VIDEO WATERMARKING BASED ON NEURAL NETWORKS

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@isimsf.rnu.tn , chokri.benamar@enis.rnu.tn and henri.nicolas@labri.fr

ABSTRACT

In this paper, we propose a novel digital video watermarking scheme based on multi resolution motion estimation and artificial neural network. A multi resolution motion estimation algorithm 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 wavelet coefficient and its neighbor's. A neural network is given to memorize the relationships between coefficients in a 3x3 block of the image. Experimental results show that embedding watermark where picture content is moving is less perceptible. Further, it shows that the proposed scheme is robust against common video processing attacks.

1. INTRODUCTION

The sudden increase in watermarking interest is most likely due to the increase in concern over copyright protection of content. With the rapid growth of the Internet and the multimedia systems in distributed environments, digital data owners can easy transfer their multimedia documents across the Internet. However, current technology does not protect their copyrights properly. This leads to wide interest of multimedia security and multimedia copyright protection and it has become a great concern to copyrights owners in recent years. In the early days, encryption and control access techniques were used to protect the ownership of media. Digital watermarking, the art of hiding information in a robust and invisible manner, has been investigated as a complementary technology [1].

There are some desirable characteristics of effective watermarking techniques, including imperceptibility, robustness, and security [2]. Imperceptibility is the degree of perceptual similarity between the original and watermarked signals. It is desirable to embed the watermark in a discreet, unobtrusive manner so that the watermark is imperceptible under casual observation. Robustness is the resilience of the watermark against manipulations such as lossy compression, linear and non-linear filtering, scaling, and cropping (Stirmark [3], dewatermarking attack [4]).Security is the ability of the watermark to resist hostile attacks. Attacks are not limited to removal of the watermark, but include watermark estimation or forgery, collusion, and ambiguity attacks. Obviously, it is desirable to have an imperceptible, robust, and secure watermarking technique.

Digital watermarking has recently been extended from still images to video. Many algorithms have already been proposed in the literature. It goes from the simple adaptation of watermarking algorithm initially designed for still images [5] to the definition of specific video watermarking scheme [6]. Open paths still remain in video watermarking. This technology is indeed in its infancy and is far from being as mature as for still images.

In this paper, we propose a novel approach to neural network watermarking for uncompressed video in the wavelet domain. The watermark is embedded multiple times into the host data. 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 and the robustness of the watermark to automated removal. A multi resolution motion estimation algorithm is adopted to preferentially allocate the watermark to coefficients containing motion. Embedding and extraction of watermark are based on the relationship between a wavelet coefficient and its neighbour's. A neural network is given to memorize the relationships among 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. Experimental results show that the embedded watermark is robust and invisible.

The rest of the paper is organized as follows: Section 2 introduces our video watermarking scheme. In section 3 we present some experimental results and section 4 concludes the paper.

2. THE PROPOSED VIDEO WATERMARKING SCHEME

The new watermarking scheme we propose is based on neural networks and multi resolution motion estimation. Figure 1 shows an overview of our watermarking process. In our scheme, a video is taken as the input, and perform multi resolution motion estimation [8] [9] [10] by a full search block matching method [7]. Then a neural network is trained to be used later in watermark embedding. Next, the first frame is watermarked as a single image.

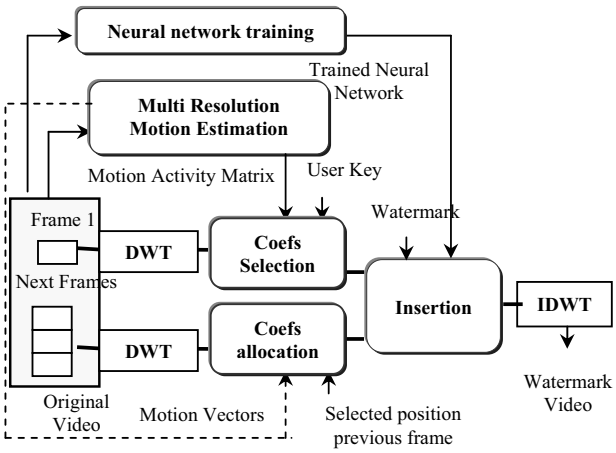

Fig1: Watermark embedding process

Afterwards, for the following frames, new positions of the watermark are obtained from motion vectors and selected positions in the previous frame. Once the new positions determined, the same embedding strategy is applied. Finally, the video is transformed back to time domain. Details are provided in next sub-sections.

