SCRAMBLING-BASED WATERMARKING FOR MPEG-4 VIDEO

Boris Vassaux, Philippe Nguyen, Severine Baudry, Patrick Bas and Jean-Marc Chassery

Boris Vassaux, Patrick Bas, Jean-Marc Chassery – Laboratoire LIS – 961 rue de la Houille Blanche
Domaine Universitaire – BP 46 – 38402 Saint Martin d'Hères Cedex - France
{Boris.Vassaux, Patrick.Bas, Jean-Marc.Chassery}@lis.inpg.fr
Philippe Nguyen, Severine Baudry – Thalès Communication – Département TSI, code X47
66 rue du fossé blanc – 92231 Genevilliers - France
{Philippe.Nguyen, Severine.Baudry}@fr.thalesgroup.com

ABSTRACT

The MPEG-4 video standard is nowadays more and more used for video compression, and for applications such as video editing, internet or wireless video communications. However the manipulation of video objects, which is one of the most interesting functionality of MPEG-4, makes many watermarking methods inefficient, especially methods which embed the digital signature in the full spatial domain. This paper presents a new video watermarking technique resisting to MPEG-4 video object manipulation. This so-called scrambling technique allows to adapt any classical spread spectrum watermarking scheme operating in the spatial domain to the MPEG-4 requirements concerning VO manipulation. Thanks to the usage of scrambling and multi-layer embedding, an efficient algorithm for video watermarking is then proposed. Our results show that this algorithm is robust to Video Object manipulations and lossy compression.

1. INTRODUCTION

The increasing development of the World Wide Web and of all the new numerical media (CD, DVD, MP3 …) gives new opportunities to pirate copyrighted supports. Watermarking is presented as a possible method to protect them and to solve the problems of copyright.
The apparition of digital video in communications sets some other problems for security : because of the size of these data, digital video must be compressed. Lossy compression techniques such as MPEG-1, MPEG-2 or MPEG-4 can have negative effects on the embedded watermark: it can be so damaged that the detection of the signature becomes impossible for the video's owner.

1.1. The MPEG-4 video standard

In the case of MPEG-4, the manipulations of Video Objects (VO) obliges to embed the watermark in each VO independently. In fact, unlike MPEG-1 or MPEG-2, the coding is content-based, i.e. single objects are coded individually. The VO contains texture information – a classical image – and shape information – a binary mask –. Each VO can be decomposed into Video Object Layers (VOL), which regroup motion information such as time and space. Then, the Video Object Planes (VOP) represent the temporal segmentations of each sequence frame.
Such a decomposition induces the fact that one VO may have a significant value ; it can represent one principal character for instance. Consequently it is very important for content owners to protect Video Objects.

1.2. Watermarking for MPEG-4 video protection

Watermarking techniques aim at solving this type of problem. Barni and al. [1] proposed a method that consists in embedding the watermark in each video object. They modify some predefined pairs of quantized DCT coefficients in the luminance blocks of pseudo-randomly selected MBs. The first step of the recovery scheme consists in summing the differences of each pair ; the value of the embedded bit is then given by the sign of the sum.
Piva and al. [2] proposed another watermarking scheme to protect MPEG-4 video streams. They embed the watermark in the wavelet transform domain and detect the signature by correlation, after having decompressed the video sequence.
Bas and al. [3] embed the watermark in the VO and use its shape to guarantee the synchronization of the signature. Principal component analysis and warping methods are used to enable the synchronization of the mark after geometric distortions.

The algorithm proposed in this paper embeds the watermark in each VO thanks to a scrambling-based method. The embedding step is similar to classical embedding methods in spatial domain but it uses a scrambling technique to gather the watermark information in the VO.
The paper is organized as follows : firstly the embedding process is presented, then the scrambling method and the recovery scheme are described, and finally some experimental results are presented to validate the method.

2. WATERMARK EMBEDDING

The embedding scheme is based on a spread-spectrum technique, more precisely the multi-layer technique [4]. The multi-layer technique can be considered as an extension of the Hartung and Girot method [5]. This technique is described below.

