THE IMPROVED ITERATIVE ALGORITHM FOR THE ENHANCEMENT OF THE EXTRACTED WATERMARK FROM VIDEO STREAMING IN A WIRELESS ENVIRONMENT

Zoran Veličković*, Zoran Milivojević**, Milojko Jevtić***

Keywords: H.264 codec, SSIM, SVD algorithm, Watermark, Wireless.

Abstract: One way to protect the copyright of the video content is inserting invisible information - the watermark in the video content. The extraction of the embedded watermark is especially difficult in wireless networks, which are a subject to packet drops. Although multimedia applications are tolerant to packet losses, it cannot be said for their extraction algorithms. In this paper, it is shown that the improved iterative algorithm for watermark extraction from the encoded video can be successfully applied in the wireless environment. The efficiency of the improved algorithm is evaluated in the simulated wireless environment. In the wireless environment with 1% of dropped packets SSIM index was corrected by 24%, while in the environment with 2% of dropped packets SSIM index was corrected by 27.7%.

1. INTRODUCTION

Significant technological advances of digital communication networks have contributed that the exchange of the multimedia content on the Internet is becoming the dominant form of network traffic [1], [2]. The easy availability and characteristics of the digital multimedia content that with copying it does not lose its quality, has led to the appearance of illegal copies and their illegal distribution. Illegal copying and distribution of the multimedia content are particularly pronounced in music and film industry.

The importance of the appearance of illegal copying is the best evidenced by the fact that the illegal film market accounts for around 35.2% of pirated multimedia contents [3]. In the
In order to protect the multimedia content from illegal copying and distribution, they have developed numerous methods based on hardware and/or software solutions [4]. Depending on the application, specialized cryptographic methods or techniques inserting invisible information – the watermark in the multimedia content is used. The cryptographic methods are applied in the protection of the multimedia content during transmission through the communication channel, while the watermark techniques are used in the protection of property rights, monitoring the formation of digital backup, access control, etc.

The process of insertion or extraction of the watermark from the encoded video is linked with a lot of problems. The process of encoding video contents is based on neglecting details from the frames in the temporal and spatial domain. Ignoring the details of the video will of course adversely affect the algorithms of watermark extraction from the encoded video. The process of extraction of the watermark from the video is even more challenging when multimedia communications take place in the interference environment. The problems of variable characteristics of the transmission medium and interference phenomena are particularly pronounced for wireless communications. The appearance of slow and fast fading in wireless communications may result in restrictions, or even a temporary ban to data access. The fading consequence in the wireless environment could be the reduction of network bandwidth because of packet loss, packet delay and packet delay variation.

The algorithms for encoding video contents are based on the imperfections of HVS (Human Visual System), so that with multimedia communications some packet loss can be tolerated. A percentage of lost packets, which can be tolerated, depends on the type of the multimedia service, and ranges from 1% to 3% [5]. Although multimedia communications are tolerant to packet loss, they are very intolerant to packet delay. Permitted package delay also depends on the type of multimedia communications and ranges from 150 ms to 400 ms [5]. Packages that have larger delay than allowed will be rejected by multimedia application. Bandwidth, packet loss, packet delay variation and packet delay are the basic parameters used to measure the technical quality of multimedia services. On the other hand, business applications are intolerant to packet loss, but they are tolerant to packet delay, so that the multimedia network node can perform resource allocation in a manner that each application is granted minimum required resources [6].

In addition to packet loss which adversely affects to the users satisfaction with network applications QoE (Eng. Quality of Experience), they negatively affects to the quality of the extracted watermark from the multimedia content. Package rejection in multimedia applications need not be manifested in significant degradation of the media content, but it has a significant impact on the watermark extraction algorithms of the multimedia content. It is clear that the missing packets negatively affect the quality of the extracted watermark. This problem is even more pronounced in video applications because they are tolerant to packet loss, which is not favorable for the extraction algorithms.

This paper discusses the quality of the extracted watermark from a protected video that has been encoded to the H.264/AVC encoder. In this paper, the protected video is considered as uncompressed video in which the watermark is inserted. This paper presents
the iterative algorithm of the watermark extraction from the set of the extracted watermarks [7]. The performance of the iterative algorithm has been evaluated in the wireless environment with simulated packet loss. It is shown that the iterative algorithm can achieve the quality of the extracted watermark in the environment without interference under certain conditions.

