Spatio-Temporal Error Concealment Scheme Using Hybrid Algorithm with Postprocessing

Dmitriy Kulikov
Department of Computational Mathematics and Cybernetics
Moscow State University, Moscow, Russia
dkulikov@graphics.cs.msu.ru

Abstract

In video streaming there is a problem of packet-loss during signal transmission. Because of this fact the quality of the video could be highly distorted. Different error concealment schemes are used at the decoder side or even after decoding process to fight this. In this paper[1] a new spatio-temporal error concealment scheme with low computational complexity and high visual and objective quality is proposed.

The algorithm exploits temporal information for finding the best candidate macroblock (MB) from previous frame to replace lost MB in current frame. After this step concealment estimation is performed. If concealment led to high discontinuities along the border of the restored block another temporal concealment algorithm is used or even spatial one if temporal concealment couldn't give high-quality result.

The algorithm was tested on 10 videos with different characteristics and results were compared to famous error concealment algorithms that work mainly in temporal domain.

Keywords: *spatio-temporal, temporal, error, concealment, video.*

1. INTRODUCTION

The progress in computational capabilities of the personal computers and new video coding standards lead to an increase in video content usage. Transmission of multimedia signal becomes more important but almost all of transmission networks are error-prone environment and there are often problems with packet-loss. Compressed video is very sensitive to any errors during transmission, especially if variable-length coding is used, because in this case errors propagate along the video frame producing visual artifacts. There are several methods to minimize the effects of damaged and lost packets [1]. There are different source based techniques for error robustness and recovery.

Postprocessing methods called error concealment try to minimize visual artifacts caused by transmission and decoding errors. These methods can be classified into spatial and temporal error concealment methods.

Spatial methods use information from neighbor known macro-blocks (MB) to restore the information in lost MB. Typically these methods are different types of interpolation, and because of it often produce blurring effects.

Temporal error concealment methods use previous (or other correctly received) frame to conceal errors in current frame using strong correlation between consequent frames in the video [1], [2], [3], [4], [6], [7]. Some of them [4], [7] use boundary matching algorithms that consider neighborhood of lost MB and try to find the similar by L_1 or L_2 norm neighborhood in previous frame. The method in [6] uses structure alignment algorithm (SAA), which considers the magnitude of gradients along the border of MB being processed and candidate MB in previous frame.

In this paper new spatio-temporal algorithm is presented for error concealment based on correlation between neighborhood area of the lost block and best candidate from previous frame. Based on the fact that human visual system is very sensitive to edges and structures [10] proposed algorithm tries to keep existing edges and does not produce new such block boundary called blocking effect.

2. PROPOSED ALGORITHM

The proposed algorithm scheme is shown at Figure 1. The corrupted frame and the mask for errors are going to the hybrid concealment method, described in section 2.1. After its work the temporal candidate MB is produced, that minimizes the error function. This candidate is used in the next function block that estimates quality of processing (this block is described in section 2.2) by calculating boundary errors for candidate block and comparing it to corresponded boundary errors for neighbors blocks. If estimated error for candidate block exceeds the threshold another temporal concealment method is used. It typically produces less boundary error. Then boundary error estimation is performed again, and if in this case error also exceeds threshold it means that temporal concealment could not produce high-quality result and spatial concealment is used for this block.

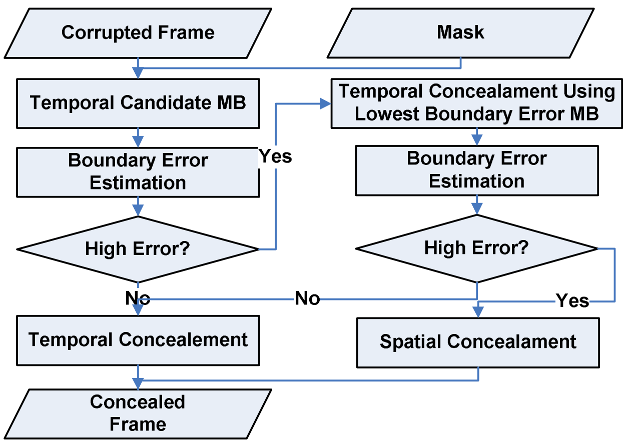


Figure 1: **Block scheme of proposed algorithm**

[1] This work was supported by RFBR grant № 07-01-00759-a.

