Robust Data Hiding Technique
for Video Error Concealment over DVB-H channel

F. De Simone, M. Carli, A. Neri
Applied Electronics Dept.
Università degli Studi Roma TRE
Rome, Italy
{carli; neri}@uniroma3.it

A. Hornsby and I. Defee
Institute of Signal Processing
Tampere University of Technology
Tampere, Finland
{adrian.hornsby; irek.defee}@tut.fi

Abstract—In this paper a method for providing additional information for error concealment during the transmission of a video file over an error-prone channel is presented. The transmission of the additional information is performed without increasing the bandwidth occupation, by hiding this information into the video itself at the encoder side. The considered transmission system is the DVB-H system. The performance of the method has been investigated by simulations and experimental tests in order to evaluate the perceived quality of the processed video and the robustness of the data hiding method against the H.264/AVC coding.

Keywords—data hiding, error concealment, H.264/AVC video coding, DVB-H transmission.

Topic area—Multimedia: Methods and Systems

I. INTRODUCTION

In spite of the introduction of error resiliency and error correction tools at the decoder and transmission level, the transmission of signals over error-prone channels leads to errors which can degrade the quality of the received content. When considering video transmission, the transport over error-prone channels can degrade the perceived quality notably [1].

Examples of the use of data hiding as an error control tool for video transmission over error-prone channel can be found in several works such as in [2], [3], [4], and [5].

In this paper we propose a scheme to transfer useful data for error concealment purpose, hidden into the video itself. At the decoder side, it is possible to detect and use the data hidden in one part of the video to conceal the errors that can occur within a next part of the received stream. In this way, the bit stream has an extra functionality for error concealment purpose, without increasing the bandwidth occupation during the file transfer.

The proposed method is different than previous works because the hidden information is directly a lower resolution version of significant frames of the video sequence. This information is hidden into the video, before the encoder. The receiver can thus use this information to conceal impulsive errors. The detected information is also useful to conceal multiple frame loss, by inserting the detected resized mark to substitute the missed frames. Our method uses the Quantization Index Modulation (QIM) embedding algorithm applied in the three dimensional Discrete Cosine Transform (3D-DCT) domain [6][7]. Thanks to the computational simplicity of the QIM detection algorithm, the embedded information can be easily extracted from the video to be used for error concealment.

In Section II, the proposed watermarking method for error concealment purpose is detailed. Simulations have been performed to analyze the visibility of the embedding method and the robustness against the H.264/AVC coding [8]. The results of these simulations are shown in Section III. Section IV presents the experimental results, obtained after the transmission of the video sequence over a real Digital Video Broadcasting Handheld (DVB-H) channel [9]. Conclusions are drawn in Sections V.

II. PROPOSED METHOD

In the proposed data hiding system, the raw video sequence is divided into groups of 8 consecutive frames, called shots. If the frame size is NrxNc, each shot can be represented by a matrix (V) with size NrxNxNf, with Nf=8.

The proposed scheme can be divided into an embedding part, detailed in subsection A, and a retrieval part, detailed in subsection B.

Useful information is extracted from the middle frame of the nth shot and it is hidden into the (n-1)th shot of the video sequence itself. Therefore, the information detected from the nth shot at the decoder side refers to the next shot to be received. The receiver can thus use this information to conceal errors in the (n+1)th shot. The middle frame of each shot is used to extract the mark because it best represents the entire shot. Indeed, the middle frame presents the minimum distance from both the first and the last frame of the shot, in terms of higher correlation.

A. Embedding Part

The overall embedding procedure can be summarized as follows:

1) The first step regards the marks extraction. A low-resolution version of the middle frame of each shot is extracted, by resizing the frame at 25% of its original dimension (from NrxNc, to N/2xNc/2 after the resizing). This low-resolution image can be expressed as an eight planes (Nr/2xNc/2x8) binary matrix. This is done by simply converting the luminance values of the image from decimal to binary representation. Each luminance value being in the range
of [0, 255]. In our experiments only the first four most significant bit planes of this binary matrix are considered. We then arrange this \( N_r/4 \times N_c/2 \times 4 \) matrix in order to obtain an \( N_r/4 \times N_c/2 \times 8 \) matrix. This matrix is now our binary mark.

2) The 3D-DCT is applied on every \( N_r \times N_c \times N_f \) shot (V) and produces a matrix of 3D-DCT coefficients (\( V_{DCT} \)), with size \( N_r \times N_c \times N_f \).

3) The binary mark, extracted from the \( n^{th} \) shot, is embedded into the \( V_{DCT} \) related to the \((n-1)^{th}\) shot (the host signal), using the QIM with encryption method [6]. Practically, the embedding algorithm modifies the 3D-DCT coefficients of the \( V_{DCT} \), quantizing them. According to the bit to hide, the quantizer chooses between one of two scalar uniform quantizers, shifted by \( \Delta/2 \). \( \Delta \) is the quantization step of the embedding. By applying the QIM embedding algorithm on \( V_{DCT} \), we obtained a matrix of watermarked coefficients (\( V'_{DCT} \)).

