

A High-Capacity Block Based Video Watermark

Stefan Thiemert
Fraunhofer IPSI
Darmstadt

Thomas Vogel
Otto-von-
Guericke-
University
Magdeburg

Jana Dittmann
Otto-von-
Guericke-
University
Magdeburg

Martin Steinebach
Fraunhofer IPSI
Darmstadt

{stefan.thiemert, martin.steinebach}@ipsi.fraunhofer.de
{thomas.vogel, jana.dittmann}@iti.cs.uni-magdeburg.de

Abstract

Robust watermarks can be used for different applications. Main aspects are robustness, transparency and capacity. In this paper we present a block based video watermark, which is robust against several image processing operations. It provides a high capacity and good transparency. The watermarking scheme enforces a relationship between block averages in groups of blocks to represent the embedded binary message. Furthermore it creates a relationship between chosen coefficients into each block to represent the message redundantly. We present several evaluation results to demonstrate the robustness and transparency of the scheme.

1. Introduction

In the last years, digital videos replaced analogue technology. They can be copied easily without taking into account loss of quality. In the same way it became difficult to protect copyrights. To solve this problem robust digital watermarks can be used. Fields of applications are e.g. identifying copyright holders (copyright watermarks) or customers of documents (fingerprint watermarks) and embedding additional information into documents (caption watermarks). Several robust watermarks for images and videos were presented in the past [1][2][7]. Three important characteristics of this type of watermarks are [3]:

- Robustness: The watermark should survive post-production operations such as lossy compression, re-encoding, AD/DA-conversion.

- Transparency: The changes by a watermarking scheme on the material should not be noticeable.
- Capacity: The capacity is the amount of information, which can be embedded into the document.

Depending on the application the watermark should satisfy several requirements. A watermark used for copyright authentication should be very robust and transparent. In this case we have a loss of capacity. A caption watermark should be very transparent with high capacity to embed the additional information. Therefore it is not necessary to be highly robust.

In this paper we present a block based video watermarking scheme, which is high transparent while robustness and capacity are competitively scalable. The scheme is robust with a capacity up to 120 bit/s. It can be used for copyright authentication and customer identification as well as for caption watermarks. In chapter 2 we describe the framework of the watermarking scheme in detail. Furthermore we pay attention to some important security aspects of the scheme. Chapter 3 presents evaluation results to demonstrate the usability of our scheme in different applications. In chapter 4 we discuss advantages and disadvantages of the watermarking scheme and propose improvements.

2. Scheme framework

The main idea of this watermarking scheme is to embed a bit of a watermarking message by enforcing a relationship into a group of blocks. This relationship can be identified in the retrieval process without using the original video. The scheme, we describe in the following sections, works on quantized DCT blocks in

compressed videos of type MPEG-1/2. The size of a luminance block is 8x8. It can also be used in other image or video types, which use 8x8 DCT blocks, e.g. JPEG. Following we describe the embedding and retrieval process in detail.

2.1 Preparation

The preparation procedure will be applied both in embedding and retrieval process. For choosing the watermark positions it starts with the detection of all edges in a video frame i [4]. The image $E(i)$ generated from edge detection will be used for extracting elementary frame properties, which are moderately robust. It is rarely possible to change or remove edges without producing visible distortions. Therefore edges in an un-watermarked frame should be similar to edges in a watermarked and possibly attacked frame. For every block we compute its edge intensity, which is the sum of the edge intensity of every pixel.

The next step splits the frame into rectangular regions. In our experiments we used 16 regions per frame. For each region we compute its edge intensity by computing the sum of edge intensities of chosen blocks. For computation we use the blocks with the highest edge intensities, which will not be used for embedding the watermark. Therefore the edge intensity of every region is the same after the embedding process. To enhance the security of the embedding and retrieval process the regions will be pseudo-randomly shuffled controlled by a secret key k . The selection of blocks and coefficients and the preprocessing of all values to embed is controlled by k . The maximal length of k depends on the resolution of the video, e.g. for a 352x288 resolution key lengths between 2^6 bit and 2^{15} bit are possible. Figure 1 shows the preparation process schematically:

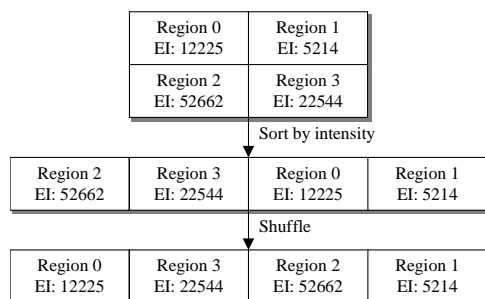


Figure 1. Preparation process

2.2 Embedding process

For embedding we build block groups, which consist of an odd number of DCT blocks. For one group we

choose blocks pseudo-randomly from different regions depending on a secret key. The embedding process combines two methods:

1. Embedding one bit into a group of blocks
2. Embedding a pattern into each block of the group representing one bit

For applying the first method we introduce the following condition:

- A group of blocks represents a 0 when the average of most of the blocks is lower than the average of the group.
- A group of blocks represents a 1 when the average of most of the blocks is higher than the average of the group.

