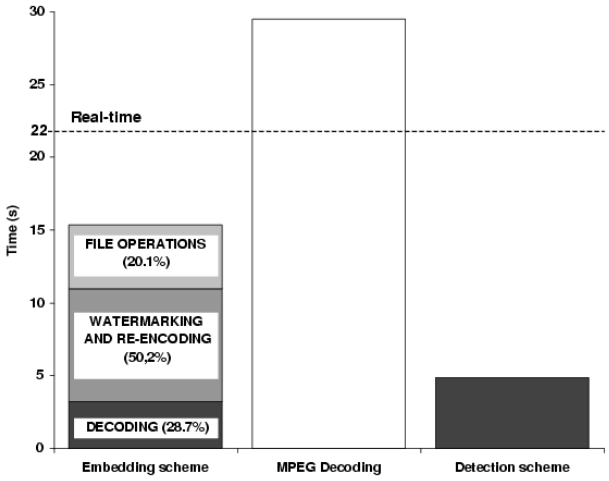

Fig. 4 Speed performance of the embedding and detection schemes.

A software simulation of the proposed embedding algorithm was implemented and executed using a Pentium III 800 MHz processor. The total execution time of the embedding scheme for the 22 sec MPEG-2 (5 Mbit/sec, PAL resolution) video sequence *sportnews* is 72\% of the real-time duration of the video sequence. Execution time is allocated in the three major operations performed for embedding: file operations (read, write headers and packets), partial decoding and partial encoding and

watermarking as shown in Fig. 4. In Fig. 4 the embedding time is also compared to the decoding time (without saving each decoded frame to a file) using the MSSG decoder software. Clearly, the embedding time is significantly shorter than the decoding and re-encoding time that would be needed in case the watermark embedding was performed in the spatial domain. Fig. 4 also presents the time required for detection using the (*detector-A*) described in Section IV. Detection time (partial I-frame decoding and detection) is only 23\% of the real-time duration of the video sequence, thus enabling the detector to be incorporated in real-time decoders/players.

In Fig. 5 the correlation metric is evaluated for 576 consecutive frames of the MPEG-2 video sequence *sportnews* using *detector-B*. As seen, the correlator output exceeds the adaptively calculated threshold for all I-frames.

In our experiments various attacks were directed on the watermarked MPEG-2 video *sportnews* with the purpose of rendering the watermark undetectable. The watermarked video underwent the following attacks: low-pass filtering, clipping, median filtering, and transcoding (full decoding and re-encoding to other bitrates). In all cases the watermark was successfully detected in all I-frames.

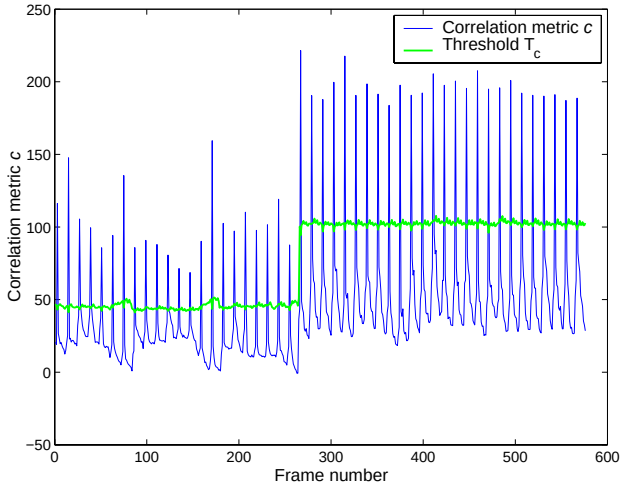

Fig. 5 Correlation metric plot for all frames of the MPEG-2 video *sportnews*.

VI. Conclusions

A novel and robust way for embedding watermarks in MPEG-1/2 multiplexed streams was presented. The proposed scheme operates directly in the compressed domain and is able to embed copyright information without causing any degradation to the quality of the video. Due to its speed, the resulting scheme is suitable for real-time content authentication applications. Experimental evaluation showed that the proposed watermarking scheme is able to withstand a variety of attacks.

VII. Bibliography

[1] G. C. Langelaar, R. L. Lagendijk, and J. Biemond, "Real-time labeling of MPEG-2 compressed video," *J. Visual Commun. Image Representation*, vol. 9, no. 4, pp. 256–270, Dec. 1998.

[2] F. Hartung and B. Girod, "Watermarking of Uncompressed and Compressed Video," *Signal Processing*, vol. 66, no. 3, pp. 283–301, May 1998.

[3] W. Zeng and B. Liu, "A Statistical Watermark Detection Technique Without Using Original Images for Resolving Rightful Ownerships of Digital Images," *IEEE Trans. Image Processing*, vol. 8, no. 11, pp. 1534–1548, Nov. 1999.

[4] B. Schneier, Applied Cryptography: Protocols, Algorithms, and Source Code in C, 2nd Edition, John Wiley & Sons, 1995.

[5] A. B. Watson, "DCT quantization matrices visually optimized for individual images," in Proc. SPIE Conf. Human Vision, Visual Processing and Digital Display IV, Feb. 1993, vol. 1913, pp. 202–216.

[6] T.-Y. Chung, M.-S. Hong, Y.-N. Oh, D.-H. Shin, and S.-H. Park, "Digital watermarking for copyright protection of MPEG-2 compressed video," IEEE Trans. Consumer Electronics, vol. 44, no. 3, Aug. 1998.

[7] D. Simitopoulos, S. A. Tsaftaris, N. V. Boulgouris, and M. G. Strintzis, "Digital watermarking of MPEG-1 and MPEG-2 streams for copyright protection," in Proc. Second USF International Workshop on Digital and Computational Video (DCV'01), Feb. 2001.

[8] A. Eleftheriadis and D. Anastasiou, "Constrained and general dynamic rate shaping of compressed digital video," in ICIP, Oct. 1995, vol. 3, pp. 396–399.