A COMPARATIVE STUDY OF DIFFERENT MODES OF PERTURBATION FOR VIDEO WATERMARKING BASED ON MOTION VECTORS

Bodo Yann[1], *Laurent Nathalie*[1] *and Dugelay Jean-Luc*[2]

[1] France Telecom R&D - DIH/HDM
4, rue du Clos Courtel - 35512
Cesson Sevigne Cedex - FRANCE
[2] Institut EURECOM, Multimedia Com. Dept.
2229 Rte des Cretes-BP 193 F-06904
Sophia Antipolis Cedex, FRANCE

ABSTRACT

In this paper, we study the robustness of different grid types used for the embedding rules described in a previous work [1]. In this algorithm, we embed a mark in the uncompressed domain by disturbing slightly the motion vectors computed by an exhaustive BMA on blocks of size N*N. The embedding rule is based on a reference grid that allows to slightly displace the marked vectors in a neighborhood of there original marked position whitout missing the watermarking information. The grid tested are a square, a circular and an angular one. To increase the robustness, we embed the mark by generating a hierarchy of motion vectors to spread the mark on the lower level associated to block N*N with N=k*n.

1. INTRODUCTION

With the development of communication systems, in particular the growth of the internet with adsl and cable, the exchange of multimedia content between people has increased. In this context, it is necessary to be able to control the life of value multimedia content. Many systems have emerged since a few years, like cryptographic system. However it does not protect agaisnt unauthorized copying after the content has been successfully transmitted and decrypted. This is the kind of protection that can be handled by watermarking. A digital watermark is in fact a piece of information inserted and hidden in the media content. This information is imperceptible to a human observer but can be easily detected by a computer. Generally the secret is based on a key and there are no a posteriori protection (for example the css system for DVD [1] has not resisted for a long time). In this context, watermarking algorithms appear to be a complementary solution which allows a full protection. The watermarking is well-suited for the protection of multimedia content. Its applications are various and numerous, they concern copyright application, fingerprinting, "smart content" application, etc. In [3], the authors propose a good introduction on the watermarking world and in [4], the authors present a good mathematical formalization of the watermarking problem. Until now, most of video watermarking systems purpose are based on the extension of still image algorithms. However they suffer from a lack of robustness. Another way is to use the intrinsic notion of the video i.e. the temporal information. Consequently, we can classify video watermarking schemes in two main categories: Still image based techniques and video-adapted techniques. Finally, few papers focus their interest on marking dynamic areas. Indeed, marking motion vectors has been first introduced by Kutter & al. in [2] in which the authors select a set of motion vectors over which we apply a parity rule to embed the mark. Zhang & al. in [5] used this principle to adapt the insertion rule by selecting the vector components that have the greatest magnitude. However, both methods suffer from serious drawbacks. Indeed, a simple filter can destroy the parity of the motion vector components. One way to hide a message in a video could consist in applying modifications on motion vectors. Several ways can be envisaged to realize this task. According to bit value to hide, the motion vector that will host the bit will be or not modified according to a priori rule. This a priori rule is in part defined by a reference grid. After having first investigated square grid [1], we are studying and comparing some other possibilities (circular and angular). The choice of the reference grid will have impact in the final results in terms of distorsion and robustness. In section 2, we remind the main principle of our algorithm and we describe the three different reference grids. Then, in section 3 we present results and comparison between the different grids. Finally, in section 4, we present our conclusion.

2. TECHNICAL DESCRIPTION

In this section, we first remind the basis of our algorithm then we present in details the different grids that could be used for our scheme. Litterature has provided only few watermarking algorithms considering temporal information as a key advantage. However, it seems natural to consider that the robustness of a mark can be greatly improved by considering the following two video properties:

- information amount (a video denotes a larger information than still images);
- motion information.

Finally, video watermarking schemes classified in 2 main categories, still image adapted technique and video adapted technique, and our scheme belongs to the second one.

