EDGE WEIGHTED SPATIO-TEMPORAL SEARCH FOR ERROR CONCEALMENT

E. S. Ng, J. Y. Tham and S. Rahardja

Media Processing Department, Institute for Infocomm Research
Agency for Science, Technology and Research (A*STAR), Singapore
esng@i2r.a-star.edu.sg, jytham@i2r.a-star.edu.sg, rsusanto@i2r.a-star.edu.sg

ABSTRACT

In temporal error concealment (EC), the sum of absolute difference (SAD) is commonly used to identify the best replacement macroblock. Even though the use of SAD ensures spatial continuity and produces visually good results, it is insufficient to ensure edge alignment. Other distortion criteria based solely on structural alignment may also perform poorly in the absence of strong edges. In this paper, we propose a spatio-temporal EC search algorithm using an edge weighted SAD distortion criterion. This distortion criterion ensures both edge alignment and spatial continuity. We assume the loss of motion information and use zero motion vector as the starting search point. We show that the proposed algorithm outperforms the use of unweighted SAD in general. Most importantly, the perceptual quality of EC is improved due to edge alignment while ensuring spatial continuity.

Index Terms— Error concealment, Edge weighted, H.264, Spatio-temporal

1. INTRODUCTION

Error concealment (EC) is essential in ensuring the quality of transmitted video signals over various networks. When transmitted over error-prone networks, errors can severely degrade the perceptual quality of video signals. This degradation is especially serious due to error propagation in video standards using motion estimation like the H.264/AVC standard [1]. Thus, EC is required to prevent this degradation.

A main category of EC is *temporal error concealment*. Temporal EC uses temporal information from closely correlated video frames to conceal errors. Simple techniques include replacing the lost macroblock (MB) with the co-located MB in the reference frame [1, 2, 3]. The best motion vector can also be selected from a list of candidate motion vectors [2, 4] through the minimisation of a distortion criterion, such as a boundary matching error (BME) measure [1]. The sum of absolute difference (SAD) between the boundary pixels [5] is commonly used in calculating BME measures. Temporal EC also employs motion estimation to select the best motion vector such as the decoder motion vector estimation (DMVE) method in [3]. These techniques tend to be computationally expensive due to the search process.

1.1. Related Works

We were particularly interested in algorithms which combined the use of spatial and temporal information as this integration of information is essential for good EC results. In [1], an enhanced error concealment with mode selection (EECMS) algorithm is proposed that selects temporal EC for both I and P frames adaptively. [6] proposes a spatio-temporal block matching algorithm (BMA) that integrates the weighted BMA temporal information and weighted spatial information measuring the spatial continuity across boundaries. In [7], an algorithm based on structural alignment is proposed which uses edge gradient as the distortion criterion and produces better results as compared to conventional techniques.

1.2. Problem Identification and Contribution

A problem associated with using SAD as the distortion criterion for side matching algorithms is the misalignment of strong edges [7]. Algorithms using structural information such as edges may also not produce good results in video frames with weak edges. A third problem is the dependency of common temporal EC algorithms on motion and coding information. In this paper, we formulate a distortion criterion, the edge weighted SAD, that accounts for both structural and temporal information. A spatio-temporal EC algorithm using this distortion criterion is proposed with several advantages:

- Ensures alignment of edges between the replacement block and its neighbouring blocks while ensuring side matching in areas with weak edges.
- Able to ensure good perceptual quality in the absence of motion information.
- Independent of coding modes when a reference frame with good correlation is available. The algorithm can thus be used for both I and P frames in the H.264/AVC standard [8].

1-4244-1437-7/07/$20.00 ©2007 IEEE

2. PROPOSED EDGE WEIGHTED SPATIO-TEMPORAL SEARCH ALGORITHM

First, we describe the EC in the H.264 reference software [9, 10] briefly for comparison with our proposed algorithm.

