Incremental-based Digital Signature with Neighbouring Block Similarity Measure for Video Authentication

Wilaiporn Kultangwattana and Nopporn Chotikakamthorn

Faculty of Information Technology &
Research Center for Communications and Information Technology,
King Mongkut's Institute of Technology Ladkrabang.
Chalongkrung Road, Bangkok 10250, Thailand
e-mail: wilaiporn17@yahoo.com, nopporn@it.kmitl.ac.th

Abstract: This paper describes a digital signature-based method for original and updated video authentication. The method uses multiple digital signatures in dealing with video data undergoing multiple change/updating. In addition, a feature based on neighbouring block similarity measure is applied to deal with certain image/video modification. The proposed method can cope with wide range of image/video tampering. It is suitable for practical use of video data, where updating may be performed by more than one legal parties. Experimental results are included with concluding remarks.

1. Introduction

Recently, the problem of video authentication has been addressed. There are two approaches, which have been suggested for achieving the task of authenticating digital video. The first one is based on the use of a digital signature technique, and the second one is based on digital watermarking techniques. A digital signature method as proposed by Ching-Yung Lin [2] can detect and localize alterations of the original video. Difference between DCT coefficients of the first image-block group and those of the second group is computed to construct a feature. Another method was proposed by Jana Dittmann [3]. The method uses edge-based feature code for digital signature so it can not detect color alterations. Another work includes that of Marc Schneider[4], which proposed a hashing method for video data. Digital watermarking method proposed by Bijan G. Mobasseri [1] has a watermark inserted into each frame of video, to detect unauthorized cut-and-splice or cut-insert-splice. Min wu [5] proposed an insertion of a watermark in a frequency domain of an I-frame.

This paper deals with the digital signature-based approach, due to the lack of enough embedding capacity when the watermarking approach is applied. The proposed method uses a feature based on similarity measure between two adjacent blocks of image. In addition, by applying the concept of incremental-based digital signature, video data undergoing multiple updating from more than on legitimated owners, can be authenticated in a hierachical manner. The paper is organized as follows. First, a general framework of authenticating digital video is described. In Section 2 Section 3 describes the proposed incremental-based digital signature for video authentication. In Section 4, some experimental results of image authentication are given. Discussion and concluding remarks are provided in Section 5.

2. Digital Signature for Video Authentication

Just as with human signatures, digital signing should be done in such a way that a signature is verifiable, non-forgible, and non-repudiable. In a processing of generating a digital signature, a private key is used to encrypt a message (or the feature or hashed data corresponding to the original image). This encrypted message is called a "digital signature". The authentication process of this message needs the public key associated with the private key used to generate a signature, to decrypt the digital signature. The message to be authenticated is then compared with the decrypted digital signature. If they are identical, then the received message is authentic.

Given the overheads of encryption and decryption, signing and verifying data can be overkill. Using a message digest, computational complexity can be greatly reduced due to data size reduction. For image and video data authentication, however, use of a hashing function is not suitable. Because, with a slight change in image data due, for example, to compression, the image/video may be regarded as invalid. A general process of digital signature method for image/video authentication, such as that described in [2] is shown in Figure 1.

When a user needs to authenticate the image or video received, a signature needs to be decrpyted and compares with the corresponding feature value extracted from the test image. If the two data sets match (or closely match), the image is said to be "authentic". The authentication process is shown in Figure 2.

