Spatio-temporal Video Error Concealment using Priority-ranked Region-matching

Yan Chen*[1], Xiaoyan Sun[2], Feng Wu[2], Zhengkai Liu[1] and Shipeng Li[2]
[1] University of Science and Technology of China, Hefei 230027
[2] Microsoft Research Asia, Beijing 100080

Abstract— When transmitted over error-prone networks, compressed video sequences may be received with errors. In this paper we propose a priority-ranked region-matching algorithm to recover the "lost" area of the decoded frames, in which both temporal and spatial correlations of the video sequence are exploited. In the proposed scheme, we first calculate the priorities of all edge pixels of the "lost" area and generate a priority-ranked region group. Then according to their priorities, the regions in the group will search their best matching regions temporally and spatially. Finally, the "lost" area is recovered progressively by the corresponding pixels in the matching regions. Experimental results show that the proposed scheme achieves higher PSNR as well as better video quality in comparison with the method adopted in H.264.

Keywords-error concealment; region matching; video concealment; block matching; spatio-temporal concealment

I. Introduction

With the explosive growth and great success of the Internet as well as the wireless network, video services over narrowband networks such as internet and mobile are becoming more and more popular. However, when transmitted over error-prone networks, video sequences may be received with "lost" macroblocks (MBs) caused by cell missing and channel errors due to which the quality of the obtained video declines. Especially, compressed video streams are extremely vulnerable to transmission errors because of the use of predictive coding and variable length coding (VLC) entropy coding schemes. In case of utilization of spatio-temporal prediction, one error not only corrupts the current decoded frame but also may propagate to succeeding frames.

Many error concealment methods have been proposed to recover the artifacts caused by transmission errors at decoder by making use of inherent correlations among spatially and/or temporally adjacent blocks [1][2][4]. Maximally smooth recovery is a typical spatial approach that exploits the neighboring information to recover the "lost" MBs by utilizing the smoothness property of image [1][3]. Moreover, content adaptable method is also proposed for error concealment [5]. One the other hand, temporal schemes are preferred for video signals to replace the damaged MBs by using temporal correlations between sequential frames. Many methods have been presented to replace the "lost" MBs by the motion compensated MBs in the previous frame with zero motion vector (MV), average or median MV of the spatially adjacent blocks [6]. Block matching and neighbor matching principles are often utilized in temporal error concealment [7][8]. A well-known algorithm is boundary-matching-based motion vector recovery, which is adopted in H.264 and described in detail in [9]. Later, Zheng et al proposed to recover the lost motion vectors by Lagrange interpolation formula [10].

More recently, hybrid algorithms have been proposed to take advantages of both spatial and temporal correlations to better recover the "lost" MBs. In [11], a concealment scheme is presented using a temporal replacement of the lost block through block matching method followed by a mesh-based transformation to fill the block content with the correctly received surrounding area. Later, a spatio-temporal interpolation for error concealment is proposed in which the edges interrupted by the "lost" area are mapped to the reference frame by contour matching together with the snake-based edge mapping [12].

In this paper, a novel error concealment method is proposed in which both spatial and temporal correlations are exploited based on the priority-ranked region-matching algorithm. The region-filling algorithm proposed in [13] is utilized in the proposed scheme to select a series of key points from the edges of the "lost" area and correspondingly results in a series of key regions centered at these points. Then, the best matching regions of the key regions are searched both in current frame and previous frame. The pixels to be concealed in the key regions are recovered with the corresponding pixels inside those matching regions. Finally, we update the edge and repeat the aforementioned processes until all pixels in the "lost" MBs are concealed. Experimental results show that the proposed scheme can recover more details and leave less block artifacts compared with the error concealment method used in H.264.

The rest of this paper is organized as follows. The proposed algorithm is described in detail in Section 2. Section 3 shows the experimental results. Finally, section 4 concludes this paper.

*This work has been done while the author is with Microsoft Research Asia.

