On Line Video Watermarking

A New Robust Approach of Video Watermarking based on Dynamic Multi-sprites Generation

Ines Bayoudh, Saoussen Ben Jabra and Ezzeddine Zagrouba
Research Team on Intelligent Systems in Image and Artificial Vision (SIIVA), RIADI laboratory, ISI, University Tunis El Manar, Abou Raihane Bayrouni 2080, Ariana, Tunisia

Keywords: Digital Video Watermarking, Multiples Sprites, LSB, Collusion, Invisibility, Robustness.

Abstract: With the development of the emerging applications, watermarking methods require a new type of constraints. In fact, robustness to malicious attacks and processing time reduction have became two important constraints which must verify a watermarking approach. In this paper, a new scheme of digital watermarking for video security is proposed. This scheme is based on dynamic multiple-sprites. These last ones allow obtaining robustness in front of collusion which presents a dangerous attack for marked video. First, original video is divided into groups of images, and then a sprite will be generated from each group. Finally, the signature will be inserted in the low bits of the obtained sprite and marked frames will be generated from the marked sprites. Experimental results show that the proposed scheme is robust against several attacks such as collusion, compression, frame suppression and transposition, and geometric attacks. In more, processing time of watermarking is reduced.

1 INTRODUCTION

The variety of digital technologies have been engendered the risk of media insecurity. In fact, digital document manipulation and transfer became easier than before and traditional methods of security like cryptography and steganography have become insufficient. Digital watermarking was appeared as a solution to protect multimedia contents from illegal manipulations. It consists to embedding an imperceptible signature into a numeric text, image, audio or video sequence, and then to try to detect the signature after an eventual distortion done on the marked data. Watermarking must verify several constraints such as invisibility of embedding, high capacity of insertion and robustness against almost of attacks. Nowadays, a new challenge appeared by emerging applications development such as broadcasting, pay TV and video conference. This challenge is the reduction of processing time.

Different techniques of video watermarking have been proposed but each of them is developped for a praticular application (Gopika and Chiddarwar, 2013), (Hood and Janwe, 2013). In fact, there are methods which care to guarantee a better robustness against almost of attacks, others try to maximize invisibilty or aim to reduce time processing. This depends on where and how the signature will be inserted. In fact, it exists mainly two insertion domains: spatial domain where signature will be embedded by modifying directly original data or frequential domain which is based on DCT, TFD or Wavelet transformation of original data to embed signature. Every insertion domain presents its own advanatges and inconvenients. In fact, spatial domain presents the best imperceptibility but it is not robust to malicious attacks. However, frequential domain allows robustness in front of many attacks but it reduces invisibilty. There is a second critera which can be used to classify watermarking methods : it is insertion mode which can be additive or substitive mode. The signature capacity depends on the insertion mode. This capacity increased when the mode is substitive and it is low in the second case.

In this paper, the proposed approach aims to maximize the compromise invisibilty, robustness against collusion with a reduction of processing time. It's why we focused on existing watermarking techniques proposed for real time applications. So, we proposed to classify them into two classes: techniques designed for real time applications and techniques robust to collusion.

Bayoudh I., Ben Jabra S. and Zagrouba E..
On Line Video Watermarking - A New Robust Approach of Video Watermarking based on Dynamic Multi-sprites Generation.
DOI: 10.5220/0005316801580165
In *Proceedings of the 10th International Conference on Computer Vision Theory and Applications* (VISAPP-2015), pages 158-165
ISBN: 978-989-758-091-8

For the first class, processing time depends on frames emission rates per second which can be 25f/s or 30f/s depending on the standard system used in each country. Considering this rate equal to 30f/s, 30 frames can be transmitted in one second it means 0.033seconds (33 milliseconds) is allocated for every frame. Thus, a real time scheme must treat a frame no more than 33 milliseconds.

In (Lee and Seo, 2013) a real time watermarking based on temporal modulation was proposed. It operates in spatial domain with an error correcting code to ameliorate robustness. Other methods are proposed in transformed domain (Maity and Kundu, 2009), (Wang et al., 2009), (Bhaisare et al., 2013) as in (Lee and Im, 2012) where authors propose to embeed signature into compressed domain to minimize the proceeding time. In fact, the video is partially compressed to generate DCT coefficients where the signature is inserted using the Quantization Index Modulation QIM. This last one consists to select the nearest byte in the coefficient from the mark to guarantee a good insertion quality. The proposed perceptual watermarking by (P.Mohanty and Kougianos, 2011) aims to develop an architectural idea which improve the treatment time using the parallel implementation.

