Biuletyn WAT
Vol. LXI, Nr 1, 2012

Real time digital video blind watermarking in Cepstrum domain

PIOTR LENARCZYK, ZBIGNIEW PIOTROWSKI

Military University of Technology, Faculty of Electronics, Telecommunication Institute,
00-908 Warsaw, Poland, 2 Kaliskiego Str., sirlevoire@interia.pl, zbigniew.piotrowski@wat.edu.pl

Abstract. The problem of real time requirements especially refers to digital video watermarking algorithms. This paper contains a description of the algorithm consisting of real time coding process in the luminance component in the spatial domain and not real time decoding process in the two dimensional Cepstrum domain. The algorithm makes the single I-frame robust against the following attacks: MJPEG compression, cropping, affine transform (two-dimensional shearing, rotation, translation). Also the Receiver Operating Curves have been performed for the above mentioned attacks. Watermarked digital stream is robust against desynchronization attack.
Keywords: watermarking, video, real time, Cepstrum

1. Introduction

One of the most important aspects of practical applications for the method of digital stream or video watermarking is the requirement for real-time processing. By meeting the above-mentioned requirement, it is possible to use an algorithm in VoIP systems [1] i.e. in stream-based applications using webcams. Another significant requirement for the digital stream watermarking algorithm is its due robustness to typical attacks aimed to distort the watermark or to make it unreadable for decoder. In addition, the algorithm should be designed to make it impossible for the most typical video processing to affect the decoding process of hidden information. Many latest video watermarking methods do not take into account the problem of affine transform attacks e.g. [2, 3, 4] and they ignore the problem of this watermarked video stream processing. There are a few proposals for watermarking algorithms [5, 6] that become partially robust to rotation or cropping, yet the problem of translation as an affine transform attack in other publications is practically omitted [7, 8, 9]. It is especially important as a simple translation attack (that even an inexperienced programmer can design easily) can make it impossible to read the embedded (hidden) information.

The beginnings of cepstral analysis date back to the 1960s, e.g. [10]. Literature shows some examples of using cepstral analysis in image processing [11, 12, 13], (used mainly for sound [14, 15], including e.g. detection of the signal echo [16]. However, so far, the cepstrum transform has not been used to watermark video streams.

The algorithm proposed embeds additional information in the cepstrum domain and decodes it by using the interdependence of 2D autocepstral feature. By using this extremely simple algorithm, it is possible to design a real-time watermarking application working even on previous generation computers and with limited processing power systems (e.g. mobile systems). In addition to this, the potential of using the algorithm in applications is extended as the decoder does not need a host (original) film to extract the embedded information. The paper presents a mathematical analysis of the embedding method described and the test results of watermark robustness to various attacks.

2. Additional information embedding and extracting algorithm description

2.1. Encoding process

The initial digital video stream is ripped into separate frames in real time according to the Group of Frames scheme (GOF) for individual format using an input buffer (it is also possible to use previously saved video files). From among them only I-frames are processed based on the watermarking method. "The initial RGB representation of I-frames matrix is changed into the matrices of luminance and chrominance. Then, to the matrix Y_I of the image Y_{YCbCr} we add the same luminance matrix translated by p_x ; p_y (missing lines resulting from the translation are copied) with values reduced by the δ coefficient. This coefficient defines the energy of embedded watermark. To reduce the luminance changes in the watermarked image, we have performed a compensation of matrix added ($1 - \delta$). It results in creating the Y_I matrix of the host image while the information bits are contained in the translation values and the sign of the translated matrix.

$$Y_{Iw}(x, y) = Y_I(x, y) \mp Y_I(x + p_x, y + p_y)\delta \pm Y_I(x, y)[1 - \delta]. \tag{1}$$

During the tests it turned out that the maximum imperceptible values of horizontal and vertical translations could not exceed the value of 6 pixels for which the average PSNR value was" [13] 37, 12 dB.