2.1 Hybrid Temporal Concealment Algorithm

Temporal concealment algorithm tries to take into account not only information of neighborhood, but also structures of the neighbors MB and how good it connects to known information across the lost MB. This algorithm performs estimation of candidate MB that can be taken with help of one of motion vectors for neighbor or performing full search of all blocks in the previous frame.

First part of hybrid algorithm is outer boundary matching algorithm (OBMA or DMVE) that is described in [4]. This part of algorithm takes into account only similarity of the known neighborhood by L_1 or L_2 norm of pixels in current and reference (often previous) frame. The main idea of this algorithm is shown at Figure 2.

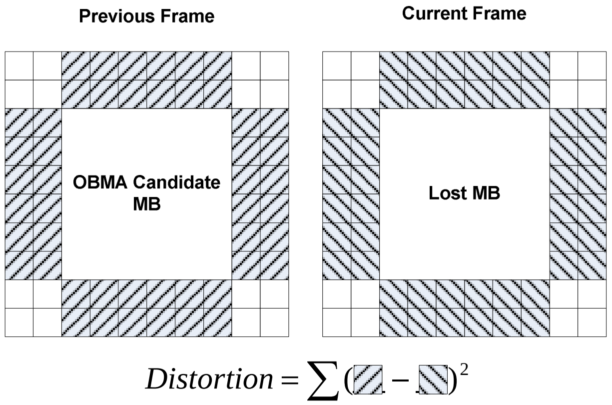
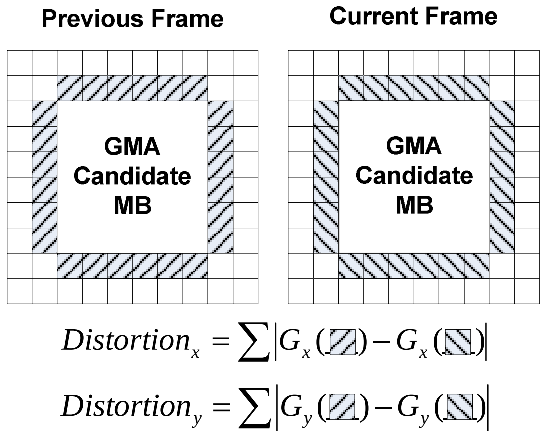


Figure 2: **OBMA Scheme**

Second part of hybrid algorithm is proposed by authors. It takes into account also similarity of structures, trespassing the lost MB, that is performed by using magnitude and direction of the gradients across the lost MB and candidate MB. It is called Gradient Matching Algorithm (GMA). To estimate similarity of the candidate MB and lost MB it calculates for every pixel of the outer boundary G_x and G_y values that are evaluated by the mean of Sobel operator for horizontal and vertical gradients. The total difference between lost MB and candidate MB is calculated as sum of $Distortion_x$ and $Distortion_y$, which are sums of absolute differences of corresponded G_x and G_y. The main idea of this algorithm is shown at Figure 3.

The goal of GMA algorithm is to find candidate macroblock in reference frame which neighborhood is close to neighborhood of lost macroblock by its structure. Proposed algorithm compares per-pixel gradient values in reference frame (candidate MB neighborhood) and current frame (lost MB neighborhood). The size of neighborhood is algorithm parameter that could have influence on quality of processing. The good balance between speed and quality of processing is reached for 16x16 macroblocks with 2-pixels outer boundary. And for outer pixels of this boundary G_x and G_y is estimated by means of Sobel operator shown at Figure 3, and for inner pixels that type of operator could not be used because of information absence in lost macroblock and because of it the modified version of Sobel operator is used which finds half-pixel shifted gradient value.



$$G_x = \frac{1}{4}\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad G_y = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Sobel Operator for outer pixel

$$G_x = \frac{1}{4}\begin{bmatrix} -1 & 1 \\ -2 & 2 \\ -1 & 1 \end{bmatrix} \quad G_y = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \\ -1 & -2 & -1 \end{bmatrix}$$

Operator for inner pixel

Figure 3: **GMA Scheme**

The idea of hybrid algorithm usage is to use both OBMA and GMA algorithm to find the temporal candidate MB. The simplest way to use these algorithms with each other is to use weighted normalized sum of distortion, calculated by means of OBMA – $Distortion_{OBMA}$ and distortion, calculated by means of GMA – $Distortion_{GMA}$.

$$Distortion = \alpha Distortion_{OBMA} + (1-\alpha)\frac{Distortion_{GMA}}{\beta}$$

where α is weighted coefficient and β is normalization coefficient.