II. CONCEALMENT ALGORITHM BASED ON PRIORITY-RANKED REGION-MATCHING

In this paper, we assume that the "lost" MBs have been detected and the corresponding MB-based status map of each frame is available for decoding. According to the status map, all correctly received MBs are decoded first and then the "lost" MBs are concealed by the proposed error concealment algorithm. Different from common blocking-matching-based error concealment methods, the error concealment scheme present in this paper enables multiple regions mapping and partial MB recovery at one concealment step. Thus, by fully making use of the correlations both inside frame and between frames, each "lost" MB will be filled partially and priority-progressively in the proposed scheme. It is introduced in detail in the following subsection.

A. The Priorities of Boundary Pixels

In the proposed error concealment scheme, the "lost" MBs are restored by first reconstructing the significant pixels and then the others. Whether a pixel is significant or not is determined by the priority algorithm presented in [13] for region-filling. Given an error MB, the priority of each edge pixel is computed. In detail, for one edge pixel p, a NxN region R_P is first selected centered at pixel P(x,y). Then the priority of P(x,y) is calculated considering two aspects, the percentage of the "known" pixels in region R_P and the gradient and normal of the edge line. Since the pixel P(x,y) belongs to the edge of the "lost" MB, it is hard to calculate its gradient $\nabla P(x,y)$ directly. Thus, in our method, we use the maximum one of its 8-adjacent gradients to approach the gradient $\nabla P(x,y)$ of P(x,y), that is

$$\nabla P(x,y) = \max\left(\nabla P(m,n)\right), \quad \text{where } x-1 \le m \le x+1, y-1 \le n \le y+1. \tag{1}$$

And $\nabla P(m,n)$ is defined as follows:

$$\nabla P(m,n) = \mathrm{Re}\left(\nabla P(m,n)\right) + \mathrm{Im}\left(\nabla P(m,n)\right) j$$

If all the five points P(m+1,n), P(m-1,n), P(m,n), P(m,n+1), P(m,n-1) are "known" pixels, then

$$\mathrm{Re}\left(\nabla P(m,n)\right) = P(m+1,n) + P(m-1,n) - 2\times P(m,n),$$

$$\mathrm{Im}\left(\nabla P(m,n)\right) = P(m,n+1) + P(m,n-1) - 2\times P(m,n),$$

else

$$\mathrm{Re}\left(\nabla P(m,n)\right) = 0,$$
$$\mathrm{Im}\left(\nabla P(m,n)\right) = 0.$$

B. The Priorities of Regions To Be Matched

As mentioned above, the "lost" MBs are recovered progressively. On the other words, the later filled blocks may utilize the former ones in concealing due to which the side effect of the false recovered pixels may propagate to other pixels. It is apparent that the filling order will play a very important role in the concealment.

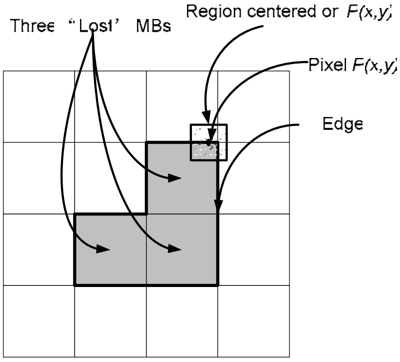

Fig.1 "lost" MBs in one frame

Fig.1 shows a frame with some "lost" MBs denoted by grey blocks. The dark line tells the edge of the "lost" area. Given the priorities of the boundary pixels, a straightforward method is to reconstruct the "lost" area once a region according to priority order of the boundary pixels. However, this method is quite sensitive to false recovered regions. In the proposed scheme, we will rank the priorities of the pixels and select multiple regions to be concealed simultaneously. The rule of the regions selection is described as follows.

First, the priorities of all n pixels belonging to the edge are calculated and a series of priorities $(P_0, P_1, \ldots, P_{n-1})$ is generated. Then, the series of priorities is rearranged in descending order of magnitude based on which a subset of region priorities $(R_0, R_1, \ldots, R_{k-1})$ is presented with the following two terms:

1) Each two pixels, R_i and R_j, can not be in a same NxN region, where $0<i<k$, $0<j<k$, $i\neq j$.
2) The priorities $R_0 > R_1 > \ldots > R_{k-1}$ are the highest ones we can select.