2.1. Neural network for video watermarking

It is well-known that neural networks perform a highly adaptive nonlinear decision function from training examples. We establish the relationship among coefficients in the discrete wavelet decomposition of the image by using the Back-Propagation Neural Networks (BPNN) model [12]. For a selected coefficient, the network is trained with its 3×3 neighbors as input vector and the value of the coefficient as output. We construct three layers BPNN with 8, 10 and 1 neurons in the input, hidden and output layer respectively. The tangent sigmoid, purelin transfer function are used for recognition (figure 2).

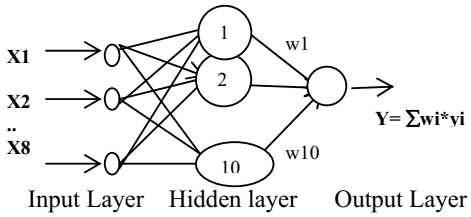

Fig 2: The Back Propagation Neural Network model

To properly train the neural network, we feed the model a variety of real life examples, called training sets. The data sets normally contain input and output data. The neural network creates connections and learns patterns based on this input and output data sets. Figure 3 illustrates how the training sets are generated. First we compute the average frame of the video. 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 center of each block is the output while the neighbor's coefficients are the input. Finally we proceed to neural network training until the specified goal or the maximum number of iteration is reached. As a result, we obtain a trained network which will be used later in the watermark embedding process.

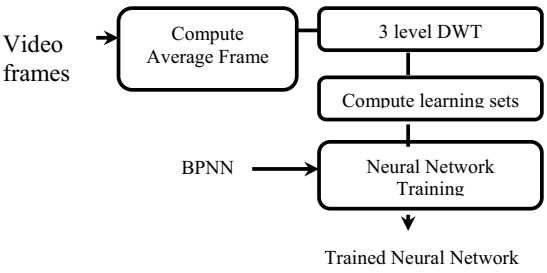

Fig 3: Neural network training process

2.2. Watermark embedding

The message bits to be embedded into the video represent usually either the ownership information or the fingerprint information. The length of the message bit is set to be 32. To enhance the performance of the detector, the message may be encoded with error correcting code [11].

The first video frame is watermarked as a single image. First, it is transformed to wavelet domain with a three level DWT. Then positions where the watermark will be embedded are selected with the help of a private embedding key (User Key). An effective block classifier is designed based on frequency localization and motion information. 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.

On the intra-frame level, middle frequency coefficients are selected. The basic idea is that the human eyes are sensitive to the low frequency noise and the quantization step of lossy compression may discard the high frequency components. On the inter-frame level, an empty two dimensional matrix is created and initialized with zeros. For convenience, this matrix is called Motion Activity Matrix. We compare all the corresponding motion vectors of two consecutives frames. If they have nonzero displacements in both directions, increase the matrix value in the corresponding position by one. Otherwise, it remains without any increasing or decreasing. This step is repeated until all consecutives frames pairs are compared. Only middle frequency coefficients that lie in region where picture content is moving are effectively selected with the help of a user secret key.

Once the locations chosen, each bit of the watermark is inserted by altering the coefficient value of the original image according to the following formula:

$$Iw(x, y) = \begin{cases} \text{Max } \{I(x, y), \sigma 1 + \delta\} \text{ if } wi = 0 \\ \text{Min } \{I(x, y), \sigma 0 - \delta\} \text{ if } wi = 1 \end{cases} \quad (1)$$

where w_i is the i^{th} bit of the watermark, I(x , y) is the original coefficient, Iw(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:

$$\sigma 1 = \begin{cases} B(x, y) & \text{If } (B(x, y) - I(x, y)) >= \delta \\ I(x, y) & \text{If } (B(x, y) - I(x, y)) < \delta \end{cases} \quad (2)$$

$$\sigma 0 = \begin{cases} I(x, y) & \text{If } (B(x, y) - I(x, y)) >= \delta \\ B(x, y) & \text{If } (B(x, y) - I(x, y)) < \delta \end{cases} \quad (3)$$

where B(x, y) is the output of the neural network. Finally, an inverse three level DWT 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. In fact, using motion vectors and selected coefficients in the previous frame, we compute the new positions of the watermark as follow:

$$nx = x + dx \quad (4)$$

$$ny = y + dy \quad (5)$$

where (x, y) denotes the coefficient position in the previous frame, (nx, ny) the new position in the current frame and (dx, dy) the x and y displacement for the block containing (x, y). Once locations of the watermark are determined, the same embedding strategy is applied.

