

ADAPTIVELY SWITCHING BETWEEN DIRECTIONAL INTERPOLATION AND REGION MATCHING FOR SPATIAL ERROR CONCEALMENT BASED ON DCT COEFFICIENTS

Yan Chen, Oscar Au, Jiantao Zhou, Chi-Wang Ho

Dept. EEE, Hong Kong University of Science and Technology

Email: {eecyan, eeau, eejtzhou, jodyho}@ust.hk

ABSTRACT

In this paper, a novel spatial error concealment algorithm, which adaptively switches between directional interpolation and region matching, is proposed. Different from the previous spatial error concealment methods, which just utilize smooth property, the algorithm exploits both smooth property and texture information to recover the lost blocks. Based on the DCT coefficients in the available neighboring MBs, the algorithm automatically analyzes whether the MB is “smooth-like” or “texture-like” and adaptively select directional interpolation or region matching to recover the lost MB. The proposed algorithm has been evaluated on H.264 reference software JM 9.0. The experimental results demonstrate that the proposed method can achieve better PSNR performance and visual quality, compared with weighted pixel average (WPA) which is adopted in H.264, directional interpolation-only and region matching-only.

1. INTRODUCTION

Error Control is very important in video communication system over unreliable transmission environment such as wireless networks and Internet. Compared with other mechanisms such as forward error correction (FEC) and automatic retransmission request (ARQ), error concealment by post-processing at the decoder is more attractive since it needs neither extra bandwidth nor transmission delays. There are two kinds of error concealment technologies, one is spatial error concealment which focuses on the inherent correlations in spatial domain; the other one is temporal error concealment which concentrates on the correlations between neighboring frames. In this paper, we just focus on spatial error concealment.

Many spatial error concealment methods have been proposed. Maximally smooth recovery is a typical one which exploits the spatial correlations to recover the lost coefficients in transform domain by making use of the smoothness property of image [1][2]. In [3], the authors utilize project-onto-convex-set (POCS) to estimate the edge information and then use directional filter to recover the lost macroblocks (MBs). However, it is too complex for video services. In order to reduce the complexity, Kwok and Sun proposed a multi-directional interpolation (MDI) algorithm

in which interpolation in pixel domain along eight possible directions is performed and both single-edge and multi-edge are considered [4]. In the latest international video coding standard (H.264/AVC) reference software, the adopted spatial error concealment method is weighted pixel average (WPA), in which the weighting factor is determined by the inverse distance between the source and destination pixels [5]. Recently, for more accurate interpolation, Zhao and Tian proposed to recover the lost MBs 4x4 block by 4x4 block using directional decision and intra prediction [6].

Most of the approaches utilize the smoothness property to recover the lost blocks based on some assumptions, which will bring on annoying blurs artifacts when the lost block is “texture-like” rather than “smooth-like”. In [7], the authors introduce a method which recovers the lost blocks pixel by pixel by combining edge-preserve interpolation and texture synthesis. It obtains fairly good results. However, it is too complex for real-time applications. In this paper, we propose a simple spatial error concealment algorithm which can adaptively switch between directional interpolation and region matching based on the decoded DCT coefficients. In the proposed algorithm, the available DCT coefficients are stored while decoding, and if there are some blocks lost, the available DCT coefficients in neighboring MBs are utilized to analyze whether the lost MB is “smooth-like” or “texture-like”. For “smooth-like” MB, directional interpolation is adopted to recover the lost MB. For “texture-like” MB, region matching is used to find similar blocks in the spatial domain 8x8 block by 8x8 block, and the total mismatch distortion is computed. If the total mismatch distortion is smaller than the preset threshold, the searched blocks are the final reconstructed blocks. Otherwise it means that there is no similar texture in the spatial domain, directional interpolation is utilized to recover the lost MB. Experimental results show that our proposed algorithm can not only improve visual quality but also PSNR performance.

