VIDEO WATERMARKING SCHEME RESISTANT TO GEOMETRIC ATTACKS

Yao Zhao[1, 2], *Reginald L. Lagendijk*[1]

[1]Information and Communication Theory Group, Faculty of Information Technology and Systems, Delft University of Technology, 2628 CD Delft, The Netherlands

[2]Institute of Information Science, Northern Jiaotong University, Beijing 100044, P.R.China

ABSTRACT

The robustness against geometrical attacks remains one of the most challenging issues in watermarking of images and video. This paper presents several improvements of the video watermarking approach presented in [1], namely (i) using a temporally low-pass watermark and (ii) synchronization to resist attacks along the temporal axis. In order to improve the watermark detection performance, we propose to use an amplitude-limiting filter and a whitening filter during the watermark extraction process. Experimental results show that the proposed techniques achieve good performance.

1. INTRODUCTION

The idea of using a robust digital watermark to detect and to trace copies of audio-visual works has led to significant interest among artists and publishers [2]. A digital watermark is an invisible mark embedded in a digital signal, image or video frame(s), which can be used for a number of purposes including copy tracing and copyright protection. In recent years a variety of approaches to watermarking of multimedia have been proposed [3-6]. Most current watermarking methods can resist attacks such as compression, filtering, enhancement, and other signal processing operations. However, even very small *geometric distortions* can already prevent the detection of a watermark. This problem is most severe in the case of blind watermarking [6].

In [1], Haitsma *et al.* proposed a novel video watermarking scheme that hides watermark signals in the mean luminance values of successive video frames. Since the mean luminance value is very insensitive to spatial geometrical operations, this watermarking scheme is robust against manipulations such as rotation, scaling, shifting. However, if individual frames are (randomly) removed or if the mean luminance values are temporally low-pass filtered, watermark detection becomes unreliable. In order to improve the robustness of the above-mentioned scheme, we propose a number of modifications of the watermark embedder and detector. In the first place, the embedder is modified such that (i) a low-pass watermark is added in order to resist temporal filtering attacks, and (ii) synchronization information is added to resist frame removal attacks. Using temporally low-pass filtered watermarks – instead of temporally white noise watermarks – also reduces temporal flicker artifacts since the human visual system is less sensitive to (very) low frequency temporal changes.

Secondly, in the detection process, we use (i) an amplitude-limiting filter and (ii) a whitening filter to improve the probability of watermark detection. In Section 2 we detail the watermark embedding process. Then, Section 3 describes the extraction process, and Section 4 presents experiments that illustrate the performance improvements achieved by the proposed techniques.

2. EMBEDDING PROCESS

The watermark embedding system is shown in Figure 1. The scheme hides meaningful information bits in the luminance mean values of video frames, denoted by $F(t)$. In order to allow for random access of the watermark, and to improve the resistance to frame removal attacks, synchronization bits are embedded alternating with the meaningful watermark bits.

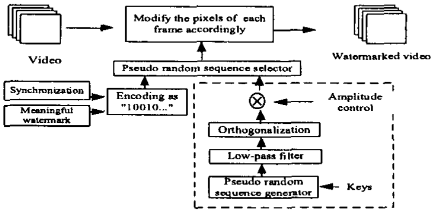


Figure 1: *Embedding of temporal watermark.*

A pseudo random sequence (*PRS*) generator is used to generate two sets of *PRSs*, namely one set for watermarking and one for synchronization according to two different keys. Also the two sets are of different lengths in order to distinguish them during detection. For both sets we have two *PRS*s, one represents bit "1", another for bit "0". The two *PRS*s are made orthogonal to each other.

0-7803-7622-6/02/$17.00 ©2002 IEEE

In order to improve the robustness against temporal filtering attacks, we use a low-pass filtered Gaussian noise watermark instead of white Gaussian noise. Furthermore, we anticipate the relatively low sensitivity of the human visual system for temporal frequencies lower than 1 Hz. The transfer function of the designed low-pass filter with cut-off frequency of approximately 1 Hz is:

$$H(Z)=\frac{1}{1-0.4505Z^{-1}-0.988Z^{-2}+0.0429Z^{-3}+0.5112Z^{-4}} \tag{1}$$

The frequency response of $H(Z)$ is shown in Figure 2.

