Watermarking-based Error Concealment in Video Communications

Taesuk Oh* ***Student member***, Yongcheol Kim* ***Regular Member***,
Chowdary Adsumilli, Sanjit Mitra ***Regular Members***

ABSTRACT

In this paper, an informed watermarking algorithm is proposed that aids in concealing packet loss errors in video communications. This watermark-based error concealment (WEC) method embeds a low resolution version of the video frame inside itself as watermark data. At the receiver, the extracted watermark is used as a reference for error concealment. The proposed DCT-based algorithm employs informed watermarking techniques in order to minimize the distortion of host frames. At the encoder, a predictive feedback loop is employed which helps to adjust the strength of the data embedding. Furthermore, the distortion of the DCT coefficients introduced in the embedding can be removed to a considerable extent, by employing bit-sign adaptivity. Simulation results on standard video sequences show that the proposed informed WEC scheme has an advantage of 3~4 dB in PSNR over non-informed WEC and that even a non-informed WEC is still superior to conventional error concealment techniques.

Key Words : watermark, informed watermarking, error concealment, video communications

Ⅰ. Introduction

Packet losses that occur during the transmission of compressed video through lossy channels produce perceivable defects over multiple frames and have a significant influence on the quality of the received video at the end user. Error concealment is a technique that detects and hides the defects in video due to packet losses and therefore is one of the key processing steps in video communications. For the detection and correction of defects, error concealment techniques typically perform computationally intensive processing of spatial and temporal data in the received video, or they depend on certain critical information from the transmitter, like synchronization markers, to identify these packet loss errors.

Watermarking has been used for security applications where authentification and malicious attack prevention have been the primary focus. A new application for watermarking has been evolving lately for the error concealment in video communications. In this paper, we will refer to such schemes as watermark-based error concealment (WEC) schemes. In this work, we propose a WEC approach where a low resolution version of a video frame is embedded in itself at the encoder in the form of watermark data.

The low resolution image is a binary image with $1/16th$ the size of the original frame. It is obtained by halftoning a second level 2D-DWT of each frame. The low resolution image is embedded in a pixel-wise 2×2 format into the mid-frequency coefficients of a full-frame DCT of the original frame. At the receiver, a correlational detector is used to extract the low resolution image, which is used as a reference to identify and conceal any packet loss errors. We validate the

※ The authors wish to acknowledge the 2005 Overseas Research Grant and the 2005 Next-generation DCP Research Fund, Ministry of Commerce, Industry and Energy for supporting this study.
* Dept. of, Electrical and Computer Eng., University of Seoul (bbole@uos.ac.kr, yckim@uos.ac.kr)
** Image Processing Laboratory, Dept. of Electrical and Computer Eng., University of California, Santa Barbara
논문번호 : KICS2006-03-110, 접수일자 : 2006년 3월 6일, 최종접수일자 : 2006년 7월 10일

use of full-frame DCT by comparing it with the block-based DCT embedding scheme.

The algorithm used in this paper has some features of informed watermarking in that the encoder makes an informed decision in the embedding process. This informed embedding makes it possible to minimize the distortion of the host frames. At the encoder, a predictive feedback loop is employed which estimates the watermark detection accuracy at the receiver. Then, the strength of the scale factor α is determined such that BER at the decoder stays under some threshold. Furthermore, some of the modified coefficients of the DCT signal are virtually free from distortion by employing bit-sign adaptivity.

The organization of this paper is as follows. Previous works on WEC are presented in Chapter 2. The processes of embedding and extracting the watermarking for error concealment are described in Chapter 3 and details of the in formed watermarking are in Chapter 4. Simulation results on standard test sequences are presented in Chapter 5. Finally, we draw a conclusion in Chapter 6.

Ⅱ. Previous Works

The concept of applying watermarking to error resilience and concealment applications is quite new and not fully developed. Typically, as a simple form of WEC for image data, certain key features were extracted from the image, encoded, and hidden in the image itself either as a resilience tool or for concealment. The concept was extended to video coding by Bartolini *et al*[3]. However, they used data hiding as a tool to increase the syntax-based error detection rate in H.263, but not for the purpose of recovering lost data.

Munadi *et al.* extended the concept of key feature extraction and embedding to inter-frame coding[4]. In their scheme, important features are embedded into the prediction error of the current frame. However, the effects on motion vectors and the loss of motion compensated errors were not addressed. Yilmaz and Alatan proposed embedding a combination of edge information, block bit-length, and parity bits into intra-frames[5]. They use *even-odd signaling of DCT coefficients* for embedding, which is minimally robust.