2.1 hierarchical process and embedding rule

The main functionality of our watermark algorithm consists in providing robust authentication of a Copyright Owner. The copyright information, composed by 8 bits, carried by the watermark could be generated by a secret key owned by the copyright owner. In fact, an exhaustive BMA estimation is applied to block n*n, with n=4. Thus, we determine a

---

[1] http://www.dvd-copy.com/

hierarchical motion vectors pyramid by averaging the fourth vectors associated to four concatenated blocks B_1, B_2, B_3, B_4. The motion vector associated to the block 8*8 represents in our approach the father block. Thus the insertion rule is applied on the averaged motion vectors associated to the blocks 8*8 (Figure 1). Then, we select a subset of motion vectors to be marked. Today, this selection is a pseudo-random one. In all performed grids, the embedding rule is defined by:

$$\forall \vec{d}_f = (d_f^x, d_f^y)^T \in \widetilde{V}_f, \vec{d}_f^W = \vec{d}_f + \widetilde{\Phi}(\alpha, \sigma_f(W), K_{\sigma_f(W)}) \quad (1)$$

where:

$$\widetilde{\Phi}(\alpha, \sigma_f(W), K_{\sigma_f(W)}) = \alpha \times \Upsilon(\sigma_f(W), K_{\sigma_f(W)}) \quad (2)$$

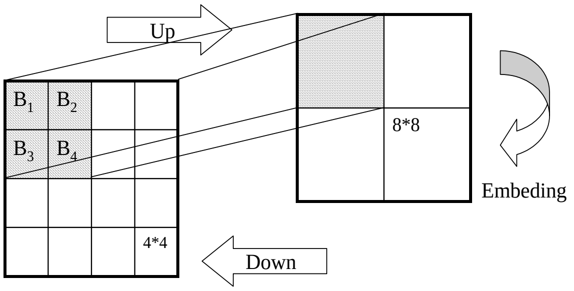
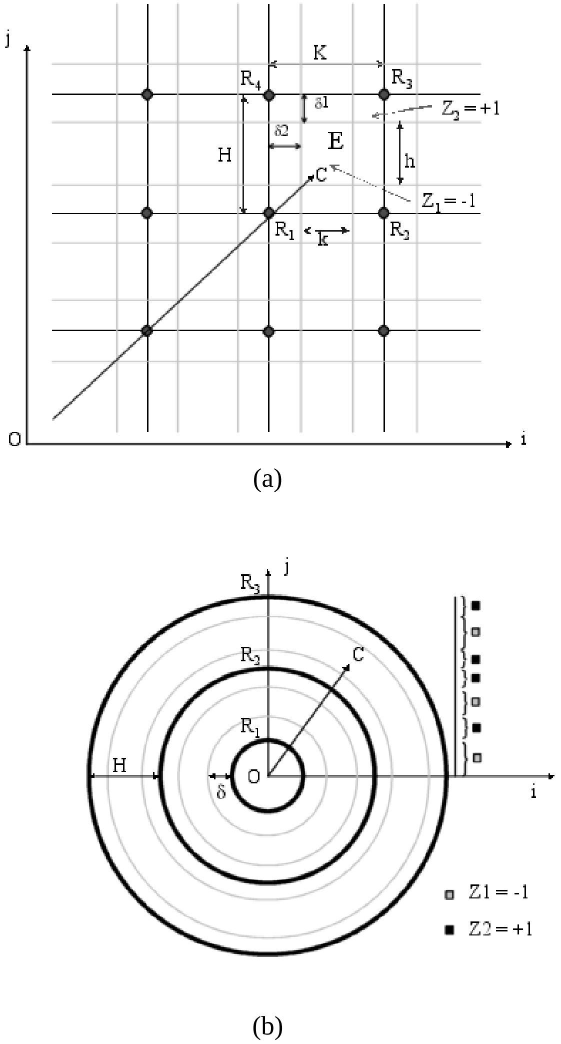


Figure 1: Hierarchical scheme

where $\widetilde{\Phi}$ and Υ are non reversible functions. To improve the robustness of this approach, the insertion rule must respect a spatial structure based on the construction of a reference grid G. This grid could be of different nature, we have performed here three different structures as illustrated on Figure 2(a) for the square grid, on Figure 2(b) for the circular one and on Figure 2(c) for the angular one. All grids are generated in the Cartesian space and are associated to a referential $(O, \overrightarrow{i}, \overrightarrow{j})$. They represents a block-based partitioning of the image compact support resulting in a set of block elements E. Let us denote by R_i the intersection points between blocks that we call *reference points*. For the circular grid, this reference is represented by circles R_i. And finally, for the angular grid, this reference is represented by the axes R_i.

