Robust Video Watermarking Scheme in Transform Domains

Ersin Elbasi, Ahmet M. Eskicioglu

Abstract—**Due to large amount of frames, similarity between frames and temporal attacks (frame dropping, frame averaging, frame swapping etc.), video watermarking process is more difficult than image watermarking. Current image watermarking methods are not adequate to solve these difficulties. We propose a novel uncompressed video watermarking system based on Hidden Markov Model (HMM) and Artificial Neural Network (ANN). The proposed watermarking scheme splits the video sequences into Group of Pictures (GOP) with HMM. Portions of the binary watermark will be embedded into each GOP with a selected transformation domain watermarking algorithm. For each GOP, ANN produces the optimal transformation algorithm. The embedding process is the standard additive algorithm in low and high frequencies in different transformation domains. This novel system increases the robustness against geometric and temporal attacks, and increases the quality of the watermarked video.**

Keywords—**Watermarking, Hidden Markov Model, Neural Network, Frequency Domain, Group of Pictures.**

I. INTRODUCTION

Digital watermarking has received increasing attention in recent years. Distribution of movies, music, and images is now faster and easier via computer technology, especially on the Internet. Hence, the content owners (e.g., movie studios and recording companies) are concerned about illegal copying of their content. Watermarking and cryptography are two standard multimedia security methods. However, cryptography is not an effective method because it does not provide permanent protection for the multimedia content after delivery to consumers. The most important properties of a watermarking system are robustness, invisibility, data capacity, and security. An embedded watermark should not introduce a significant degree of distortion in the cover multimedia element.

Ersin Elbasi: The Scientific & Technological Research Council of Turkey, ersin.elbasi@tubitak.gov.tr
Ahmet M. Eskicioglu: The City University of New York, eskicioglu@sci.brooklyn.cuny.edu

Robustness is the resistance of the watermark against normal A/V processes or intentional attacks. Data capacity refers to the amount of data that can be embedded without affecting perceptual transparency. The security of a watermark can be defined to be the ability to thwart hostile attacks such as unauthorized removal, unauthorized embedding, and unauthorized detection. There are basically two approaches to embed a watermark: spatial domain and transform domain (e.g., DCT, DFT, or DWT) [1,2]. In the spatial domain, the watermark is embedded by modifying the pixel values in the original cover image. Transform domain watermarking is similar to spatial domain watermarking; in this case, the transform coefficients are modified.

Most of the current video watermarking techniques are based on the image watermarking methods [3], and directly applied to uncompressed or compressed video sequences [4]. However, these methods are not sufficient for copyright protection in video data. Video watermarking has a number of issues, and image based algorithms could not solve these problems. Embedding large amount of data, redundancy between frames, and robustness against temporal attacks (e.g., frame averaging, frame dropping, and frame swapping) are some of the main problems in video applications [5]. Neither embedding the same watermark to each frame, nor embedding different watermarks in every frame of the video would be robust against all types of common attacks.

An important criterion for classifying watermark schemes is the type of information needed by the detector.

- Non-Blind Schemes [6]: Both the original image and the secret key(s) are needed in detection.
- Semi-Blind Schemes [7]: The secret key(s) and the watermark bit sequence are needed.
- Blind Schemes [8,9]: Only the secret key(s).

To provide the necessary properties (robustness, invisibility, data capacity, and security), we developed a new watermarking method in video sequences. The proposed algorithm decomposes the binary visible watermark into m sub images, and embeds each into GOPs. Best time durations for GOPs will be calculated using Hidden Markov

Model (HMM) and Neural Network decides best frequency domain for each GOP [10]. This method provides robustness against common geometric and temporal attacks. Figure 1 shows general structure of proposed method.

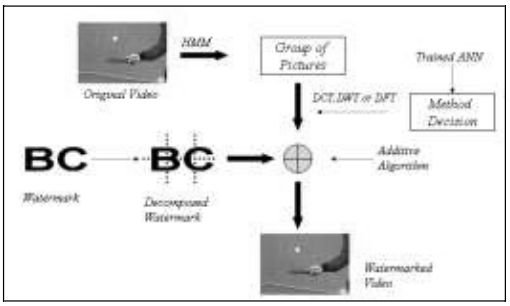

Figure 1. General Structure

II. PREPROCESS

1. In the system, we use *N*x*M* binary watermark. Our purpose is to embed different watermarks into different GOPs. For this purpose, we divide the current binary watermark into m equal parts. This is a simple process and each portion of the watermark will be embedded into low and high frequencies of each frame in the video sequence.