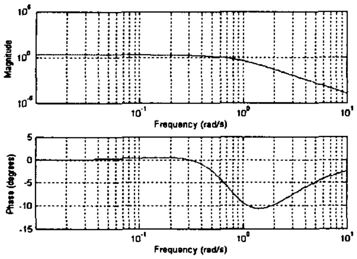


Figure 2: *Frequency response of designed low-pass filter.*

The construction, embedding and extraction processes of *PRS*s for synchronization and watermark bits are essentially the same. Therefore, without loss of generality, we use *P*(*t*) to denote a low-pass filtered *PRS* that represents either synchronization or watermark information. With $P_i(t)$ ($i=0,1$) we denote the *PRS* associated with bit "0" and "1". For the constructed $P_i(t)$ we have the following properties:

$$E[P_i(t)]=0\,,\; i=0,1 \qquad E[P_1(t).P_0(t)]=0 \tag{2}$$

In order to embed the watermark *PRS* values, we modify the mean luminance value of individual frames as follows:

$$F^P(x,y,t)=F(x,y,t)+P(t) \tag{3}$$

Here $F(x,y,t)$ denotes the luminance of pixel (*x*,*y*) in frame *t*. We limit ourselves to embedding watermark bits in the luminance component of a video sequence.

3. EXTRACTION PROCESS

Figure 3 illustrates the watermark extraction process. After calculating the mean luminance values of individual frames of the received video, denoted by $F^P(t)$, detection takes place by correlation with the *PRS*s used at the embedder.

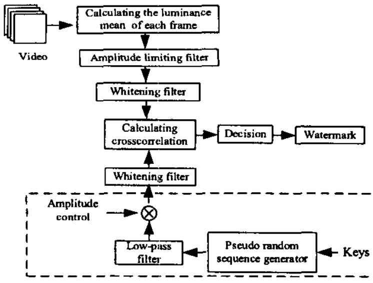

Video
Calculating the luminance mean of each frame
Amplitude limiting filter
Whitening filter
Calculating crosscorrelation
Decision
Watermark
Whitening filter
Amplitude control
Low-pass filter
Pseudo random sequence generator
Keys

Figure 3: *Watermark extraction.*

Watermark detection by correlation of the embedded *PRS*s and video frame mean luminance is less reliable than results obtained for spatial image or audio watermarks: an actual video usually consists of different shots. Different shots usually have different luminance, as illustrated in Figure 4(a). Consequently the luminance over time does not form a wide-sense stationary sequence, which decreases the detection performance. In order to combat this problem, we use an *amplitude-limiting filter* to suppress the luminance jump between two shots while maintaining the watermark signal. The proposed filter is defined as:

$$F_L^P(t)=H(F^P(t))$$
$$=\begin{cases} F_L^P(t-1)+(F^P(t)-F^P(t-1)), \text{if} \left|F^P(t)-F^P(t-1)\right| \le \alpha\cdot Am \\ F_L^P(t-1), \quad \text{otherwise} \end{cases} \tag{4}$$

Here *Am* is the amplitude of the watermark, and α is a tuning parameter (typically $\alpha = 2$). The filter has the property that it passes small signals undisturbed while it decreases the signal discontinuities (see Figure 4(b)).

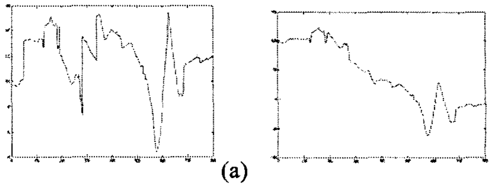


Figure 4: *(a) Mean luminance sequence; (b) Processed version using an amplitude-limiting filter.*

As shown in Figure 3, the cross-correlation function between the luminance sequence and watermark is used to detect the presence of watermark:

$$R_{F_L^P,P}(0)=E(F_L^P(t)P(t)]=E[(F_L(t)+P_L(t))P(t)]=E[(F_L(t)+P(t))P(t)]$$
$$=E[F_L(t)P(t)]+E[P(t)^2]=R_{F_L,P}(0)+E[P(t)^2]=E[P(t)^2]\cdot(\frac{R_{F_L,P}(0)}{E[P(t)^2]}+1) \tag{5}$$

Since *P(t)* is a given signal we find:

$$\frac{R_{F_L^P,P}(0)}{E[P(t)^2]}=\frac{R_{F_L,P}(0)}{E[P(t)^2]}+1 \tag{6}$$

We define the *normalized cross-correlation* as

$$R'_{F_L^P,P}(0) = \frac{R_{F_L^P,P}(0)}{E[P(t)^2]} \tag{7}$$

$$R'_{F_L,P}(0) = \frac{R_{F_L,P}(0)}{E[P(t)^2]} \tag{8}$$

Since the original mean luminance sequence and the watermark are independent, theoretically Equation (8) equals 0, and therefore it is easy to detect the watermark. However, in practice, we use a finite number of data points to estimate (8). Therefore realizations of (8) are generally non-zero, and have only zero mean. The variance of the estimator of (8) then determines the detection performance.

