A VIDEO WATERMARKING SCHEME RESISTANT TO GEOMETRIC TRANSFORMATIONS

Maher EL'ARBI, Chokri BEN AMAR[1] and Henri NICOLAS[2]
[1]REsearch Group on Intelligent Machines (REGIM)
University of Sfax, National Engineering School of Sfax, B.P. W, 3038, Sfax, TUNISIA
[2]Laboratoire Bordelais de Recherche en Informatique (LABRI)
Domaine Universitaire, 351, cours de la Libération 33405 Talence Cedex FRANCE
maher.elarbi@gmail.com , chokri.benamar@ieee.org and henri.nicolas@labri.fr

ABSTRACT

This paper describes a blind video watermarking system invariant to geometrical attacks. Our scheme embeds different parts of a single watermark into different shots of a video under the wavelet domain. A multi resolution motion estimation algorithm (MRME) is adopted to preferentially allocate the watermark to coefficients containing motion. In addition, embedding and extraction of the watermark are based on the relationship between a coefficient and its neighbors. Experimental results show that inserting watermark where picture content is moving is less perceptible. Further, it shows that the proposed scheme is robust against common video processing attacks.

Index Terms: Invisibility, Motion, Wavelet Network, Watermark Strength.

1. INTRODUCTION

One of the most difficult problems in digital video watermarking is watermark recovery in the presence of geometric attacks like frame shift, cropping, scaling, rotation, and change of aspect ratio, especially when some of these are combined together [1,2]. The recovery problem is compounded for video since it must be carried out blind due to the difficulty of storing the original. In this case, for the typical spread spectrum (SS) watermarking system, blind retrieval is performed via cross-correlation between the marked video and the secret pseudo-noise (PN) sequence used to spread the watermark at the embedding stage. Recovery is straightforward given perfect synchronization between the attacked video and the PN sequence, but is difficult when geometric attacks destroy the synchronization. In this case it is possible to perform some form of sliding correlation in order to re-establish synchronization i.e. multiple cross correlations over a specified search space. Unfortunately the search space grows very quickly, making it difficult to recover the watermark in a reasonable time.

We have recently described a technique for video watermarking which combines neural networks and motion estimation in the wavelet domain [7]. In this paper, we propose a complete watermarking system that combines wavelet networks and motion estimation in the wavelet domain. Our scheme embeds different parts of a single watermark into different shots of a video under the wavelet domain. Each part of the watermark is allocated according to the motion intensity contained in the shot. Motionless shots are watermarked with small watermark. Whereas motioned shots are embedded with longer watermarks.

A multi resolution motion estimation algorithm is adopted to preferentially allocate the watermark to coefficients containing motion. Further, a wavelet network is given to memorize the relationships between the coefficients in a 3x3 block. The watermark detection process does not require the original video. The multi-frame based extraction strategy ensures that the watermark can be extracted correctly from a very short sequence of video. Individual frames extracted from the video sequence also contain watermark information.

The rest of the paper is organized as follows: Section 2 presents relation between motion and visibility. Section 3 introduces our video watermarking system. In section 4 we present some experimental results and section 5 concludes the paper.

2. RELATIONSHIP BETWEEN INVISIBILITY AND MOTION

It has been established from literature on the perception of motion [4] and from inspection of the watermark in video sequences that the relationship between the motion of an object and the visibility of the watermark in the object is shaped according to the curve in figure 1[10].

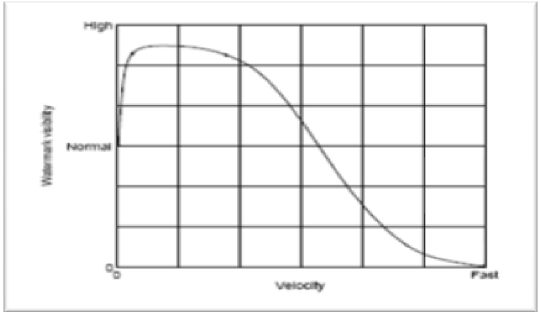


Figure 1: Relation between motion and visibility

When the object is at rest, the visibility of the watermark in the object is normal. When the object moves at a small velocity, the visibility of the watermark sharply increases to a level high above normal visibility. It remains high until the object moves so fast that the eyes can no longer track the object or keeps it in good focus.