The second chapter shows the H.264/AVC encoder with its specificities, as well as a set of tools (profiles) that determines the quality of encoding. The third chapter presents reliable algorithms for insertion and extraction of the watermarks from the uncoded video based on SVD decomposition. This chapter also presents an improved algorithm for correcting the quality of the extracted watermark regardless of the selected encoding profile. The fourth chapter presents the results in a simulated interference environment. The quality of the extracted watermarks is compared to the objective quality index SSIM [8]. In the fifth chapter, certain conclusions based on the obtained results are derived.

2. THE H.264/AVC CODER

By its nature, a video presents a series of correlated images in the temporal and spatial domain [9]. The compression algorithms applied to the H.264/AVC standard are based on the elimination of redundant information from a video [10]. The content prediction of the current frame is performed based on one or more previous or future frames, called reference frames. Reliability prediction is achieved by compensating movements between the reference and current frame. Thus, there are different frames, type I (intra), P (inter), B (bidirectional), SP and SI, which use one or more (past or future) reference frames. Specially encoded frames SP and SI are used to provide transitions from one bit rate to another if it requires the network condition. These frames provide smooth transitions between frames of the same video sequence or transitions between different sequences. SP and SI frames are designed to prevent the enormous increase of the network flow when frames are skipped or the video sequence is changed. SP frames within a video sequence define transition points, whereas SI frames are used to change the entire video sequence.

The H.264/AVC macroblock is defined as a region in the image determined with $16 \times 16$ pixels, which is a basic unit for compensation movements. A powerful mechanism for exploring the similarities in the current image or images which precede or follow is the basis of the prediction models of the H.264 encoder.

Anticipating the contents of certain parts of the image based on the perceived similarities, it is possible to form a "residual frame" with considerably less data. The consequence of this approach can be neglect of fine details in the frame, which will have a negative effect on the inserted mark. The result of this approach is the variable quality of the video, thus extracted mark, especially at lower bit rates and environments with interferences.

In previous works [7], [14] the iterative algorithm of the inserted watermark in the encoded video H.264/AVC encoder in the environment without interference is described.

This paper presents the results of the watermark repairs in the interference environment. A method for error concealment of the H.264 decoder that replaces the missing parts of the current frame to the relevant parts of the frame that precede it or follow is tested.

In order to enable the H.264 standard in the widest possible set of devices (different quality and formats), profiles, which are determined by a set of tools to generate
compressed video stream, are formed [10], [11]. The H.264/AVC standard defines the following profiles: Baseline, Extended, Main, and High. Main profile includes I, P and B frames, as well as tools for enhancing the resistance of encoded stream of errors. The Extended Profile is a superset of the Baseline Profile with attached B, SP and SI frames. Error tolerance is improved by forming multiple Data Partition (DP) in one frame. The baseline profile includes I and P frames, and some basic tools for resistance errors: Flexible Macroblock Ordering (FMO), Arbitrary Slice Ordering (ASO) and Redundant Slices (RS), and entropy coding CAVLC (Context-Adaptive Variable-Length Coding). Main profiles are designed for television and consumer entertainment electronics. An extended profile is effectively applicable to stream HD video. For professional use, the standard version of the H.264/AVC encoder is extended by adding new coding tools. This is an expanded version known as FRExt [11].

3. ALGORITHMS FOR INSERTING/EXTRACTING AND THE IMPROVEMENT OF THE WATERMARK QUALITY

Two major classes of algorithms for inserting watermarks in the multimedia content are used [12]. One class of algorithms is based on the insertion of the watermark in the spatial domain, while the second class of algorithms is based on modifying the coefficients in a transformation domain. In the case that the first class of algorithms is applied to the video, the watermark is hidden in the values of the luminance component of spatially distributed pixel images. The second class of algorithms is based on the modification of the transformation coefficients of a video content based on the transformation coefficients of the watermark. The transformation coefficients can be calculated by using DCT (Discrete Cosine Transform), FFT (Fast Fourier Transform) or SVD-a (Singular Value Decomposition). From the modified transformation coefficients, inserted information can be extracted by an inverse procedure. Based on the extracted information, inserted pictures are formed and that proves the ownership of the multimedia content.

Before encoding the H.264 encoder, in each frame of the uncompressed video, watermark is inserted. Inserting a watermark in the video depends on insertion factors \(\alpha(0,1)\), while the coding profile is defined by a set of parameters of H.264 encoder.

A. The reliable SVD algorithm

In this paper, the algorithm based on SVD decomposition [13], [14], which eliminates the false positive problem [15] for inserting a watermark in each video is used. Applying this insertion algorithm to each frame of video sequences, protected video is obtained. The input and output parameters of the algorithm for the watermark insertion [16], [17], [18] are:

**Input:**
- A series of matrices \(A_{m\times n}\) that represent the uncoded frames of video sequences.
- The matrix \(W_{m\times n}\) represents the image – the watermark that will be is inserted into the video.
- Insertion factor \(0 < \alpha < 1\).