With average we mean the average of all AC-values in a block respectively in a block group. Similar conditions are used e.g. in [6] but are based on a correlation between coefficients.

We describe the embedding process for embedding a 1 into a block group. Embedding a 0 is similar. If the given condition for a 1 is already fulfilled the algorithm builds the next block group. Otherwise the average of that block will be increased, whose average is the highest of all block averages under the group average. To increase the block average we use chosen middle frequency coefficients. For preserving the average of the group we decrease the average of another block. We change chosen middle frequency coefficients of the block with the lowest average in the opposite way as we have done with the first block. So the group average keeps constant. Figure 2 shows the first method of the embedding process:

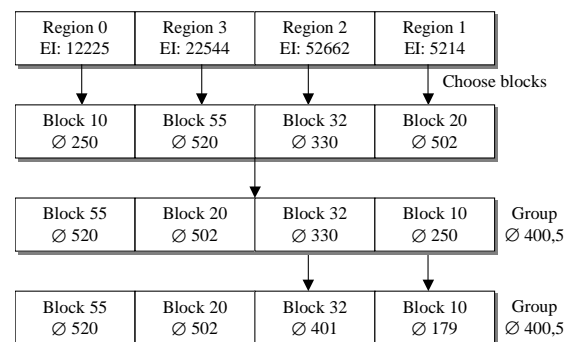


Figure 2. Embedding into block group

If more than one block needs to be changed to enforce the relationship into the group we repeat the described procedure with the remaining block nearest to the group average, which has to be modified and the block with the second lowest average.

The second method manipulates chosen coefficients in every block of the group so that the relationship between those coefficients represents the embedded bit.

In contrast to the approach proposed in [5] we do less modification to one DCT block. In contrast to Zhao and Koch we can abandon invalid patterns because the algorithm uses each block for embedding. In addition our approach is not limited to three coefficients. The number of coefficients used for embedding are chosen by the user.

An example: We use four coefficients of a block. If the coefficients should represent a *1* the following conditions have to be fulfilled:

$$|Coefficient0 - Coefficient1| \bmod 2 = 1 \quad (1)$$

$$|Coefficient0 - Coefficient2| \bmod 2 = 0 \quad (2)$$

$$|Coefficient0 - Coefficient3| \bmod 2 = 1 \quad (3)$$

The vector in those coefficients is therefore *101*. In the case of embedding a *0* the vector could be e.g. *011*. If these conditions are not fulfilled only a few of the coefficients have to be changed. Figure 3 shows an example of that method. Here we have the coefficients 8, 52, 13 and 30. As can be noticed three of the coefficients does not fulfill the requirements. To make only minor changes we change the coefficient value 8 to 9.

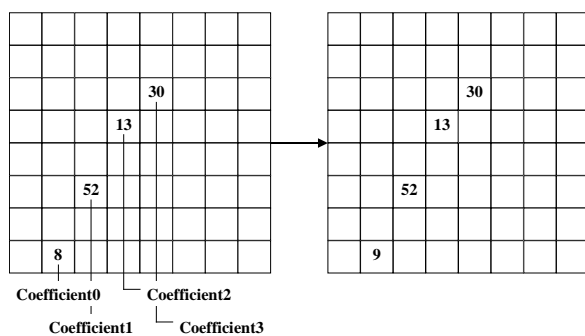


Figure 3. Embedding into block

For preserving the average of the block group it is reasonable to use the same coefficients for embedding the pattern and increasing the block average.

2.3 Retrieval process

The retrieval process starts with the preparation procedure and with building the block groups. Afterwards, for every block group, it analyses the average relationship. If the average of most of the blocks is higher than the group average the first result of the retrieval process in that group is a *1*, otherwise a *0*.

In the second part the algorithm detects the coefficient relationship in each block (eq. 1-3). The detected vector will be compared with reference

vectors for an embedded *1* (*101*) and an embedded *0* (*011*). The reference vector with the smallest Hamming distance to the retrieved vector is the detected vector. The Hamming distance describes the number of bits of the detected vector, which are different from the bits in the reference vector. Afterwards the algorithm uses a majority decision between the blocks in the group to generate the second result.

Finally the algorithm makes a decision depending on the computed results, which bit was embedded into a group. The weighting of the two results can be adjusted by the user. Because the relationship between the coefficients is more fragile it is useful to set a higher weight to the relationship between the averages of the blocks.

3. Evaluation

For evaluation ten 352x288 pixel video clips had been marked using a watermarking capacity of 60 bits per frame. All videos were encoded with MPEG with three different bitrates. We concentrated on the aspects transparency and robustness.

3.1 Transparency evaluation

Since perceptibility is mainly a subjective impression by the human visual system, seven persons were asked for visual evaluation. Furthermore the mean Peak Signal to Noise Ratio (PSNR) was calculated over all frames of each video to get additionally an objective measure. Figure 4 shows the results of the evaluation averaged over all test persons. The lower the value the better was the visual quality. The PSNR values are given in Figure 5.

