

DIGITAL WATERMARKING OF MPEG-1 AND MPEG-2 MULTIPLEXED STREAMS FOR COPYRIGHT PROTECTION

D. Simitopoulos S. A. Tsiftaris N. V. Boulgouris M. G. Strintzis

Informatics and Telematics Institute
1, Kyvernidou str.
Thessaloniki 546 39, Greece
e-mail: dsim@iti.gr

ABSTRACT

In this paper, a new technique for watermarking of MPEG-1/2 compressed video streams is proposed. The watermarking scheme operates directly in the domain of MPEG-1/2 program streams. Perceptual models are used during the embedding process in order to preserve the quality of the video. The detection of the watermark is performed in the compressed domain without requiring the original video. 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 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 and uncontrollable. In order to protect multimedia content from their unauthorized trading, digital watermarking techniques have been introduced. These techniques embed the digital signature of the copyright holder in an image [1], audio or video [2] signal. In the case of unauthorized copying of the multimedia data, the copyright holder can prove the ownership of the data.

Numerous watermarking techniques have been proposed. However, only few of them treat the problem of compressed domain watermarking for images and video [3]. Especially for video, watermarking of MPEG-1/2 program streams is almost ignored by most approaches. These streams are multiplexed streams that contain at least two elementary

This work was supported by the EU IST Project "ASPIS" and the Greek GSRT Project "PANORAMA".

streams, an audio and a video elementary stream. Thus, it is very important to develop a watermarking scheme that works not only in the compressed domain but also with multiplexed streams as input to the scheme.

In this paper, a compressed domain watermarking scheme is presented which is able to perform watermark embedding in MPEG-1/2 program streams. The resulting scheme is fast and suitable for real-time content authentication applications.

II. VIDEO WATERMARKING SYSTEM REQUIREMENTS

In all watermarking systems the watermark should be imperceptible and it should also be robust against attacks such as compression, cropping, filtering and geometric transformations. Apart from the above, video watermarking systems have the following additional requirements:

- **Compressed domain processing.** A video watermarking system must operate in the compressed domain since most of the times it is impractical or infeasible to decompress and then recompress the video data.
- **Fast embedding/detection.** A video watermarking system must be very fast since video data are usually huge.
- **Blind detection.** The system should not use the original video for the detection of the watermark. This is very important not only because of the concerns raised in [4] about using the original data in the detection process, but also because it is impractical to keep all original sequences in addition to the watermarked ones.

In the ensuing sections, an MPEG-1/2 watermarking system is described which meets the above requirements.

III. PREPROCESSING OF MPEG-1/2 PROGRAM STREAMS

MPEG-1/2 streams are multiplexed streams that contain at least two elementary streams i.e. an audio and a video elementary stream. The video watermarking system should be able to cope with multiplexed streams. It should be able to discern one elementary stream from the other, watermark the video stream (or possibly also the audio stream), and then multiplex again the two elementary streams in one stream. While these operations are performed, the original stream size should undertake only minor alterations (a few bytes). Taking into account all these, a compressed domain watermarking system operating with multiplexed streams was developed.

An obvious approach to MPEG-2 watermarking would be to use the following procedure. The original stream is de-multiplexed to its constituent elementary streams (video and audio). Subsequently the video elementary stream is processed in order to embed the

watermark. Finally the resulting watermarked video elementary stream and the audio elementary stream are again multiplexed to produce the final MPEG-2 stream. The above process is unacceptable in terms of complexity and speed.

The demand for low complexity leads us to pursue a technique that does not fully de-multiplex the stream before the watermark embedding, but instead deals with the multiplexed stream itself. First the video elementary stream packets are detected in the mul-