The rest of the paper is organized as follows. In section 2, directional interpolation and region matching are first introduced. Then we analyze the correlation between DCT coefficients and character of the MB and propose a novel spatial error concealment which can adaptively choose directional interpolation and region matching. Experimental results are shown to demonstrate the advantage of the

proposed method in section 3. Finally, we conclude this paper in section 4.

2. ADAPTIVELY SWITCHING BETWEEN DIRECTIONAL INTERPOLATION AND REGION MATCHING

2.1. Directional Interpolation

The directional interpolation algorithm tries to exploit spatial correlations from the available neighboring MBs to enforce the smooth property to recover the lost MBs. First the most possible edge direction of the lost MB is estimated by using the gradient measures to four layers of pixels of the available neighboring MBs as shown in Fig. 1. The local edge gradient is computed by applying Sobel operators to the target pixel.

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (1)$$

For the pixel $x(i, j)$, the gradient components are defined as follows,

$$g_x = x_{i+1,j-1} + 2x_{i+1,j} + x_{i+1,j+1} - x_{i-1,j-1} - 2x_{i-1,j} - x_{i-1,j+1} \quad (2)$$

$$g_y = x_{i-1,j-1} + 2x_{i,j-1} + x_{i+1,j-1} - x_{i-1,j+1} - 2x_{i,j+1} - x_{i+1,j+1}$$

The magnitude and direction of the gradient are given by,

$$|G(x, y)| = \sqrt{g_x^2 + g_y^2}; \theta(x, y) = \arctan\left(\frac{g_y}{g_x}\right) \quad (3)$$

For each target pixel shown in Fig. 1, the magnitude and the direction are computed. Then the direction will be classified to one of the eight directions as shown in Fig. 2. And the corresponding counter will increase with the gradient magnitude. If all the counters are smaller than a threshold, it means that the MB is a flat MB, WPA is enough to reconstruct the lost MB. Or else the direction with the largest counter result which characterizes the strongest edge will be selected for directional interpolation. Then the weighted interpolation in the selected direction is taken on each pixel in the lost MB. As shown in Fig. 3, according to the selected direction, $Y(x, y)$ should be interpolated using $Y_1(i, j)$ and $Y_2(i, j)$. However, if the neighboring MB predMB7 is lost and not yet concealed, then predMB0 is utilized, and $Y(x, y)$ is interpolated using $Y_1(i, j)$ and $Y_3(i, j)$. The weighting factor is considered as the inverse distance between the estimated pixel $Y(i, j)$ and the target pixel $Y(x, y)$.

$$Y(x, y) = \frac{\frac{1}{D(Y, Y_1)} Y_1(i, j) + \frac{1}{D(Y, Y_3)} Y_3(i, j)}{\frac{1}{D(Y, Y_1)} + \frac{1}{D(Y, Y_3)}} \quad (4)$$

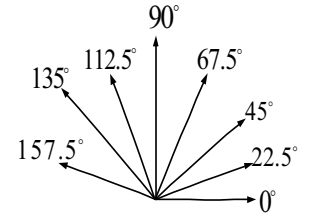
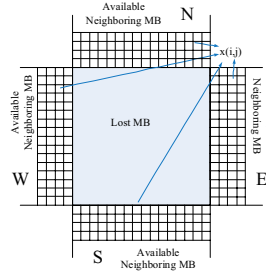


Fig. 1 Pixels to make edge decision. Fig. 2. Eight edge directions

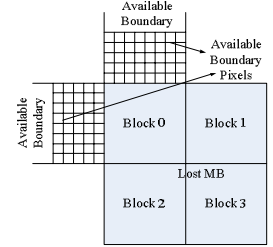
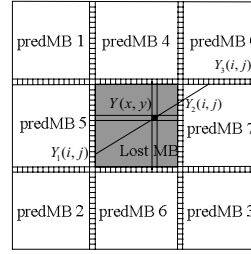