From that point, the visibility of the watermark decreases towards zero visibility. Exact data on this curve could not be found in literature, nor could it be established whether the direction of the motion is of influence on the visibility of the watermark.

3. THE VIDEO WATERMARKING ALGORITHM

In our scheme, an input video is split into audio and video stream and only the video stream undergoes watermarking. An overview of the watermarking procedures is shown in figure 2.

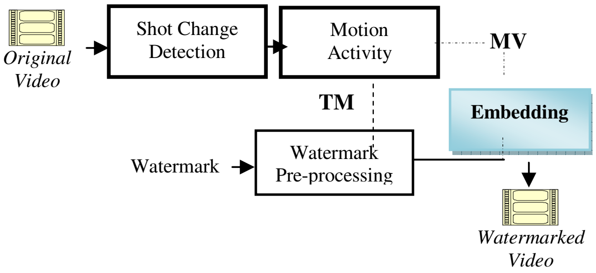


Figure 2: The proposed watermarking algorithm

First, scene changes are detected from the video by applying histogram difference method [5] on the video stream. Then a motion activity analysis is performed to obtain a motion intensity repartition in the video. On the other hand, a watermark is divided into different parts with variable length according to the amount of motion contained in the shot where it would be embedded. Details are provided in next sub sections.

3.1. Video Pre-processing

Our objective is to produce for each frame not only the coefficients containing motion, but also to track them while moving. To reach such a goal, a Motion Activity Matrix (MAM) is created to memorize for each coefficient: its position (x, y) in the frame, the frame where it stops moving and its motion Intensity in the shot.

Shot changes detected from the video and independent watermarks are embedded in frames of different shots. As shown in figure 4, watermark **m1** is used for the first shot. When there is a shot change, another watermark **m2** is used for the next shot. The watermark for each shot is chosen according to the motion intensity of the shot.

Let V be a video shot composed of N frames. A Motion Activity Matrix is obtained for each frame according to the following steps:

Step 1: Initialize Motion Activity of the coefficient (CM) to 0

Step 2: Get the corresponding motion vector and calculate its norm according to the equation.

Norm = $\sqrt{dx^2 + dy^2}$

Where dx and dy denote the amount of refinement.

Step 3: If norm = 0 (the coefficient stops moving) then save the coefficient position, the frame where it stops moving and its motion activity and go to the next coefficients. Otherwise, update the CM of the coefficient according to the following formula:

CM=CM+Norm

And go to the next frame.

These steps are repeated until all coefficients in the frame are treated.

3.2. Watermark Embedding

The new watermarking scheme we propose is based on Discrete Wavelet Transform and multi resolution motion estimation. Figure 3 shows an overview of our watermark embedding process.

3.2.1. Adaptive selection of embedding regions

Before actually embedding the watermark information, it takes full advantage of both intra-frame and inter-frame information of video content to select embedding regions adaptively and thus guarantees the perceptual invisibility of the watermark. In order to achieve such a goal, two criteria for the selection of embedding regions are introduced here: motion detection and detail detection mechanisms.

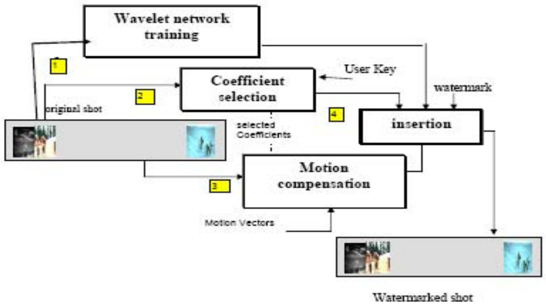


Figure 3: Watermark embedding process

On the intra-frame level, coefficients of LL i.e. low frequency sub-band) are not watermarked, as video energy is concentrated on lower frequency wavelet coefficients. If they are altered, it will affect perceptual quality. Furthermore, coefficients of HH (i.e. high frequency sub-band) are also not watermarked. It can make the watermark survive MPEG lossy compression as lossy compression removes the details of the image. Only middle frequency wavelet coefficients of the frame are watermarked.