**Output:**
- A series of matrices \(A_{m\times n}^w\) that represent protected video.
The algorithm for the insertion watermark in one frame of the video is displayed in six steps:

Step I1: Perform DCT transformation of the uncoded frame $A_{m \times n}$:

$$A_{DCT} = dct(A);$$  \hspace{1cm} (1)

Step I2: Apply the SVD decomposition of the matrix $A_{DCT}$ and determine the matrices $U$, $S$ and $V^T$:

$$A_{DCT} = USV^T,$$  \hspace{1cm} (2)

where $U$ and $V$ are orthogonal matrices of dimensions $m \times m$ and $n \times n$, respectively, $S$ diagonal matrix of dimension $m \times n$ with elements that represent singular values. The columns of the matrix $U$ are called left singular vectors, while the columns of the matrix $V$ in the right singular vectors.

Step I3: Perform the DCT transformation of the watermark $W_{m \times n}$:

$$W_{DCT} = dct(W);$$  \hspace{1cm} (3)

Step I4: Apply the SVD decomposition of the matrix $W_{DCT}$ and calculate the principal components $w_a$ (multiplication the left singular vector and singular values of the watermark) and $V_w^T$:

$$W_{DCT} = U_wS_wV_w^T = w_aV_w^T;$$  \hspace{1cm} (4)

Step I5: Calculate a matrix $S_1$ by adding the principal components $w_a$ on a diagonal matrix $S$ with insertion factor $\alpha$:

$$S_1 = S + \alpha w_a;$$  \hspace{1cm} (5)

Step I6: Calculate a watermark image $w$ with a new matrix $S_1$ and inverse DCT transformation:

$$w = idct(US_SV^T).$$  \hspace{1cm} (6)

Input or output parameters of the algorithm to extract the watermark are:

Input:
- A video with an inserted watermark, a series of matrices $A^*_w$,
- An original video, a series of matrices $A$,
- Insertion factor $\alpha$.

Output:
- An extracted watermark $W^*$. 
Watermark extraction from the frame, which is potentially different from the original due to the rounding and superimposed noise algorithm is performed to extract the watermark in seven E steps:

**Step E1:** Apply the DCT transformation of the matrix $A_w^*$:

$$ A_{DCT}^* = dct(A_w^*); $$

(7)

**Step E2:** Apply the DCT transformation of the matrix $A$:

$$ A_{DCT} = dct(A); $$

(8)

**Step E3:** Apply the SVD decomposition of the matrix $A_{DCT}$ and determine the matrices $U$, $S$ and $V^T$:

$$ A_{DCT} = USV^T; $$

(9)

**Step E4:** Calculate the difference matrix $A_1$:

$$ A_1 = A_{DCT} - A_{DCT}; $$

(10)

**Step E5:** Calculate the DCT transformation of the watermark $W$:

$$ W_{DCT} = dct(W); $$

(11)

**Step E6:** Apply the SVD decomposition of the matrix $W_{DCT}$ and determine the matrix $V_w^T$:

$$ W_{DCT} = U_w S_w V_w^T = A_w V_w^T; $$

(12)

**Step E7:** Construct the watermark $W^*$ applying the inverse DCT transformation:

$$ W^* = idct\left(\frac{1}{\alpha} U^{-1} A_1 \left(U^T \right)^{-1} V_w^T \right). $$

(13)

The application details of modified algorithm insertion and watermark extraction can be found in the previous papers of the authors [7], [14].

**B. The improved iterative algorithm for the enhancement of watermark extraction**

The quality variations of decoded video effects are the consequences of the variable multimedia content, in numerical rounding algorithms insertion/extraction of the watermark and disturbances caused on the transmission path. A consequence of variable video quality is, also, variable quality of the extracted watermark, so that in some situations it is necessary to apply algorithms to repair the quality of the extracted watermark [7].