The α coefficient controls the influence of different parts of algorithm, and it depends on dispersion of pixels values around the lost MB. The higher dispersion is a good signal that neighbor MB has strong edges or texture, so it is better to use GMA than OBMA and α is lower. But if neighbor pixels have low dispersion it means that there are no edges near the MB, so it is better to use OBMA, because it finds candidate, that has similar neighborhood with considering edges, so α is higher.

2.2 Boundary Error Estimation

The human visual system is very sensible to different discontinuities in smooth image, and because of it to reduce visual artifacts and to prevent blocking effect the next estimation system is used for postprocessing. For every border of the concealed MB (left, right, tip and bottom borders) average errors is calculated. For MB with coordinates *(x,y)* these errors will be:

$$E_{left} = \frac{\sum_{i=0}^{N}(I(x,y) - I(x,y+i))^2}{N} \quad E_{right} = \frac{\sum_{i=0}^{N}(I(x+N,y) - I(x+N,y+i))^2}{N}$$

$$E_{top} = \frac{\sum_{j=0}^{N}(I(x,y) - I(x+j,y))^2}{N} \quad E_{bottom} = \frac{\sum_{j=0}^{N}(I(x,y+N) - I(x+j,y+N))^2}{N},$$

Where $I(x,y)$ – brightness for pixel with coordinates (x,y) N – block size of MB

To compare these errors with corresponding errors for neighbor blocks (population) it is better to use not only difference between average values, but Z-test like methods to consider the standard deviation of the population. The formula for calculation *z-score* for the Z-test is

$$z = \frac{x-\mu}{SE}, \text{where } SE = \frac{\sigma}{\sqrt{n}}$$

Where x – is the mean of the sample, E_{left}, E_{right}, E_{top} or E_{bottom}
μ – is the mean of the population,
σ – the standard deviation of the population,
n – the size of the sample, block size of MB

The main idea of this is shown at Figure 4 for estimating left boundary. The return value for the estimation could be maximal z-score or average value for all boundaries.

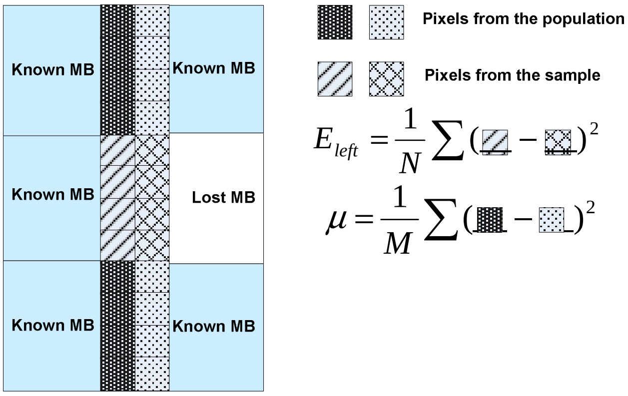


Figure 4: **Boundary Error Estimation Scheme for Left Boundary**

If the return value for this estimation exceeds the pre-defined threshold it means that concealed MB has discontinuities along its borders. In this case algorithm performs temporal concealment with minimizing error across the boundary – boundary matching algorithm (BMA), which is fully described in [4]. The main idea of this algorithm is shown at Figure 5. The principal difference to OBMA scheme is distortion calculation – BMA uses difference between pixels from inner boundary of candidate MB and outer boundary for lost MB.