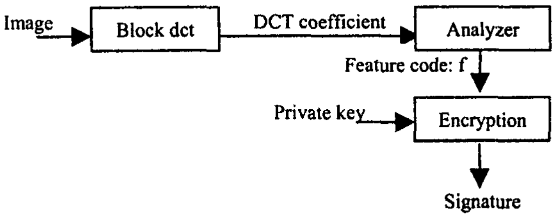

Figure 1. Digital Signature Generation

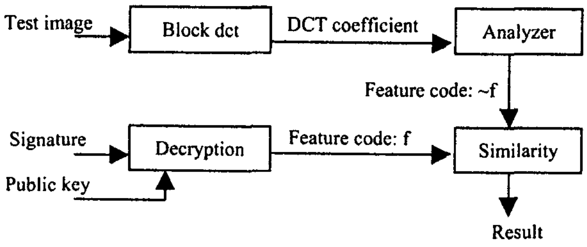

Figure 2. Authentication Process

3. Incremental-based Digital Signature for Video Authentication

The method described here is based on the digital signature approach, as detailed in [2]. However, unlike the method in [2], similarity between each block to the adjacent one on the right, and to another block below, is used (see. Figure 3). In addition, we introduce the concept of incremental-based digital signature. The method was inspired by [6, 7], but its main principle bears no relationship with [6, 7]. In this scheme, for the video I-frame, digital signature is generated from an extracted feature with complementary information (which includes information regarding to day, month, year, hour, minute, second, frame number, and frame rate). For the video P-frame and B-frame, digital signature is generated from the feature obtained as the difference between the feature corresponding to the nearest I-frame and P-frame.

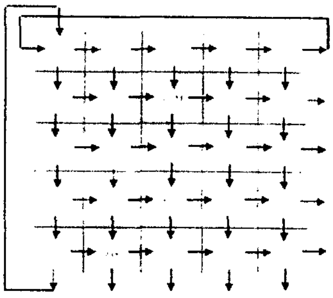

Figure 3. Use of neighbouring blocks for feature extraction

In feature extraction process, an image (or video frame) is divided into blocks of size 8×8. For each block pair, says block (*i,k*) and (*i,k*+1), a local histogram equalization is performed. From the pair of equalized blocks, DC coefficient (mean) of the block (*i,k*), denoted by $f_{i,k}$, is computed. Normalized difference between the coefficients corresponding to the two adjacent blocks is then calculated as given by

$$\tilde{f}_{i,k}^{r} = \frac{f_{i,k} - f_{i,k+1}}{f_{i,k} + f_{i,k+1}} \qquad (1)$$

for the feature corresponding to the block on the right, and

$$\tilde{f}_{i,k}^{b} = \frac{f_{i,k} - f_{i+1,k}}{f_{i,k} + f_{i+1,k}} \qquad (1)$$

for the feature corresponding to the block below the current one. For the blocks at the image edge, the adjacent blocks are chosen as described in Figure 3. Next, $\tilde{f}_{i,k}^{r}$ and $\tilde{f}_{i,k}^{b}$ are quantized, so that they can be represented by a finite number of binary bits. Here, two bits are used for representing each feature. Therefore, there are at most 4 different quantization levels. In this case, the quantized data is given by

$$\breve{f}_{i,k}^{*} = \begin{cases} 1 & \tilde{f}_{i,k}^{*} \geq \alpha+\beta \\ 2 & \alpha-\beta < \tilde{f}_{i,k}^{*} < \alpha+\beta \\ 0 & -\alpha+\beta \leq \tilde{f}_{i,k}^{*} \leq \alpha-\beta \\ 2 & -\alpha-\beta < \tilde{f}_{i,k}^{*} < -\alpha+\beta \\ -1 & \tilde{f}_{i,k}^{*} \leq -\alpha-\beta \end{cases} \qquad (2)$$

where $\tilde{f}_{i,k}^{*}$ ($\breve{f}_{i,k}^{*}$) represents either $\tilde{f}_{i,k}^{r}$ ($\breve{f}_{i,k}^{r}$) or $\tilde{f}_{i,k}^{b}$ ($\breve{f}_{i,k}^{b}$). From Eq. (2), α is a threshold parameter for quantization. It should be chosen such that the quantized data represents well the difference between image blocks. One criterion for choosing α is to select the parameter such that the probabilities of $\breve{f}_{i,k}^{*}$ having values 0, 1, and –1, are approximately the same. In addition, from Eq. (2), β is a parameter used to create an unknown band. This allows for the detection process to be reliably performed with the image or video suffered from noise and distortion.