Concerning the collusion attack, it became a major threat for video security. In fact, the attacker tries to predict the original document exploiting the temporal redundancy in the sequence video. There were two types of collusion. The first one estimates the embedded mark then eliminates it from the document. This method is possible when the hacker has a lot of different marked images using same signature. For the second type of collusion, the attacker must have a large combination of similar images which are marked differently and he averages them to destroy the signature and obtain the unmarked video. To resolve collusion problem, movement and redundancy must be reduced (Dorr and Dugelay, 2005). So, the better idea to obtain robustness against this attack consists to embedding signature into mosaic image generated from the video. In fact, the mosaic presents an accumulation of dispersed information into the sequence where the same point repeated along the video will be represented by a single reference in the panoramic image. Therefore, all similar pixels are marked with the same way. Koubaa et al (Koubaa et al., 2012) proposed to insert the signature with a wavelet watermarking into the feature regions selected from the mosaic image. Bayoudh et al (Bayoudh et al., 2013) also used the mosaic image generated from the original video to embed the signature. The insertion is done using frequential domain and the Krawtchouk moments to improve robustness and invisibility of the signature. Kerbiche et al (Kerbiche et al., 2012) proposed another watermarking based on mosaic image. They selected feature regions where the objects move. First, the wavelet transform is applied on the selected region. Then, obtained frequencies are marked using two different approaches to verify the best compromise invisibility/robustness. These techniques present a good robustness but they require an important processing time allocated to generate the mosaic. So, they can't be applied for real time applications.

In this paper, a new video watermarking scheme is proposed. It aims to resist to almost of attacks with a higher capacity, imperceptibility and time processing reduction. In fact, the new approach is based on dynamic multisprites generation to guarentee a robustness against collusion attacks with a processing time minimization. In more, the signature embedding is based on spatial domain which allows obtainning high invisibility and an independency from standard of compression. Finally, the proposed algorithm is blind. This last criterion permits to reduce processing time at detection step. The remaining sections of this paper are organized as follows. The proposed watermarking is described in section 2 and the experimental results are provided in section 3. Finally, section 4 summaries the proposed work and presents the future work.

2 PROPOSED WATERMARKING

The main idea of the proposed scheme is to embed signature into mosaic image generated from original video to guarantee the robustness against collusion attacks. Different recent techniques are based on panoramic video image to resist collusion (Koubaa et al., 2012), (Kerbiche et al., 2012), (Bayoudh et al., 2013). For all these techniques, the first step consists to generate mosaic from different images of original video sequence. Then, the obtained mosaic will be marked using a spatial or frequential embedding. Every technique try to improve robustness against different attacks with a high invisibility but the mainly limitation of these techniques is processing time which is important. In fact, the generation of a panoramic image from the whole video sequence requires an important processing time with an eventual distortion of the reconstructed video sequence. To resolve these two problems, multi-sprites are proposed in this paper. The use of multi-sprites will offer a reduction of processing time of mosaic generation and will reduce the distortion effect due to the video reconstruction from the mosaic image.

2.1 Multi-sprites Generation

Multi-sprites generation from a video sequence can be static (standard) or dynamic.

2.1.1 Standard Multiples Sprites Generation

For a single sprite, a single image is generated from the whole sequence. Obtained image presents a large size with a hard distortion. In Figure 1 we show an example of a single sprite generated from the sequence ”Stefan” which contains 300 images. The generation

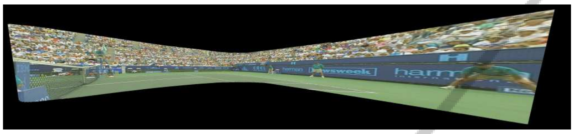


Figure 1: Single sprite generated from ”Stefan” sequence.