Fig. 3. Directional interpolation. Fig. 4. Region matching

2.2. Region Matching

Region matching algorithm is a variation of the boundary matching algorithm which is adopted as the temporal error concealment method in H.264. The basic concept of region matching algorithm is to search the most similar region in the searching range using known boundary pixels. As shown in Fig. 4, the lost MB is recovered 8x8 block by 8x8 block. For each 8x8 block, five layers of the available boundary pixels are used to search the most similar block in the current frame. The most similar block is defined as follows,

$$R(\tilde{q}) = \arg \min_{R(q) \in \Phi} Distortion(R(q), R(p)) \quad (5)$$

Here $R(p)$ indicates the block to be matched, $R(q)$ represents the candidate matching block, $R(\tilde{q})$ is the resulted matching block, and Φ is the searching range. In the formula, the distortion function is defined as the sum of absolute differences between the available boundary pixels of the matched block and the matching block.

The total distortion of finding the four most similar blocks is defined as the sum of the mismatch distortion of each 8x8 block.

$$Distortion_{total} = \sum_{i=0}^3 Distortion(R_i(\tilde{q}), R_i(p)) \quad (6)$$

If $Distortion_{total}$ is too large, it means that region matching can not find a very similar MB from the searching range.

2.3. The Proposed Algorithm

Let $x(i, j)$ and $x'(i, j)$ denote the expected pixel in the lost MB and the pixel reconstructed by directional interpolation method respectively. We can define an energy function to

measure how well the directional interpolation performs. The energy function is the sum of square differences between $x(i, j)$ and $x'(i, j)$ in the lost MB as follows:

$$E = \sum_j \sum_i (x(i, j) - x'(i, j))^2 \quad (7)$$

When the energy is small, it tells us that the directional interpolation performs well. But if the energy is large, it may tell that the directional interpolation fails to well-reconstruct the lost MB. More interestingly, this energy function is similar to the residue energy function in the predictive coding scheme.

$$E_{residue} = \sum_j \sum_i (x(i, j) - x_{pred}(i, j))^2 \quad (8)$$

Since $x'(i, j)$ is near to $x_{pred}(i, j)$ in intra-prediction, the energy function is near to the residue energy. According to Parsvel's formula, the energy in pixel domain is equal to the energy in frequency domain. So,

$$E \cong E_{residue} = \sum_v \sum_u F^2(u, v) \quad (9)$$

Here, $F(u, v)$ is the DCT coefficient in frequency domain. That is to say, we can use the DCT coefficients to detect whether the directional interpolation performs well or not. Since the information of the target MB is totally lost, the coefficients of the available neighboring MBs are utilized. In the proposed algorithm, we use the sum of available DCT coefficients of south, north, east and west neighboring MBs, as shown in Fig. 1.

$$\begin{aligned} E_{total} &= \sum_{k \in \{S, N, E, W\}} E_{residue}(k) \\ &= \sum_{k \in \{S, N, E, W\}} \sum_v \sum_u F_k^2(u, v) \end{aligned} \quad (10)$$

So, the procedures of the proposed algorithm are described as follows:

- 1) First, for each lost MB, E_{total} is computed.
- 2) If $E_{total} \leq T_1$, the directional interpolation is enabled to recover the lost MB.
- 3) Otherwise region matching algorithm is adopted.
 - i. If $Distortion_{total} \leq T_2$, the MB recovered by the region matching algorithm is the final result.
 - ii. Else it means that the region matching algorithm performs badly and directional interpolation is utilized to recover the lost MB.
- 4) Repeat step 1), 2), 3) until all the lost MBs are recovered.