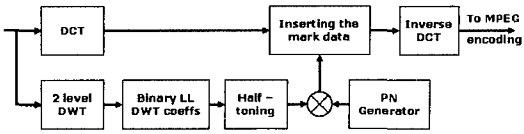


Fig.1. Block diagram of the embedding algorithm.

Informed watermarking has also been studied as a part of watermarking research[6]. A set of concealment techniques that do not use watermarking while giving similar high levels of performance have also been proposed[7]. However, WEC methods proved to give an improvement over the other techniques. A comparison of these techniques with the proposed one is provided in [1].

The problems with existing techniques are that (1) only one or a few selected set of key features are used for embedding. These features may not necessarily follow the loss characteristics of the channel employed and so are not effective for error resilience, and (2) they often use frequency domain such as DCT to encode the key data to be embedded. However, if losses occur on the DC or a set of the first few AC coefficients, the loss to the extracted reference would be significant and therefore may lead to reduction in concealment performance. Our proposed technique avoids these problems by embedding a half-toned low resolution version of the whole reference image. This way, loss or errors in the data will have smaller and local effects on the reconstructed video.

Ⅲ. Watermark-based Error Concealment

3.1 Watermark Embedding

The watermarking scheme used in this paper is a modified version of the Cox's algorithm[8]. In this work, Harr discrete wavelet transform(DWT) and halftoning techniques are used to generate the low resolution image, which is the watermark data to be embedded. This way, the DWT approximation image is transformed into binary values before being embedded. The block diagram of the embedding algorithm is shown in Fig. 1. A second level 2D-DWT is performed on the video frame to obtain an image that is $1/16th$ the size of the original frame. A half-toned image, the marker, is then generated from the reduced image. One marker is used for each key frame in the video. Each pixel of the marker is embedded 4 times in a 2×2 matrix format. Details are as follows.

Let the k-th frame be f_k with $m\times n$ pixels. The $2nd$ level DWT approximation image, α_k has a size of , which is halftoned into a binary-valued marker, m_k, with the same size of $m/4\times n/4$. The dithering technique in the halftoning is Floyd-Steinberg error diffusion algorithm[9]. Each pixel of the marker is then repeated in a 2×2 format to form w_k which has a size $m/2\times n/2$. Then, the final watermark data, $\widetilde{w}_k$, is generated as follows ($\bullet *$ represents element-by-element multiplication).

$$\widetilde{w}_k = w_k \bullet * p_k \quad w_k(i,j) \in \{-1,+1\}$$

p_k is a sized array of zero-mean unit-variance Gaussian.

Let x_k be the DCT coefficients of the luminance channel of the frame. The watermark, $\widetilde{w}_k$ is then scaled by a factor α, and then added to a mid-frequency set of coefficients, starting at the initial frequencies of (Δ_1,Δ_2). The modified DCT coefficients y_k are obtained by

$$y_k(i+\Delta_1, i+\Delta_2) = x_k(i+\Delta_1, i+\Delta_2) + a \bullet \widetilde{w}_k(i,j)$$

$$\begin{pmatrix} 0 \le i < m/2,\ 0 \le j < n/2 \\ \Delta_1 \in [0,\ m/2],\ \Delta_2 \in [0,\ n/2] \end{pmatrix} \qquad (1)$$

where i and j correspond to the pixel location in the watermark data and also the DCT coefficients. Finally, y_k is inverse transformed, encoded in compressed form and then transmitted.

3.2 Watermark Extraction

The extraction of the watermark employs a

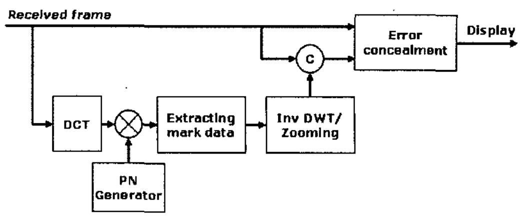


Fig.2. Block diagram of the retrieval algorithm.

correlator receiver and the block diagram is shown in Fig. 2. The received (noisy) DCT coefficients of the luminance channel, $\widetilde{y}_k$ are multiplied by the same pseudo noise array p_k and then summed for each 2×2 block. Then the binary marker is extracted by taking the polarity of the sum.