2. HMM based watermarking works with probabilistic distribution in video sequence. There are two methods to calculate probabilities: The first method is the Naïve Bayes Classifier after feature extraction, and the second method is the overall performance calculation for each frame, which indicates accuracy in watermarking based on the robustness and quality of the watermarked frame. Naïve Bayes Classifier predicts class membership probabilities, such as the probability that a given image belongs to class "*Watermark Present"* or "*Watermark Absent.*" We use some statistical attributes such as number of selected coefficients, mean, variance, range of the coefficients, etc. as a feature set after transforming frames to DWT [11,12].

In the second method, we calculate probability for each frame, which shows accuracy of the current embedding algorithm in that frame. There are several performance measurements in watermarking research. We want to provide invisibility, robustness and data capacity requirements. PSNR is only for invisibility measurement, and SR should be used for robustness measurement. There is no measurement defined for all three requirements. In this work, we define a measurement for overall performance.

$$OP = a_1 \times P_1 + ... + a_k \times P_k$$

The OP is the overall performance, and a_k is the distribution probability and P_k is the calculated accuracy probability. Each P indicates different measurements such as success in invisibility, data capacity, cropping attack, rotation attack, and other attacks. Based on the application we can increase the number of the Ps.

$$\bigtimes_{i=1}^{k} a_i = 1$$

There are two optimization problems in video preprocessing:

1. Optimal matching between watermark portions and group of pictures.
2. Optimal time duration decision for each GOPs.

A preprocessed set of watermarks is prepared in watermark preprocessing, $W = W_1, W_2, ... W_N$, where W_i is the portion of the single visual binary watermark. The extracted feature vectors of each frame is another input for the HMM algorithm, $F = F_{(1,t)} = F_1, F_2, ..., F_t$, where F_i is the feature vector for frame *i*, and *t* is the number of the frame in video sequences. The optimal criteria is to find the sequence of frames most likely to produce optimal watermarking with the watermark portion.

3. *ANN Training:* In training, set of pictures have taken from 4 different video sequences. Based on the objective and subjective evaluation, for each picture is classified as DCT, DFT or DWT [13,14]. Training evaluation techniques are (in both watermarked image quality and resistant against to common attacks):

a. Subjective Evaluation
b. PSNR [1]
c. M-SVD

Extracted features are trained using Backpropagation algorithm using following formulas.

$$in_i = \sum_j W_{j,i} \times a_j = W_i \times a_i$$

$$a_i \leftarrow g(in_i) = g(\sum_j W_{j,i} \times a_j) = g(W_i \times a_i)$$

Where *in* is the input feature vector, *W* is weight and *g* is the sigmoid function.

III. EMBEDDING AND EXTRACTION

Binary watermark embedding procedure is given as follow [6]:

Watermark Embedding Algorithm:

For each frame in GOP, the following embedding procedure is applied.

a. Convert the *N*x*M* RGB frame to YUV.
b. Compute the frequency domain of the luminance layer (Y) for each frame.
c. Modify the coefficients V_{ij} in the lower and higher bands in all frames.
d. $V_{w,ij} = V_{i,j}^{k} + \alpha_k \, . \, W_{i,j}$, where $i = 1,\ldots,n$; $j = 1,\ldots,m$; k=1,2.
e. Apply the inverse transformation to obtain the watermark cover frame I_w for each frame.

Watermark Detection Algorithm:

There are some pieces of information stored during the embedding process such as durations of the group of frames. Based on this information, we apply the following procedure to extract the watermark [6].

1. Split video sequence to group of frames.
2. Apply the following algorithm in frames.
   a. Convert *N*x*M* RGB frames to YUV.
   b. Compute the frequency domain of the luminance layer (Y) for each frame.
   c. Extract the binary visual watermark from the low and high bands.
   d. $W_{ij}^{*} = (V_{w,ij}^{*k} - V_{ij}^{k}) / \alpha_k$ where $i = 1,\ldots,n$; $j = 1,\ldots,m$.
   e. If $W_{ij}^{*} > T$, then $W_{ij}^{*} = 1$ else $W_{ij}^{*} = 0$, where T is the threshold between 0 and 1.
3. Combine all watermark portions, and output the watermark.

IV. EXPERIMENTAL RESULTS

We validated the proposed algorithm using the tennis, akiyo and flower garden video sequences. The video sequence is about 20-40 seconds, and the frame sizes is 240x352.

Figure 2 shows the composed of sub watermarks. SR values show that the proposed algorithm increases the performance.