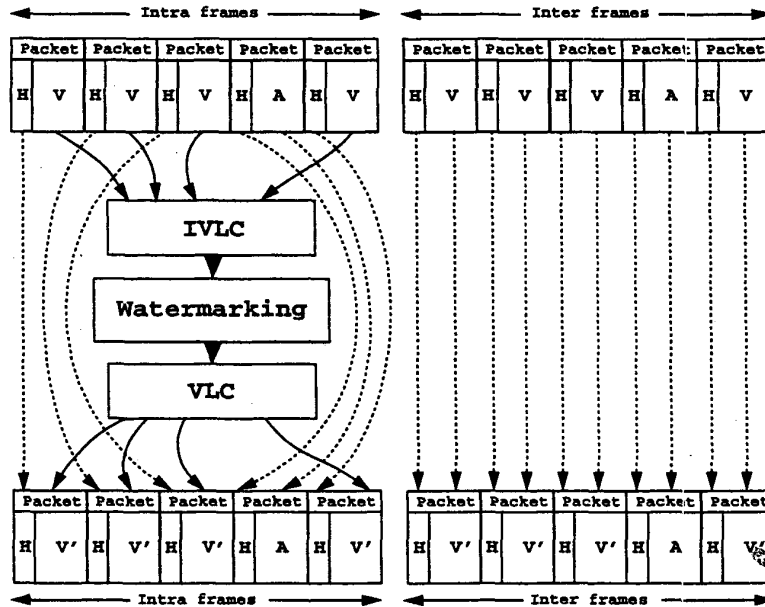


Fig. 1. Operations of the proposed scheme on a MPEG-1/2 stream (V: encoded video data, A: encoded audio data, H: elementary stream packet header, Packet: elementary stream packet, V': watermarked encoded video data).

tiplexed stream. The headers of these packets are left intact. The encoded video data are extracted from the video packets and inverse Huffman is performed in order to obtain the quantized DCT coefficients. 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 Huffman encoded. The video encoded data are partitioned so that they can fit into video packets that use the original headers. Audio packets do not undertake any alterations. Basically the stream structure remains unaffected and only the video packets that contain coded I-frame data are altered. Furthermore, choosing to watermark only the I-frames reduces the complexity and increases the speed of the system. This approach improves dramatically the overall performance of the system and leads to a significant reduction of drift error. In addition, due to the predictive coding (from the intra coded frame) used for B and P frames, a dissipated, but still detectable, version of the

watermark carries over to all frames in a GOP.

IV. PERCEPTUAL WATERMARKING OF VIDEO IN THE QUANTIZED DOMAIN

The watermark sequence consists of the binary information -1, 1 and has zero mean and unitary standard deviation. 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 random number generator is seeded with the result of a hash process. This process consists of the MD5 algorithm that produces a 128 bit key from a meaningful message and another hash function that accepts the 128 bit key as input and produces a 32 bit integer suitable for seeding.

The proposed watermarking scheme alters only the quantized AC coefficients of a luminance block $Q_{\kappa,\lambda}(i, j)$ (where κ is the index of the current macroblock, λ is the index of the block within the current macroblock and i, j are indexes indicating the position of the current coefficient in an 8×8 DCT block) and leaves the chrominance block unaffected. In order to make the watermark as imperceptible and robust as possible, perceptual analysis and block classification techniques [5, 6] are applied on the quantized domain to select which coefficients are best for watermarking. Each selected coefficient is added to the product of the watermark coefficient with the corresponding parameter that resulted from the perceptual analysis. This addition is only possible if both the block classification mask and perceptual analysis indicate that this will not degrade visual quality of the video (see Fig. 2). Each quantized luminance block is classified with respect to its energy distribution by using four classification masks. The possible types of classified blocks are flat, diagonal edge, horizontal edge, vertical edge and textured block. The result of the procedure is a binary mask $C_{\kappa,\lambda}(i, j)$ that indicates the best coefficients to be altered without reducing the visual quality:

$$C_{\kappa,\lambda}(i, j) = \begin{cases} 0 & \text{no alteration} \\ 1 & \text{alteration} \end{cases}$$

where $0 \leq \kappa \leq \text{total picture macroblocks}$, $0 \leq \lambda \leq 3$ and $i, j \in [0, 7]$.

The perceptual model used is a transformation of the perceptual model proposed by Watson [7]. He introduced a measure $T''(m, n)$ that determines the maximum just noticeable distortion for each DCT coefficient. This model has been transformed in order to be applicable to the domain of quantized DCT coefficients.

Let w_x and w_y be the horizontal and vertical width of a pixel in 1/16 pixels/degree of visual angle for 48.7 cm viewing distance. Then the horizontal and vertical spatial frequencies are $f_{m,0} = \frac{m}{2Nw_x}$, $f_{0,n} = \frac{n}{2Nw_y}$, where m, n are the line and column indexes respectively and $N \times N$ are the DCT block dimensions. Let also

$$\log T(m, n) = \log \left(\frac{T \min(f_{m,0}^2 + f_{0,n}^2)^2}{(f_{m,0}^2 + f_{0,n}^2)^2 - 4(1-r)f_{m,0}^2 f_{0,n}^2} \right)$$