2.1. Embedding process

The first step consists in dividing the image into blocks of equal size ; the number of blocks corresponds to the number of bits of the message that we want to include into the image.
In a second time, a binary sequence (composed by {+1} and {–1}, and with a null average) is generated using a secret key. Such a sequence will be named 'S' :

$$S: \quad S(k) = +/- 1, \quad \sum_k S(k) = 0$$

Finally the watermark, which will be embedded in the original image, is composed by a succession of blocks equal to +S or –S : +S if the bit of the message is equal to +1, -S else. The watermark is then added to the image to perform the embedding phase illustrated on Figure 1.

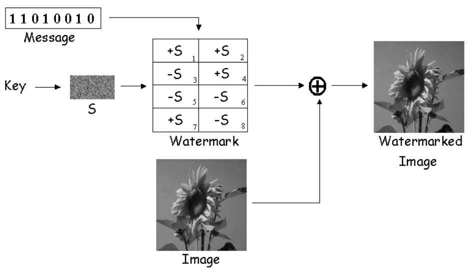

Figure 1 – Classical spread spectrum embedding scheme

To detect the message in the watermarked image, the correlation between S and each block of the watermarked image is calculated. If the result is higher than zero, we will affect a bit value equal to 1 to this block, in the contrary case we will affect 0.

2.2 The multi-layer method

In communication, the CDMA technique suggests to mix several signals at the emission by adding them. The CDMA in watermarking consists here in superposing (and adding) several watermarks to elaborate the definitive watermark. For instance, to embed 64 bits into a 512x512 image, we won't use 64 blocks sized by 64x64 but 2 superposed layers of 32 blocks sized by 128x64, or 4 superposed layers of blocks sized by 128x128. Figure 2 shows the different layers

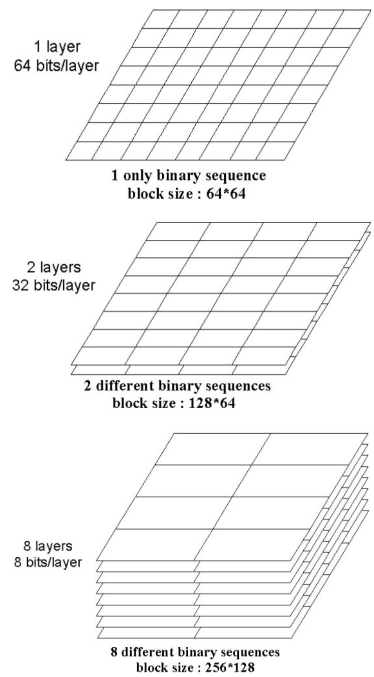

Figure 2 – Principle of the multi-layer technique,
64 bits embedded in a 512x512 image using 1,2 or 8 layers

The visual impact of the multi-layer technique is not so important. For a 8 layer method, the probability to obtain +/-8 is minimal by comparison to the probability to have 0 or +/-2. The most frequent coefficients for a 8-layer method are –2, 0 and +2, and they represent the less penalizing coefficients for the image.
Finally to fit the watermark to the characteristics of each image of the video sequence, we use a classical psycho visual weighting mask resulting from the computation of image local variance. This mask depends on each image and contributes to attenuate the most penalizing coefficients of the watermark.
The detection of the watermark is performed by computing the correlation for each slice between the watermark and the watermarked image. [4]

However, as the MPEG-4 video standard is VO-based, the watermark information has to be present in each VO. This is the subject of the next section.

3. SCRAMBLING METHOD

The goal of this section is to adapt the algorithm described in section 2 to requirements of the MPEG-4 standard, such as the VO manipulations. The principle is to mark each VO delimited by the segmentation masks.

3.1. Definition of scrambling

Scrambling techniques are often used in security applications such as cryptography. Image scrambling consists in spreading the pixels of an image randomly, and in a one-to-one way, in the same image. The spreading rules are defined thanks to a secret key.
In the case of MPEG-4, scrambling can be used to spread VO pixels on the full surface of the frames. Consequently, it is also possible to use the algorithm of section 2 by embedding the watermark in the scrambled frames.

3.2. Scrambling to mark a VO

Because of MPEG-4, the information of the VO's shape may be loosed. After the compression MPEG-4, the segmentation masks may not have exactly the same shape.