Figure 3 shows the extracted watermark portions and the composed watermark after common attacks. In the second column, there is another extracted watermark given which uses the standard additive algorithm instead of the proposed video watermarking algorithm.

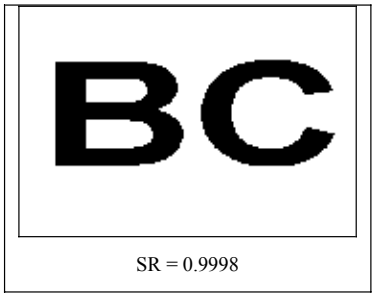

Figure 2. Composed watermark

Artificial neural network based watermarking transformation classification gives very promising results. Testing results with 80 frames taken from 4 different video sequences gives more than 90% accuracy. Table 1 shows the accuracy for each video sequences.

Table 1. Accuracy for video sequences

Video Sequence	Accuracy (%)
1	92.3
2	88.5
3	91.2
4	94.3

V. CONCLUSION

The most important properties of a video watermarking system are robustness, invisibility, data capacity, and security. There is a significant amount of research in video base watermarking, but the problem is still an open research area because of a number of challenging problems: embedding large amount of data, redundancy between frames, and robustness against temporal attacks (e.g., frame averaging, frame dropping, and frame swapping). To solve these problems, we proposed a novel HMM-ANN based algorithm. Our scheme embeds different parts of a single binary watermark into different Group of Pictures of a video sequence under the frequency domain. The proposed Hidden Markov Model obtains optimal matching and time durations in video sequences. Basically, an iterative HMM based Viterbi algorithm will be applied to decrease complexity. In the detection process, some information from the embedding process is used such as durations of GOPs. Experimental results show that this new algorithm has higher visual video fidelity and robustness against all kinds of attacks. It is very difficult to change, guess or remove the watermark with this proposed algorithm.

REFERENCES

[1] R. Dugad, K. Ratakonda and N. Ahuja, "A New Wavelet-Based Scheme for Watermarking Images," Proceedings of 1998 International Conference on Image Processing (ICIP '98), Vol. 2, Chicago, IL, October 4-7, 1998, pp. 419-423.

[2] C. Hsu, J. Wu, "DCT-Based Watermarking for Video," IEEE Transaction on Consumer Electronics, Vol. 44, No. 1, February 1998, pp. 206-216

[3] G. Doerr, J. Dugelay, "A Guide Tour of Video Watermarking," Signal Processing: Image Communication, Vol. 18, No.4 , April 2003, pp. 263-282.

[4] F. Hartung, B. Girod, "Digital Watermarking of Raw and Compressed Video," Proc. European EOS/SPIE Symposium on Advanced Imaging and Network Technologies, Berlin, Germany, October 1996.

[5] Pik-Wah Chan, Michael R. Lyu and Roland T. Chin, "Copyright Protection on the Web: A Hybrid Digital Video Watermarking Scheme," Proceedings of the 13th International World Wide Web Conference (WWW '04), New York , May 17-22, 2004 , pp.354-355.

[6] P. Tao and A. M. Eskicioglu, "A Robust Multiple Watermarking Scheme in the DWT Domain," Optics East 2004 Symposium, Internet Multimedia Management Systems V Conference, Philadelphia, PA, October 25-28, 2004, pp. 133-144.

[7] E. Elbasi and A. M. Eskicioglu, "MPEG-1 Video Semi-Blind Watermarking Algorithm in the DWT Domain," IEEE International Symposium on Broadband Multimedia Systems and Broadcasting 2006, Las Vegas, NV, April 6-7, 2006.

[8] H. Wang, Z. Lu, J. Pan, S. Sun, "Robust Blind Video Watermarking with Adaptive Embedding Mechanism," International Journal of Innovative Computing, Information and Control, Vol. 1, Number 2, June 2005.

[9] Pik-Wah Chan and Michael R. Lyu, "Digital Video Watermarking with a Genetic Algorithm," Proceedings International Conference on Digital Archives Technologies Technologies (ICDAT '05), Taipei, Taiwan, June 16-17, 2005, pp. 139-153.

[10] Rabiner, L.R, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceedings of the IEEE, Volume 77, Issue 2, Feb 1989 Page(s):257 – 286.

[11] Shaohui L, Hongxun Y, Wen G, "Neural Network Based Steganalysis in Still Images," Multimedia and Expo, 2003. ICME'03, Vol. 2 , 2003.