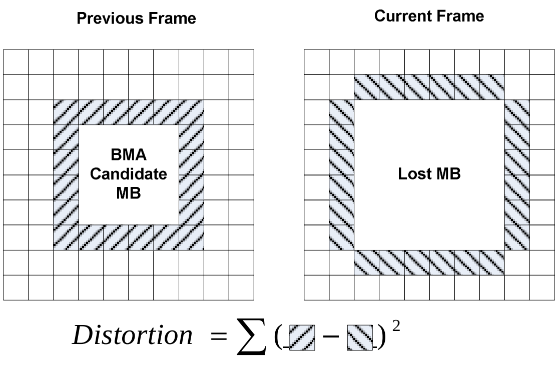


Figure 5: **BMA Scheme**

2.3 Spatial and Temporal Concealment

If boundary estimation score exceeds threshold even after BMA that means that there was a scene change or strong motion in the frame, so temporal methods could not find correct candidate MB in the previous frame. In this case spatial concealment method can be used. Spatial method often has more computational complexity, but because of low percentage of frame with need of spatial method using this method could be high-quality and very complex for calculation, such as for example [5], [8]. Proposed algorithm uses spatial interpolation as spatial concealment, when for unknown pixel the weighted sum of four known pixel is used. The main idea is shown at Figure 6. This method has low computational complexity and produces good quality. It is used in different video coding standards as recommended spatial error concealment even for H.264/MPEG-4 AVC coding standard [9].

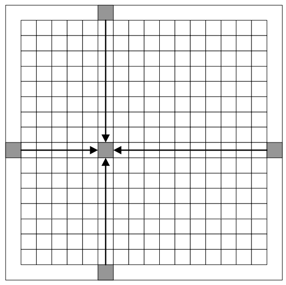


Figure 6: **Spatial Interpolation Scheme**

To minimizes blocking effect during candidate MB insertion in temporal concealment proposed algorithm use blending around inserting block with 1-2 pixel line width.

3. EXPERIMENTAL RESULTS

For experiment 10 CIF video sequences with different types of motion were chosen and proposed algorithm was compared to well-known temporal algorithms BMA, OBMA and SAA. For objective comparison PSNR (peak signal-to-noise ratio) as quality metric was used. Figure 7 shows the visual difference between OBMA and proposed algorithm. It is clearly seen the one of OBMA problems – when finding incorrect block for concealment in one frame if other frames have zero like motion OBMA keeps this block during all frames until motion.

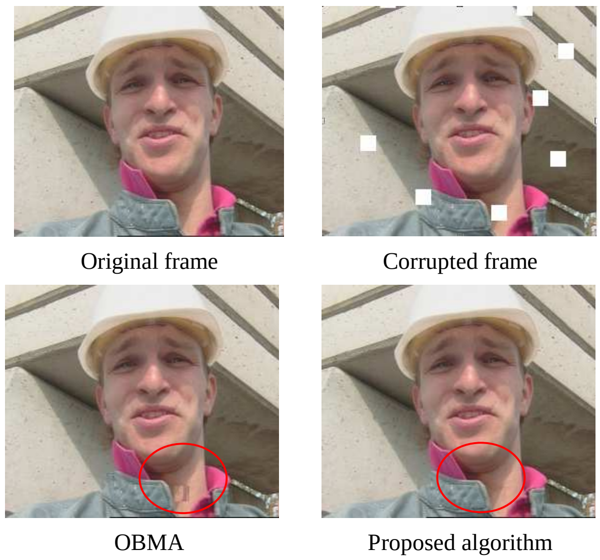


Figure 7: **Visual Comparison between OBMA and proposed algorithm**

For proposed algorithm fixed α weighted coefficient was used. Next table shows results for all sequences for BMA, OBMA, SAA and proposed algorithm. The proposed algorithm shows better PSNR value for all tested sequences.

VIDEO SEQUENCE	BMA	OBMA	SAA	Proposed Algorithm
akiyo	39,94	38,69	37,80	**39,97**
coastguard	38,66	36,07	36,12	**38,82**
crew	41,70	39,37	40,24	**41,70**
flower	33,04	33,92	31,15	**33,85**
foreman	41,59	40,50	39,59	**42,53**
mobile	32,89	28,33	32,20	**33,35**
stefan	36,69	36,94	34,53	**37,03**
susi	42,44	40,37	40,52	**42,85**
tens	35,03	31,83	32,85	**35,26**
waterfall	40,01	37,03	38,23	**40,35**

Table 1. Y-PSNR values for test set

The total enhancement between proposed algorithm and is better for BMA, OBMA and SAA algorithm is 0.03 to 0.95 dB. This value is shown at Figure 8.

