Digital Watermarking for the Authentication of AVS data

M. Barni[a], F. Bartolini[b], J. Fridrich[c], A. Piva[b]

[a] *Dip. di Ingegneria dell'Informazione, Università di Siena,*
e-mail barni@dii.unisi.it

[b] *Dip. di Ingegneria Elettronica, Università di Firenze,*

[c] *State University of New York in Binghamton and*
Mission Research Corporation

1 Introduction

A problem raised by the availability of increasingly powerful signal processing tools is the lost of *credibility* of digital images and video sequences, due to the ease of modifying their content without leaving any visible traces of the modification. Such a problem assumes great importance in Advanced Video-based Surveillance (AVS) applications where image data virtually has no value as proof in front of a court of law, since doubts would always exist about the possibility that it had been tampered to intentionally incriminate or exculpate the defendant.

Digital watermarking has been proposed as mean to guarantee the reliability of digital data such as images or audio files [1, 2]. Of course, according to the particular application scenario watermarking is used in, different requirements have to be fulfilled in terms of perceptual invisibility, robustness, decoding mechanism and so on [3]. It is the purpose of this work to discuss the possibility of using watermarking to ensure the integrity of the image data produced by AVS systems, and to outline some possible solutions based on the state of the art in the field.

2 Digital watermarking for image data authentication

To analyze the most important issues in AVS data authentication, we will adopt the simplified AVS model reported in figure 1. One or more digital video cameras frame the scene of interest and send the corresponding data to a central unit which is in charge of analyzing it and generating an alarm or pre-alarm signal whenever a suspect situation is detected. When necessary, stored images can be used as a proof in front of a court of law. For them to be legally credible, however, some means must be envisaged to avoid the tampering of their content, or at least, to detect it when it occurs. In order to reduce the possibility of data tampering to a minimum, authentication tools must act as early as possible in the acquisition-transmission-storage chain, that

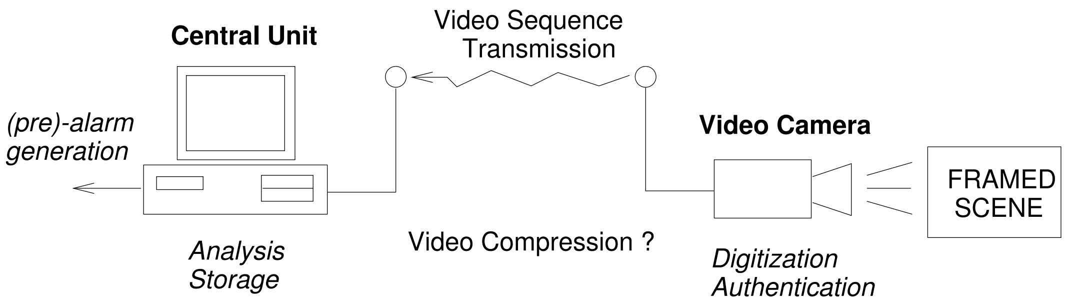

Figure 1: Simplified sketch of a generic AVS.

is immediately after the acquisition of data in digital form. In other words, it is up to the camera itself to prepare data so that, at a later time, it is possible to understand whether the video sequence has been tampered or not.

Though cryptography may provide adequate protection in many practical cases, the development of alternative strategies to ensure data integrity are desirable to deal with some potential weaknesses of cryptography.

Cryptography works by adding some data to the video sequence, be the additional data in the form of a header or a file. This may not be desirable in some applications, where the transmission of separate data is not allowed. Another drawback of cryptography is that malicious (e.g. data tampering) and innocuous (data compression) modifications are treated in the same way.

Watermarking-based authentication represents a feasible solution to the above problems. The information is inserted in such a way that any non-authorized modifications result in a non-recoverable lost or alteration of the watermark. The impossibility to recover the watermark is here taken as an evidence that data has been tampered. Two main differences with respect to cryptography can be outlined: no header or separate file is used to ensure data integrity, distinction is possible between different types of modification. Another major benefit achievable through watermarking is that it allows to localize the modification occurred to the image or the sequence, or even to recover the original image content.

3 Analysis of requirements

From a general perspective, the main requirements watermarking algorithms must obey can be summarized as follows (we only refer to data authentication applications) [2]:

1. the watermark must be invisible to a human observer;
2. the watermark is altered (removed) by non-authorized (assumed as malicious) processing;
3. the watermark survives innocuous manipulations, e.g. image compression or zooming;
4. it is difficult for non-authorized persons to insert a fake watermark;
5. the watermark is easily extracted by authorized persons;
6. the extracted watermark should give an indication of where the data has been manipulated.

The above requirements assume a particular meaning when AVS is considered. This is especially true with regard to invisibility, since, often, the end user of an AVS system is not a human observer, but an automated inspection system. As to points 2 and 3, ideally the watermark should survive all innocuous manipulations and be altered by any malicious modifications. In AVS applications the necessity to distinguish between innocuous and malicious processing only comes into play if data compression is performed by the central unit. Under such an assumption, the watermark is inserted before compression and therefore must survive it. All the other kinds of processing can be regarded as not allowed. As in most applications, the requirement that it should be very difficult to insert a fake watermark, must be strictly satisfied if the watermark (item 3) is to have any value. The easiest and safest way to obtain it, is through the use of key-based encryption; for example, if the watermark is signed with the private key of the video camera it is impossible to reproduce it, while anyone is allowed to read it. In any case, the use of a secret key hardwired within the video camera, without which valid watermarking is not possible, seems the only viable solution. Localization of the modification (requirement 6) is not mandatory in many AVS applications, though an indication of where alterations have taken place may sometimes be useful.

