ERROR CONCEALMENT TECHNIQUES IN MPEG-2

Sofia Tsekeridou and Ioannis Pitas

Department of Informatics
Aristotle University of Thessaloniki
Box 451, 54006 Thessaloniki, Greece
e-mail: *pitas@zeus.csd.auth.gr*

ABSTRACT

The MPEG-2 compression algorithm, due to the use of variable length coding, is very sensitive to channel disturbances. A single bit error during transmission leads to noticeable degradation of the decoded sequence quality in that part or entire slice information is lost until the next resynchronization point is reached. Error concealment (EC) methods are employed at the decoder side to deal with this problem. A novel concealment scheme is proposed in this paper that uses spatial or, when available, temporal information to reconstruct the corrupted frames. The concealment strategy is embedded in the MPEG-2 decoder model to avoid further concealment of propagation errors. The proposed scheme is proved to be of better performance compared to that achieved by other error concealment (EC) methods. Furthermore, the quality of the concealed sequence seems to ameliorate with time.

1. INTRODUCTION

A single bit error during the transmission of MPEG-2 compressed bitstreams results in noticeable degradation of the decoded sequence at the decoder side in that partial or entire slice information is lost until the next resynchronization point. Bitstream errors in anchor frames propagate to the subsequent frames in the coding order inside a group of pictures (GOP), due to predictive coding, thus resulting in the so-called propagation errors. While the locations of bitstream errors are sensed by the decoder in the way described in [1], the locations of propagation errors have to be estimated, when concealment is performed after the decoding of the entire sequence. A simple method of locating partially or entirely damaged MBs due to error propagation is described.

Different EC methods have been proposed for intra- and inter-coded frames (I and P/B frames respectively). They exploit spatial or temporal redundancy or both, when available. Concealment is performed by spatial interpolation methods (applied mainly to I-frames, since no motion information is available for those) [1], temporal replacement methods [1], which consider motion information only when available, or by combining both spatial and temporal methods [4]. For the latter case, the decision on which method to use is based on measures indicating the degree of local spatial detail or local motion. Regions with high local motion are concealed by spatial interpolation techniques, whereas regions with high local spatial detail are concealed by temporal replacement methods. Spatial EC methods result in blurring and in poor continuation of lines and edges in directions other than vertical. Temporal EC methods result in noticeable shifts, when motion is not considered (which is the case for I-frames). The use of concealment motion vectors for I-frames improves the quality of the concealed I-frames but adds a 0.7% overhead on the total bitrate. Furthermore, such information could be lost as well during transmission. The combination of both spatial and temporal methods is proved to be of better performance with a small increase in complexity.

A novel EC scheme is presented in this paper consisting of the *Spatial Split-Match EC method*, used for the concealment of the first I-frame of an MPEG-2 coded video sequence, since no temporal information is available for that frame, and the *Forward-Backward Temporal Block Matching EC method*, used for the concealment of the other frames of the sequence (both intra- and inter-coded ones). The *Spatial Split-Match EC method* uses spatial information from existing top and bottom MBs aiming at concealing the lost MBs in such a way so that the transition between existing and reconstructed lost regions in a frame is smoothly performed. The *Forward-Backward Temporal Block Matching EC method* exploits reconstructed temporal information from previously decoded frames. The proposed scheme is embedded in the decoder model. In this way, only bitstream error concealment is required. The results obtained indicate satisfactory performance of both methods. Temporal shifts or spatial blurring are eliminated at a great extent. The processing time is kept small in order to ensure real time implementation. The improved visual quality of the video sequence justifies the observed performance especially as time passes, i.e., the frame index number increases.