Each selected motion vector of $\widetilde{V}_f$ is first projected on G and this projection serves to compute its associated reference point. The extremity of the projected motion vector $\vec{OC}$ belongs to a block E of G from which intersection points can be deduced. The reference point of the motion vector is the one which is the nearest of the extremity of the vector according to the L^2 distance. In the example of the Figure 2, the reference point of $\overrightarrow{d_f}$ is R_1.
Then, to embed the mark, we generate two areas Z_1 and Z_2 corresponding respectively to the bit -1 and $+1$. The different paremeters, H, h, K, k, δ_1 and δ_2 for the square grid, H and δ for the two other grids are determined in order to have the same area covered by Z_1 and Z_2. These two areas Z_1 and Z_2 drive the mark embedding rule (Figure 2 (a), 2(b) and 2(c)).

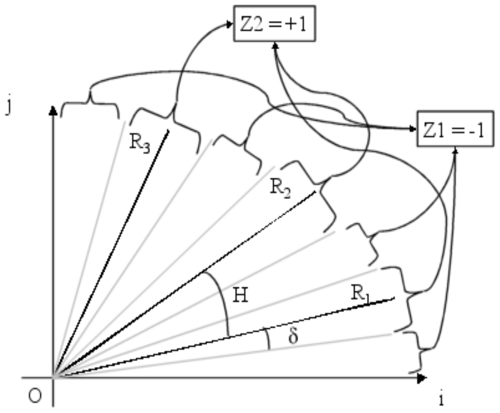


(a)

(b)

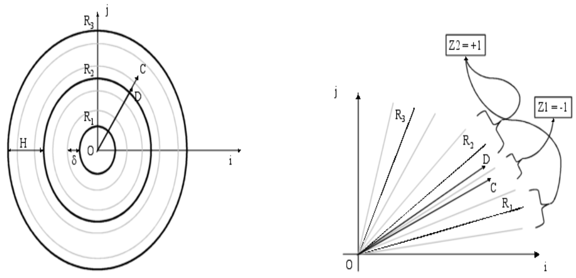
Then the motion vector is modified according to symmetry rules. In all cases if the motion vector is in the right area (the bit to be embedded), we do not modified the motion vector. If the motion is not in the right area, we perform a central symmetry for the circular grid as shown on Figure 3(a), an axial symmetry for the angular grid as shown on Figure 3(b) where *OC* is the original vector and where *OD* is the marked one. For more details on the square grid, see [1].

According to this consideration we can say that for the circular grid, we keep the orientation of the motion vector, only the magnitude changes. Contrary to the angular grid, where the magnitude is kept, and the orientation is changed. Thus for the square grid the previous cases are used. Finally, the choice of the symmetry must minimize the distortion of $\overrightarrow{d_f}$.
The modification realized on the highest level is then applied to the lowest one (down step). By this way, this approach allows us to create redundancy and a spreading of the mark in the insertion phase, the watermarking scheme is then consequently more robust. To end, we perform a motion compensation. This step can be either performed on all of the blocks or either only on marked blocks and completed by original blocks. The second approach allows us to avoid artifacts generated by only exploiting motion estimator, and in the same



(c)

Figure 2: Construction of a reference grid to embed a watermark on motion vectors



(a) (b)

Figure 3: Computation of the marked vector

time increases the robustness of the detection process.

2.2 Extraction

The extraction steps correspond to the dual of the embedding process for the three different grids. Thus to detect the mark, we only had to apply algorithm1.