$$\widehat{m}_k(i,j) = \begin{cases} 1 & \text{if } \lambda_k(i,j) \ge 0 \\ 0 & \text{if } \lambda_k(i,j) < 0 \end{cases}$$

$$\lambda_k(i,j) = \sum_{i'=2i-1,\, j'=2j-1}^{2i,\, 2j} \widetilde{y}_k(i'+\Delta_1, j'+\Delta_2) \bullet p_k(i',j')$$

$$(0 \le i < m/4,\ \ 0 \le j < n/4) \qquad (2)$$

An inverse halftoning algorithm proposed by Xiong[10] is applied to $\widehat{m}_k$ to obtain an estimate of the low-resolution approximation image, $\widehat{a}_k$. A 2-D inverse DWT is performed on $\widehat{a}_k$ to obtain an intermediate resolution image, $b_k(m/2\times n/2)$. Then, b_k is zoomed by a factor of 2 by up-sampling and passing through a lowpass filter to obtain a reference image, $g_k(m\times n)$ size. An estimate of the current frame, $\widehat{f}_k$ is obtained by decompressing the received (watermarked) video data packets. The reference g_k is compared with $\widehat{f}_k$ to detect and conceal the corrupted areas of $\widehat{f}_k$. More detailes on WEC operations can be found in [1,2].

Ⅳ. Informed Watermarking Algorithm

Data embedding in this work is based on informed watermarking. The strength of the embedded watermark varies adaptively by employing a predictive feedback loop. This adaptivity highly decreases not only the BER of the extracted watermark but also the perceivable distortion in the video introduced by the watermarking process. Two variations of the informed methods are explained in detail herein.

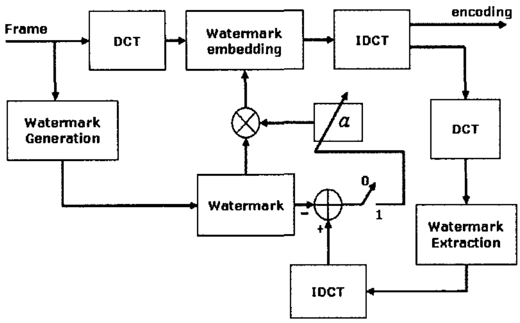


Fig.3. Feedback-based watermark embedding model.

4.1 Predicted Watermark Detection

The embedder has a predictive detector which is connected in a feedback loop to adapt to the strength of the embedding host signal, as shown in Fig. 3. This allows the values of the scale factor α to vary such that the probability of error in detecting the watermark at the receiver is minimized. Starting from a small fixed value, α continues to vary in incremental steps until the watermarked bit is extracted correctly in the predictive detector for a target level of packet errors. When it is correctly extracted, the value of α is fixed for that coefficient and therefore eq. (1) can be restated as follows.

$$\begin{aligned} & y_k(i+\Delta_1, j+\Delta_2) \\ &= x_k(i+\Delta_1, j+\Delta_2)+a_k(i,j) \cdot \widetilde{w_k}(i,j) \end{aligned} \tag{3}$$

The advantages of this informed WEC method over the other WEC methods discussed in Chapter 3 are threefold: (1) the perceivable watermarking defects are negligible due to the adaptive scaling of the strength of the embedded bits, (2) the BER values are much lower in the informed technique when compared to the non-informed WEC methods, and (3) a higher level of compression can be achieved at the codec as the entropy of the embedded video frame is not as high in case of the informed WEC technique since smaller α implies a reduction in the modification of the host coefficients.

However, one of the conceivable disadvantages of this technique would be that the complexity of the embedding module will be slightly higher. Nevertheless, this is compensated by the performance improvement at the receiver.

4.2 Watermark Bit-Sign Adaptivity

In this variation, the watermark is embedded by adapting the host frame coefficient with the sign of the pseudo-random number. A watermark bit of +1, for example, is embedded by modifying the sign of the coefficient to be in accordance with the sign of $p_k(i,j)$. This could be alternatively considered as a form of multiplicative embedding since the embedded frame y_k is obtained by multiplying the host frame f_k with the sign of embedded bit and that of the pseudo-noise image:

$$\begin{aligned} & y_k(i+\Delta_1, j+\Delta_2) \\ &= x_k(i+\Delta_1, j+\Delta_2) \cdot (a_k(i,j) \cdot \widetilde{w_k}(i,j)) \\ &= x_k(i+\Delta_1, j+\Delta_2) \cdot (a_k(i,j) \cdot m_k . p_k(i,j)) \end{aligned} \tag{4}$$