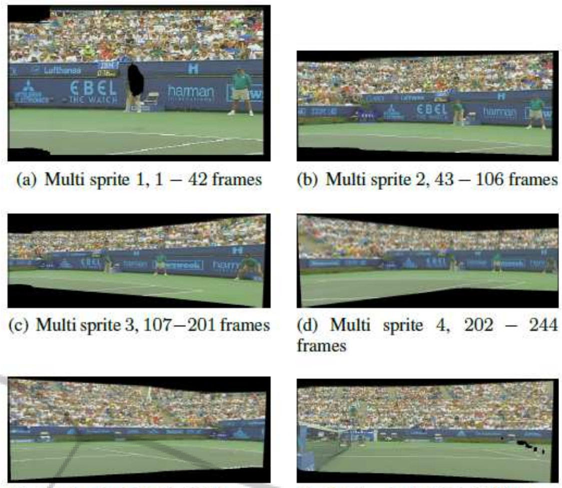
of a single mosaic consists to choosing the reference image, and then, each image will be transformed by this reference and warped in the same space to construct the mosaic.

For multi-sprites case, multiple mosaics are generated from the original video sequence. In fact, the video will be subdivided into parts depending on the camera parameters (Krutz and Glantz, 2008). Figure 2 presents the different sprites obtained from ”Stefan” sequence using the estimation proposed in (Krutz and Glantz, 2008) with parameters variation. We can conclude that sprites number depends on the parameters used to subdivide the sequence. So, multi-sprites generation scheme is based on selecting the segmentation classifier then from every segment the standard generator of single mosaic will be adopted.

As the goal of the proposed approach is to minimize the processing time, we propose to replace the step of choosing segmentation classifier by a standard fragmentation without treatment. We propose to select an independent part from video each second, and from every part a sprite will be generated. However, the time to build the mosaic is also significant. This is caused by the algorithm of mosaic generation. In fact, to generate a mosaic we should save all images, then for every new image the reference can be changed so, all the mosaic will be generated using the whole set of images. To resolve these problems, a new sprite generation scheme is proposed in this paper. This scheme will be adapted to the watermarking constraints.

2.1.2 Dynamic Multiple-sprites Generation

The contribution of this proposed approach is to use dynamic multi-sprites to embed signature. This allows reducing processing time. This allows reducing



(a) Multi sprite 1, $1-42$ frames (b) Multi sprite 2, $43-106$ frames (c) Multi sprite 3, $107-201$ frames (d) Multi sprite 4, $202-244$ frames (e) Multi sprite 5, $245-261$ frames (f) Multi sprite 6, $262-297$ frames

Figure 2: Multiple sprites generated from ”Stefan” sequence.

processing time. In more, to verify robustness against collusion, dynamic aspect will be proposed at mosaic generation step.

For a static mosaic we need all the sequence to generate the panoramic image. It means that we can’t produce the mosaic while the whole set of frames is not obtained. The mosaic will be created by manipulating all the images relatively to the reference.

However the generation of a dynamic mosaic depends only on the received images and not all the sequence. So, every current image will be treated in function of previous received images d didn’t need the achievement of the sequence. Dynamic mosaic or mosaics on line are proposed in literature (chafik Bakkay et al., 2011), (Kuo and Chen, 2009).

2.1.3 Proposed Multiples Sprites Generation

The proposed multi-sprites generation scheme can be composed to the following steps:

1. Giving an original video sequence, it will be divided in n seconds

2. Each i seconds, the first received frame will be considered as the reference image of the i^{th} mosaic

3. The first frame of each second is considered as the current mosaic

4. Every new received frame is transformed based on the reference image

5. The mosaic will be updated based on the current mosaic and the current received image

6. Steps 4 and 5 will be repeated every second.

In more, the proposed approach improves mosaic updating step to increase robustness and to reduce processing time. In fact, when a new image is received, Sift parameters are calculated to determine

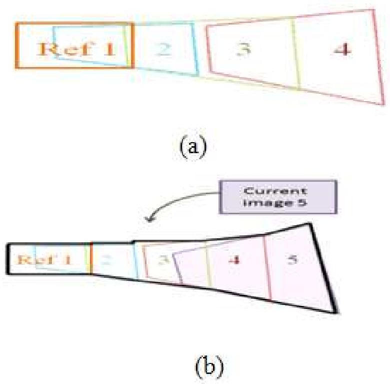


Figure 3: Updating step (a): the current generated mosaic (b): the new mosaic generated based on the current image.

the applied transformation between this image and the reference image. Therefore, the image is transformed at the same space of the latest generated mosaic. Finally, to obtain the new mosaic, only the transformed image and the current mosaic will be used. Figure 3 shows the updating step of the current mosaic based on the current image.

