Video Watermarking Synchronization Based on Profile Statistics

Shih-Wei Sun and Pao-Chi Chang
National Central University

ABSTRACT

A novel temporal synchronization method for video watermarking by matching the profile statistics. The profile statistics, represented by the characteristic parameters such as position mean and variance in the x- and y-directions, of a frame in a video sequence can easily be calculated and sent as side information to the receiver. At the receiving end, temporal attacks such as transposition, dropping, and insertion can be detected by comparing side information and characteristic parameters calculated from the received video. The simulation results show that the proposed method can successfully re-synchronize the attacked video back to the original format with accuracy from 72.41% to 98.15% for various video sequences based on single frame matching. After the voting process, the GOP detection accuracy can be improved to the range of from 96.30% to 100%.

INTRODUCTION

Many video watermarking techniques were proposed in recent years, but most are relatively weak to temporal attacks. Although video watermarking techniques that utilize the video-specific characteristics are under development, the watermark synchronization for temporal distortion is a challenging problem [1]. However, since no video watermarking standards are determined, it is desirable to develop a re-synchronization technique that can fit most video watermarking systems. Therefore, we propose a temporal re-synchronization method independent of the video watermarking system.

Conventional video watermarking systems use time-invariant key schedules. Lin et al. proposed a temporal redundancy key-based video watermarking synchronization system [2, 3] that uses a periodic watermark key to provide higher security. In their work, a pixel value of a certain location is chosen as the feature used for synchronization. They declare that if the feature is robust, their key schedule can be detected successfully. However, once the video sequence is under attack, the feature they proposed seems weak and may lose synchronization. Therefore, the watermark cannot be extracted successfully.

Feature extraction is an important topic in face detection. In [4], knowledge-based methods for finding robust feature extraction were surveyed. Yang and Huang used a multi-resolution hierarchy-based method for face detection [5]. The multi-resolution hierarchy of images is created by averaging and sub-sampling the image to create the center of the face. Kotropoulos and Pitas [6] presented a rule-based localization method by calculating the horizontal and vertical projections of the image.

The local minimums of the horizontal and vertical profiles are recorded as the feature points in the input image. The 86.5% successful detection ratio shows the feature seems robust.

In our proposed method, the profile-based feature is enhanced by including the statistics that contain the first- and second-order moments of the profiles in x- and y-directions, which can represent the coarse scale of an image or a frame. The feature is used to resynchronize the video to the original temporal order. Consequently, the watermark can be successfully extracted from the synchronized frames and keys.

Fig. 1. Akiyo

Authors' Current Addresses:
S-W. Sun and P-C. Chang, Electrical Engineering, National Central University, Chung-Li, Taiwan.