Then, there are still many insignificant regions whose priorities are nearly equal to zero, which might cause false recoveries. Such errors would propagate to other pixels. To remove these trivial priorities, two adaptive thresholds h_1 and h_2 are adopted to filter the region priority group. Here,

$$h_1 = R_0/k, \tag{2}$$

$$h_2 = \left(\sum_{0\le i<m} R_i\right)/m. \tag{3}$$

Accordingly, we firstly get rid of the priorities which are smaller than h_1 from the region priority group and result in a new priority set $R_0 > R_1 > \ldots > R_{m-1}$, where $R_{m-1} \ge h_1$. Secondly, the average value of the new priority group,

h_2, is utilized to generate the ranked-priority region group G(R) with priorities $R_0 > R_1 > \ldots > R_{s-1}$, where $R_{s-1} \geq h_2$.

C. Ranked-Priority Region Matching

For each NxN region in the ranked-priority region group, the region matching is performed both temporally and spatially. In other words, the most similar region is searched from current frame and the previous frame. Here the most similar region is defined as follows:

$$\psi(\hat{q}) = \arg\min_{\psi(q)\in\varphi} D(\psi(q), \psi(p)) \quad (4)$$

Here, $\psi(p)$ indicates the region to be matched, $\psi(q)$ represents the candidate matching region, $\psi(\hat{q})$ is the resulted matching region of $\psi(q)$ and φ is the searching range. In formula (4), $D(\psi(q), \psi(p))$ is defined as the sum of square differences of the "known" pixels in the two regions. Notice that all pixels should be "known" pixels in case of the spatial candidate regions.

D. Region recovery

Having found the most similar region, the "lost" pixels of each matched region $\psi(p)$ are recovered with the pixels inside $\psi(\hat{q})$ at corresponding positions. All "known" pixels (including the previously filled ones) are unchanged.

E. Update and Repeat

After all the priority-ranked regions are recovered, the "lost" area and its edge are updated. Then all the four steps mentioned above are repeated until the entire "lost" area is concealed.

III. EXPERIMENTAL RESULTS

The proposed error concealment scheme is evaluated based on the H.264 codec. The performance of the proposed method is compared with the error concealment method adopted in H.264 [2][8][9]. The JM 8.2 software is used in this experiment. The sequence Foreman in CIF format is encoded at 30Hz. In this experiment, I frame is coded every ten frames and no B frame is used. The quantization parameter is set to be 28.

In the proposed scheme, the search range of region mapping in current frame is 64, while the search range of temporal region mapping is 32.

From the third frame, a number of MBs in every frame of the video sequence are dropped according to the error pattern of 20% loss rate, which are assumed to be caused by transmission errors. Then, the corrupted frame, which is the reference of succeeding frame, is concealed by the proposed algorithm as well as the H.264 method respectively. The experimental results are shown in Fig.2, Fig.3 and Fig.4.

As shown in Fig.2, the PSNR of each recovered frame of the proposed scheme and the H.264 method are presented. Our scheme can always provide higher PSNR compared with the H.264 method.

The visual quality of the reconstructed frames is also evaluated, as shown in Fig.3 and Fig.4, where (a) and (b) are the undamaged and damaged video frames, respectively. Fig. 3-(c) and Fig. 4-(c) show the images reconstructed using weighted pixel value averaging for intra pictures [2] and boundary-matching-based motion vector recovery for inter pictures [8] respectively, which are adopted in H.264. Fig.3-(d) and Fig.4-(d) show the results generated by the proposed algorithm.

It can be seen that the concealed frames provided by the proposed scheme are more smooth and with more details and less artifacts. The visual quality of the proposed scheme is much better than that of the scheme adopted in H.264.