For a video I-frame, the resulting quantized data $\breve{f}_{i,k}^{r}$ and $\breve{f}_{i,k}^{b}$ are combined with those from other blocks to construct a digital signature. For a video P-frame and B-frame, $\breve{f}_{i,k}^{r}$ (and $\breve{f}_{i,k}^{b}$) is compared with the corresponding $\breve{f}_{i,k}^{r}$ (and $\breve{f}_{i,k}^{b}$) of the I-frame. Difference between the two is then used for signature creation. Use of feature obtained as a difference between I-frame and P- or B-frame can reduce the number of bits required to code a feature. The same process is applied for the case where there is a need to create a signature of an image/video which is a result of rightful modification to an original content. Here, a variable-length coding scheme is used to code the difference of I-frame block and that of P-frame block (as well as the difference between blocks from the original and the rightful modified image/video). For example, in our experiment, if the two blocks are considered identical (the difference is below a certain threshold), a single bit of value '0' is used to code the difference. On the other hand, if the difference is above the threshold, two data bits are used. In this latter case, the first bit set as '1', while the second bit value reflects the direction of the difference.

Figure 4 describes how signatures are created for the case of multiple image/video updating.

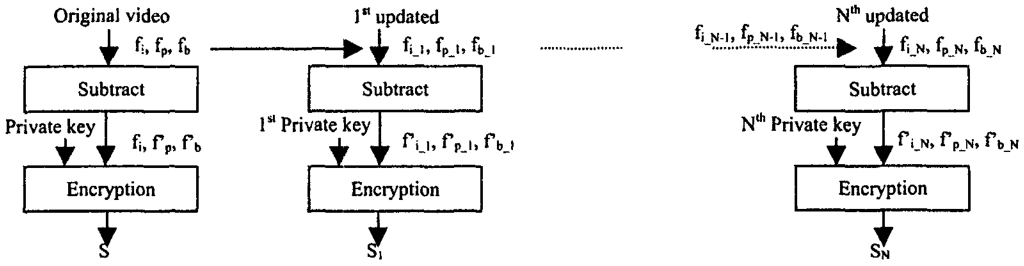

Figure 4. Updated digital signature

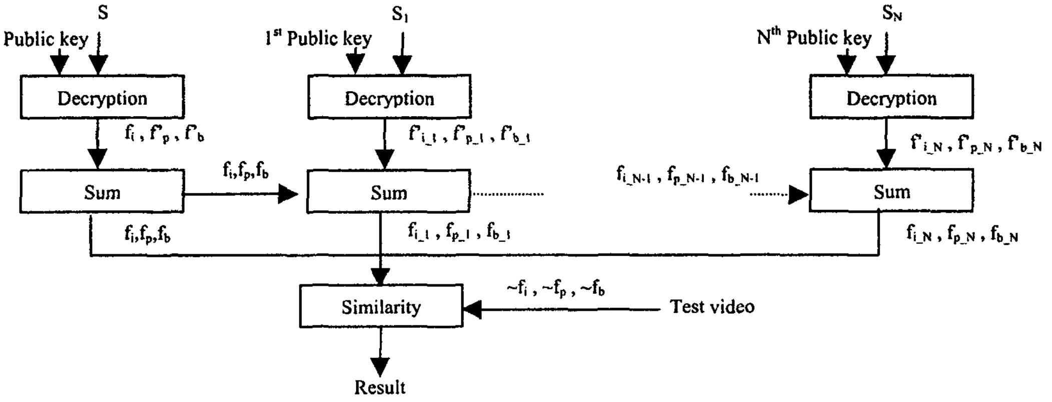

Figure 5. Authentication process

To authenticate an image/video content, a signature corresponding to each image frame is first decoded. For the I-frame case, the decoded signature is used to compare with the feature extracted from the image/video frame under consideration. For the P-frame and B-frame, the feature difference between that of the frame under consideration and the corresponding one of the I-frame must be first computed. The resulting feature difference is then used to compare with the decoded signature corresponding to the same block pair. In any case, any block pair of which the extracted feature is different from that of the decoded signature is regarded as dissimilar. The exception is when the decoded signature corresponding to any block pair has its value, which represents an unknown case (=2, see Eq. 2). In this special case, the two feature sets corresponding to that block pair are always considered identical. The authentication process for multiple updating is shown in Figure 5.