Based on a presentation at Carnahan 2003.
0885/8985/04/ $17.00 © 2004 IEEE

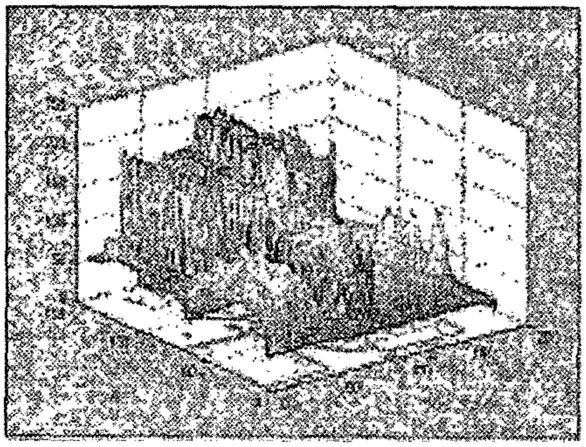

Fig. 2. 3-D Meshing Function

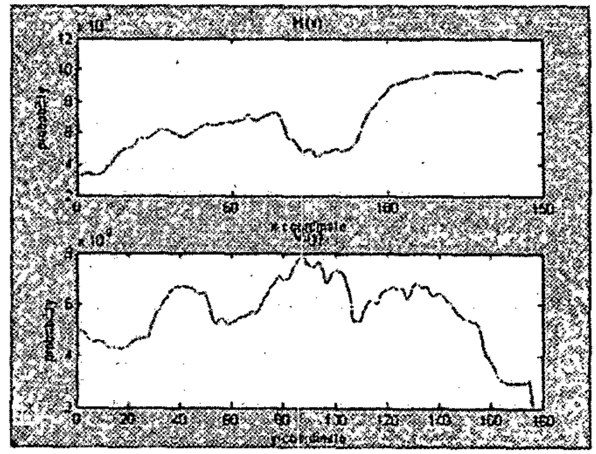

Fig. 3. The Profiles

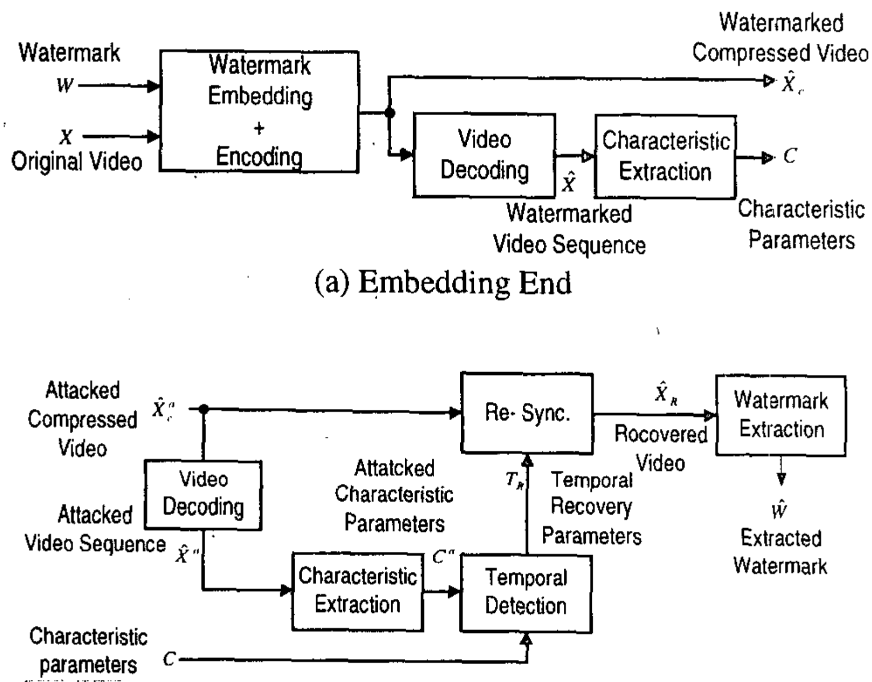

Fig. 4. The Proposed System Architecture

The outline of the paper is as follows: the proposed profile statistics are described in the next section; system architecture is described in the like-titled section; followed by a discussion on the temporal detection procedure; next the experimental results are described; and finally, conclusions are presented.

PROFILE STATISTICS

In digital watermarking, the content of an image or a frame will possibly be attacked and become different from the originals. The feature extracted from the profile statistics of a frame should be robust enough under certain kinds of attack.

Profile of an Image

Here is an example that shows how to get the profiles from a video frame. Figure 1 is the first frame of the video source Akiyo. Furthermore, the pixel intensity can be represented by the amplitude axis of the 3-D meshing function, the image can be transformed from a 2-D model to a 3-D model, as shown in Figure 2. By summing the meshing function along the x- and y-directions, we obtain the profiles of the image, as shown in Figure 3, where the upper part represents the profile along the x-direction, while the lower part shows the profile in the y-direction. The profiles of the image intensity I(x, y) can be derived in two directions – x and y.

Table 1. Frame Synchronization without Voting

Class	Accuracy	Transposition	Dropping	Insertion
A	Akiyo	75.56	72.59	72.41
B	Foreman	98.15	93.70	95.17
C	Stefan	84.44	81.85	83.45

Table 2. GOP Synchronization with Voting

Class	Accuracy	Transposition	Dropping	Insertion
A	Akiyo	100	96.30	96.55
B	Foreman	100	93.30	96.55
C	Stefan	100	96.30	96.55