A large set of watermarks relatively poor quality extracted from each frame is the available algorithm to improve the watermark quality. The basic idea of an iterative algorithm for correcting the extracted watermark quality consists in averaging luminance
component watermarks on a selected subset of extracted marks. The elements of the selected subset are the extracted watermarks with the highest SSIM (Structural Similarity Index) quality index [8]. The SSIM index is determined as follows:

\[
SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)},
\]

where \(\mu, \sigma\) are the mean value and standard deviation, respectively, of the luminance component of the original \((x)\) and extracted watermark \((y)\). All parameters for the local statistics are taken from [8]. \(C_1\) and \(C_2\) are positive stabilization constants (15):

\[
C_1 = (0.01 \cdot L)^2, \quad C_2 = (0.03 \cdot L)^2, \quad L = 255.
\]

The first averaging of the luminance component of the watermark is performed between the extracted watermarks with the highest quality mean SSIM index. If the value of the mean SSIM index of the resulting watermark is greater than the watermark with the highest quality, so the resulting watermark is accepted as the best, and it continues to participate in the averaging process. The following steps illustrate the pseudocode of the improved algorithm for correction the quality of the extracted watermark:

1. Extracting the watermark from each frame of the decoded video.
2. Calculating the mean SSIM index for each extracted watermark.
3. Creating the descending order watermarks from the largest to smallest mean SSIM index.
4. Creating a new watermark from the average value of luminant components of two watermarks with the highest mean SSIM index.
5. Determination the mean SSIM index of a new formed watermark.
6. If the mean SSIM index of newly created watermark is larger than the mean SSIM's watermarks from which it was developed, it should be used in further calculations, or reject it.
7. If all watermarks are not processed - jump to Step 4, otherwise jump to Step 8.
8. The improved video watermark is newly created – the averaged video watermark.
9. Filtering newly formed watermark with the \textit{wiener2} filter available in the most of the numerical mathematics programs.

Unlike the iterative algorithm for the quality watermark improvement that was presented in the previous works [7], [14] step 9 which performs further filtering watermark to remove additive noise generated by averaging luminant components of watermark has been introduced. The \textit{wiener2} method is based on statistics estimated from a local neighborhood of each pixel. A neighborhood of size \(2\times2\) pixels is used, and the additive white Gaussian noise power is assumed. In this way SSIM index of extracted watermark is considerably increased thus providing an improved algorithm for correcting the quality of the watermark.
4. REVIEW AND ANALYSIS OF RESULTS

In this paper the central part of the famous picture "Lena" in resolution 352 × 288 pixels was used as a watermark. Watermark was inserted in the first 50 frames of the famous video stream "Foreman.cif" by means of the reliable SVD algorithm, while the constant insertion factor $\alpha = 0.05$ was elected. The value of the insertion factor is empirically elected as a compromise between the quality of the extracted watermark and video. Greater value of the insertion factor $\alpha$ provides higher quality of the extracted watermark but causes unwanted artifacts in video sequences. Several methods for determining the optimal value of this parameter are introduced, but they are very processor-intensive and are not the subject of this study. In the literature, the insertion factor can be found in the range from 0.01 to 0.1 [18].

For encoding and decoding of the protected video sequence JM reference software [19], [20] was used. It is the official edition of ITU (International Telecommunication Union) in version 18.4 FRext. Protected video encoding was done by using tools from the Main profile, with a GoP (Group of Pictures) of 12 frames, one reference I-frame and one B frame. Testing the improved algorithm for correcting the extracted watermark quality from the encoded video was done in the wireless environment [21]. Two tests were performed. In the first test the wireless environment was simulated in which 1% of packets was dropped, while the latter rejected 2% of packets from the H.264 video stream.

Figures from 1 to 4 show decoded frames of the protected video in the noisy wireless environment. In order to preserve basic information needed for decoding video streams, the rejection of the package from the beginning of the video stream is prevented. In the frames where the packets are rejected, the H.264 decoder committed the replacement of the missing parts in the damaged frames from frames that precede or follow, and correctly received. Adjusted and correctly received frames make the decoded video at reception.

Fig. 1. The frames 31 (left) and 32 (right) of decoded protected video sequence streamed in a wireless environment with 1% of dropped packets.
Z. Veličković, Z. Milivojević, M. Jevtović: The improved iterative algorithm for the enhancement of the extracted watermark from video streaming in a wireless ...
Fig. 1 shows the layout of the decoded protected frames with numbers 31 and 32 received in a wireless environment with 1% of dropped packets. From Fig. 1 it can be observed that there are no noticeable artifacts in the displayed frames.

Fig. 2 shows the appearance of the decoded protected frames with numbers 34 and 35 received in a wireless environment with 1% of dropped packets. On the pictures, artifacts created as a result of replacing the dropped packets are observed.

Fig. 3 shows the frames with numbers 31 and 32 which are decoded in a wireless environment with 2% of dropped packets. Now, artifacts on both shown frames are noticeable as opposed to the ones shown in Fig. 1. This is a consequence of doubling a number of dropped packets.