3. EXPERIMENTAL RESULTS

The proposed error concealment algorithm is evaluated based on the H.264 codec. The JM9.0 reference software is used in the experiment. We compare the performance of the proposed algorithm with three algorithms: weighted pixels averaging (WPA), directional interpolation-only (DI) and region matching-only (RM). The foreman sequence in QCIF format is encoded at 30HZ frame rate. In order to better compare the performance of the proposed algorithm with

the other three algorithms which all focus on intra frame concealment, all frames are encoded as I frame. Slice mode is enabled. Each slice contains a row of MBs. The quantization parameters at 20, 24, 28, 32 and 36 are tested. For each frame, a number of slices are randomly dropped to simulate the transmission errors according to the error pattern. The packet loss rates at 1%, 2%, 5% and 10% are tested in the experiment. And the corrupted frames are concealed by the proposed algorithm, WPA, DI and RM, respectively. The searching range for RM and the proposed algorithm is 64x64. The experimental results are shown in Fig. 5, 6, 7 and Table 1.

As shown in Fig. 5, the foreman sequence in QCIF format is encoded at different QP while the packet loss rate is set to be 2%. The average PSNR of the corrupted frames, which are concealed by different methods, are presented. We can see that our algorithm can always perform better. In Fig. 6, the PSNR performances at different packet loss rates are shown with the QP fixed at 28. The proposed algorithm still achieves better results all the time.

The visual quality of the reconstructed frames is also evaluated, as shown in Fig. 7, where (a) and (b) are the original and error-mask video frames, respectively. Fig 7(c), (d), (e), (f) show the results concealed by WPA, DI, RM and the proposed algorithm, respectively. The reason DI can not reconstruct the background edge is that DI only considers edge direction from up, down, left and right blocks. It can be seen that the proposed algorithm well-combines the advantages of the directional interpolation and region matching and the concealed frame is more smooth and has less artifacts than the other three algorithms.

In table 1, the average time for decoding and concealing one frame at different packet loss rates are shown with the QP fixed at 28. The speed is evaluated on 1.86G Pentium-M Notebook. We can see that even when the lost rate is 10%, the average time for decoding and concealing one frame is acceptable. We should notice that no fast motion estimation but full search is adopted in the experiment. And no optimization is used.

4. CONCLUSIONS

In this paper, a novel spatial error concealment, which can adaptively switch between directional interpolation and region matching, is proposed. The algorithm utilizes the DCT coefficients to automatically analyze whether the lost MB is "smooth-like" or "texture-like" and adaptively selects directional interpolation or region matching to conceal the corrupted frame. Since both the smooth property and texture information in spatial domain are exploited, the algorithm performs well and can recover both "smooth-like" and "texture-like" MBs. Compared with WPA, DI-only or RM-only, the simulate results demonstrate that the proposed algorithm can improve both the visual quality and the PSNR performance of the reconstructed video.

5. ACKNOWLEDGEMENT

This work has been supported by the Innovation and Technology Commission of the Hong Kong Special Administrative Region, China (project no. ITS/122/03 and project no. GHP/033/05).

6. REFERENCES

- [1] Y. Wang, Q. Zhu, "Error Control and Concealment for Video Communication: A Review", *Proceedings of the IEEE*, vol. 86, no. 5, pp. 974-997, May 1998
- [2] Y. Wang, Q. -F. Zhu, and L. Shaw, "Maximally smooth image recovery in transform coding," *IEEE Trans. Commun.*, vol. 41, pp. 1544-1551, Oct. 1993.
- [3] Huifang Sun and Wilson Kwok, "Concealment of Damaged Block Transform Coded Images Using Projections onto Convex Sets," *IEEE Trans. Image Processing*, pp470-477, Apr. 1995.
- [4] Wilson Kwok and Huifang Sun, "Multidirectional interpolation for spatial error concealment," *IEEE Trans. Consumer Electron.*, vol. 39, pp455-460, Aug. 1993.
- [5] Y. K. Wang, M. M. Hannuksela, V. Varsa, A. Hourunranta, and M. Gabbouj, "The error concealment feature in the H.26L test model," *Proceeding of ICIP*, pp. 729-732, 2002.
- [6] Y. Zhao, D. Tian, "Spatial Error Concealment Based on Directional Decision and Intra Prediction", *Proceeding of ISCAS*, pp. 2899 -2902, 2005.
- [7] S. Belfiore, M. Grangetto, E. Magli, G. Olmo, "An edge and texture preserving algorithm for video error concealment", *IEEE Workshop on Multimedia Signal Processing*, pp. 121-124, 2002.