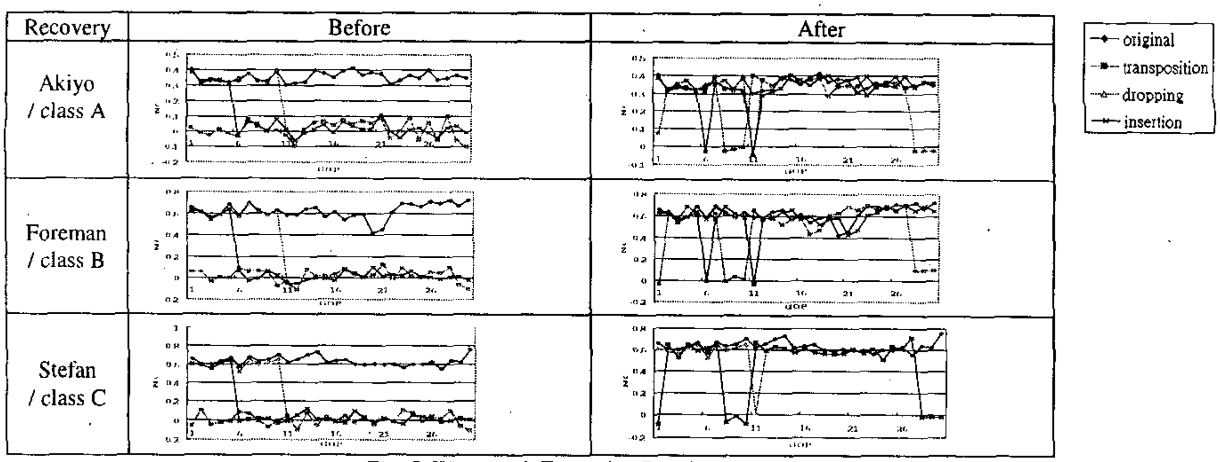

Fig. 5. Watermark Extraction Results

Moments of the Profiles

Since the characteristics of a frame will be sent as side information for synchronization, the number of bits representing the characteristics should be kept as few as possible. In our proposed method, the mean and variance that represent the statistics of the first- and second-order moments are used in evaluating probability characteristics [7].

Furthermore, the characteristic vector of the original image C and the attacked C^a can be presented.

In addition, the similarity, evaluated by the Normalized Correlation (N.C.) [8], can be calculated.

SYSTEM ARCHITECTURE

We propose a temporal synchronization scheme for a video watermarking system. The method can be applied to most of the existing video watermarking schemes to enhance the robustness, especially against temporal attacks. The block diagram is depicted in Figure 4.

Figure 4A shows the embedding end of the proposed watermarking system. The original video sequence X is first embedded by the watermark W and compressed to produce a watermarked compressed video χ_c. After the video decoding process, the video can be recovered to the watermarked sequence χ in the pixel domain, which contains spatial characteristics. Based on the analysis of profiles in the previous section, the characteristics parameters C can be extracted from each frame and recorded as the prior knowledge [9].

Figure 4B shows the extracting end of the proposed watermarking system. The input is the attacked compressed video. The possible attacks include compression, filtering, noising attacks, especially temporal attacks, such as frame dropping, frame insertion, and transposition. Because temporal attacks may affect the video significantly, the video may not be recoverable because of out-of-synchronization. Therefore, we propose a method that offers a re-synchronization technique to recover the attacked image to the proper position in temporal axis. The attacked video sequence in the pixel domain must be obtained first by video decoding. The characteristic extraction process is similar to that in the embedding end, and then characteristic parameters set C^* is extracted. By comparing the two sets of characteristic parameters, C received from the side information and C^* computed from the received video, the temporal recovery parameters T_R can be obtained and used for re-synchronization. Finally, the re-synchronized video X_R is ready for extracting the watermark W.

TEMPORAL DETECTION

The main purpose of this method is to re-synchronize the temporal attacked video to the proper time relation for accurate watermark extraction. The procedure using a simple set of characteristic parameters for temporal re-synchronization is detailed below.

The parameter C_i represents the characteristics of frame i. The original video characteristics CV and the attacked video characteristics CV^* can be formed by the set of C_i and C_i respectively.

By comparing the characteristics contained in CV and CV^*, the received frames can be re-synchronized. The frame index re-synchronization is determined by the nearest neighborhood rule in which the attacked frame should be fit to the most similar frame in the original video.