In some cases the bandwidth of the transmission channel does not allow the transmission of the non-compressed video sequence, but at the same time the video camera may not be powerful enough to compress itself the sequence. To overcome the problems due to the limited processing capabilities at the camera site, a PC station may be added right after the camera output; however, as PC station are usually not tamper-proof, the signature of data must still be performed inside the camera. As a consequence watermark embedding should not be much computational demanding. A sketch of a simple AVS system in which watermarking is used to ensure data integrity is reported in figure 2.

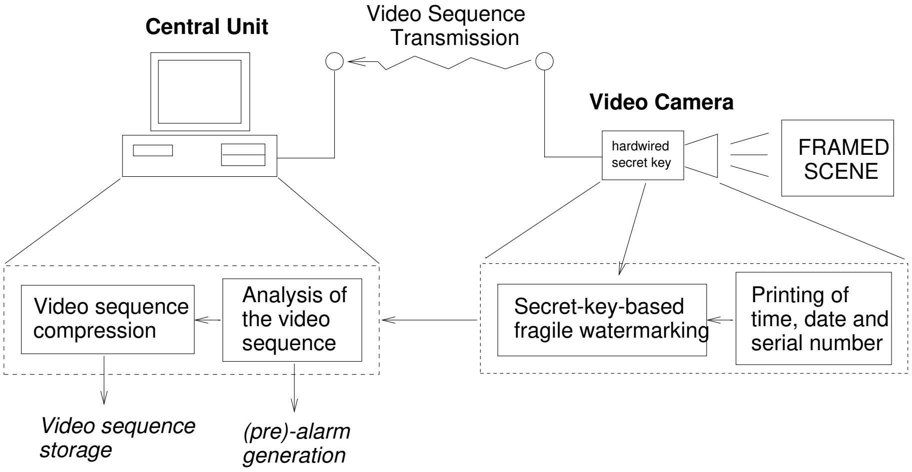

Figure 2: Outline of a watermark-based authentication system for AVS applications. The channel is assumed to permit the transmission of the video sequence in non-compressed form. Note that in some cases the central unit may be required to check for the integrity of data before analyzing it for possible alarm generation.

4 A possible solution

In the final version of the paper, after a brief description of the state of the art in the field of fragile image and video watermarking, an algorithm will be presented which, based on image self-embedding, permits the partial recovery of tampered information.

The idea of embedding a low quality version of the data into the data itself enables not only detection of areas that have been tampered or damaged, but also recovering some of the missing information. The robustness of self-embedded data may vary according to the application, even if there is an obvious trade-off between the robustness of the self-embedded data and the quality of the marked image/video. In the paper a system is presented for the authentication of still images (be they a single image frame or a still image sent for AVS purposes). According to the proposed system, each image is first divided into 8×8 blocks. Setting the LSB of pixels to zero, a DCT is calculated of each block. The DCT matrix is then quantized with the quantization matrix corresponding to a 50% quality JPEG. The resulting quantized matrix is then encoded using 64 bits and the code inserted into the LSBs of a distant 8×8 block. On average, the watermarking process modifies 50% of pixels by one gray level. The quality of the reconstructed image is somewhat worse than 50% quality JPEG. If two LSBs are used for inserting the code, 128 bits can be used instead of 64. For most blocks, this enables encoding almost all quantized DCT coefficients. Thus, the quality of the reconstructed image is roughly equivalent to a 50% JPEG compression. A somewhat more robust algorithm, which is based on differential encoding, is also proposed in the paper. The difference between neighboring pixels are adjusted so that they mimic a decreased color depth approximation of the same image shifted by a third of the image dimension in a random direction. This provides some robustness to moderate noise addition while maintaining the quality of the embedded image quite reasonable (16 color approximation of the original). The embedded image gradually degrades with noise adding. The visual quality of the recovered image is still acceptable after adding noise of amplitude of 2 gray levels, and after moderate JPEG compression.

References

[1] G. L. Friedman, "The trustworthy digital camera: restoring credibility to the photographic image," *IEEE Trans. on Consumer Electronics*, vol. 39, pp. 905–910, November 1993.

[2] F. Mintzer, G. W. Braudaway, and M. M. Yeung, "Effective and ineffective digital watermarks," in *Proc. ICIP'97, IEEE Int. Conf. on Image Processing*, Santa Barbara, CA, October 1997, vol. III, pp. 223–226.

[3] A. Piva, M. Barni, F. Bartolini, and V. Cappellini, "Application-driven requirements for digital watermarking technology," in *Proc. EMMSEC'98, European Multimedia Microprocessor System and Electronic Commerce Conference and Exhibition*, Bordeaux, France, September, 28-30 1998, pp. 513–520.