TABLE I AVERAGE CONCEALING TIME OF ALL FRAMES AT DIFFERENT PACKET LOSS RATES.

Foreman QP = 28	Packet Loss Rate			
	1%	2%	5%	10%
WPA (ms)	15.6	16.2	16.8	17.8
DI (ms)	16	16.8	17.8	19.2
RM (ms)	25.6	44.2	49.6	75
Proposed (ms)	25	43.2	48	72

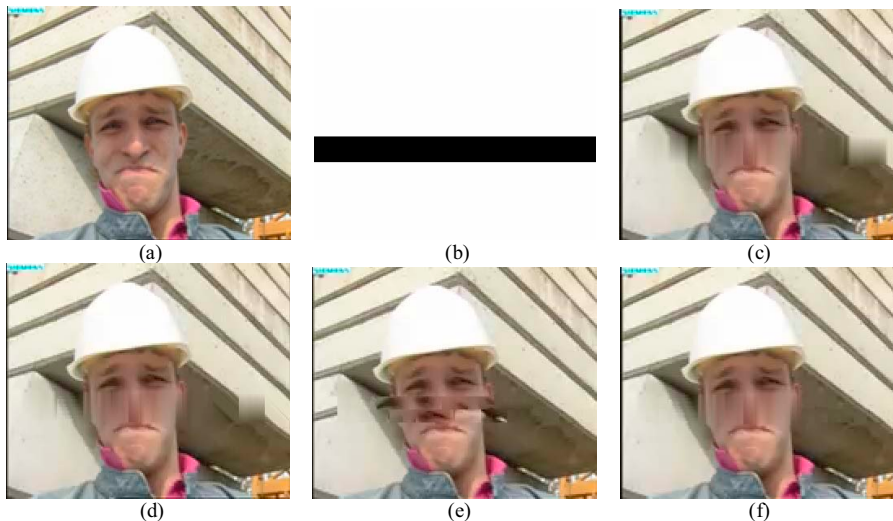


Fig. 7 The 38th frame of Foreman with random slices dropped for 2% packet loss rate while the QP is fixed at 28. (a) original frame; (b) error-mask frame; (c) concealed with WPA. (d) concealed with DI. (e) concealed with RM. (f) concealed with the proposed algorithm.

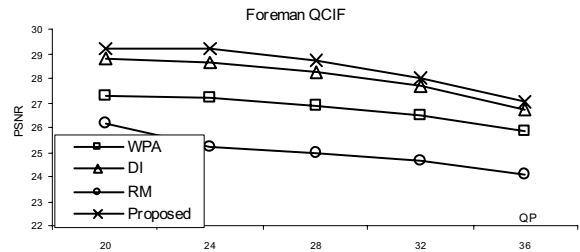


Fig. 5. Average PSNR of all corrupted frames at different QP when the packet loss rate is set to be 2%.

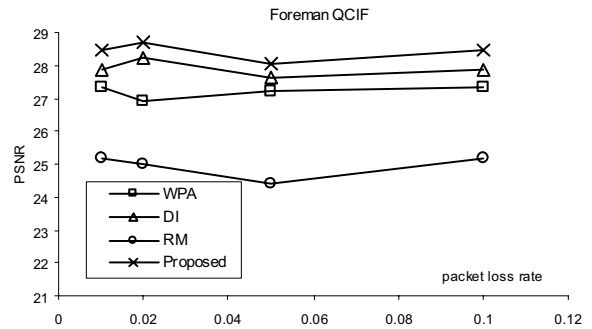


Fig. 6. Average PSNR of all corrupted frames at different packet loss rate when the QP is fixed at 28.