From detection theory it follows that correlation detectors are optimum in the case of a linear time-invariant, frequency non-disperse, additive white Gaussian noise channels [7]. However, in our system, the luminance values of subsequent frames are naturally highly correlated and the low-pass filtered watermark is also correlated. In order to improve the detector performance, i.e. to decrease the variance of the estimator of (8), we use a *whitening filter* prior to correlation. We use a parametric first-order FIR whitening filter:

$$G(z) = 1 - aZ^{-1} \tag{9}$$

We propose two approaches to the problem of choosing a suitable value for the parameter a. The first, simple but sub-optimal, approach is as follows. Since the weak watermark signal is usually carried by a strong host signal, we use the filter $G(z)$ to decrease the energy of the original temporal luminance signal as much as possible. After whitening, the luminance signal (after the amplitude-limiting filter) $F_L(t)$ becomes:

$$F_{LW}(t) = F_L(t) - aF_L(t-1) \tag{10}$$

of which the energy is:

$$\begin{aligned} E[F_{LW}(t)^2] &= E[(F_L(t) - aF_L(t-1))^2] \\ &= (1+a^2)R_{F_L}(0) - 2aR_{F_L}(1) \end{aligned} \tag{11}$$

If we now solve $\frac{dE[F_{LW}(n)^2]}{da} = 0$, we get:

$$a = \frac{R_{F_L}(1)}{R_{F_L}(0)} \tag{12}$$

Clearly this approach is sub-optimal because only properties of host video signal are considered. Our second approach aims at finding the optimal value of a taking into account properties of both video signal and watermark. We here summarize the results and conclusions: detailed information will be presented in [8]. The autocorrelation function of video frame mean value signals can be reasonably well modeled as an exponentially decaying function. For variance-normalized signals $F_L(t)$ and $P(t)$ the autocorrelation functions become $\rho_{F_L}(\Delta) = \alpha^{|\Delta|}$ and $\rho_P(\Delta) = \beta^{|\Delta|}$. Using these models we can now find optimal expressions for a. Since the resulting expressions are fairly complicated, we to present graphs for the relationship between the optimal value of a, and different values of the parameters α and β (see Figure 5).

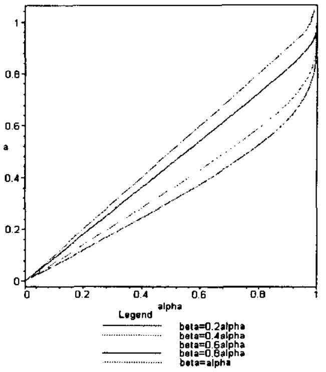


Figure 5: *Optimal a as a function of the parameters* α *and* β.

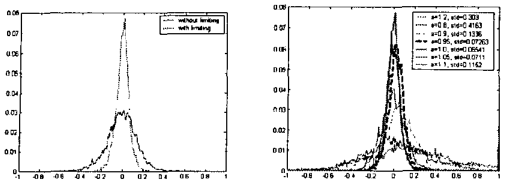
We draw the following conclusions:

- If $\alpha = \beta$, the optimal parameter for a is $a_{opt} = \alpha$. This result is also what one would expect intuitively.
- If $\alpha = 1$, the optimal parameter for a is $a_{opt} = 1$ irrespective of the value of β. That means that if the autocorrelation function of $F_L(t)$ is constant, the optimal whitening filter is $G(z) = 1 - Z^{-1}$. Many practical cases of interest approach can be well approximated by this special case.

4. EXPERIMENTAL RESULTS

The average luminance of the video frames of the test video we used is shown in Figure 4(a).

Experiment 1. The experiment tests the efficiency of the amplitude-limiting filter. The detection performances are evaluated with and without an amplitude-limiting filter, respectively. The practical realizations of Equation (8) are shown in Figure 6(a). Clearly, with the amplitude-limiting filter the variance of the estimator is smaller, implying a better detection performance