All test persons recognized no or only small perceptible artifacts in the marked videos in comparison to their un-watermarked clips. More precisely, the higher the bitrate of the marked video was the better the transparency of the scheme had been rated by the test persons. The PSNR values affirm the subjective observation of the test persons. The clips encoded with the highest bitrates achieved the highest PSNR values during evaluation. Therefore we conclude that the transparency increases with the bitrate used for encoding.

3.2 Robustness evaluation

To evaluate the robustness all clips have been examined in the following experiments:

- Adding a fixed value to each coefficient
- Adding white noise

- Re-encoding and compression
- Scaling

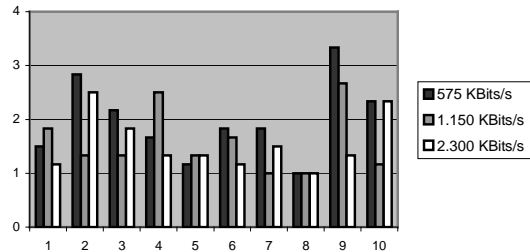


Figure 4. Results of visual evaluation

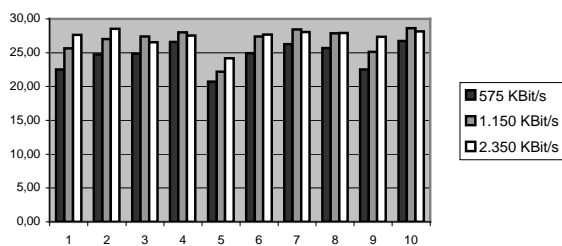


Figure 5. PSNR values

To get comparable results we used the Error Bit Rate (EBR). Results are shown in Figure 6. Except for scaling the EBR is below 7% for all tested videos, e.g. for adding white noise the EBR is constantly 0%. We used the scaled video directly to detect the watermark. Resizing the frames would lead to a definitely lower EBR. Another manipulation, cropping, was not evaluated. However, using the edge detection for generating arbitrarily regions should also provide robustness against this kind of manipulation. Overall we can determine that the results demonstrate a good robustness to basic image processing operations in conjunction with high capacity and non-perceptibility.

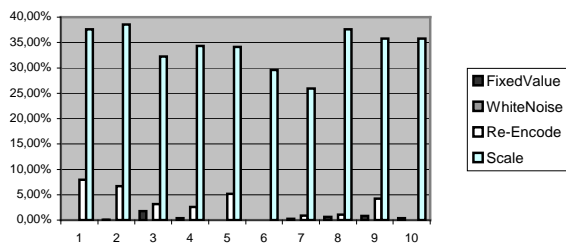


Figure 6. Error Bit Rates

4. Discussion and Conclusion

We present a block based watermarking scheme enforcing relationships between block averages to represent binary messages. Evaluation results show that the scheme is robust against several image processing

operations while providing high capacity and good transparency. The evaluation also shows possible challenges for future optimizations, e.g. with respect to scaling. There are many additional parameters not mentioned in this paper, which affect the properties of our scheme (e.g. size of block group, coefficients per block). Only a few combinations of this parameters have been evaluated yet. Hence, finding optimal parameter sets is another challenging task.

5. Acknowledgements

The information in this document is provided as is, and no guarantee or warranty is given or implied that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. The work described in this paper has been supported in part by the European Commission through the IST Programme under Contract IST-2002-507932 ECRYPT.

6. References

- [1] J. Fridrich, "Methods for data hiding", Center for Intelligent Systems & Department of Systems Science and Industrial Engineering, SUNY Binghamton, working paper, 1997
- [2] G.C. Langelaar, R.L. Lagendijk, J. Biemond: "Watermarking by DCT Coefficient Removal: A Statistical Approach to Optimal Parameter Settings", Security and Watermarking of Multimedia Contents, San Jose, CA, January 1999
- [3] Ingemar J. Cox, Matthew L. Miller, Jeffrey A. Bloom: "Digital Watermarking" Morgan Kaufmann Publishers, ISBN: 1-55860-714-5, 2001
- [4] W. K. Pratt, "Digital Image Processing", Ed. John Wiley and Sons, New York, USA, 1991
- [5] J. Zhao, E. Koch, "Embedding Robust Labels into Images for Copyright Protection", Proc. International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Technologies, Vienna, Austria, 21-25 August, 1995.
- [6] J. Dittmann, M. Stabenau, R. Steinmetz, „Robust MPEG Video Watermarking Technologies”, Proc. of ACM Multimedia '98, The 6th ACM International Multimedia Conference, Bristol, England, pp. 71-80, 1998
- [7] Jana Dittmann, Enrico Hauer, Claus Vielhauer, Jörg Schwenk, Eva Saar: „Customer Identification for MPEG Video based on Digital Fingerprinting”, Second IEEE Pacific-Rim Conference on Multimedia, Beijing, China, 2001