On the intra-frame level, we start by watermarking the first frame of the shot as a single frame. Using Motion activity matrix of the current fame, we select coefficients having maximum motion intensity (if not enough coefficients are found and using the embedding key K, we select other coefficients even if they are at rest). Moreover, in the remaining frames, we perform motion compensation in order to obtain the new positions of the watermark. If any coefficient stops moving, we drop it form the list of candidates and select new coefficients from the MAM of the current frame using the same procedure described above.

3.2.2. Watermark Embedding Strategy

The first video frame is watermarked as a single image. First, a color transformation from the RGB color space to the YCRCB one followed by a three level DWT is performed. Then positions where the watermark will be embedded are selected with the help of a private embedding key (User Key).

Each bit of the watermark is inserted by altering the coefficient value of the original image according to the following formula:

if $w_i = 0$ Then $I_W(x, y) = \text{Max}\{I(x, y), \sigma 1 + \delta\}$
Elise $I_W(x, y) = \text{Min}\{I(x, y), \sigma 0 - \delta\}$

where w_i is the i^{th} bit of the watermark, I (x, y) is the original coefficient, I_W (x, y) is the watermarked coefficient and δ is the embedding strength, its value determines the watermark power and σ1, σ0 are determined as follow:

If $(B(x, y) - I(x, y)) >= \delta$ Then $\sigma 1 = B(x, y)$
Else $\sigma 1 = I(x, y)$
If $(B(x, y) - I(x, y)) >= \delta$ Then $\sigma 0 = I(x, y)$
Else $\sigma 0 = B(x, y)$

Where B (x, y) is the output of the wavelet network. Then, an inverse three level DWT followed by an inverse color transformation is performed to obtain the watermarked frame.

For the remaining frames, the same embedding process is performed but the positions of the watermark are obtained differently. Using motion vectors and selected coefficients in the previous frame, we compute the new positions of the watermark as follow:

Step1: compute new position for the selected coefficients in the first frame by motion compensation.

$$nx = x + dx$$
$$ny = y + dy$$

where (x, y) denotes the coefficient position in the first frame, (nx, ny) the new position in the current frame and (dx, dy) the amount of refinement.

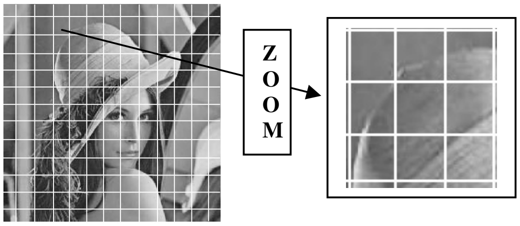
Step2: find coefficients that stops moving.

Step3: replace them by new coefficients from the current frame.

Using this algorithm we guaranties that the watermark is always carried by the same coefficient. Furthermore, if any coefficient stops moving, new coefficients are selected on the current frame and thus allowing that the watermark is always embedded in motioned coefficients.

3.2.3. Wavelet network training

We establish the relationship between the coefficients around a coefficient I by using the wavelet networks model [3]. For a selected coefficient I (x, y), the network is trained with its 3×3 neighbor, as input vector and the value of coefficient I(x,y) as output value.



We construct three layers Wavelet Network with 8, 10 and 1 wavelet in the input, hidden and output layer respectively, and the **Slog1** [8] mother wavelet is used for recognition. The wavelet network can approach the relationship and memorize it between the original image and the watermarked one.

To train the network, we must feed it with data sets. Three approaches can be adopted: use a wavelet network for each frame, use a wavelet network for each shot or use a single wavelet network for the entire video. In our method, we adopt the second approach. Figure 4 illustrates how the training sets are generated.

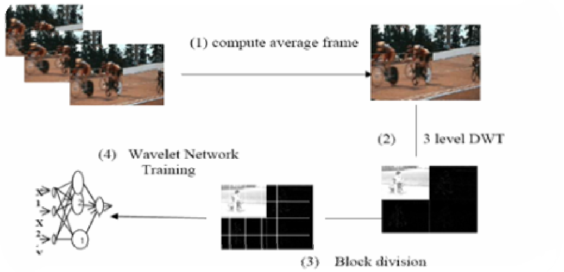


Figure 4: Wavelet Network Training

First we compute the average frame of each shot. Then the resulting frame is transformed to the wavelet domain with a three level DWT. Afterwards; it is divided into non-overlapping 3x3 blocks. The centre coefficient of each block is the output while the neighbor's coefficients are the input. Finally we proceed to wavelet network training until the specified goal or the maximum number of iteration reached. As a result, we obtain a trained network, which will be used later in the watermark embedding process.