Algorithm 1 Mark detection algorithm

```
for f= 1 to $N$ { //$N$ denotes the video frame number
    for i = 1 to $k_f$ {
        if $\overrightarrow{d_f^i} \in Z_1$ then $\sigma_f^i(W) = -1$;
        else if $\overrightarrow{d_f^i} \in Z_2$ then $\sigma_f^i(W) = +1$;}}
```

Once a candidate mark $\widetilde{W}$ is detected by the Algorithm 1, we must decide if it corresponds to the real embedded mark W. For this purpose, we compute the correlation C_f at frame f between $\widetilde{W}$ and W by the following recursion:

$$C_f = \frac{C_{f-1} \times (f-1) + (1 - \frac{d(\widetilde{W},W)}{N})}{f} \tag{3}$$

where $d(\widetilde{W},W)$ denotes the Hamming distance between $\widetilde{W}$ and W and N is the mark length.

If $C_f \geq \theta$, where θ is a pre-defined correlation threshold, $\widetilde{W}$ is considered to correspond to W.

3. RESULTS

The above watermarking system has been tested on various videos. In this section, we give some results obtained on the well-known sequences Stefan (100 frames) and Ping-pong (250 frames). These sequences are in the YUV format and their size are 288*352 (CIF format). We have conducted some experiments to compare the robustness of the different grids. For this purpose, we have performed some of the classical sequence manipulations including Divx[2] lossy compression (version 3 and 5), blurring with a uniform kernel and rotation. The compression ratio used for experiments was 1:29 for the Divx3 codec, and 1:43 for the divx5 codec.

The correlation results obtained with these attacks on the Stefan sequence are plotted on Figure 4 and on Figure 5 for the Ping-pong sequence. Let us recall that the correlation level for a frame index f tells us if the mark has been detected in f. On this figure, the correlation treshold θ has been set to $\theta = 0.875$. These results show that the mark is well detected for all cases when there is no attack. But when we apply attacks on the sequences, only the square grid give us good results. The two other grids seems to be not stable. By analyzing these results, we can conclude that the square grid is the best one for our system. We can suppose that for particular cases, like sequence with more fast motion, the two other grids could gives good results. We have to check for this hypothesis.

4. CONCLUSION

In this paper, we have analysed the behavior of our system with different type of grids (square, circular and angular). As shown in the results, the square grid gives us the best results. However, according to the structure of the circular and the angular grid, we can suppose that for specific attacks like scaling or other rotation (with a higher angle), this two grids should be more robust. Indeed, for the circular grid, according to its properties, we can suppose that it will be more robust agaisnt rotation. For the angular grid, it will probably be more robust against scaling. For this reason, we have to perform more test with different attacks and different video content (with more motion for example). Even if this hypothesis is true, in general the best results are obtained by the square grid and for this reason we choosed this one for our system. Finally, we have to test an other type of grid, an hexagonal one, wich seems to be geometricaly more suitable than square one.

REFERENCES

[1] Bodo Y., Laurent N., and Dugelay J-L. Watermarking video, hierarchical embedding in motion vectors. *ICIP 2003*, Barcelonna, Spain September 14-17, 2003.

[2] Kutter M. Jordan F. and Ebrahimi T. Proposal of a Watermarking Technique for Hiding/Retrieving Datas in Compressed and Decompressed Video. In *ISO/IEC/JTC1/SC29/WG11/MPEG97/M2281*, 1997.