Then, there occurs a change of the frame form — I luminance and chrominance matrix into the matrices of red green and blue colours and recoding watermarked frames into a watermarked chosen format of video stream by using embedding information redundancy, for the 11 I-frames, corresponding to 55 bits of embedding information, where 40 bits contain encoded data (keys are distributed beyond the watermarking system), whereas 15 constitutes its preamble. The normalized data payload coefficient [17] equals 125 b/s (at 25 fps). The parameters of group I-frames are reconfigurable. The encoding process occurs in real time.

2.2. Decoding process

The first and continuous step is to rip the watermarked stream into separate I-frames (it is also possible to use the decoder to process watermarked video files saved). "Then, a 2D discrete Fourier transform of I-frame Y_W luminance is performed, where the Fourier transform base matrices possess in their lines 1D orthogonal vectors. Another step is to transform $Y_{WDFT}(k, l)$ into the 2D cepstrum domain (the two dimensional autocepstrum function of the matrix $Y_{WDFT}(k, l)$:

$$Y_{cepstr} = (IDFT(\ln(|Y_{WDFT}(k,l)|))), \tag{2}$$

where IDFT is a two-dimensional inverse discrete Fourier transform. If in the matrix $Y_W(x, y)$ there is a translated luminance copy $Y_I(x+p_x, y+p_y)$, then the coefficient with cepstral coordinates (quefrencies) that equals to translation $Y_{(p_x,p_y)} \leftrightarrow Y_{cepstr(p_x,p_y)}$ reaches considerably higher values than the matrix mean calculated in the 2D Cepstrum domain (excluding the low-frequency components p_x, $p_y < 4$). If the value of this coefficient exceeds the threshold then, the picture I_W is interpreted as watermarked. On one hand there is a limitation of maximum translation to 6 pixels (the changes become perceptible for insufficiently large pictures), on the other hand there is also a limitation for minimum translation to 3 pixels. It results from the property of non-linear Cepstrum transform that translates the low-frequency Fourier spectrum components into the low 2D quefrencies that cannot be exceeded by cepstral coefficient value (responsible for the translated luminance copy). Additionally, the sign of added and translated matrix $Y_I(x+p_x, y+p_y)$ determines the phase of the coefficient $Y_{cepstr(p_x,p_y)}$. As a result, we receive a matrix in a 2D cepstral space with 4 rows and columns, preserving the coefficient phase and providing data payload of the method at a level of $L_{pinf} = 5$ bits. The hidden information is defined by the spatial position $Y_{cepstr(p_x,p_y)} > \tau$ and the coefficient sign" [13]. If the subsequent information sequence values being decoded based on the first I-frames correspond to the preamble sequence, next frames are decoded. The next 8 watermarked I-frames contain embedded information that is a subject for decoding them. It is the decoder limit τ that protects the system against incomplete extracted additional bits vector errors.

3. Experiments

The experiments have been carried out using the following hardware and software:

Table 1

Used hardware and software

Processor	Intel Mobile Core 2 Duo T5500 2x1.66GHz
Main board	i945PM
Memory	DDR2 3072Mb
Graphics card	NVIDIA GeForce Go 7300
Operating system Basic	XP Professional SP2 + VMware Workstation 6.5.1
Operating system on the VMware Workstation	Linux Mandriva core 2.6.27
Webcam	creative syntek ov511 USB 2.0 Camera

The experiments have been carried out using intercepted streams coming from watermarked webcam video streams.

The experiments involved a real-time watermarking of specific I-frame matrices, before their saving into a file. Then, they were subjected to offline attacks in the form of specific types of digital video stream processing, such as: MJPEG lossy compression, cropping, adding noise to the watermarked I-frame, affine transform attacks and desynchronization. They were decoded one after another, in accordance with the description provided in subsection 2.2. The tests used a one hundred I-frame stream intercepted from the internet camera. In the case of marking a specific decoder threshold, a 1,000 unwatermarked I-frame stream was used. The results include a description of the decoder's robustness against the following types of attack: two-dimensional shearing, cropping, lossy compression, noise, rotation and desynchronization.