4) The three dimensional Inverse Discrete Cosine Transform (3D-IDCT) is applied on \( V'_{DCT} \) matrix. The watermarked shot (\( V' \)) ready to be transmitted is thus obtained.

B. Retrieval Part

To detect the useful hidden information at the receiver side, the inverse process of the embedding procedure is performed:

1) The decoder considers groups of 8 consecutive frames (shots) and applies the 3D-DCT on each shot so that the \( V''_{DCT} \) matrix is obtained. This matrix can differ from the \( V'_{DCT} \) because of transmission errors and losses.

2) The embedded mark is detected from each shot by applying the QIM detection algorithm [6]. Knowing the parameters of the embedding algorithm, the mark can be detected from \( V''_{DCT} \) as a \( N_r/4 \times N_c/2 \times 8 \) binary matrix. These parameters are the values of the quantization step \( \Delta \), the key used for the encryption and the zone of the embedding.

3) The values of the \( N_r/4 \times N_c/2 \times 8 \) binary matrix are rearranged using the inverse process used in step 1) of the embedding procedure. The four bit-planes matrix \( N_r/2 \times N_c/2 \times 4 \) is obtained. The conversion from binary to decimal representation is then performed, considering a \( N_r/2 \times N_c/2 \times 8 \) binary matrix. The first four most significant bit-planes of this eight planes matrix are composed of the \( N_r/2 \times N_c/2 \times 4 \) binary matrix. The other four bit-planes, the less significant, are set to zero. After the binary to decimal conversion a \( N_r/2 \times N_c/2 \) image is obtained and only a resizing is necessary to obtain the detected mark as a \( N_r \times N_c \) image.

III. SIMULATIONS

According to the QIM algorithm, if the quantization step \( \Delta \) increases, the quantization error increases. Therefore the quality of the video decreases. On the other hand, increasing \( \Delta \) increases the distance between the two quantizers. Thus the robustness of the embedded mark increases. A trade-off between visibility and robustness of the embedding method has to be defined, setting the best range of values for the embedding parameters. The embedding zone can be considered as a parameter too. We consider the data hiding in three different zones of the matrix of 3D-DCT coefficients \( V_{DCT} \): the left-high zone, corresponding to the low frequencies 3D-DCT coefficients, the middle zone, corresponding to the middle frequencies 3D-DCT coefficients, and the right-low zone, corresponding to the higher frequencies 3D-DCT coefficients. The effect of the data hiding is the most visible when the information is hidden in the low frequency zone of the 3D-DCT coefficients matrix. The data hiding in higher frequencies coefficients, has less effect on the perceived video quality, because the human visual system is more sensitive to the low frequencies.

Simulations have been performed to test the visibility of the data hiding method and its robustness against the H.264/AVC coding. The results of these simulations are described in the following subsections. In our simulation we have used a raw gray level video with frame rate 15 fr/s and format QCIF standard.

A. Objective Evaluation of the Loss in Perceived Quality on the Video Sequence

Three different metrics have been used: the Peak Signal to Noise Ratio (PSNR) and the Weighted PSNR (WPSNR), as average values considering all the frames of the original and the watermarked video sequence, and the Visual Quality Metric (VQM) developed by the ITS/NTIA [10].

The WPSNR metric gives the best results, according to the subjective perception evaluations. The WPSNR has been computed as follows:

\[ WPSNR(dB) = 10 \log_{10} \max(F)^2 \]  

\[ \frac{\text{max}(F)^2}{\| NVF(F' - F) \|^2} \]  

where the noise visibility function NVF is 1 in flat regions and 0 in textured regions and edges [11].

The average WPSNR plot in Fig. 1 shows that increasing \( \Delta \) degrades the video quality. It also clearly shows that the data hiding in the high zone of the 3D-DCT coefficients matrix causes the most visible impairment on the video. Considering the data hiding in the low zone of the matrix, the impairment produced by the data hiding becomes annoying only when \( \Delta > 120 \). In our simulations we also noticed that enlarging the screen size to visualize the video increases the perceived impairment. Since we consider the DVB-H scenario, our application involves mobile devices with limited size of the screen. So it is possible to neglect this screen size limitation. However, for the same \( \Delta \) value and the same screen size, better results in terms of perceived quality of the watermarked video are obtained considering higher resolution video sequences.