2.3. Watermark detection

Authorized detection of the hidden information can be easily accomplished by using the watermarked video and the user secret key. All the watermarked frames are transformed to the wavelet domain by a three level DWT. Once the locations of the watermark are determined, a straightforward extraction strategy may be applied. According to the model of BPNN and the secret key the masked watermark can be retrieved as follows:

$$W = \begin{cases} 1 & \text{if } Iw(x, y) > Bw(x, y) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where Iw(x, y), Bw(x, y) are the watermarked and the output of BPNN coefficient, respectively.

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 from different frames. This reduces the effect if the attack is carried out at some designed frames. After extracting the watermark, similarity measurements of the extracted watermark w' and the referenced watermark w is used for objective judgment of the extraction fidelity and it is defined as:

$$NC = \frac{\sum_{i=1}^{n} wi * w'i}{\sqrt{\sum_{i=1}^{n} w'i^2 * \sum_{i=1}^{n} wi^2}} \quad (7)$$

which is the cross correlation normalized by the reference watermark energy to give unity as the peak correlation (NC). If NC is greater than a pre-specified threshold, then the watermark is determined to be present in the image. Otherwise, it is not. Experimental results suggest a value of 0.4 should be used for the test video.

3. EXPERIMENTA RESULTS

Football sequence (176x144) is utilized during simulations, the message "mark" is used as a watermark, the user key is set to 1234 and the training goal for the neural network is 0.00025. Only the first 100 frames of the sequence are used and the watermark is embedded only into the Y component. Typical sample from original frame from the video are shown in figure 4. The NC is computed to be 1, which shows the exactly extraction.

Fig. 4: (a) Original frame (b) Watermarked frame

To evaluate the performance of the watermarking scheme, several experiments have been done. They include experiments with various dropping ratio, the experiment with various number of frames colluded, the experiment with various quality factor of MPEG compression. Another DWT based watermarking scheme which embed the watermark into wavelet domain in fix coefficients m under same conditions, is used to compare with the proposed watermarking scheme. The NC values are retrieved when the watermarked video is under different attacks. The experimental results are described in details in the following.

3.1. Visibility Measure

Table 1 illustrates the PSNR values of some watermarked frames.

	Frame 1	Frame 2	Frame 3	Frame 4	Frame 5
Moving	45,2	45,9	45,9	40,8	41,9
Fix	33,4	33,2	33,2	31,8	30,8

Table 1 PSNR values for football video

The LL sub-band is not involved in watermark embedding which makes the distortion, due to watermark embedding, invisible. As we can see, the distortion due to watermarking where picture contents are moving is less perceptible.

3.2. Robustness to lossy compression attack

One of the possible operations on video is a lossy coding stage applied for the purpose of storage and transmission of digital video at low bit rates. In the testing process, the bit rate is decreased until 224 kbps. It is observed that the watermark may survive until this bit rate. The correlation results are presented in figure 5.

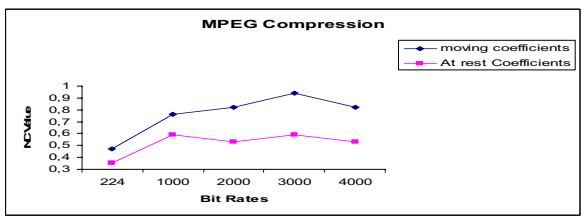

Fig 5: NC values against MPEG compression

Both methods satisfy the following condition. Higher frequency part of the video frame and high frequency sub-band DWT coefficients (HH) of video frame are not watermarked. This approach makes the watermark survive MPEG lossy compression, as lossy compression removes the details of the image. The performances of the proposed scheme are significantly improved by using motion compensation.

3.3. Robustness to frame dropping attack

Some other distortions, which are based on temporal characteristics of the digital video, are temporal cropping, frame dropping and frame averaging. An attacker can maintain the visual quality of the digital video by dropping some frames of the video and/or replacing then by frame interpolation. For the frame dropping case, a number of frames are dropped from the video. Figure 6 illustrates the NC value against the amount of frames dropped.