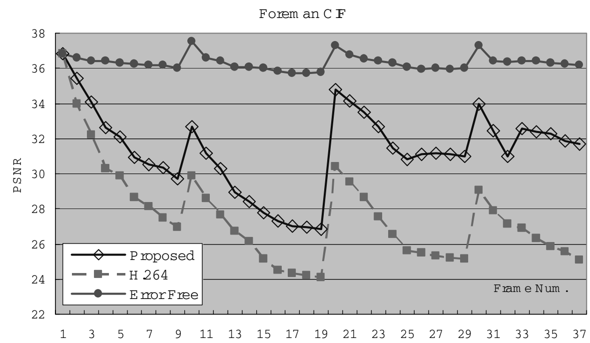

Fig.2 "Foreman" sequence PSNR performance comparison vs. frame number at 20% loss rate

IV. CONCLUSIONS

In this paper, a spatio-temporal error concealment algorithm based on priority-ranked region-matching is proposed. Rather than recovering the MBs in "lost" area one by one, the proposed scheme is able to conceal the "lost" MBs from different directions according to the priorities of the edge pixels. It means that multiple regions can be recovered once in the proposed method. Thus, the block artifacts are effectively reduced and more boundary information can be restored. Moreover, the region mapping is performed both inside current frame and in previous frame. Therefore, more details and less blurs are presented. Compared with the error concealment method adopted in H.264, the simulation results demonstrate that the algorithm we propose can improve both visual quality and PSNR performance of the recovered video sequences.

It should be noted that the proposed algorithm can be used not only for H.264, but also for other video compression methods based on motion compensation.

In this paper, only current frame and previous frame are used in the proposed region-matching. For better reconstructed results, more reference frames can be utilized to achieve even better performance.

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] P.Salama, N.B.Shroff, and E.J.Delp, "Error concealment in encoded video," Image Recovery Techniques for Image Compression Applications. Norwell, MA: Kluwer, 1998.

[3] 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.

[4] S.C. Hsia, "An edge-oriented spatial interpolation for consecutive block error concealment," IEEE Signal Processing Letters, vol. 11, pp. 577-580, June 2004.

[5] R.F. Zhang, Y.H. Zhou, X.D. Huang, "Content-Adaptive spatial error concealment for video communication," IEEE Trans. On Consumer Electronics, vol. 50, pp. 335-341, Feb. 2004.

[6] P. Haskell and D. Messerschmitt, "Resynchronization of motion compensated video affected by ATM cell loss," Proc. ICASSP'92, San Francisco, CA, Mar. 1992, pp. III545-548.

[7] S. Tsekeridou and I. pitas, "MPEG-2 error concealment based on block-matching principles" IEEE Trans. Circuits and System for Video Technology, Vol. 10, pp. 646-658, June 2000.

[8] W. M. Lam, A. R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors," Proc. ICASSP, vol.5, Mar. 1993, pp. 417-420.

[9] Y.-K. Wang, M. M. Hannuksela, V. Varsa, A. Hourunranta, And M. Gabbouj, "The error concealment feature in the H.26L test model," Proc. ICIP, vol. 2, pp. 729-732, Sept. 2002.

[10] J.H. Zheng, L.P. Chau, "A temporal error concealment algorithm for H.264 using Lagrange interpolation," ISCAS'04, vol. 2, pp. 23-26, May 2004.

[11] L. Atzori, G.B. De Natale Francesco, C. Perra, "A spatio-temporal concealment technique using boundary matching algorithm and mesh-based warping (BMA-MBW)," IEEE Trans. On Multimedia, vol. 3, pp. 326-338, Sep. 2001.

[12] F. De Natale and A. Matera, "Video Error Concealment using Spatio-Temporal Interpolation with Snakes." International Symposium on Control, Communications and Signal Processing, pp. 83-86, March, 2004

[13] A. Criminisi, P. Pérez, and K. Toyama, "Object Removal by Exemplar-Based Inpainting," IEEE Proc. CVPR, 2:721-728, 2003

(a) (b) (c) (d)

Fig.3 The 30th frame of Foreman at 20% loss rate (a) error-free frame; (b) error-damaged frame; (c) concealed with the method adopted in H.264; (d) concealed with the proposed algorithm.

(a) (b) (c) (d)

Fig.4 The 3th frame of Foreman at 20% loss rate (a) error-free frame; (b) error-damaged frame; (c) concealed with the method adopted in H.264; (d) concealed with the proposed algorithm