3.3. Watermark Extraction

Authorized detection of the hidden information can be easily accomplished by using the watermarked video and the user secret key. The video stream is processed to get the video watermark. The watermark is first extracted from the first frame. According to the model of wavelet network the masked watermark can be retrieved as follows:

$$W = \begin{cases} 1 & \text{if } I_W(x, y) > B_W(x, y) \\ 0 & \text{Otherwise} \end{cases}$$

Where $I_W(x, y)$, $B_W(x, y)$ are the watermarked coefficient and the output of the Wavelet Network, respectively.

Using motion vectors, the new locations of the watermark in the remaining frames are determined. The same straightforward extraction strategy may be applied. If enough scenes are found and all parts of the watermark are collected, the original large watermark image can be reconstructed.

As an identical watermark is used for all frames, multiple copies of the watermark may be obtained. The watermark is recovered by averaging the extracted watermarks out of different frames. This reduces the effect if the attack is carried out at some designed frames.

4. EXPERIMENTAL RESULTS

To evaluate the performance of the new video watermarking scheme, several experiments have been done. They include frame shift, cropping, scaling, rotation, and change of aspect ratio. To compare with the proposed approach two other scheme are used. The first embeds different parts of a single watermark into different shots of a video and the second adopt a full video frames watermarking approach under the wavelet domain. A video clip with 300 frames of size 176x144 is used in our experiment. The NC is computed to be 1, which shows the exactly extraction. The watermark is an 8x8 binary image. The experimental results are described in details in the following.

4.1. Evaluation of the PSNR

The PSNR measures the signal to noise ratio of the watermarked video; thus, we can evaluate the video fidelity accordingly.

Video	The proposed Method	Scene based Method	All Frames Method
Shot 1	54,3366	47,764	36,3301
Shot 2	54,7224	48,1565	36,0595
Shot 3	54,6589	46,874	36,9283
Shot 4	55,0115	48,7575	36,7784

Table 1: Average PSNR for different shots

From the values depicted in table 1, it is clear that the proposed algorithm can successfully reduces the video frame distortion due to watermark embedding. Moreover, the fact that the watermark is embedded in motioned region, the perception of the distortion is reduced due to the proprieties of the human visual system in perception of motion [4]

4.2. Experiment with Aspect Ratio Change

Figure 5 shows the NC values of the extracted watermarks with different aspect ratio.

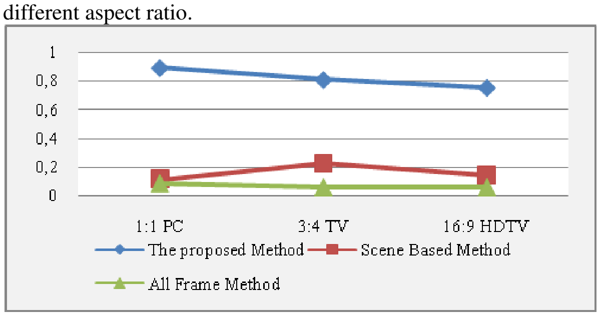


Figure 5: NC Values under Aspect ratio change

4.3. Experiment with rotation attack

Figure 6 shows the NC values of the extracted watermarks after a rotation with different degrees.

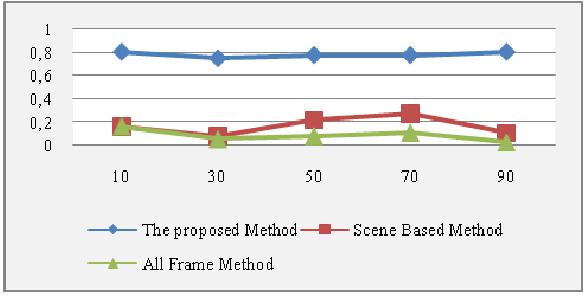


Figure 6: NC Value under rotation attack

4.4. Experiment with resize attack

Figure 7 shows the NC values of the extracted watermarks under resize and flip attacks.