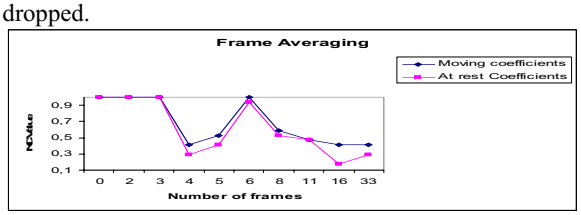

Fig 6: NC values against frame averaging

All frames are embedded with the same watermark. This prevents attackers from removing the watermark by frame dropping. If they try to remove the watermark, they need to remove the whole trunk of frames and this would lead to a significant damage to the video. Here also, the approach with moving coefficient gives better results.

3.4. Robustness to frame interpolation attack

For the frame interpolation case, we replace a number of frames by the average of the two neighbouring frames i.e. F (2n) = (F (2n-1) +F (2n+1))/2. Figure 7 illustrates the NC value against the amount of frames averaged.

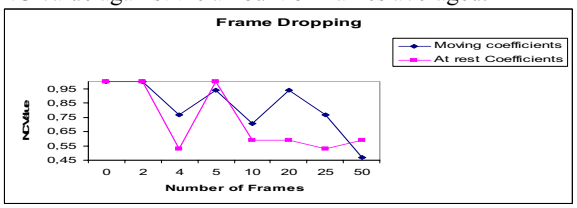

Fig 7: NC values against frame dropping

Identical watermark used within all frames in the same coefficients, can prevent attackers from removing the watermark by comparing and averaging the frames statistically. Embedding in moving coefficient can improve the performances of the watermarking algorithm because moving objects are difficult to estimate.

4. CONCLUSION AND FUTURE WORKS

In this paper, we have proposed a video watermarking **a**lgorithm which combines Neural networks with motion estimation in the wavelet domain. Experimental results show that embedding watermark where picture content is moving is less perceptible. Further, it shows that the proposed scheme is robust against common video processing attacks. Further studies will be carried to complete with the proposed approach. We have not tried yet all kinds of attacks such as noise pollution, channel error and so on. In addition, in this paper we wanted to demonstrate that embedding in motioned coefficients can give better performances. A complete watermarking system will be developed.

5. REFERENCES

[1] F. Petitcolas, R-J. Anderson and M-G. Kuhn, "Information hiding a survey", Proceeding of the IEEE (Special Issue on Protection of Multimedia Content), vol.87 no.7 pp.1062-1078, July 1999.

[2] R. Barnett, "Digital watermarking: Applications, techniques, and challenges", Journal of Electronics and Communication Engineering, Vol.11 n°4, pp: 173–183, August 1999.

[3] F. Petitcolas, R. Anderson and M. Kuhn, "Attacks on copyright marking systems", in Proceeding of the Second International Workshop on Information Hiding, Lecture Notes in Computer Science, Vol. 1525, Springer, Berlin, pp. 218–238 , 1999.

[4] C. Rey, G. Doerr, J.L. Dugelay and G. Csurka, "toward generic image dewatermarking" In Proceeding of the IEEE International Conference on Image Processing, Vol.3, pp. 633–636, 2002.

[5] D. Mitchell, S-B. Zhu and A.H. Tewfik, "Multi resolution Scene-Based Video Watermarking using perceptual Models", IEEE Journal on Selected Areas in Communications, Vol. 16, No. 4, May 1998.

[6]A. Alattaar ,M. Celik et E. Lin, "watermarking low bit-rate advanced simple profile MPEG-4 bitstreams" in proceedings of the international conference on Acoustics, speech and Signal Processing, Hong Kong 2003.

[7] C. H. Cheung and L. M. Po, "Novel Cross-diamond-hexagonal Search Algorithms for Fast Block Motion Estimation," *IEEE Trans. on Multimedia*, vol. 7, No. 1, pp. 16 - 22, Feb 2005

[8] F-G. Meyer, A. Averbuch and R-R. Coifman "motion compensation of wavelet coefficients for very low bit rate video coding", proceeding of the IEEE international conference on image processing. Vol 3, pp 638-641, 1997.

[9] 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.

[10] D.-W. Sun and J. Yoo, "A Fast Motion Estimation using Characteristics of Wavelet Coefficients," The Journal of The Korean Institute of Communication Sciences, vol.28, no.4C, pp. 397-405, Apr. 2003.

[11] D. Johnson, "error correcting codes: hamming codes", April 15, 2005. Available at: http://cnx.rice.edu/content/m0097/2.24.

[12] M. Shiang Hwang , C. Chang and K. Hwang "Digital watermarking of images using neural networks", Journal of Electronic Imaging Volume 9, Issue 4, pp. 548-555, October 2000.