Besides the advantage of processing time reduction, the proposed algorithm allows a memory space reduction by minimizing the used parameters for the mosaic generation. In fact, contrary to standard mosaic updating which depends on the whole set of images, in the proposed generation only the latest mosaic and the current image are needed for updating.

2.1.4 Static and Dynamic Multi-sprites Generation Comparison

Figures 4 and 5 show respectively sprites generated from test original videos using static generation, and the results of dynamic multiple- sprites generation. Due to the averaging way used to generate static mosaic, the moved objects are eliminated whereas they exist for dynamic mosaic.

Table 1 shows average time allowed to generate a sprite. This last one is progressively reduced in case of dynamic mosaic.

Table 1: Average time oh the generation of a sprite in seconds.

Approach	Stefan	Soccer	Granguardia
Static	25.07 s	119.91 s	82.49 s
Dynamic	16.85 s	70.67 s	49.2 s

2.2 Proposed Embedding Scheme

The proposed scheme aims to obtain a new watermarking scheme for real time application with a high

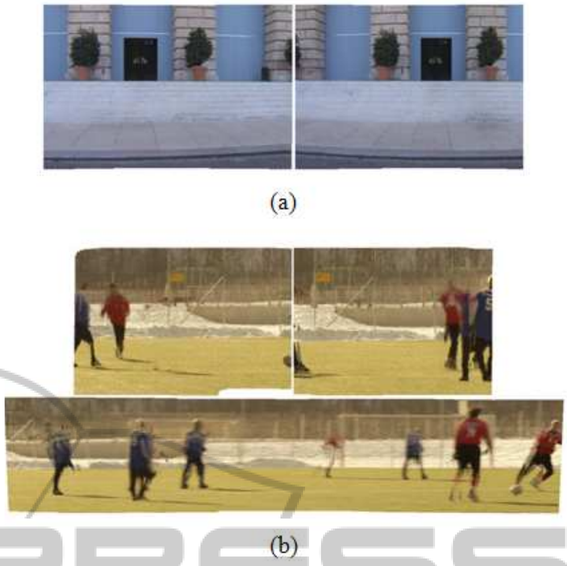


Figure 4: Examples of Static multiple-sprites generated from (a) ”Granguardia” (b)”Soccer” sequence.

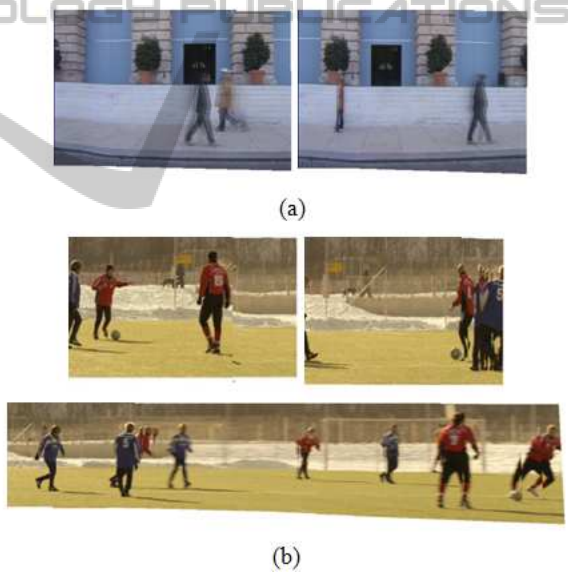


Figure 5: Examples of dynamic multiple-sprites generated from (a) ”Granguardia” (b)”Soccer” sequence.

invisibility and robustness against almost attacks especially collusion attacks. Based on dynamic multiples sprites generation, the proposed approach guarantees robustness against collusion although a reduction of treatment time. The proposed watermarking scheme is presented in Figure 6 and it will be repeated each second for the continuous set of images. It can be decomposed in different steps:

1. Dynamic sprites are generated from the received frames every second

2. Embedding scheme is applied to the final sprite

obtained
3. Marked images set will be reconstructed from the marked sprite
4. All steps will be repeated until sequence emission achievement.

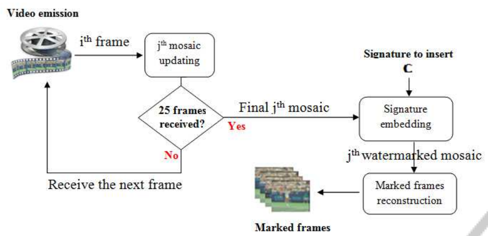


Figure 6: Proposed watermarking scheme.