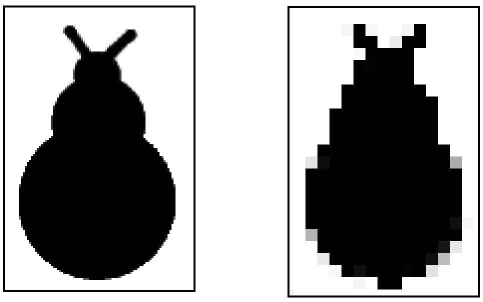

Figure 3 – An example of lossy compression for a segmentation mask

Thanks to scrambling, the VO pixels are dispersed in the full image as it is shown on Figure 4.

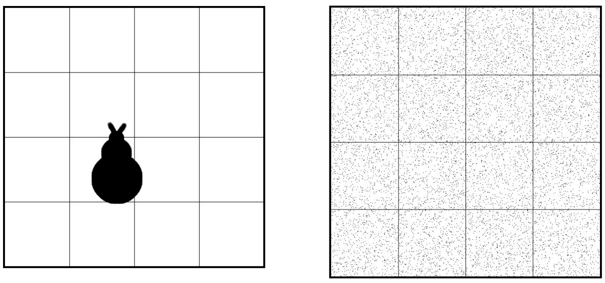

Figure 4 - Pixels of the VO are spread by scrambling

The watermark is then embedded in this scrambled image. Finally each frame of the video is unscrambled to form the watermarked video (with the same secret key as in the scrambling). As a consequence each VO contains a part of the information embedded in each block of the mark.
If we call :

$$\alpha = \frac{\text{number of pixels in the VO}}{\text{total number of pixels in the image}}$$

There will be on average α% of the pixels of each block in the VO. For the detection process, it is just necessary to scramble the compressed images of the video sequence (always with the same key). The correlation is then calculated for each block. Only α% of the pixels are indeed "watermarked"; the detection process is not optimal but experiment shows that redundancy is sufficient to allow correct watermark retrieval (see section 5).
Finally, this method has three major advantages :
. Firstly, each VO can be watermarked separately.
. Secondly the lossy compression for the segmentation masks does not obstruct the detection. One major advantage of scrambling utilization is the fact that without scrambling technique, classical methods (consisting in mapping the VO pixels in a rectangular matrix and then in marking it) are not efficient. If one pixel of the segmentation mask is deleted, the whole detection process may fail because of the desynchronisation. With the scrambling technique, this constraint is removed.
. Thirdly, this method allows to adapt easily any classical spread spectrum watermarking scheme operating in the spatial domain to the MPEG-4 format. With this method the watermark is embedded in the scrambled version of the video, in the same way as before.

3.3. Summary schemes

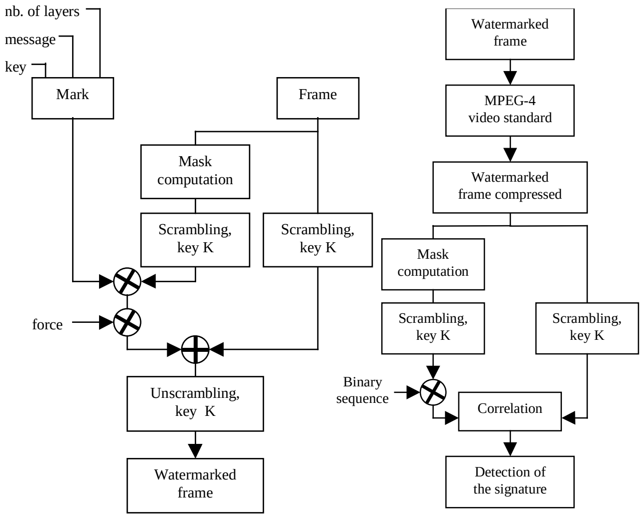

Figure 5 – Embedding and detection schemes

Figure 5 represents the embedding and detection schemes of the method, aiming at watermarking one VO in a video sequence resisting to MPEG-4.
The next section will present the results of this method.