[12] Ersin Elbasi, Ahmet M. Eskicioglu, "Naïve Bayes Classifier Based Watermark Detection in Wavelet Transform", Int. Worksop on Content Represention, Classification and Security September 11-13, 2006, Istanbul.

[13] K. Mehrotra, C. K. Mohan, S. Ranka, "Elements of Artificial Neural Network." MIT Press, pp. 70-94, 2000.

[14] Shaohui L, Hongxun Y, Wen G, "Neural network based steganalysis in still images," Multimedia and Expo, 2003. ICME'03, Vol. 2 , 2003.

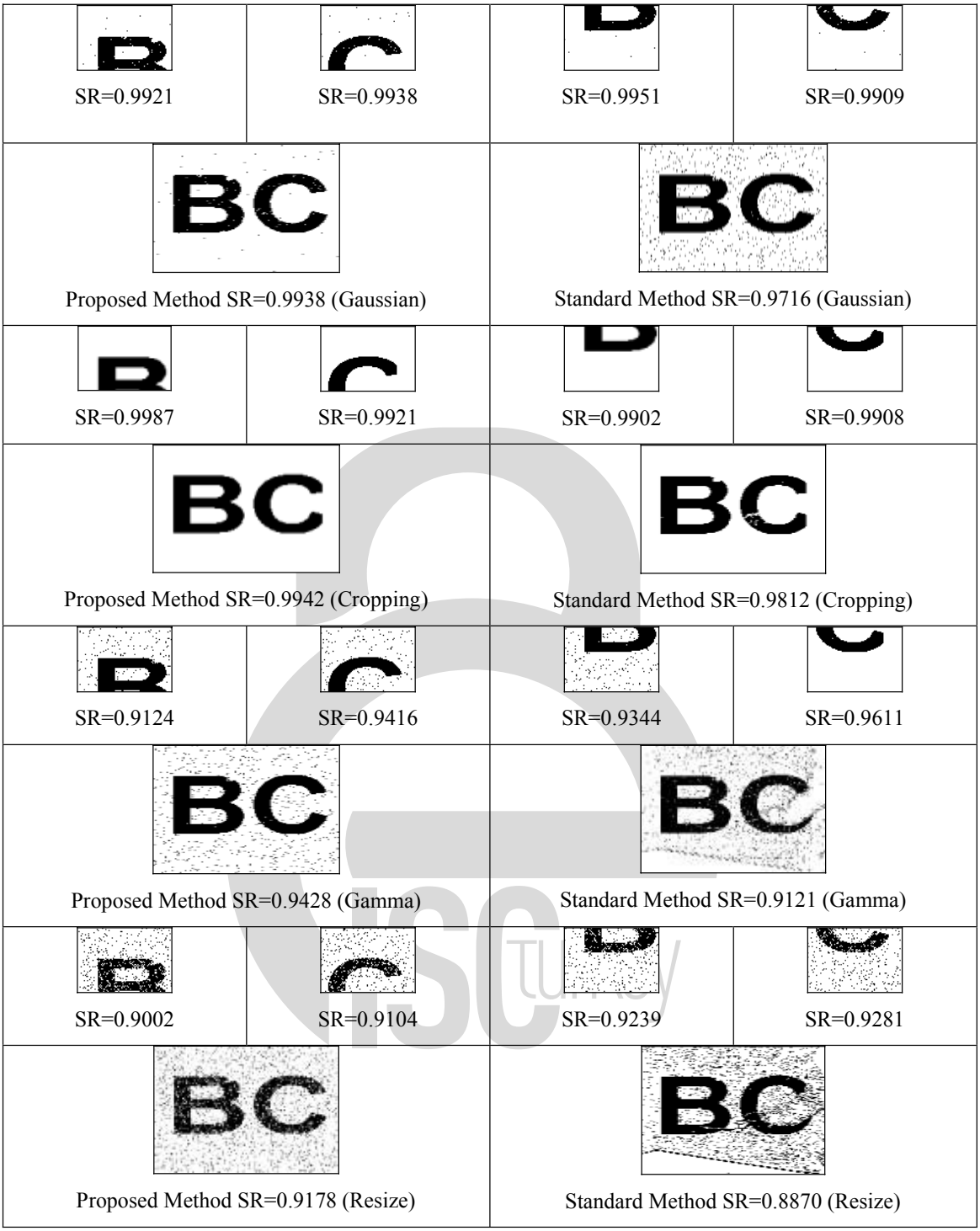

Figure 3. Extraction results after some common attacks