2.1. Error Concealment in JM 12.0

For a lost MB, the temporal EC algorithm in the H.264 reference software JM 12.0 [9, 10] compares the availability of motion vectors from neighbouring macroblocks (MBs) with a threshold. Above the threshold, the unweighted SAD of each candidate MB is calculated. The MB corresponding to the motion vector minimising the SAD is selected to replace the lost MB. This algorithm depends heavily on the availability of motion vectors and is restricted to P frames. Edges may not be aligned due to the use of unweighted SAD as the distortion criterion. In erroneous I frames, spatial EC is used due to the lack of motion information. The use of bilinear interpolation [9, 10] for spatial EC leads to a blurring effect.

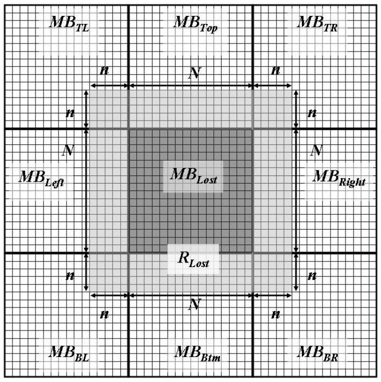

Fig. 1: Locations of neighbouring macroblocks (MBs) around the lost MB, MB_{Lost}, and the region R_{Lost}. R_{Lost} is identified by the light grey box in the middle, which consists of pixel values of the neighbouring MBs and MB_{Lost} as well.

2.2. Splitting the edge and non-edge components

Here, we formulate a distortion criterion that integrates both spatial and temporal information, called the edge weighted SAD, which accounts for the alignment of edges between the lost MB and its neighbours while ensuring smooth matching of the sides in non-edge areas. To elaborate, we define a frame with errors to be F_{Lost}. We consider a lost MB, MB_{Lost}, in F_{Lost}. The neighbouring MBs of MB_{Lost} are shown in Fig. 1. We define a region R_{Lost} consisting n lines of pixels around MB_{Lost}. R_{Lost} will be of size $(N+2n)\times(N+2n)$ with $N\times N$ the size of MB_{Lost}. We then perform edge detection using Sobel operators on R_{Lost} to obtain $G_{Lost}(i,j)$, the gradient of each pixel (i,j) in R_{Lost}. When a neighbouring MB is erroneous, we ignore the MB in the EC process.

Based on $G_{Lost}(i,j)$, a binary map of R_{Lost}, defined as B_{Lost}, is generated with $1s$ indicating pixels corresponding to strong edges, and $0s$ indicating pixels corresponding to weak or non-edges. B_{Lost} is used to generate R_{Lost1} with pixel values corresponding to non-edges in R_{Lost}. Similarly, R_{Lost2} consists of pixel values corresponding to edges in R_{Lost}. This process is illustrated in Fig. 2.

2.3. Spatio-temporal search using edge weighted SAD

A full search in the reference frame F_{Ref}, spanning a search window of $(N+k)\times(N+k)$ centred around MB_{Lost}, is conducted with $k > n$ to identify the best motion vector and replacement MB. We define a candidate MB as MB_x and a region similar to R_{Lost}, called R_x. Using B_{Lost}, we define R_{x1} and R_{x2} by splitting R_x into non-edge and edge pixels respectively. With these parameters, we define the edge weighted SAD distortion criterion, SAD_{EW}, as in (1). Our algorithm compares the external pixels of MB_{Lost} with MB_x, similar to the external boundary matching error (EBME) in [1].

$$\begin{aligned} SAD_{EW} &= \alpha SAD_{Edge} + (1-\alpha) SAD_{Norm} \\ SAD_{Edge} &= \sum_{i,j}^{N+2n} |R_{Lost2}(i,j) - R_{x2}(i,j)| \\ SAD_{Norm} &= \sum_{i,j}^{N+2n} |R_{Lost1}(i,j) - R_{x1}(i,j)|) \qquad (1) \end{aligned}$$

The weight α depends on the strength of edges in R_{Lost}. For $\sum R_{Lost2} > \sum R_{Lost1}$, we weigh SAD_{Edge} favourably using (2). Otherwise, we use (3). The weights change SAD_{EW} adaptively to ensure smooth matching across boundaries while maintaining good edge alignment. The MB that minimises this distortion criterion in the search window is selected. However, since a full search is computationally expensive, we adopt the unrestricted diamond search algorithm in [11] to reduce computational time, while removing the need for a search window.