Setting the detection threshold is a classical decision estimation problem. If the threshold is set to be too high, it creates missed detection more often, while setting the threshold too low results in more false alarm.

4. Experiment Results

Experiment have been carried out to test the performance of the proposed method. The test image (frame) size is 240×320. We divided it into blocks of size 8×8. In the experiment, a local histogram equalization is performed with the window size of 16×16. In addition, $\alpha = 0.12$ and $\beta = 0.05$ were chosen. Figure 6 shows a video sequence used.

By using the proposed incremental-based method, it has been found that the obtained feature required less bit to encode to code. As a result, on average, the signature size is reduced by 33.74% per frame.

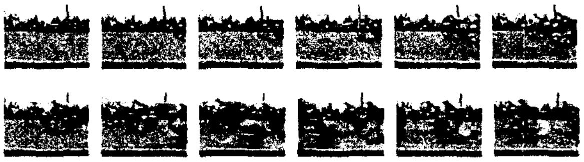

Figure 6. Sequence of a video clip.

Next, various image manipulations were performed. The results are given below.

Cropping: Parts of the image on the right and bottom were cropped is shown in Figure 7(b), as compared with the orignal one shown in Figure 7(a). Borders of blocks of which their relation to neighbouring ones differ from those of the original were marked by a white line in the figure.

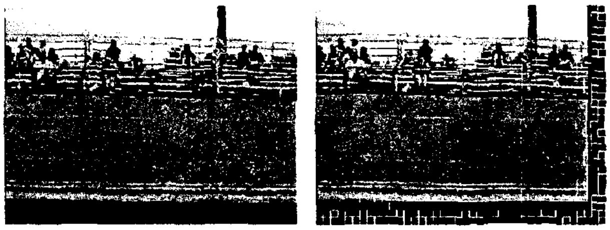

Figure 7(a). Original image (Frame 1), (b) Cropped image

Brightness adjustment : Figure 8(a) shows original image with 20% increase in brightness level and Figure 8 (b) shows original image with 20% decrease in brightness level. The percentage of feature mismatches found for the case of Figure 8(a) is 1.67%. and 3.17% for Figure 8(b).

The errors found are due to pixel intensity saturation as a result of brightness adjustment.

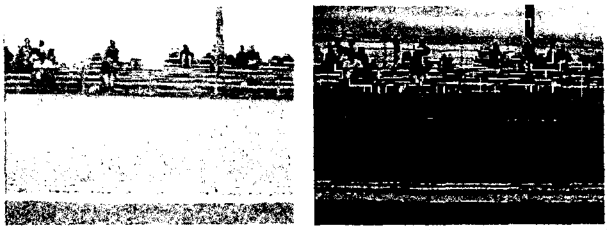

Figure 8. Original image: (a) with 20% increase in brightness level, (b) with 20% decrease in brightness level

Contrast adjustment: Figure 9(a) shows original image with a global histogram equalization. The percentage of feature mismatches is 0.08 %.

JPEG compression: Figure 9(b) shows 60% JPEG compression image. The number of feature mismatches found is 0.96%.

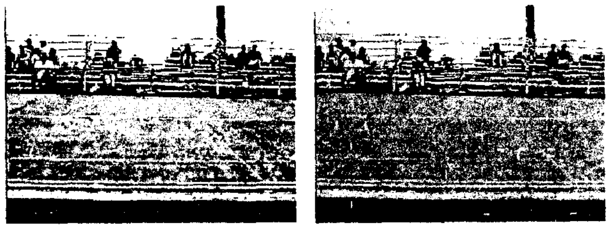

Figure 9. Original image: (a) with global histogram equalization, (b) with 60% JPEG compression

Noise: Figure 10(a) shows image adds "salt & pepper" noise, where the noise density is 0.04. And Figure 10(b) shows image adds "gaussian" SNR = 67.41 dB. In Figure 10(b) the percentage of feature mismatches is 1.83%.