At the detector, the value of the watermark bit is extracted in a correlation receiver similar to (2) from the lossy received signal, $\widetilde{y_k}$.

$$\widehat{m_k}(i,j) = \begin{cases} 1 & \text{if } \lambda_k(i,j) \geq 0 \\ 0 & \text{if } \lambda_k(i,j) < 0 \end{cases}$$

$$\lambda_k(i,j) = \sum_{i'=2i-1, j'=2j-1}^{2i, 2j} \widetilde{y_k}(i'+\Delta_1, j'+\Delta_2) \cdot p_k(i', j') \tag{2}$$

$$(0 \leq i < m/4, \quad 0 \leq j < n/4) \tag{5}$$

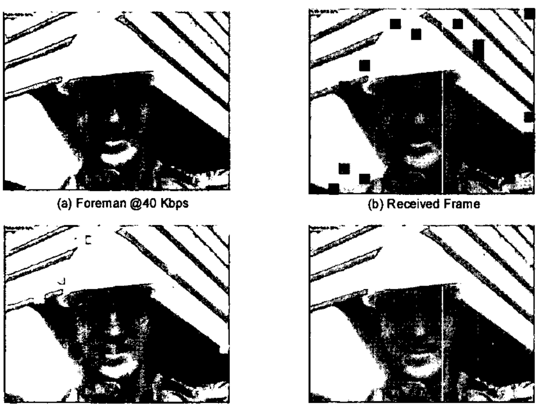
This technique has some advantageous features as follows. In addition to the convenience of tandem operation with the feedback-based informed WEC algorithm, it reduces the host frame distortion: Once the watermark is correctly detected, the coefficient in the host frame can return to some value with its original polarity by simply multiplying the pseudo noise matrix with the embedded coefficients.

V. Experimental Results by Simulations

A sample set of CIF resolution (288x352) videos is considered for simulation and the watermark is inserted in the central AC frequencies of the full frame DCT. NS-2 simulator is used to generate packet losses with a two-state Gilbert-Elliot Gaussian packet loss model with predefined mean and variance. The packet size was fixed at a macro block (MB) size for a given video transmission and no retransmissions were allowed.

Fig. 4shows the performance of WEC on the Foreman video sequence. The received frame is obtained for a mean packet loss probability of 0.12 % and variance 0.02%. The value of α was fixed at 0.6 for the non-informed WEC case. Fig. 4(b) shows the lossy received frame while Fig. 4(c) shows the error concealed (EC) frame. Since the watermark is a low resolution version, most of the high frequency information is not retained. Therefore, the reference frame looks "smooth" and the EC frame is patchy. This effect is reduced if the EC frame is locally-scaled based on the neighboring luminance values. We call the resulting frame as locally-scaled error concealed (LEC) frame. Fig. 4(d) represents the LEC frame combined with bit-sign adaptivity.

Table 1 shows the performance of the proposed algorithms over various video sequences. As seen,



(a) Foreman @40 Kbps (b) Received Frame (c) EC frame (d) Bit-sign LEC

Fig.4. The results for the 36-th frame of Foreman in CIF resolution compressed to 40 Kbps. The PSNR values are (b)14.26 dB (c) 27.63 dB (d)36.25 dB, respectively.

Table 1. Performance of the proposed algorithms and variations. PSNR (in dB) for a fixed mean loss 0.15% and variance 0.025%

		Non-informed WEC		Informed WEC		
Video	Received	EC	LEC	EC	LEC	Bit-Sign
Akiyo	16.1248	28.9114	31.5345	31.9824	34.7662	35.5219
Foreman	14.8279	26.6248	31.4983	32.6751	34.3655	35.5787
Table tennis	15.8202	27.4553	30.1257	31.6767	32.9044	35.0286
Flower	14.5233	26.4553	30.0476	30.8902	32.0967	33.3261
Football	15.0728	27.3546	30.5126	31.0207	33.2548	35.3652
Paris	14.6905	27.0817	30.9213	31.0576	32.9247	33.6563
Tampete	14.4857	27.6707	30.1732	30.5498	33.1634	35.2458
Highway	15.0253	28.7346	32.4018	32.9790	34.3862	34.2892

for all cases except Flower and Highway, the bit-sign adaptive informed WEC gave higher PSNR values of about 3 dB over the non-informed WEC technique.