3.1. PSNR

To measure the differences between two digital images we use the term of peak signal to noise ratio (PSNR) expressed in a logarithmic scale. It is defined as:

$$PSNR = 20\log_{10}\left(\frac{I_p}{RMS}\right)[db], \tag{3}$$

$$RMS = \frac{1}{XY}\sum_{i=1}^{N}\sum_{j=1}^{M}([I(i,j) - I_W(i,j)]^2), \tag{4}$$

where: X, Y — spatial resolution of pictures compared: host and watermarked images;
I — peak value (number of colour quantization levels);
RMS — root mean square.

3.2. False-positive tests

The tests began with defining a probability of false-positive (FP) rate. The detection theory defines that the error occurs when a decoder interprets a picture as watermarked, while actually it does not contain any embedded information. The false-positive probability P_{FP} of the threshold method is defined as [18]:

$$P_{FP} = P(D_{\max} > \tau), \tag{5}$$

where: $D_{\max}$ — decoder decisions for unwatermarked pictures;
τ — decoder threshold equals 18.

The definition test consisted in testing 1000 separate I-frames and checking whether the decoder had recognized them as watermarked. Zero false-positive probability was obtained for decoder threshold $\tau = 18$ and this value was assumed for the algorithm.

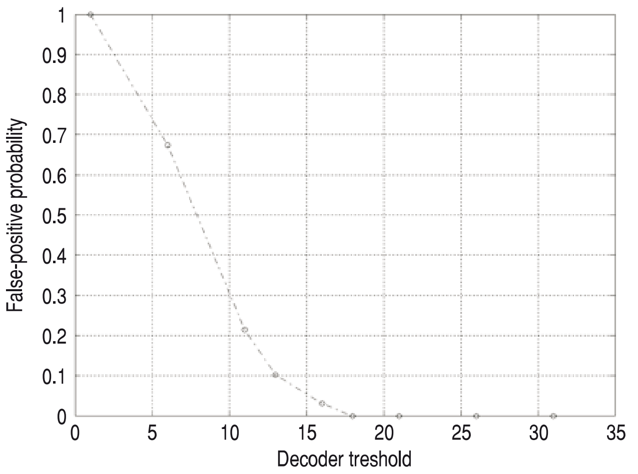

Fig. 1. The false-positive probability of the decoder for unwatermarked pictures

3.3. Invisibility

Invisibility is one of the critical requirements for the watermark algorithms (invisibility, data payload, robustness). In the method implemented, the sign δ value was selected to make the watermark invisible in the method described (PSNR not lower than 35 dB). For $\delta = 0.78$ average PSNR (for 100 pictures) after adding the translated luminance copy amounting to 37.12 dB.

3.4. Effectiveness

The effectiveness of the method is based on ROC curves showing the true-positive probability as a function of false-positive probability connected with the chosen decoder threshold τ. The attacks of translation, shearing and cropping have a low impact on the decoding process in the cepstrum domain, as shown in Fig. 3.

3.5. Robustness

The tests of robustness cover the attacks of rotation, shearing, translation, cropping, lossy compression, desynchronization and adding noise to the signal. The sections include the description of each attack. The red line on graphs marks curve approximations.

3.5.1. Lossy compression

The MJPEG standard makes it possible to increase the digital film compression factor by compromising its quality (Fig. 1c). Below we present a BER curve as a function of quantization matrix scaling coefficient.