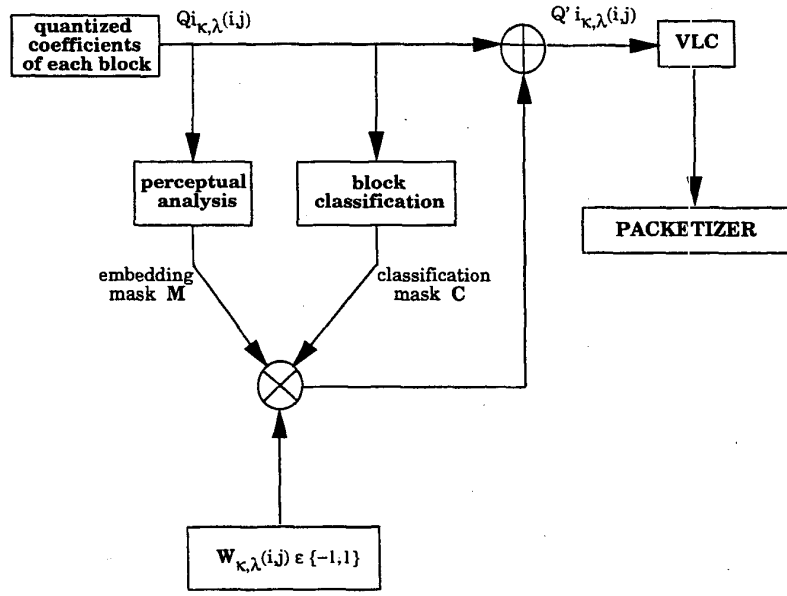


Fig. 2. Watermark embedding scheme.

$$+K \left(\log \sqrt{f_{m,0}^2 + f_{0,n}^2} - \log f_{\min} \right)^2$$

where $K = 1.728$, $f_{\min} = 3.68$ cycles/degree, $T_{\min} = 1.1548$, $a_T = 0.749$. Then

$$T'(m, n) = T(m, n) \left(\frac{Q_{\kappa,\lambda}(0, 0)}{\overline{X_{0,0}}} \right)^{a_T},$$

where $Q_{\kappa,\lambda}(0, 0)$ denotes the quantized DC coefficient of each block and $\overline{X_{0,0}}$ is the quantized DC coefficient corresponding to the mean luminance of an 8×8 block. For the case of 8-bit color depth, $\overline{X_{0,0}} = 32$. Finally,

$$T''(m, n) = T'(m, n) / \sqrt{2}, \quad m = 0 \text{ or } n = 0$$

By incorporating the above perceptual model the embedding decision rule is

$$T_1 M_{\kappa,\lambda}(i, j) \leq |Q_{\kappa,\lambda}(i, j)| < T_2 M_{\kappa,\lambda}(i, j)$$

where $M_{\kappa,\lambda}(i, j)$ is the corresponding perceptual mask T'' for each block in a macroblock and T_1, T_2 are experimentally defined parameters. Then for the quantized coefficients that satisfy the above inequality, the watermark is inserted as follows

$$Q'_{\kappa,\lambda}(i, j) = Q_{\kappa,\lambda}(i, j) + C_{\kappa,\lambda}(i, j) M_{\kappa,\lambda}(i, j) W_{\kappa,\lambda}(i, j)$$

V. DETECTION

The detection of the watermark is performed without using the original data. The original meaningful message that produces the watermark sequence is needed in order to check if the specified watermark sequence exists in a copy of the watermarked video. Then, a statistical approach is taken similar to that analyzed in [4].

Huffman decoding is first performed to get the quantized DCT coefficients. Then the block classification and perceptual analysis procedures are performed as described in the previous section in order to derive the $C_{\kappa,\lambda}(i, j)$ and $M_{\kappa,\lambda}(i, j)$ for each block in a macroblock. Similar with the embedding decision rule, the coefficients that are expected to be watermarked are determined using the following rule

$$T_3 M_{\kappa,\lambda}(i, j) \leq |Q_{\kappa,\lambda}(i, j)| < T_4 M_{\kappa,\lambda}(i, j)$$

where T_3 and T_4 are experimentally defined parameters. The above rule is applied for those quantized DCT coefficients $Q_{\kappa,\lambda}(i, j)$ for which $C_{\kappa,\lambda}(i, j)$ is equal to 1. Let $\{X\}$ be the set of the N coefficients $Q_{\kappa,\lambda}(i, j)$ that satisfy the above equation. Each member of the above set is multiplied with the corresponding watermark coefficient $W_{\kappa,\lambda}(i, j)$ resulting to the data set $\{X_W\}$. Then the statistical characteristics (mean and variance) of the data set $\{X_W\}$ are calculated

$$\begin{aligned} \text{mean} &= E\{X_W\} = \frac{1}{N} \sum_{l=0}^{N-1} X_W(l) \\ \text{var} &= E\{(X_W - \text{mean})^2\} = \\ &= \frac{1}{N} \sum_{l=0}^{N-1} (X_W(l) - \text{mean})^2 \end{aligned}$$

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

$$q = \frac{\text{mean}\sqrt{N}}{\text{var}}$$

If this metric exceeds a predefined threshold T_q , the examined frame is considered watermarked. The statistical correlation metric q_{mean} for an entire sequence is calculated by averaging the mean value of q' for all I-frames. If q_{mean} is larger than T_q then the entire sequence is considered watermarked.