2. ERROR DETECTION

Sometimes, information loss occurs in communication channels. In the case of transmission of highly compressed MPEG-2 video sequences, this leads to significant errors observed at the receiver. These errors are either:

(a) **Bitstream errors**, caused by direct loss of part of or the entire compressed bitstream of a coded MB, and result in the loss of the following slice information, if default resynchronization has been adopted, and

(b) **Propagation errors**, which corrupt P- and B-frames only, due to the additional use of motion-compensated information during their decoding. Thus, bitstream or propagation errors in anchor frames propagate in subsequent P- or B-frames inside the same GOP.

Bitstream error detection is performed in the way proposed in [1]. Consequently, the locations of bitstream errors are available at the decoder. The knowledge of them in anchor frames (I- or P-frames) enables the detection of propagation errors in subsequent P- or B-frames of the same GOP. At the locations of propagation errors all information is available at the decoder (i.e. motion vectors, MB coding modes, prediction errors), since no bitstream loss has occurred during transmission at these locations. If the motion vector and coding mode of a MB in the current P- or B-frame are such that, for the decoding of this MB, information is obtained from a damaged MB in the reference frame, then the current MB is most probably partially or entirely corrupted due to propagation of the reference frame errors. It is also noted that propagation errors in P-frames could propagate in subsequently decoded B-frames of the same GOP.

3. SPLIT-MATCH EC METHOD

The best possible reconstruction of an I-frame is most desirable because imperfect reconstruction results in the propagation of imperfect concealment errors to all frames in the respective GOP and finally in the degradation of the video sequence visual quality. Since for the first frame (I-frame) of an MPEG-2 coded sequence, no temporal information is available, the *Split-Match* EC method has been added in the concealment scheme to deal with this case.

Given the location of a lost MB in a certain slice, the algorithm first examines if the neighbouring MBs at the top and bottom slices are available. If the immediate top or bottom ones are also lost, the algorithm searches for the next available ones. In such a case, the region to be concealed becomes larger in the vertical direction. For this reason, the search region defined in the fourth step of the algorithm is enlarged, aiming at attempting to fully reconstruct lines or edges in almost all directions. The algorithm consists of four steps, as shown in Figure 1. It performs matching of existing top and bottom neighbouring blocks. The mean absolute difference (MAD) between those, located in a predefined search region, is estimated. That combination of blocks (referred to as the “best match”) that leads to the minimum MAD is used to recover the image content of the lost image part between them. The matching procedure initially tries to match large neighbouring blocks, proceeds by splitting the initial blocks into smaller ones and performs matching between the new ones, if $MAD_i < T_i$, $i = 1, .., 4$, is not satisfied, where T_i are predefined thresholds, or until a minimum block size has been reached, e.g. $4{\times}4$. The initialization of the algorithm with large block matching ensures that big homogeneous or textured regions will be reconstructed directly, thus avoiding processing delays or computational complexity. On the other hand, the satisfactory reconstruction of smaller homogeneous or textured regions, regions with small details, lines and edges, requires the matching of smaller blocks. As soon as a best match has been found, concealment is performed by copying the image content of the best matched blocks into the lost region between them in the direction of the match. Due to initial concealment by copying, blocking artifacts in the concealed areas are unavoidable. In order to smooth this blocking effect, the postprocessing technique of [2] is used. This method performs only on the concealed areas.