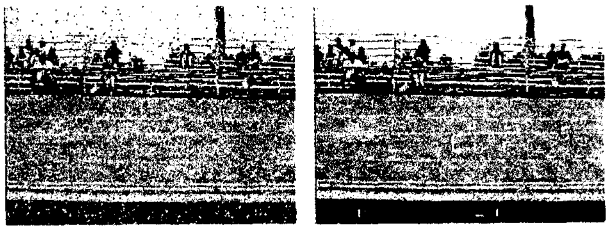

Figure 10. Original image: (a) adds salt & pepper noise, (b) adds gaussian noise

Updated and Manipulation: Figure 11(a) shows updated image. The digital signature is generated from the different feature, obtained as difference between the feature of the original image and updated image. Length of feature is 4800 bits and length of different feature is 2424 bits so, it decrease 49.50%. Experiment is made by manipulating the updated image, is shown in Figure 11(b). The authentication result, when compared with the updated image is shown in Figure 12(a). The authentication result, when compared with the original image is shown in Figure 12(b).

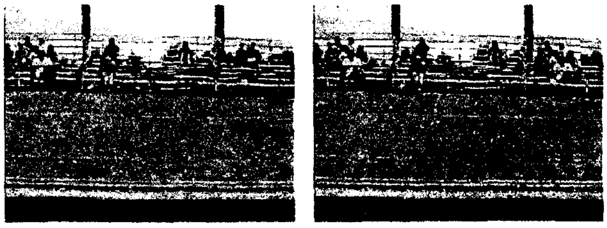

Figure 11. (a) Updated image, (b) Manipulated image

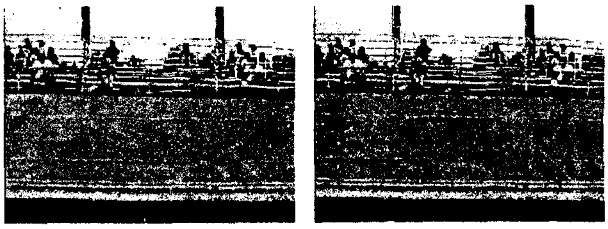

Figure 12. Authentication result: (a) compare with updated image, (b) compare with original image

5. Conclusion

In this paper, the video authentication method based on relative similarity of neighboring block features has been proposed. The proposed scheme has the advantages of being able to identify a modified portion of image/video, as compared with either the original image/video or an updated image/video. The use of incremental-based signature approach reduces the size of a signature. In addition, it allows for P-frames and B-frames of the original video, and video frames from the updated or modified video, to be treated in a unified manner.

References

[1] G. Bijan Mobassseri, J. Michale Sieffert, J. Richard Simard. "Content Authentication and Tamper detection in Digital Video," *Proc. Int. Conf. on Image Processing,* Vol.1, pp.458-461, 2000.

[2] Ching-Yung Lin. "Watermarking and Digital Signature Techniques for Multimedia Authentication and Copyright Protection," *Doctor of Philosophy in the Graduate School of Arts and Science, Columbia University,* 2000.

[3] Jana Dittmann, Arnd Steinmetz, Ralf Steinmetz. "Content-based Digital Signature for Motion Pictures Authentication and Content-Fragile Watermark," *IEEE Int. Conf. on Multimedia Computing and Systems,* Vol.2, pp.209-213, 1999.

[4] Marc Schneider and Shih-Fu Chang. "A Robust Content Base Digital Signature For Image Authentication," *Proc. Int. Conf. on Image Processing,* Vol.3, pp.227-230, 1996.

[5] Min Wu and Bede Liu. "Watermarking For Image Authentication," *Proc. Int. Conf. on Image Processing,* Vol.2, pp. 437-441, 1998.

[6] M. Bellare, O.Goldreich, S.Goldwasser. "Incremental Cryptography:The Case of Hashing and Signing," *Crypto'94, Lecture Note in Computer Science,* Vol. 839, pp.216-233, 1994.

[7] M. Fischlin. "Lower bounds for the signature size of incremental schemes." *In 38th Annual Symposium on Foundations of Computer Science,* pp. 438-447, 1997.