Moreover, video compression and attacks are often processed based on the group of pictures (GOP). In the case that the watermark embedding procedure uses the same key for all frames in a group of pictures (GOP) and the temporal distortion is also based on GOP, the re-synchronization accuracy can be further improved by a voting process. The re-synchronization accuracy can be relaxed from a frame to a GOP. The GOP index for frame i is determined by the majority rule voting results of all frames in the same GOP. The resulting index sequence V^* represents the re-synchronized GOP indices for all frames j, j=0,1,2, δ , M−1. Finally, the temporal recovery parameter T_R is sent to the re-sync block to produce the recovered video, then the watermark can be extracted successfully.

EXPERIMENTAL RESULTS

In the experiments, the video watermarking system presented in [10] is used as an experimental system in the embedding and extracting ends. Our proposed re-synchronization method is used with the watermarking system improving the robustness against temporal distortion, especially on transposition, dropping, and insertion.

In the simulation, the video sources Akiyo, Foreman, and Stefan of class A, B, and C are tested. The frame detection results without voting are shown in Table 1. The accuracy of frame detection is from 72.41% to 98.15%. After the voting process, the GOP detection accuracy, shown as in Table 2, is improved to the range from 96.30% to 100%. The overall performance in Figure 5 shows that, after resynchronization following temporal attacks, the N.C. values of watermark detection are improved from 0 to 0.4 or even more. In other words, the un-extractable watermark in the original watermarking system becomes recognizable with the proposed method incorporated.

CONCLUSIONS

In our proposed method, we have developed a novel video watermarking synchronization technique based on profile statistics. The experimental results show that our technique can efficiently improve the watermark extraction results because the characteristics of profile statistics are suitable for re-synchronization. The technique can also improve the robustness for video watermarking against temporal attacks, such as transposition, dropping, and insertion.

ACKNOWLEDGEMENT

Special thanks to C.S. Lu, J.R. Chen, H.Y. Liao and K.C. Fan, at the Institute of Information Science, Sinica, and National Central University, Taiwan, to provide the watermarking programs in [10].

REFERENCES

[1] I.J. Cox and Matt L. Miller, 2001,
Electronic Watermarking: The First 50 Years,
Proc. IEEE Fourth Workshop on Multimedia Signal Processing, pp. 225-230, 2001.

[2] E. Lin, C. Podilchuk, T. Talker and E. Delp, January 2001,
Streaming Video and Rate Scalable Compression: What are the Challenges for Watermarking?"
Proc. SPIE Security and Watermarking of Multimedia Contents III, Vol. 4314, pp. 116-127, San Jose, January 2001.

[3] E. Lin and E. Delp, January 2002,
Temporal Synchronization in Video Watermarking,
Proc. SPIE Security and Watermarking of Multimedia Contents IV, Vol. 4675, pp. 478-490, San Jose, January 2002.

[4] M.H. Yang, D.J. Krieman and N. Ahuja, January 2002,
Detection faces in images: a survey,
IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 24, No. 1 , pp. 34 -58, January 2002.

[5] G. Yang and T.S. Huang, 1994,
Human face detection in complex background,
Pattern Recognition, Vol. 27, No. 1, pp. 53-63, 1994.

[6] C. Kotropoulos and I. Pitas, 1997,
Rule-based face detection in frontal views,
Proc. Int'l Conf. Acoustics, Speech and Signal Processing, Vol. 4, pp. 2537-2540, 1997.

[7] Henry Stark and John W. Woods, 2002,
Probability and random processes with applications to signal processing,
3rd ed., Prentice Hall, New Jersey, 2002.

[8] I.J. Cox, M.L. Miller and J.A. Bloom, 2002,
Digital watermarking,
Morgan Kaufmann Publishers, l/e, USA, 2002.

[9] C.S Lu, H.Y. Liao and M. Kutter, 2002,
Denoising and copy attacks resilient watermarking by exploiting knowledge at detector,
IEEE Trans. on Image Processing, Vol. 11, No. 3, pp. 280-292, 2002.

[10] C.S. Lu, J.R. Chen, H.Y. Liao and K.C. Fan, August 11-15, 2002,
Realtime MPEG2 video watermarking in the VLC domain,
Proc. 16th IAPR Int. Conf. on Pattern Recognition, Quebec, Canada, Vol. II, August 11-15, 2002.