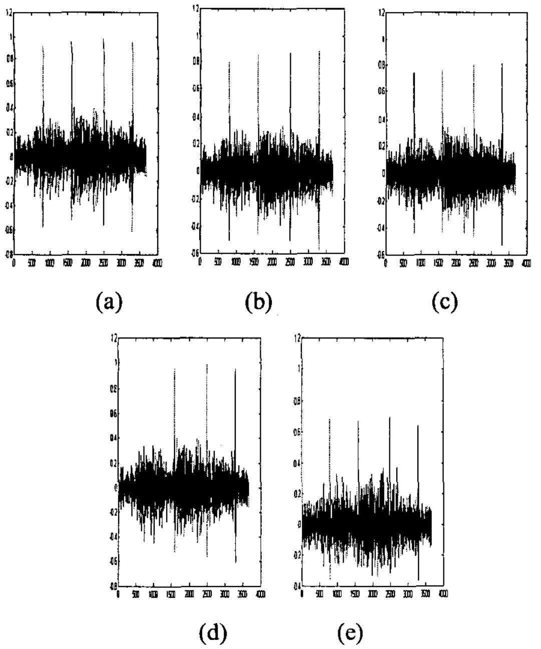


(a) (b)

Figure 6: *(a) Histogram of realizations of Equation (8) with and without an amplitude-limiting filter, where Am=10, N=800, a=1. (b) Histogram of realizations of Equation (8) using the whitening filter with varying value of a, where Am=10, N=800.*

Experiment 2. The experiment evaluates the efficiency of Equation (12) and Figure 5. The practical realizations of Equation (8) using the whitening filter $G(z) = 1 - aZ^{-1}$ with varying value of a are shown in Figure 6(b). The standard deviations of the estimates are also listed in the figure. The value $a=1$ achieves best performance; this is close to the calculated value, a=0.997 by Equation (12).

Experiment 3. Figure 7(a) shows the cross-correlation without attacks. The peaks in the Figure 7(a) indicate the location of watermark bits. Figure 7(b) shows the cross-correlation with a frame *spatial shift* attack. In the experiment, every 180×144 sized frame is shifted 10×10 pixels and the shifted area is replaced with zero-valued pixels. Figure 7(c) shows the cross-correlation with frame rotating attack. In the experiment, every frame is rotated by 45 degrees. Clearly the proposed scheme can resist these geometrical attacks.



(a) (b) (c)

(d) (e)

Figure 7: *(a) The cross-correlation without attacks; (b) The cross-correlation with frame spatial shift attacks; (c) The cross-correlation with frame rotating attacks;(d) The cross-correlation with frame removal attacks;(e) The cross-correlation with temporal filtering attacks;*

We also test the robustness against frame removal and temporal filtering attacks. The embedded frame no. 1201 is removed and Figure 7(d) shows the cross-correlation. Comparing with Figure 7(a), we find that in Figure 7(d), the second peak disappears because of the removal. However, the third one appears again and then we can detect the meaningful watermark. We also filter the embedded luminance sequence with a low-pass filter $H(Z) = 1 + 0.5Z^{-1}$, Figure 7(e) shows the cross-correlation. Even though the peaks are a little lower, we still can successfully detect the synchronization and watermark.

5. ACKNOWLEDGEMENTS

This work has been jointly funded by a Fellowship Grant of TU-Delft, the EU CertiMark project, Huo Yingdong Yong Teachers Foundation and NSFC of China (Grant No. 69802001 and 60172062).

6. REFERENCES

[1] J. Haitsma and T.Kalker, "A watermarking scheme for digital cinema," *Proceedings of International Conference for Image Processing*, pp. 487-489, 2001.

[2] K. Tanaka, Y. Nakamura, and K. Matsui, "Embedding secret information into dithered multilevel image," *Proc. of 1990 IEEE Military Communications Conf.* pp. 216-220, Sept. 1990.

[3] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *Technical Report 95-10*, NEC Research Institute, Princeton, NJ, USA, 1995.

[4] M-S. Hsieh, D-C. Tseng, and Y-H. Huang, "Hiding Digital Watermarks Using Multiresolution Wavelet Transform," *IEEE Transactions on Industrial Electronics*, Vol. 48, No. 5, October 2001.

[5] P. Bassia, I. Pitas, N. Nikolaidis, "Robust audio watermarking in the time domain", *IEEE Transactions on Multimedia*, Vol. 3, No. 2, 2001.

[6] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Attacks on copy-right marking systems," *Proc. Workshop Information Hiding*, Portland, Apil 1998.

[7] G. Depovere, T. Kalker, and J.-P. Linnartz, "Improved watermark detection reliability using filtering before correlation", *Proc. of Int. Conf. On Image Processing*, pp. 430-434, 1998.

[8] Yao Zhao, Reginald L. Lagendijk, "Optimal whitening filter and its application to video watermarking", submitted to *IEEE Trans. on Image Processsing*, May 2002.