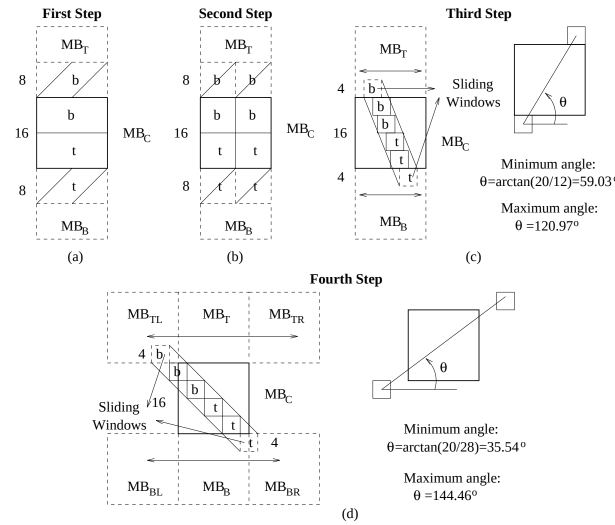

Figure 1: Graphical presentation of the Split-Match EC algorithm. The edge or line directions that can be reconstructed are indicated. MB_T, MB_B, MB_{TL}, MB_{TR}, MB_{BL}, MB_{BR} are the surrounding 16×16 MBs to the current lost MB_C, located in the top (T) and bottom (B) slice, respectively.

An alternative attempt has been made to conceal the pixels inside the in-between region in the direction of the best match for the third and fourth steps of the algorithm by applying a simplified form of anisotropic diffusion [3]. One reason for considering concealment by diffusion in the third and fourth steps of the algorithm is the fact that the blocking artifacts produced by copying during the execution of these steps are more noticeable. The diffusion process eliminates the artifacts and thus prevents use of a postprocessing method. To determine the pixels belonging to the region in the direction of the best matched blocks the process illustrated in Figure 2 is followed. Those pixels for which

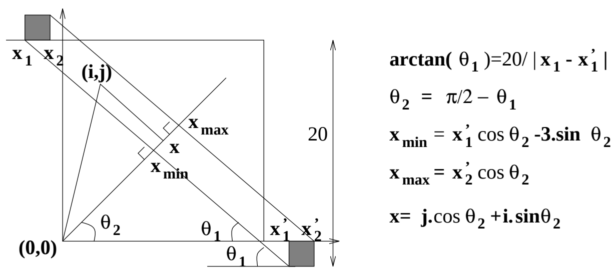

Figure 2: Graphical presentation of the method for the determination of pixels to be concealed located in the direction of the best matched blocks.

the inequality: $x_{min} \leq x \leq x_{max}$ is valid are concealed by employing anisotropic diffusion between them and the border pixels of the best matched blocks. The image intensity $I_{i,j}$ at pixel (i,j) is determined by iteratively employing the expression:

$$I_{i,j}^{t+1} = I_{i,j}^{t} + \lambda [c_T \nabla_T I + c_B \nabla_B I]_{i,j}^{t} \quad (1)$$

where $\nabla_{T/B} I_{i,j}$ is the intensity difference between the considered pixel (i,j) and the respective border one in the top / bottom block of the best match. It is evident that the diffusion process is performed in the direction of the match between the vertical neighbours only, since horizontal neighbours are lost. The parameter λ controls the stability of the numerical presentation (1) of the diffusion equation and is set equal to 0.15. The conduction coefficients $c_T(i,j)$ and $c_B(i,j)$ are determined by the following function:

$$c_{T/B} = 1 - \frac{d_{T/B}}{d_{max}} e^{-\nabla_{T/B} I_{i,j}/K} \quad (2)$$

$d_{T/B}$ represents the vertical distance between the current pixel and the top/bottom border pixel employed in the diffusion process and d_{max} is its maximum value. The constant K is used in order to reduce the effect of large intensity differences and ensure smooth intensity diffusion. The purpose for introducing vertical distances in expression (2) is to ensure larger conduction coefficients for less distanced pixels and vice versa. It has been found that only a few iterations (around 10-20) are necessary for total concealment. Thus no additional delays are added in the concealment scheme.

The first two steps of the Split-Match algorithm are performed once, whereas the other two are performed iteratively until total MB concealment. The processing time is kept small because the introduction of several steps prevents useless computations, in that only those necessary for current MB concealment are executed.

4. FORWARD-BACKWARD TEMPORAL BLOCK MATCHING EC METHOD

The concealment scheme that is proposed is embedded in the MPEG-2 decoder. The decoder receives the compressed bitstreams, performs the decoding and then the error concealment processor steps in to process the possibly damaged decoded frame. If this frame is an I- or P-frame, it is stored, after processing, in the past or future frame buffer. Thus, propagation errors are prevented leaving only the bitstream errors to be dealt with. The concealment of bitstream errors in frames for which temporal information is available is performed by the *Forward-Backward Temporal Block Matching EC method*.