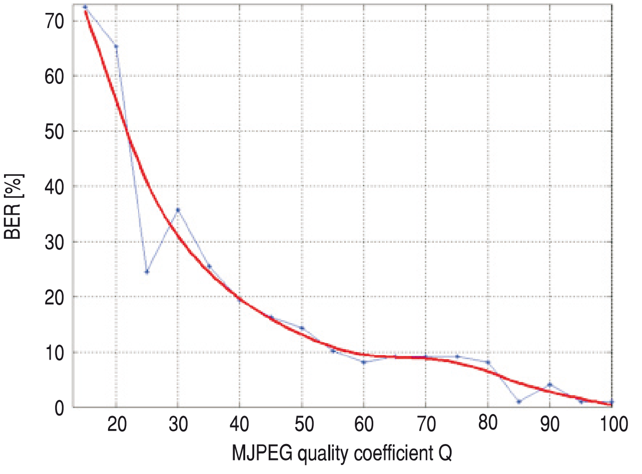

Fig. 2. BER graph as a function of MJPEG compression factor

3.5.2. Cropping

It results from the fact that the watermark (reduced energy luminance copy) is added to the whole luminance matrix and its cropping has a low impact on the decoding process. The unique advantage of algorithm is its robustness to strong cropping up to 10% and it provides 76% true-positive probability at a 1.02% false-positive rate.

3.5.3. Translation

Robustness to a translation attack, the same as for cropping and shearing results from using the cepstral transformation. Even an two-dimensional translation of 250% of each dimension processing causes no errors (BER equals 0%) as shown in Fig. 3.

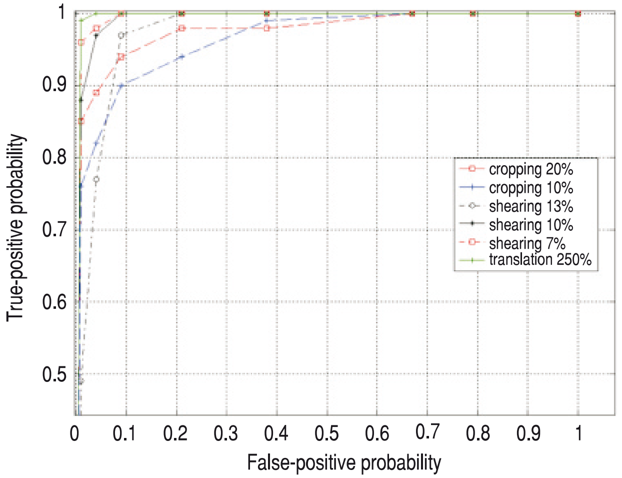

Fig. 3. ROC graph showing the dependence on the grade of cropping

3.5.4. Shearing

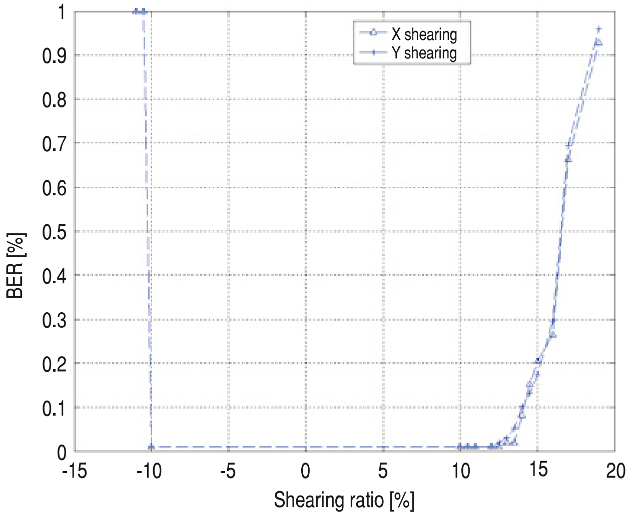

Fig. 4. BER graph showing the dependence on the grade of shearing

3.5.5. Rotation

For rotation attacks there is performed ROC analysis.