Fig. 4 shows the frames with numbers 34 and 35 which are decoded in the same wireless environment. In these frames, significant distortion of the image content, which can be attributed to the progressive propagation of errors through the frames, can be seen.

Fig. 5. The SSIM index of extracted watermarks in a wireless environment a) by 1%, or b) 2% of dropped packets for each frame from the video "Foreman".
Fig. 5 shows SSIM index of extracted watermarks from video that make the collection over which is applied an improved iterative algorithm to repair quality. Fig. 5 shows the values of SSIM parameters in the wireless environment a) with 1% of dropped packets and b) 2% of dropped packets for each frame.

Drastically lowering SSIM index can clearly be seen from 32 to 35 frames in the wireless environment with 1% dropped packets, or from the 24th frame in the wireless environment with 2% of dropped packets.

In the wireless environment with 2% of dropped packets, methods for error concealment insert one additional inter-frame. This extra frame has a positive effect on man's perception of a video, but results in a low SSIM index of all frames until the end of the video sequence.

Fig. 6. Progressive correction of the SSIM index by applying the improved iterative algorithm in a wireless environment with a) 1% of dropped packets and b) 2% of dropped packets.
Fig. 6 shows the progressive correction of the SSIM index by applying the improved iterative algorithm for enhancement the quality of the extracted watermark from the video streamed in a wireless environment with a) 1% of dropped packets and b) 2% of dropped packets.

In the wireless environment with 1% dropped packets previously published the iterative algorithm for correcting the quality of the extracted watermark from the video [14] was recorded SSIM index of 0.55182 (a total of 18 iterations), and the proposed improved iterative algorithm achieved SSIM index of 0.68422 (a total of 19 iterations).

![Iteration #18, ssim =0.55182, DCT-SVD: Extr. watermark, α=0.05](image)

**a)**

![Iteration #11, ssim =0.508, DCT-SVD: Extr. watermark, α=0.05](image)

**b)**

Fig. 7. The appearance of the extracted watermarks obtained by the iterative algorithm for the quality improvement of the encoded video from the wireless environment with a) 1% of dropped packets, or b) 2% of dropped packets.
Fig. 7 shows the appearance of the extracted watermarks obtained by the iterative algorithm for quality improvement of video streaming in a wireless environment a) with 1% of dropped packets SSIM = 0.55182 and b) with 2% of dropped packets SSIM = 0.508.

In the wireless environment with 2% of dropped packets the iterative algorithm for correcting the watermark quality is achieved SSIM index of 0.508 (11 iterations), while the improved iterative algorithm is achieved SSIM index of 0.64856 (12 iterations).

Fig. 8. The appearance of the extracted watermarks obtained by the iterative algorithm for the quality improvement of the encoded video from the wireless environment with a) 1% of dropped packets, or b) 2% of dropped packets.
Fig. 8 shows the appearance of the extracted watermarks obtained by the improved iterative algorithm for correcting the quality of the encoded video in the wireless environment a) with 1% of dropped packets SSIM = 0.68422, and b) with 2% of dropped packets SSIM = 0.64856.

Based on the presented data derived from experiments it can be concluded that the improved iterative algorithm for correcting the extracted watermark from the video showed significantly better results, and can be efficiently used in wireless environments that are not immune to packet reject.

5. CONCLUSIONS

In order to prove the ownership of the video content, the reliable algorithms for insertion or extraction of the watermarks from a video have been developed. In the process of coding of the protected video, the degradation of the video occurs, and it has a negative impact on the extraction of the inserted watermark. The process of extracting the watermark is even more complex in the wireless environment where the rejection of packets is common. The paper shows that the quality of the extracted watermark from the video in the environments with noise can be improved by using the iterative algorithm.

Thus, the SSIM index of the extracted watermark in the environment with 1% rejected package amounted to 0.55182, while the environment with 2% of discarded packets was 0.508. By applying the improved algorithm for extracting the watermark from the encoded video the SSIM index of the extracted watermark is increased, so that in the wireless environment with 1% dropped packets amounted to 0.68422, while the wireless environment with 2% of dropped packets was 0.64586. Applying the improved iterative algorithm for correcting the quality of the extracted watermark from a video stream in the wireless environment with 1% dropped packets the SSIM index is corrected by 24%, while in the wireless environment with 2% of dropped packets SSIM index is corrected by 27.7%.

REFERENCES

Z. Veličković, Z. Milivojević, M. Jevtić: The improved iterative algorithm for the enhancement of the extracted watermark from video streaming in a wireless ...


