영상통신에서의 패킷 오류 은닉에 활용되는 워터마킹 기법

Application of watermarking to error concealment in video communications

오태석, 김용철
Tae Suk Oh, Yong Cheol Kim

Abstract - An informed watermarking algorithm is proposed in this work 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 has features of informed watermarking in order to minimize the distortion of the host frame. At the encoder, a predictive feedback loop is employed which helps to adjusts the strength of the scale factor. Furthermore, some of the modified coefficients of the DCT signal are virtually free from distortion by employing bit-sign adaptivity. The performance of the detector is qualitatively analyzed for error concealment applications where full-frame DCT embedding proved to be more advantageous.

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

1. Introduction

Watermarking has been used for security applications where authentification and malicious attack prevention have been the primary focus[1]. 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 an efficient WEC approach where a low resolution version of a video frame is embedded in itself at the encoder in the form of watermark data. This informed embedding makes it possible to minimize the distortion of the host frame. 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 is determined such that BER at the decoder is sustained under some threshold. Furthermore, 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. The processes of embedding and extracting the watermarking are described in Chapter 2 and details of the in formed watermarking in Chapter 3. Comparison of full-frame DCT vs. block-based DCT is provided in Chapter 4. Simulation results in Chapter 5. Finally, we draw a conclusion in Chapter 6.

2. Watermark-based Error Concealment

2.1 Watermark Embedding

The watermarking scheme used in this paper is a modified version of the Cox's algorithm[2]. 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 size image. One marker is used for each key frame in the video. Each pixel of the marker is embedded 4 times in a matrix format. Details are as follows.

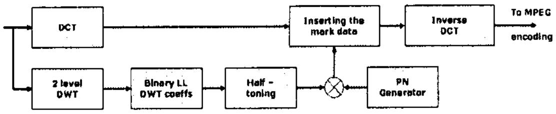


Fig.1. Block diagram of the embedding algorithm.

Let the $k-th$ frame be f_k with $m \times n$ pixels. The $2nd$ level DWT approximation image, a_k has a size of $m/4 \times n/4$, 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[3]. 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, $\tilde{w}_k$, is generated as follows ($\cdot *$represents element-by-element multiplication). $\tilde{w}_k = w_k \cdot * p_k \quad w_k(i,j) \in \{-1,+1\}$

p_k is zero-mean unit-variance Gaussian. Let the DCT coefficients of the luminance channel of the frame x_k. The watermark, $\tilde{w}_k$ is then scaled by a factor, a, and then added to a mid-frequency set of coefficients in x_k, starting at the initial frequencies of (Δ_1,Δ_2). The resulting coefficients y_k are obtained by

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

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

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

2.2 Watermark Extraction

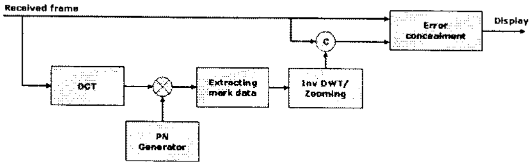


Fig.2. Block diagram of the retrieval algorithm.

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

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

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

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

An inverse halftoning algorithm proposed by Xiong[4] is applied to $\hat{m}_k$ to obtain an estimate of the low-resolution approximation image($\hat{a}_k$). 2-D inverse DWT is performed on $\hat{a}_k$ 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($g_k(m \times n)$). An estimate of the current frame($\hat{f}_k$) is obtained by decompressing the received(watermarked) video data packets. The reference g_k is compared with $\hat{f}_k$ to detect and conceal the corrupted areas of the image $\hat{f}_k$. More details on WEC operations can be found in [5].

3. Informed Watermarking Algorithm

3.1 Predicted Watermark Detection

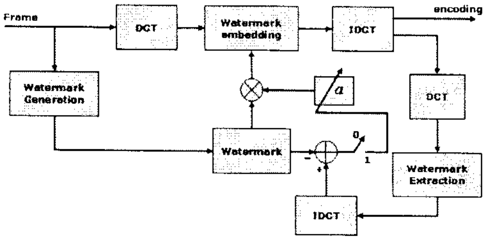


Fig.3. Feedback-based watermark embedding model.