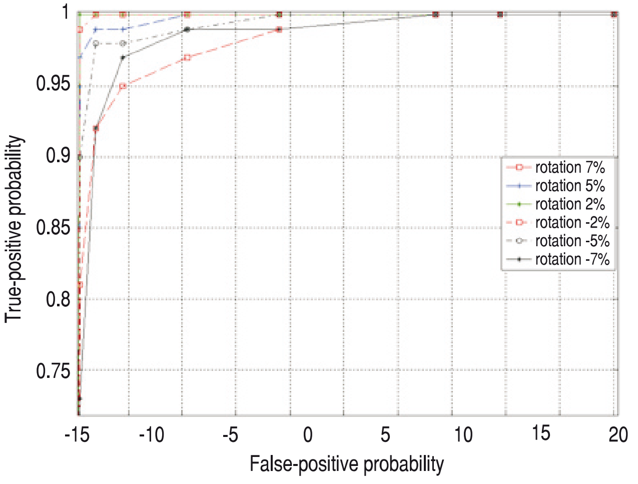

Fig. 5. ROC graph showing the dependence on the grade of rotation

3.5.6. Noise

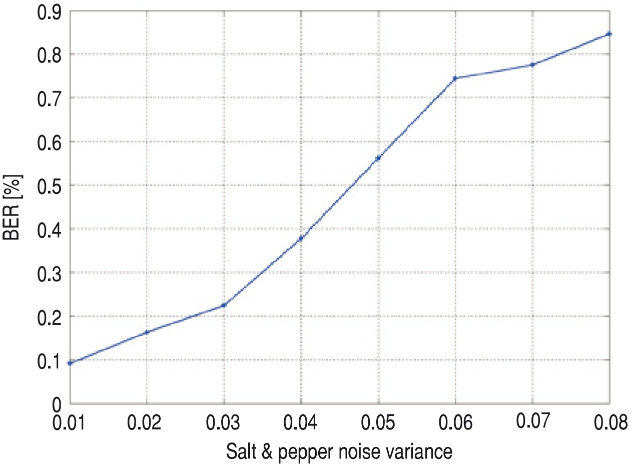

Fig. 6. BER graph showing the dependence on salt & pepper noise variance

3.5.7. Video stream desynchronization

Watermarked video streams are robust to desynchronization attacks and film fragment sequence swapping, because hidden information is embedded in 11 I-frame groups. If the attacks do not distort the films e.g. during defragmentation throughout its whole period (longer than 2 s) in intervals shorter than 440 ms (at 25 fps), BER equals 0%.

3.6. Example

Fig. 7. Original I-frame

Fig. 8. Watermarked I-frame, PSNR equals 36.9032 dB, additional information vector: [10101]

3. Conclusions

The paper presents a mathematical description of the algorithm for embedding additional information in the spatial domain. The process of watermark extraction occurs in two-dimensional cepstrum domain. It contains the results of tests covering both video I-frames and the whole video stream fragments (lossy compression, cropping, translation, shearing, rotation). One of the special advantages of the method suggested is its unique easiness of implementing the encoder for the algorithm described, non-requiring high computing power (I-frame exclusion, RGB conversion — YCbCr, double transcription of two-dimensional Y matrix, YCbCr conversion — RGB, I-frames inclusion). In addition, the algorithm to a large extent is independent of the host stream or video file. The main disadvantage of the algorithm is the decoder computing complexity that makes it difficult to develop a real-time processing program.

Future investigations will cover the implementation of Bose and Ray-Chaudhuri code for the described algorithm.

Received February 17 2011, revised May 2011.

REFERENCES

[1] Hui Tian, Ke Zhou, Hong Jiang, Jin Liu, Yongfeng Huang, Dan Feng, *An M-Sequence Based Steganography Model for Voice over IP*, Wuhan, China, Communications, ICC '09, 2009, 1-5.

[2] Hu, Lujian, Wang, Rangding, *Recoding-Resisted I-Frame Blind Watermarking Algorithm for H.264/AVC*, Multimedia Technology (ICMT), 2010, 1-5.

[3] Gao, Jie, Zhang, Chunfeng, *A content authentication scheme of H.264 video based on compressed*, Computer Application and System Modelling (ICCASM), 2010, V14-394–V14-397.

[4] Yule Yuan, Yong Zhao, Kunpeng Wang, Xuhui Bai, *A robust video watermarking algorithm based on sliding window for AVS*, Digital Content, Multimedia Technology and its Applications (IDC), 2010, 7-10.

[5] Li Jing, *A Novel Scheme of Robust and Blind Video Watermarking*, Information Technology and Applications, IFITA '09, 2009, 430-434.