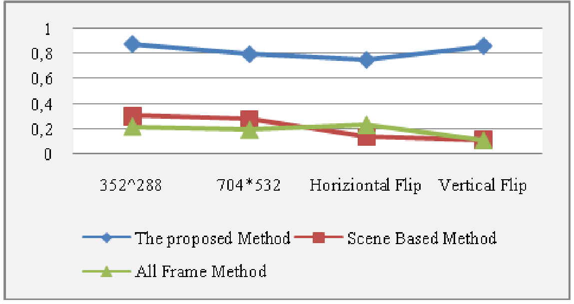


Figure 7: NC Values under resize

4.5. Experiment with cropping attack

Figure 8shows the NC values of the extracted watermarks with different resolutions.

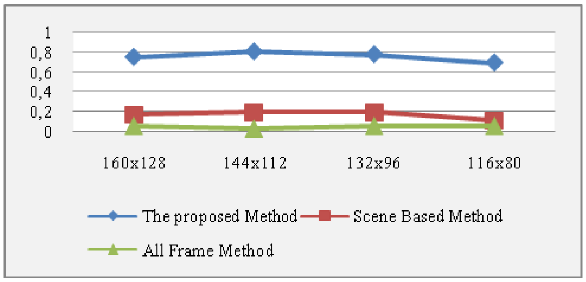


Figure 8: NC Values under cropping

Other attacks such as frame dropping and frame dropping were tested. In particular, the even frames the video were deleted and replaced with the average of the preceding and following frames. The watermark is still found in the newly derived (i.e., even) frames.

It's clear from the experimental results that the proposed algorithm is very robust against common geometric transformations such as rotation, horizontal flip, vertical flip, resize and so on. The NC values are much better in the proposed algorithm than those computed in the scene based and full frame algorithm. The positions where to embed the watermark are obtained according to motion activity and not to their physical positions in the frame. Moreover, the wavelet network used in watermark embedding to memorize the relationship between coefficients of the same block and motioned region embedding enhance the performances of the algorithm.

5. CONCLUSION AND FUTURE WORKS

In this paper, the techniques of wavelet networks and motion estimation have successfully been incorporated into video watermarking to develop a novel watermarking for digital video. The proposed method has effectively employed a wavelet network for memorizing the relationships between the original video and the watermarked video. Moreover, the method makes the watermark imperceptible via exploiting the motion characteristics of the video. Finally, the experimental results - illustrate that the method significantly possesses robustness to be immune against common attacks of digital video.

6. REFERENCES

[1] F. Petitcolas, R. Anderson, and M. Kuhn, "Attacks on copyright marking systems," Proceedings of the Second International Workshop in Information Hiding, April 14-17, 1998, Portland, OR, pp. 218-238.

[2] E. Lin and E. Delp, "Temporal synchronization in video watermarking," Proceedings of the SPIE Security and Watermarking of Multimedia Contents IV, vol. 4675, January 21-24, 2002, San Jose, CA, pp. 478-490.

[3] M. El'arbi, C.Ben Amar and A.Alimi, "A novel image watermarking algorithm based on wavelet networks" proceeding of the 3rd International Symposium on Image Video Communications over fixed and mobile networks (ISIVC 2006), September 13-15th 2006, Hammamet, Tunisia (to appear)

[4] K-R. Boff, L. Kaufman, and J-P. Thomas. "Handbook of Perception and Human Performance", volume I. Wiley-Interscience, 1986.

[5] J Breezy and L Rowe, "Comparison of video shot boundary detection techniques", In SPIE Conf. Storage & Retrieval for Image & Video Databases, volume 2670, pages 170–179, 1996.

[6] J.W Bae, S.H. Lee and J.S. Yoo "an efficient wavelet based motion estimation algorithm", IEICE transaction INF & SYST Vol E88-D,NO1,January 2005

[7] M. El'arbi, C.Ben Amar and H.Nicolas « A video watermarking algorithm based on neural network » proceeding of the IEEE International Conference on Multimedia & Expo (ICME 2006), pp 1577-1580, July 9-12 2006, Toronto, Ontaria, Canada.

[8] S. Haykin, Adaptive Filter Theory, 2nd Ed., Prentice Hall, Inc., Englewood Cliffs, NJ, 1991.

[9] Steven Schimmel, "Motion sensitive video watermarking: final report," August 2001 *Master's Thesis, Delft University of Technology, The Netherlands. Published as Philips Research Technical Report 2001/825, 2001.*