$$\alpha = 1 - \frac{\sum_{i,j}^{N+2n} R_{Lost1}}{2\sum_{i,j}^{N+2n} R_{Lost2}} \qquad (2)$$

$$\alpha = 1 - \frac{\sum_{i,j}^{N+2n} R_{Lost2}}{2\sum_{i,j}^{N+2n} R_{Lost1}} \qquad (3)$$

We have thus a distortion criterion, SAD_{EW}, that ensures matching of edges and spatial continuity. The proposed spatio-temporal search algorithm is expected to produce concealed frames of better perceptual quality as compared to conventional algorithms. This algorithm considers the loss of motion information and uses zero motion vector as the candidate for both I and P frames when a suitable reference frame is available.

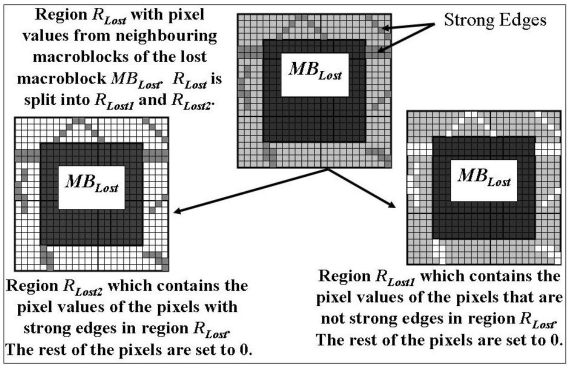

Fig. 2: Process of splitting the pixel values in region R_{Lost} of the lost macroblock MB_{Lost}, based on the binary map B_{Lost} indicating strong edges.

3. SIMULATION RESULTS

We introduced errors of varying error rates by randomly dropping slices from coded sequences using Flexible Macroblock Ordering (FMO) with a dispersed pattern. Simulations are carried out using the baseline profile of JM 12.0 on 2 sets of sequences with $N = 16$ and $n = 4$. To demonstrate the advantages of our algorithm under loss of motion information, we select zero motion vector as the starting search point.

Fig. 3 shows selected frames of our simulations. As highlighted, SAD_{EW} ensures better edge alignment as compared to the unweighted SAD used in JM 12.0. Table 1 shows that our proposed algorithm, in general, outperforms the EC algorithm in JM 12.0, with average improvements of up to $2.36dB$. More importantly, the proposed algorithm produces results of better perceptual quality due to superior edge alignment as seen in Fig. 3(g)-(j). Bilinear interpolation used for erroneous I frames accounts for the larger differences in PSNR values for Set 2 in Table 1.

Lastly, Table 1 shows that the proposed algorithm takes only an average of up to $0.24s$ more when implemented in JM 12.0. As the implemented algorithms are not optimised during simulations, an analysis of computational time using an optimised decoder will be more meaningful.

4. CONCLUSION

We have developed a spatio-temporal EC fast search algorithm that ensures the alignment of edges and spatial continuity for erroneous video frames. We formulated an edge weighted SAD distortion criterion that balances between edge alignment and spatial continuity in EC. Our simulations have shown that the proposed algorithm produces better quantitative results with significant average improvements of up to $2.36dB$ over EC in JM 12.0. More importantly, the proposed algorithm also results in better perceptual quality due to superior edge alignment, thus demonstrating the advantages of the edge weighted SAD.

To further enhance the performance of the proposed algorithm, we can explore an integrated EC algorithm that uses both spatial and temporal EC for I and P frames. Currently, the use of zero candidate motion vector limits the scope of the proposed algorithm. Thus, we can make use of candidate motion vectors as a starting point for searching such that results in sequences with large motion will be further improved.

5. REFERENCES

[1] D. Agrafiotis, D.R. Bull, and C.N. Canagarajah, "Enhanced error concealment with mode selection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, pp. 960–973, Aug 2006.

[2] Y. Wang and Q. Zhu, "Error control and concealment for video communication: A review," *Proc. IEEE*, vol. 86, no. 5, pp. 974–994, May 1998.

[3] J. Zhang, J.F. Arnold, and M.R. Frater, "A cell-loss concealment technique for mpeg-2 coded video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, pp. 659–665, Jun 2000.