[6] Ke Niu, Xiaoyuan Yang, Li Xiang, *Hybrid quasi-3D DWT/DCT and SVD video watermarking*, Software Engineering and Service Sciences (ICSESS), 2010, 588-591.

[7] Luo Wei, *A improved video watermarking scheme based on spread-spectrum technique*, Networking and Digital Society (ICNDS), 2010, 511-514.

[8] Tien-Ying Kuo, Yi-Chung Lo, *A hybrid scheme of robust and fragile watermarking for H.264/AVC video*, Broadband Multimedia Systems and Broadcasting (BMSB), 2010, 1-6.

[9] Liu Guang-qi, Zheng Xiao-shi, Zhao Yan-ling, Li Na, *A robust digital video watermark algorithm based on DCT domain*, Computer Application and System Modelling (ICCASM), 2010, V2-202–V2-205.

[10] B.P. Bogert, M.J.R. Healy, J.W. Tukey, *The Quefrency Alanysis of Time Series for Echoes: Cepstrum, Pseudo Autocovariance, Cross-Cepstrum and Saphe Cracking*, Proceedings of the Symposium on Time Series Analysis (M. Rosenblatt, Ed) Chapter 15, New York: Wiley, 1963, 209-243.

[11] Min-rui Zhang, Yangmei Zhang, *Colour Image Watermarking Algorithm in Cepstrum Domain*, Intelligence and Security Informatics, ISI '09, 2009, 226-228.

[12] Zhang Min-rui, Lu Chen-hong, Yi Ke-chu, *Cepstrum digital image watermarking algorithm*, Industrial Electronics, 1, 2004, 283-286.

[13] P. Lenarczyk, Z. Piotrowski, *Hybrid blind digital image watermarking in Cepstrum and DCT domain*, International Conference on Multimedia Information Networking and Security, MINES 2010, 356-361.

[14] Yanling Zhao, Xiaoshi Zheng, Na Li, Lufang Liao, Chengzhong Yang, *A robust audio digital watermarking in complex cepstrum domain based on chaos theory against STIRMARK*, Intelligent Control and Automation (WCICA), 2010 6398-6401.

[15] Xiaoming Zhang, Zhaoyang Yu, *Cepstrum-Based Audio Watermarking Algorithm against the A/D and D/A Attacks*, Information Assurance and Security, 2009, 740-743.

[16] Chen Ning, Zhu Jie, *A Novel Echo Detection Scheme based on Autocorrelation of Power Cepstrum*, Communications and Networking in China, CHINACOM '07, 2007, 554-558.

[17] Z. Piotrowski, *Drift Correction Modulation scheme for digital audio watermarking*, Proceedings 2010 Second International conference on Multimedia Information Networking and Security, MINES 2010, 392-397.

[18] C. Lin, M. Wu, Y.M. Lui, J.A. Bloom, M.L. Miller, I.J. Cox, *Rotation, Scale, and Translation Resilient Public Watermarking for Images*, IEEE Transactions on Image Processing, 10, 5, 767-782.

P. LENARCZYK, Z. PIOTROWSKI

Ślepe znakowanie cyfrowych strumieni wideo w czasie rzeczywistym w przestrzeni cepstralnej

Streszczenie. Wymóg pracy w czasie rzeczywistym jest szczególnie ważny w znakowaniu wodnym cyfrowych strumieni wideo. Artykuł zawiera opis algorytmu znakującego w czasie rzeczywistym macierzy luminancji w dziedzinie przestrzeni. Proces dekodowania odbywa się w dwuwymiarowej

przestrzeni Cepstralnej. Metoda znakująca uodparnia pojedynczą I-klatkę na ataki: kompresji MJPEG, przycinanie, przekształcenia afiniczne (dwuwymiarowe skręcenie, rotację, translację). Ponadto algorytm jest odporny na atak desynchronizacji. Dla przeprowadzonych ataków obliczono Charakterystyki Operacyjne ukazujące efektywność metody.
Słowa kluczowe: znakowanie wodne, wideo, czas rzeczywisty, Cepstrum