VI. Conclusions

We presented an application of informed watermarking algorithm to video error concealment. This WEC approach used full-frame DCT to embed a low resolution version of the video frame in itself. The extracted watermark was used for error concealing in the lossy received video frame. The algorithm employed a feedback loop to predict the values of the extracted watermark bits, thereby reducing the overall BER of detected watermark at the receiver. Bit-sign coefficient modifications were also presented.

Based on the obtained results, we conclude that the informed watermarking algorithm gave better performance not only in terms of higher PSNR values but also in terms of reduced BER values. The bit-sign variation proved to reduce the perceivable defects introduced by watermarking process, when compared to the non-informed WEC method. Finally, we analyzed the advantages of full-frame DCT embedding scheme over the block-based scheme.

References

[1] Adsumilli, C.B., Farias, M.C., Mitra, S.K., Carli, M, "A robust error concealment tech-

nique using data hiding for image and video transmission over lossy channels," Vol. 15, No. 11, IEEE Trans. Circuits and Systems for Video Technology, Nov. 2005, pp.1394-1406.

[2] Adsumilli, C.B.; Mitra, S.K.; Kim, Y.C., "Detector Performance Analysis of Watermark-Based Error Concealment in Image Communications,"Proc. of ICIP, IEEE, Sept. 2005 pp.:916-919.

[3] Bartolini, F. et al. A data hiding approach for correcting errors in H.263 video transmitted over a noisy channel. Proc. IEEE Workshop on Multimedia Signal Processing (2001) pp. 65-70.

[4] Munadi, K., Kurosaki, M., Kiya, H., " Error concealment using digital watermarking technique for interframe video coding," Proc. Intl. Tech. Conf. on Circuits and Systems, Computers, and Communications (2002) pp. 599-602.

[5] Yilmazi, A., Alatan, A., "Error concealment of video sequences by data hiding," Proc. of ICIP, IEEE, 2003, pp.679-682.

[6] Miller, M.L., Cox, I.J., Bloom, J.A., "Informed embedding: Exploiting image and detector information during watermark insertion," Proc. of ICIP, IEEE, 2000.

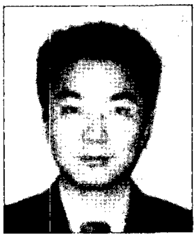
[7] Salama, P. et al, "Error concealment techniques for encoded video streams," Proc. ICIP, IEEE, 1995, pp.9-12.

[8] Cox, I., Kilian, J., Leighton, F., Shamoon, T., "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Processing 6(12) (1997) pp.1673-1687.

[9] Floyd, R.W., Steinberg, L., "An adpative algorithm for spatial gray-scale," Proc. Society of Information Display 17(2) (1976) pp.75-78.

[10] Xiong, Z., Orchird, M.T., Ramachandrani, K., "Inverse half-toning using wavelets," IEEE Trans. Image Processing 8(10) (1999) pp.1479-1483.

Taesuk Oh Student member



2001 B.S (Electronics Information communication Coeducation) Chongju University

2003 M.S. (Electrical and Computer Engineering) University of Seoul.

2004~ Ph.D. student in Dept. of ECE, University of Seoul

Yongcheol Kim Regular member

1981 B.S. (Electronics Engineering) Seoul National University

1983 M.S. (Electrical Engineering) KAIST, Seoul

1983~1986 Team Leader of Spread-Spectrum Radio in LG Electric

1993 Ph.D. (Electrical and Computer Engineering) University of Southern California

1993~1996 Team Leader of Computer Software Team in LG Innotek

1996~ Professor in Dept of ECE., University of Seoul

Chowdary Adsumilli Regular member

1999 B. Tech (Electronics Engeering) Jawaharlal Nehru University, India

2001 MS (Electrical and Computer Engineering) University of Wiscon, Madison

2005 Ph.D. (Electrical and Computer Engineering) University of California, Santa Barbara

2005~ Citrix Systems, Inc

Sanjit Mitra Regular member

1953 B.Sc. (Physics) Utkal University, India

1960 MS (Electrical Engineering) University of California, Berkeley

1962 Ph.D. (Electrical Engineering) University of California, Berkeley

1977~ Professor at Department of ECE, University of California, Santa Barbara