B. Evaluation of the Robustness of the Data Hiding Method against the H.264/AVC coding

We have evaluated the distortion introduced by the H.264/AVC coding on the detected information (the mark). The baseline profile of the H.264/AVC coding has been considered, varying the quantization parameter (QP) and the bit-rate of the output video.
sequence [12][13]. The robustness of the embedding method has been evaluated varying the embedding quantization step \( \Delta \). Regarding the embedding zone, it is a parameter that highly affects the quality of the watermarked sequence. Meanwhile, its influence on the robustness of the embedding method against the coding can be neglected. Therefore, the data hiding in the low zone of the 3D-DCT coefficients matrix has been considered, because it affects less the quality of the video.

Fig. 2 shows the plot of the PSNR between the original and the detected mark after the H.264/AVC coding (QP=26). It also shows the plot of the average WPSNR between the frames of the original and the watermarked video sequences, against \( \Delta \). The first shot of the video sequence has been considered. Fig. 3 shows the detected marks for \( \Delta = 70 \) and \( \Delta = 120 \), and the first frame of the watermarked sequence in these two cases. The size of the output video file after the coding is 380 KB (bit-rate 460kbps), for \( \Delta = 70 \), and 490 KB (bit-rate 600kbps) for \( \Delta = 120 \). The original file size is 7.3 MB (bit-rate: 9125kbps).

The quantization parameter of the coding is the critical parameter regarding the robustness of the embedding method. In particular, for the same output bit-rate of the coding, the quality of the detected marks strongly decreases if the QP is not constant. The performed simulations show that the minimum value of \( \Delta \) for which the embedding is robust enough to the coding strictly depends on the compression rate, e.g. on the QP. Moreover, the QP should be constant during the H.264/AVC encoding.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The robustness of the proposed method has been evaluated in terms of quality of the detected information at the receiver side, after the transmission over real error-prone DVB-H channel. The DVB network of Tampere University of Technology, Institute of Signal Processing, has been used to perform all the experiments. The watermarked video file has been loaded on a VideoLAN server, coded using the H.264/MPEG4 AVC encoder (using x264 library loaded on VideoLAN), encapsulated in MPEG TS, and transmitted over the IP network by sending it to a multicast address [14]. Using an external IP feed, the stream has been sent to the DVB-H encapsulator and then sent to the air over the DVB-T/H channel. At the receiver side, a laptop with DVB-T/H receiver and an antenna capable to receive the DVB signal has been used. The received file has been stored at the receiver side and decoding has been performed using VideoLAN. All the experiments have been done considering a grey levels raw avi file in QVGA format, frame rate 25fr/s, bit-rate 46084kbps, and length in time 42 seconds. Several transmissions have been performed, with different bit-rate of the encoded file and different error-resilience conditions, e.g. with and without MPE-FEC and in different weather conditions [9].

Examples of information detected from one shot received with errors on macroblocks level are shown in Fig. 4. The embedding has been done in the low zone with \( \Delta = 120 \), and the sequence has been coded using MPEG4 coding with output bit-rate of 512kbps (Fig. 4a), and H.264/AVC coding with output bit-rate of 256kbps (QP=26) (Fig. 4c). The MPE-FEC was disabled. Fig. 4b and Fig. 4d show further improved versions of the detected marks of Fig. 4a and Fig. 4c, obtained by applying a median filter on the detected mark.
Fig. 5 shows an example of results obtained by exploiting the information detected at the decoder side. The information detected from the n-th shot of the received video sequence is used to conceal errors within a frame of the (n+1)th shot. The fourth frame of the (n+1)th shot before the concealment is shown in Fig. 5a. The same frame after the concealment is shown in Fig. 5b and Fig. 5c. The information used for the concealment is the detected and filtered picture shown in Fig. 4d.

Figure 4: (a) Information detected from one shot received with errors at macroblocks level and after MPEG4 coding (output bitrate 512kbps); (b) picture (a) after applying a median filter; (c) information detected from one shot received with errors at bitrate 256kbps, QP=26; (d) picture (c) after applying a median filter.

Figure 5: (a) Example of a received frame of the (n+1)th shot with macroblocks errors; (b) example of error concealment applied to this frame by exploiting the information detected from the n-th shot; (c) frame after the error concealment.

V. CONCLUSIONS

In this paper we have presented a video watermarking scheme for error concealment purpose over error-prone video transmission. According to the performed simulations and to the experiments on running DVB-H system, the results outline that the method exhibits high robustness against attacks, still maintaining perceptual invisibility. Thus, the method provides a bit stream with an extra functionality for error concealment purpose without increasing the bandwidth occupation. The additional information detected at the receiver side shows excellent quality. This information can be thus really useful for the concealment of errors within the received video.

Future work will concentrate on further enhancing the embedding technique, in order to achieve robustness of the method independently from the compression ratio and without increasing its visibility. Furthermore, the development and the implementation of a complete error concealment algorithm at the decoder side could be a future step.

REFERENCES