2.2.1 Sprite Generation

The first step of the proposed approach consists to generate the on-line sprite every second. To optimize the sprite generation time, we proposed to improve the way of mosaic update. In fact, instead of using averaging of the different frames calculation to construct the mosaic, we propose to maximize the latest version of mosaic with the most new received frame. This allows reducing memory space, updating time as well as robustness. Figure 7 presents the mark to insert and Figure 8 shows mark detection results using dynamic mosaic update by averaging (a) and detection results using proposed mosaic updating (b). These results show that the detection is more performed in the second case.

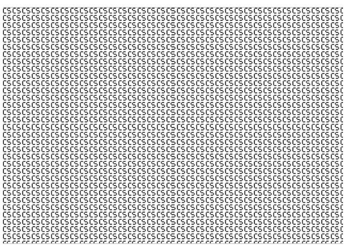


Figure 7: Original mark.

2.2.2 Sprite Conversion to YUV Space

After sprite generation, it will be converted to the YUV space. This space is described by its three components: Y which presents the luminance in image and U, V which present image chrominance (blue and red colors).

The Y component is selected to be marked because it represents the more important information on the image. In more, almost of compression techniques modify chrominance components so the luminance will not be modified and then the embedded signature will resist to compression manipulation.

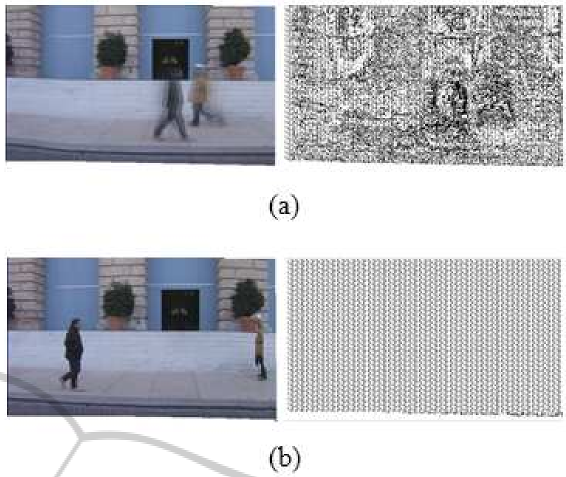


Figure 8: Mark extraction, (a) using dynamic generation with averaging, (b) using proposed dynamic generation.

2.2.3 Spatial Embedding Scheme

The signature embedding scheme contains two steps as shown in Figure 9. First, signature spreading is applied based on sprite size. Then, LSB (Low Significant Bit) algorithm is applied to insert the mark. The low bit of every pixel in the luminance component is substitute with the corresponding bit from the spread mark to obtain the marked image. Finally, the marked sprite is reconverted to the RGB space and marked frames will be reconstructed. The LSB scheme is applied thanks to its high invisibility. In addition, the use of a substitution mode of insertion allows obtaining blind detection and increases the mark capacity.

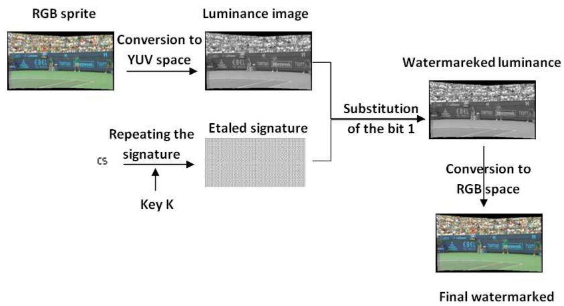


Figure 9: Insertion scheme.

2.3 Detection Stage

To detect the signature from the watermarked video, different steps are applied where the three first steps are similar to those of insertion scheme. In fact, different sprites are generated from the continuous frames received each second. Then, every final generated sprite is converted to the YUV space and the luminance component is selected to extract the signature from the last significant bit. Finally, the spread mark is extracted and divided into a set of signatures where the one which presents the best correlation is selected.

3 EXPERIMENTAL RESULTS AND EVALUATION

To evaluate the proposed approach it was applied on three test videos. These videos are ”Stefan”, ”Soccer” and ”Granguardia” (Figure 10). Each video presents different criteria such as scene frame number, movement frequency and background texture (Table 2).

Figure 10: Sequences tests, (a) ”Stefan”, (b) ”Soccer”, and (c) ”Granguardia”.

Table 2: Test videos characteristics.