[4] D. Kim, S. Yang, and J. Jeong, "A new temporal error concealment method for h.264 using adaptive block sizes," *Proc. ICIP 2005*, vol. 3, pp. 928–931, Sept 2005.

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

[6] Y. Chen, O. Au, C. Ho, and J. Zhou, "Spatio-temporal boundary matching algorithm for temporal error concealment," *Proc. ISCAS 2006*, May 2006.

[7] A.S. Bopardikar, O.I. Hillestad, and A. Perkis, "A temporal error concealment algorithm based on structural alignment for packet video," *Proc. SPCOM 2004*, pp. 373–377, Dec 2004.

[8] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 560–576, July 2003.

[9] "H.264/AVC Reference Software JM 12.0," http://iphome.hhi.de/suehring/tml/.

[10] Y.K. Wang, M.M. Hannuksela, V.V.A. Hourunranta, and M. Gabbouj, "The error concealment feature in the h.26l test model," *Proc. ICIP 2002*, vol. 2, pp. 729–732, 2002.

[11] J.Y. Tham, S. Ranaganath, M. Ranaganath, and A.A. Kassim, "A novel unrestricted center-based diamond search algorithm for block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 369–375, Aug 1998.

(a) (b) PSNR = $31.76dB$ (c) PSNR = $32.32dB$

(d) (e) PSNR = $25.93dB$ (f) PSNR = $30.91dB$

(g) (h) (i) (j)

Fig. 3: Error concealment results. (a) P frame with slice error with 25% error probability. (b) JM 12.0. (c) Proposed algorithm. (d) P frame with slice error with 25% error probability. (e) JM 12.0. (f) Proposed algorithm. (g) Edge misalignment in (b). (h) Improved perceptual quality in (c). (i) Edge misalignment in (e). (j) Improved perceptual quality in (f).

Table 1: Results for 2 sets of video sequences of 100 frames coded differently for comparison. JM is the algorithm in the reference software, EW the proposed algorithm. The differences between the average luminance component PSNR values (dB) and decoding time (s) of EW and JM are shown.

Error	Foreman.cif				Paris.cif				Mobile.cif			
	JM		EW		JM		EW		JM		EW	
(%)	(dB)	(s)	(dB)	(s)	(dB)	(s)	(dB)	(s)	(dB)	(s)	(dB)	(s)
Set 1. Results of sequences with the first frame being coded as I frame and the rest as P frames												
10	30.19	18.86	-0.08	+0.61	35.58	17.05	+1.32	+0.17	23.07	21.44	+0.16	-0.17
15	28.35	19.37	+0.42	-0.68	36.08	16.94	+0.97	+0.22	21.81	21.75	-0.12	-0.17
20	26.99	19.53	+0.85	-0.45	33.95	16.53	+1.51	+0.81	19.28	21.70	+0.83	-0.79
25	27.60	18.95	+0.72	-0.12	33.62	16.47	+1.86	+1.18	20.08	21.81	+0.34	-0.56
30	25.58	19.88	+0.94	0.98	32.67	16.61	+1.49	+2.45	18.99	22.88	-0.09	-0.88
Set 2. Results of sequences with an I frame inserted every 10 frames and the rest as P frames												
10	30.01	18.72	+1.86	-0.11	25.43	17.99	+6.17	+0.53	22.93	20.95	+1.95	-0.39
15	28.45	19.60	+1.75	-0.58	26.18	18.06	+3.62	+0.39	22.66	21.45	+0.65	-0.06
20	28.63	19.44	+2.02	-0.20	26.16	17.72	+3.04	+0.44	21.86	21.26	+0.94	-0.43
25	27.22	19.02	+2.73	+0.18	24.56	18.16	+4.68	+0.81	21.40	21.66	+0.99	-0.40
30	26.98	19.88	+1.50	+1.07	24.81	17.99	+3.01	+1.45	21.05	22.88	+0.44	+0.84

Set 1. Average difference in PSNR = +0.74	Set 1. Average difference in time = +0.17
Set 2. Average difference in PSNR = +2.36	Set 2. Average difference in time = +0.24