The embedder has a predictive detector which is connected in a feedback loop to adapt to the strength of the embedded signal, as shown in Fig. 3.

This allows the values of the scale factor a to vary such that the probability of error in detecting the watermark at the receiver is minimized. Starting from a small fixed value, a 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 a is fixed for that coefficient and therefore eq. (1) can be restated as follows.

$$y_k(i+\Delta_1, j+\Delta_2) \tag{3}$$
$$= x_k(i+\Delta_1, j+\Delta_2) + a_k(i,j) \bullet \tilde{w}_k(i,j)$$

3.2 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:

$$y_k(i+\Delta_1, j+\Delta_2) \tag{4}$$
$$= x_k(i+\Delta_1, j+\Delta_2) \bullet (a_k(i,j) \bullet \tilde{w}_k(i,j))$$

it reduces the parent frame distortion: Once the watermark is correctly detected, the coefficient in the host frame can return to the value with its original polarity by simply multiplying the pseudonoise matrix with the embedded coefficients. Furthermore, the detection of the embedded bits can be performed based on comparison between the embedded coefficients and the host coefficients at the encoder.

4. Full frame DCT vs. block-based DCT embedding

There are two distinct advantages of using a full frame DCT over a block-based DCT for embedding.

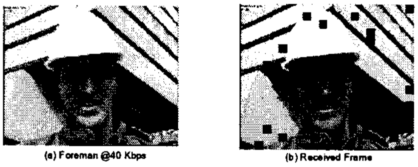
The first advantage is evident in that, in the full frame DCT case, we don't need to embed a single bit multiple times to get the same level of decoding efficiency. Since we use only 4 coefficients (2×2 set) for embedding each bit, instead of the central fold of 8×8 AC coefficients, we can embed more information in fewer coefficients than the conventional block-based embedding. The second advantage, let us first consider the distribution of AC coefficients in both full frame and block-based DCT cases. The coefficients in the DCT are a sum of cosine-weighted pixel values. According to the central limit theorem, the greater the number of terms in the summation, the smaller the standard deviation of the normalized sum. Since the coefficients in the full frame DCT are a weighted sum of a much larger number of pixel values than those in block-based DCT, the deviation in the sum of coefficients throughout the AC mid-coefficients is much less than the deviation in the sum of AC mid-coefficients of the individual blocks in block-based DCT. We can expect to have a higher accuracy of the detection of the incoming bits with full-frame DCT when compared to the block-based DCT.

5. Experimental Results by Simulations

A sample set of CIF resolution 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. 4 shows 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 a 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 with bit-sign adaptivity. Here, it should be noted that the bit-sign variation is used on top of informed WEC. Table 1 shows the performance of the proposed algorithms over various video sequences. As seen, 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.


(a) Foreman @40 Kbps (b) Received Frame

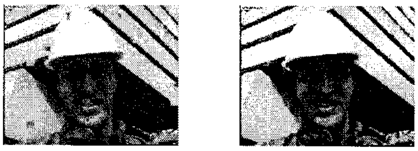

(c) EC frame (d) Bit-sign LEC

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

6. Conclusions

We presented an informed watermarking algorithm for the application of 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 the lossy received video frame. The algorithm employed a feedback loop to predict beforehand the values of the extracted watermark bits, thereby reducing the overall BER of the 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] Kalkar, T., "Considerations on watermarking security," Proc. IEEE Workshop on Multimedia Signal Processing (2001) pp.201-206

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

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

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

[5]Adsumilli, C.B., Farias, M.C., Mitra, S.K., Carli, M, " A robust error concealment technique 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

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

1981 : B.S. (Electronics Engineering) Seoul National University
1983 : M.S. (Electrical Engineering) KAIST, Seoul
1983~1986 : Team Leader of Spread-Spectrum Radio R&D Team in LG Electric Co.
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