---

[2] http://www.divx.com/

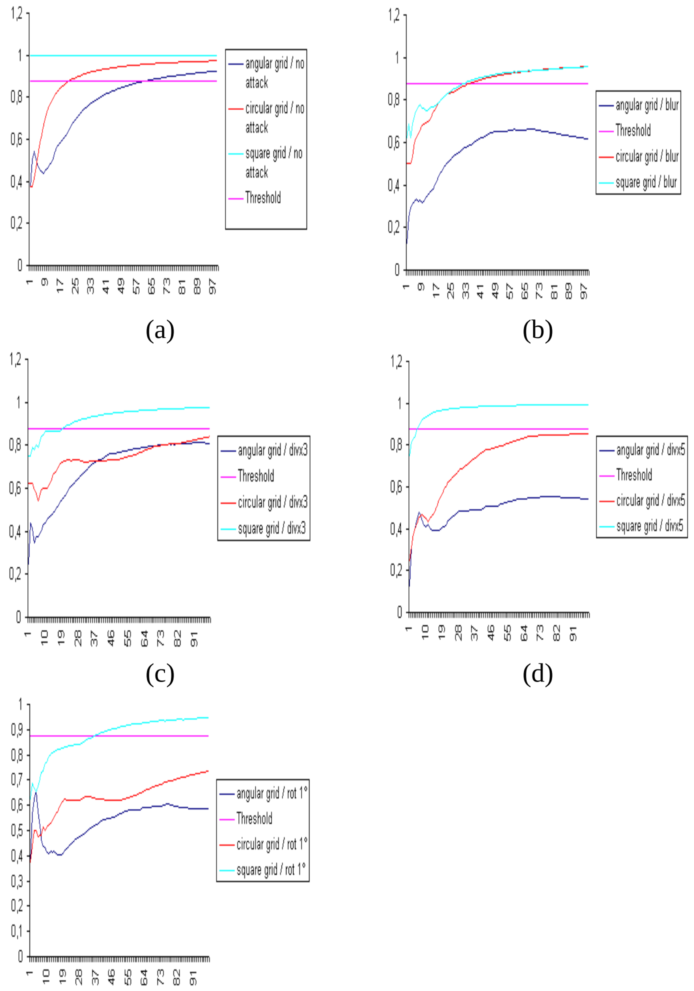


Figure 4: correlation score for stefan sequence

[3] Katzenbeisser S. and Petitcolas F.A.P. Information Hiding. *Artech House*, London 2000.

[4] Pitas I. and Voyatzis G. The use of watermarks in the protection of digital multimedia products. *Proceedings of the IEEE*, 87:1197–1207, 1999.

[5] Li J. Zhang J. and Zhang L. Video Watermark Technique in Motion Vector. In *XIV Brazilian Symp. on Computer Graphics and Image Processing*, pages 179–182, Florianopolis - Brazil, October 2001.

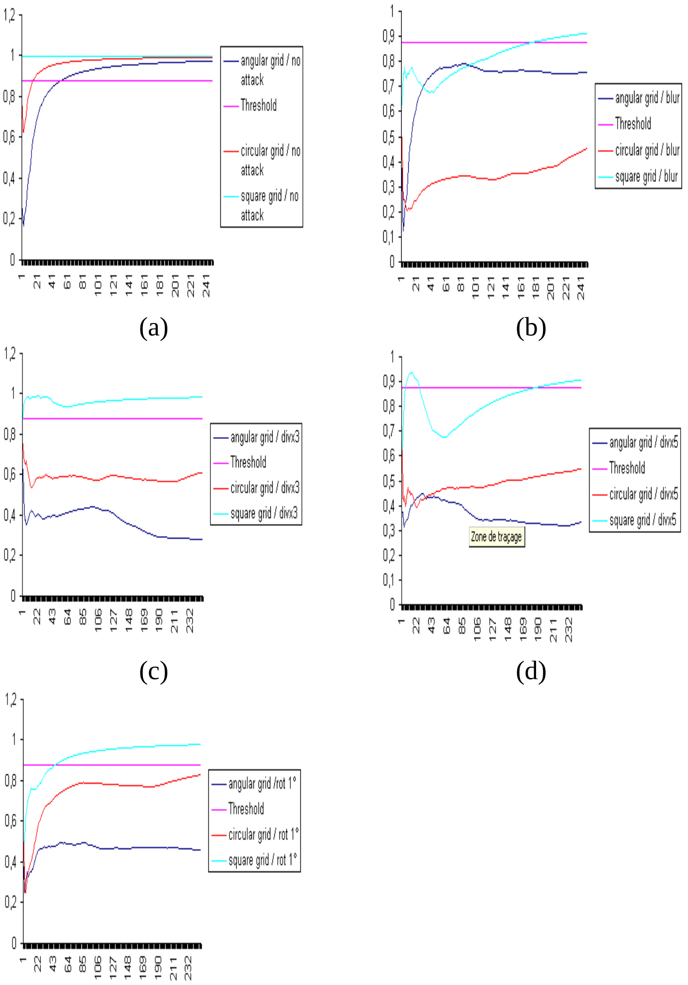


Figure 5: correlation score for pingpong sequence