VI. EXPERIMENTAL RESULTS

A software simulation of the proposed algorithm was implemented and executed on a computer using a Pentium III processor. The video sequences were MPEG-1/2 video compressed at a variety of rates (constant or variable) and resolutions. The total execution time

was below real-time for half resolution sequences and close to three times the real-time for full resolution sequences.

The watermarking algorithm selects the set of AC coefficients of an I-frame that can be altered without degrading the visual quality of the video frames. Although, only a portion of the total number of DCT coefficients is altered, because of the visual quality constraint, the detection results show that the statistical detection is very effective and accurate even if the watermarked sequence is manipulated in a number of ways.

Two correlation metric curves for the 48 I-frames of the MPEG-2 videosequence *spokesman* (5 Mbits/sec constant bitrate) are shown in Fig. 3. The first curve shows the correlation metric for a watermarked videosequence when the copyright owner's watermark W is used during the detection process. The second curve shows the correlation metric for the watermarked videosequence when a random watermark sequence W' is used for the detection. As seen, using the proposed watermarking system, the actual copyright owner can be clearly identified since watermarks provided by others that claim copyright ownership do not correlate with the content.

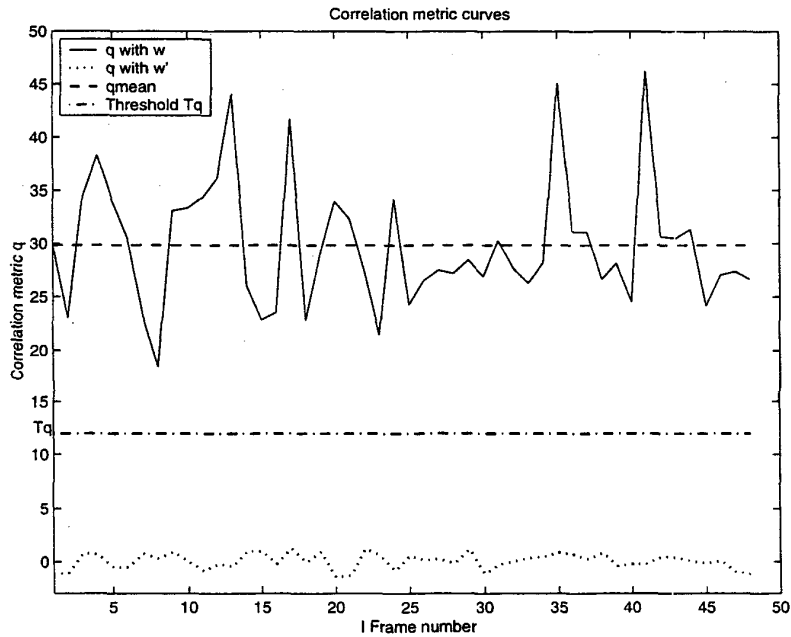


Fig. 3. Detector output for 48 consecutive I-frames of the *spokesman* MPEG-2 stream when the owner's watermark and a random sequence is used for the detection.

Although the proposed method is applied only on I-frames, the embedded watermark carries over to GOP inter-frames due to the motion compensation process. This is particularly useful in the case of *decoding and re-encoding* attacks, in which intra-frames may

become inter-frames and vice-versa. Detailed results about the robustness of the proposed scheme under this kind of attacks will be provided in the camera-ready paper.

Apart from being very effective and reliable, the detection procedure used in the proposed scheme 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 real-time decoders/players that accommodate immediate content authentication.

VII. CONCLUSION

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 very suitable for real-time content authentication applications.

VIII. REFERENCES

- [1] D. Tzovaras, N. Karagiannis, and M. G. Strintzis, "Robust image watermarking in the sub-band or discrete cosine transform domain," in *Proc. 9th European Signal Processing Conf. (EUSIPCO'98)*, Rhodes, Greece, Sep. 1998, pp. 2285–2288.
- [2] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," *Proc. IEEE*, vol. 87, no. 7, July 1999.
- [3] F. Hartung and B. Girod, "Watermarking of Uncompressed and Compressed Video," *Signal Processing*, vol. 66, no. 3, pp. 283–301, May 1998.
- [4] 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.
- [5] K. R. Rao and J. J. Hwang, *Techniques and standards for Image, Video and Audio coding*, Prentice Hall PTR, 1996.
- [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] 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.