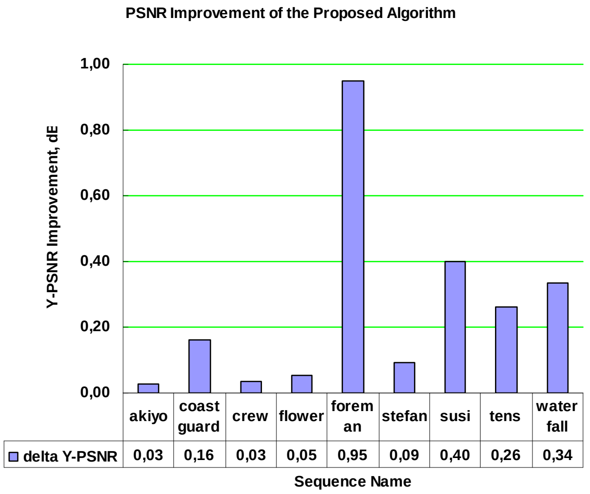


Figure 8: **Y-PSNR Improvement for Proposed Algorithm Comparing to the Best of BMA, OBMA, SAA**

4. CONCLUSION

In this paper new spatio-temporal algorithm for error concealment was presented. It could be used after or while decoding stage after packet-loss during video transmission. This algorithm shows better quality than well-known algorithm boundary matching or structural alignment. The proposed algorithm could also be used for video sequences restoration for removing scratches, spots or unwanted objects from the video or other real-time video processing.

5. REFERENCES

[1] Y. Wang and Q.-F. Zhu, "Error control and concealment for video communication: A review," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 974–997, May 1998.

[2] M. C. Hong, H. Schwab, L. P. Kondi, and A. K. Katsaggelos, "Error concelament algorithms for compressed video," *Signal Processing: Image Communication*, vol. 14, no. 6-8, pp. 473–492, May 1999.

[3] W. M. Lam, A. R. Reibman, and B. Liu, "Recovery of lost or erroneaously received motion vectors," in *IEEE International Conference on Acoustics, Speech and Signal Processing* , Minneapolis, MN, USA, May 1993, pp. V–417 – V420.

[4] J. Zhang, J. F. Arnold, and M. R. Frater, "A cell-loss concealment technique for mpeg-2 coded video," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 10, no. 4, pp. 659-665, 2000.

[5] A. Kaup, K. Meisinger, and T. Aach, "Frequency selective signal extrapolation with applications to error concealment in image communication," *Int. J. Electron. Commun.*, vol. 59, pp. 147-156, June 2005.

[6] A. S. Bopardikar, Inge, and A. Perkis, "A temporal error concealment algorithm based on structural alignment for packet video," in *Signal Processing and Communications, 2004. SPCOM '04. 2004 International Conference on*, 2004, pp. 373-377.

[7] T. H. Tsai, Y. X. Lee, and Y. F. Lin, "Video error concealment techniques using progressive interpolation and boundary matching algorithm," in *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on*, vol. 5, 2004, pp. V-433-V-436 Vol.5.

[8] Bertalmio, M., Sapiro, G., Caselles, V., and Ballester, C. 2000. Image inpainting. In *Proceedings of the 27th Annual Conference on Computer Graphics and interactive Techniques* International Conference on Computer Graphics and Interactive Techniques. ACM Press/Addison-Wesley Publishing Co., New York, NY, 417-424.

[9] V. Varsa, M. M. Hannuksela, and Y. K. Wang,"Nonnormative error concealment algorithms", *ITU-T VCEG-N62*, 2001.

[10] Z. Wang, A. C. Bovik, , H. R. Sheik, and E. P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 1, January 2004.

About the author

Dmitriy Kulikov is Ph.D. student at the Laboratory of Computer Graphics and Multimedia at Moscow State University, Department of Computational Mathematics and Cybernetics. His contact email is dkulikov@graphics.cs.msu.ru