Criteria	Stefan	Soccer	Granguardia
Number	300	150	50
Movement	Fast	Variable	Slow
Texture	High	Mixed	Unifom
Size	176x144	352x288	324x276

Figure 11: Watermarked frame of sequences tests, (a) ”Stefan”, (b) ”Soccer”, (c) ”Granguardia”.

Proposed approach evaluation is tested based on three criteria: invisibility, robustness and time of proceeding.

3.1 Invisibility

Figure 11 shows that there is no difference between original and marked frames. To prove this invisibility, PSNR (Peak Signal to Noise Rapport) and the RMSE (Root Mean Square Error) between the original and the marked video frames are measured (Table 3). The lowest mean value of PSNR is 70.613db and the highest is under 82.797db and RMSE are 0.018 and 0.076. The best values of invisibility are obtained for ”Granguardia” sequence and the wrothest are obtained for ”Stefan” sequence. In fact, this is due to the frames texture which is uniform in case of ”Granguardia”.

Table 3: The Average Value of PSNR (db) and RMSE.

Criteria	Stefan	Soccer	Granguardia
PSNR	82.797	79.888	70.613
RMSE	0.018	0.028	0.076

3.2 Robustness

To evaluate the robustness of the proposed approach, various attacks are applied to the mark video and then the correlation is calculated between the extracted and the original signatures.

Two types of attacks are tested: the usual attacks such as geometric attacks, filtering, noises, compression and temporal attacks such as frame suppression, insertion and transposition and finally, collusion attacks which aim to estimate the signature and remove it from the marked sequence. Two binary signatures are used in evaluation tests. These two signatures are shown in Figure 12 where (a) presents the copyright signature with a low density and a small size of 12x9 and (b) refers to the second signature characterized by a high density with a large size of 63x24.

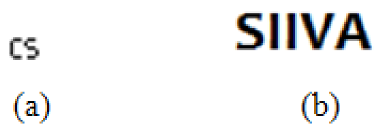


Figure 12: used signatures (a) low capacity signature, (b) high capacity signature.

Obtained results are described in Table 4 where the extracted marks with the correlation value are presented. The proposed watermarking is robust against several usual and temporal attacks and even collusion attacks.

First, geometric attacks are tested as translation, scaling, and rotation with various angles. In addition, we have applied different filters and noises such as means filter, high and low pass filter, also Gaussian and salt and pepper noise. The proposed approach succeeds to detect signature after all these attacks. This is obtained thanks to the invariance of the points used to embed signature in front of these attacks.

The proposed watermarking is robust against mpeg compression until 500kb/s. This can be explained by choosing the luminance components to embed signature. For temporal attacks frames trans-

Table 4: Robustness results.

Attacks	Correlation	Parameters
O attack	1	-
Collusion	1	-
Compression	1	500kb/s
Cropping	1	3 windows
Frame changes	1	20f/s,30f/s
Rotation	1	180,90,45
Gaussian noise	0.399	0.01,0.001
Salt pepper	1	0.1,0.01
Gaussian filter	1	3x3,5x5/0.5,0.9
Medium filter	0.441	3x3,5x5
High pass filter	0.596	-

pose and drop, insertion of unmarked frames and of a frame that don't make part of the considered sequence are tested. After application of all these attacks, the proposed watermarking can detect signature. In addition, the mark is detected after modifying the frames rate between 20f/s to 30f/s. This is obtained thanks to the use of sprite where same points in different images are marked similarly.

Furthermore, detection succeeds after a cropping attack thanks to multi-sprite use. In fact, in case of cutting a part of video, if the signature is not detected from the first sprite, it can be extracted from other sprites.

3.3 Processing Time Comparison

The proposed approach aims to improve the compromise between invisibility, robustness against usual and temporal attacks, and processing time reduction. To achieve this goal, we proposed to adopt sprite generation to real time watermarking. In fact, the best solution proposed in literature to resist collusion is to insert the signature into the mosaic image generated from the video sequence. Instead static mosaic generation cannot be applied at real time applications for many raisons. In fact, static generation depends on the whole sequence, and then the standard generation can take a long time. As a solution, multiple-sprites are proposed to reduce the time of waiting to complete the all sequence. Besides the amelioration proposed using the online generation decreases also the time of generation of a sprite as described in Table 1.