5. RESULTS

The algorithm is tested on eight video sequences. The principle consists in embedding 16 bits in the principal VO of the video sequence. The size of the slices is 352*288. The sequences comprise 300 frames, except CalculatriceCif which comprises only 50 frames. For each frame, the bit error rate is computed (i.e. the probability to have detected a false bit).
Each time, we pay attention to have an acceptable wPSNR [6] to guarantee watermark's invisibility.

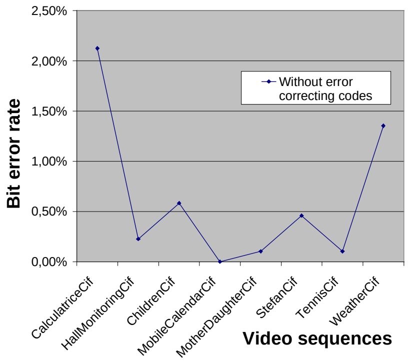

Figure 6 – Results without error correcting codes

Figure 6 shows that the algorithm is efficient although channel coding and error correcting codes are not used.
The result differences between the sequences can be easily explained by the variations in size for each VO : for instance, in the case of Calculatricecif the VO surface represents 7% of the full surface of the slice whereas for StefanCif the number is 29 % (and the detection is about 4 times better).
To improve the results, error correcting codes have been implemented. Results are given on figure 7.

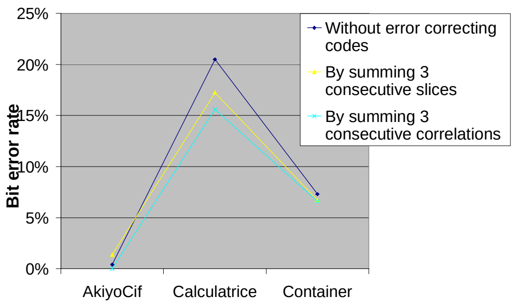

Figure 7 – Improvements thanks to error correcting codes

To improve the performances of the algorithm, two error correcting codes are used : the first consists in summing three consecutive slices of the sequence and then making the correlation. The second consists in summing the three consecutive correlation computation for each embedded bit.
The tests are made by embedding 64 bits in three video sequences (that's the reason why inversion probability is bigger). The graphic on figure 7 shows that using error correcting codes improves the results.

6. CONCLUSION

The method, presented in the paper, proposes a new video watermarking scheme compliant with MPEG-4 : in particular, the method enables to watermark Video Objects. A scrambling technique is used to resist to an eventual lossy compression of the segmentation mask due to MPEG-4. This scrambling technique allows also to adapt easily any classical spread spectrum watermarking scheme operating in the spatial domain to the MPEG-4 format. By completing the scheme with the multi-layer technique and an error correcting code, an efficient video watermarking method is then obtained. In perspective, an effort has to be done to counter geometric attacks by binding the watermark embedding with the image content for instance using methods similar to [3].

7. REFERENCES

[1] M. Barni, F. Bartolini, V. Capelini, and N. Checacci, "Object Watermarking for MPEG-4 video streams copyright protection", *in IST/SPIE's : Security and Watermarking of Multimedia Content II*, SPIE Proceedings, San Jose, CA, vol 3971, pp. 465-476, 2000

[2] A. Piva, R. Caldelli, and A. De Rosa, "A DWT-based object watermarking system for the MPEG-4 video streams", *in IEEE/ICIP'00*, Vancouver, Canada, pp. 5-9, 2000

[3] P. Bas, and B. Macq, "A new video-object watermarking scheme robust to object manipulation", *in ICIP 2001*, Thessaloniki, October 2001

[4] B. Vassaux, P. Bas, and J. M. Chassery "A new cdma technique for digital image watermarking, enhancing capacity of insertion and robustness", *in ICIP 2001*, Thessaloniki, October 2001

[5] F. Hartung & B.Girod. "Watermarking of uncompressed and compressed video." *Signal Processing*, VOL. 66 (NO.3) : 283-233, May 1998

[6] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun. "Generalized watermarking attack based on watermark estimation and perceptual remodulation." *In Proceedings of SPIE*, San Jose, CA, USA, January 2000