In Table 5, the proposed approach is compared with two existing approaches. The first one is based on static mosaic to embed signature (Koubaa,2012) and the second is developed for real time application (Lee, 2013). Obtained results show that the proposed approach allows obtaining the best compromise invisibility, robustness and processing time reduction.

Table 5: Comparative study.

Approach	PSNR	Robustness	Time
Proposed	65-88	Geometric, temporal attacks filters, noises and collusion	Reduced
Mosaic based	45	Rotation, collusion compression	High
Real time based	48	Geometric attacks	Real time

4 CONCLUSIONS

This paper presents a new video watermarking approach based on on-line multiple-sprites. These last ones allow obtaining robustness to collusion with an important reduction of processing time. In fact, original video is divided into groups of images, and then a sprite will be generated from each group. Finally, the signature will be inserted in the low bits of the obtained sprite and marked frames will be generated from the marked sprites. In sprite generation step, the mosaic updating stage is improved to minimize processing time. In addition, signature is embedded in the luminance components of YUV space to enhance the robustness against compression. The insertion is done with substitutive mode which allows reducing detection time. In more, extraction is blind and this permits to decrease the memory needed to save original video. Experimental results show a high invisibility with robustness against the most important attacks including collusion attacks. Also, we have reduced the time of treatment comparing with others existing methods. In future work we will try to improve the proposed approach by proposing a faster sprite generator.

REFERENCES

Bayoudh, I., Jabra, S. B., and Zagrouba, E. (2013). A robust video watermarking based on krawtchouk moment. In *TAIMA'13, 8me dition des ateliers sur le traitement et l'analyse de l'information Mthodes et Applications*.

Bhaisare, S., Karode, A., and Suralkar, S. (2013). Significance research review on real time digital video watermarking system for video authentication. In *CEEE'13, Advances in Computer, Electronics and Electrical Engineering*.

chafik Bakkay, M., Barhoumi, W., and Zagrouba, E. (2011).

Mise jour dynamique de l'image de rfrence pour l'optimisation du rsum vido en ligne par multiple mosaque'. In *TAIMA'11, 7me dition des ateliers sur le traitement et l'analyse de l'information Mthodes et Applications*.

Dorr, G. and Dugelay, J. (2005). Problmatique de la collusion en tatouage vido attaques et ripostes. In *CORESA'05, Colloque Compression et Reprsentation des signaux Audiovisuels*.

Gopika, V. and Chiddarwar, G. (2013). Review paper on video watermarking techniques. In *IJSRP'13, In International Journal of Scientific and Research Publication*.

Hood, A. and Janwe, N. (2013). Robust video watermarking techniques and attacks on watermark-a review. In *IJCTT'13, International Journal of computer Trends and Technology*.

Kerbiche, A., Jabra, S. B., and Zagrouba, E. (2012). Tatouage video robuste bas sur les rgions d'intrt. In *CARI'12, 11me Colloque Africain sur la Recherche en Informatique et Mathmatiques Appliques*.

Koubaa, M., Amar, C., M.Elarbi, and H.Nicolas (2012). Collusion, mpeg4 compression and frame dropping resistant video watermarking. In *MTA'12, Springer Multimedia tools and applications*.

Krutz, A. and Glantz, A. (2008). Multiple background sprite generation using camera motion characterization for object-based video coding. In *3DTV'08, The True Vision Capture, Transmission and Display of 3D Video*.

Kuo, I. and Chen, L. (2009). A fast multi-sprite generator with near-optimum coding bit-rate. In *IJPRAI'09, International Journal of Pattern Recognition and Artificial Intelligence*.

Lee, M. and Im, D. (2012). Real time video watermarking system on the compressed domain for high-definition video contents: Practical issues. In *DSP'12, Digital Signal Processing*.

Lee, S. and Seo, D. (2013). Robust video watermarking based on temporal modulation with error correcting code. In *IJSA'13, In International Journal of Security and its Application*.

Maity, P. and Kundu, K. (2009). Dual purpose fwt domain spread spectrum image watermarking in real time. In *CEE'09, In Computers and Electrical Engineering*.

P.Mohanty, S. and Kougianos, E. (2011). Real time perceptual watermarking architectures for video broadcasting. In *JSS'11, The Journal of Systems and Software*.

Wang, J., Liu, C., and Masilela, M. (2009). A real time video watermarking system with buffer sharing for video-on-demand service. In *CEE'09, Computer and Electrical Engineering*.