The proposed method exploits the temporal redundancy of an image sequence. The assumption of a smooth and uniform motion for adjacent blocks is adopted. Temporal block matching is performed between the available top and/or bottom MBs adjacent to the current lost one and equally sized blocks located in a predefined search region of size $N \times M$ in the reference frame/frames, its center being the lost MB. The search region is bigger if the matching is performed between far-distanced frames and decreases if the frames in question are closer in time. The purpose of this procedure is to avoid useless computations and reduce the time delay. Special attention is paid so that the matched MBs are vertically aligned with vertical distance equal to the height of the lost region. In this way, when the best match is found based on the MAD minimization, the region between the 16×16 blocks of the best match is copied to the current lost MB. The matching process is illustrated in Figure 3. For I- and P-frames, the matching is performed between the current and the past reference frames, while for B-frames, the matching is performed between the current and the past as well as the future reference frames. No additional memory requirements are introduced because the reference frames used are previously reconstructed and stored in the past or future frame buffers of the decoder. Even for I-frames, the previously concealed P-frame already exists in the past frame buffer to be used for the decoding of the next in line B-frames.

5. SIMULATION RESULTS

The Flower Garden sequence (125 frames) has been used to test the performance of the proposed scheme. Objec-

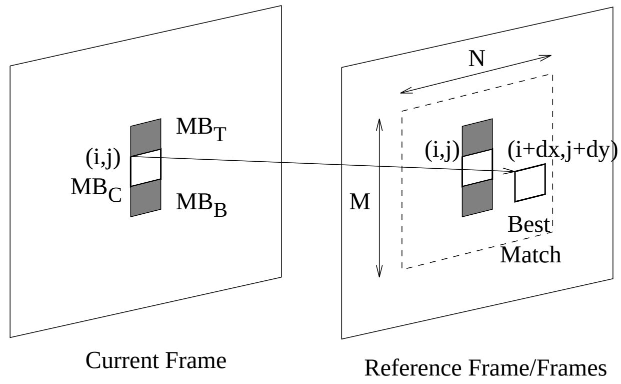

Figure 3: Temporal Matching of the available top and/or bottom MBs (MB_T, MB_B) between current and reference frame/frames. The in-between 16×16 region is copied to the current lost MB.

tive comparison is performed by estimating $PSNR$ values, whereas subjective one is achieved through the perception of the visual quality of the processed frames.

The Split-Match EC method (S/M EC) is compared with the method proposed in [1] (SVI EC). Results are presented in Table 1, for all three color components Y, U and V. While

Table 1: $PSNR$ measured on Y,U,V components of Frame 0 processed by the S/M EC ((1): copying, (2): diffusion) and the SVI EC of [1].

Method	$PSNR$		
	Y	U	V
Error Fr	32.2193	34.3306	36.0781
S/M EC (1)	26.3157	33.0859	35.2774
S/M EC (2)	26.4613	33.1087	35.2965
SVI EC	26.1087	33.1572	35.1536
Erroneous	15.7172	30.1665	32.9893

the proposed technique is only slightly better than the one in [1] concerning $PSNR$, the visual quality of the processed image is proved to be significantly better than that achieved by the latter method (Figure 4). Edges or lines other than vertical ones are better reconstructed and the texture continuation is smoother. The proposed method fails only in regions whose reconstruction is hardly possible if only spatial information is used, as it is the case in similar methods [1]. One disadvantage of the proposed method are the blocking artifacts caused by block copying into the lost areas. These however are eliminated by the diffusion scheme presented earlier, preventing thus use of a postprocessing method. The processing time remains small, because processing is performed only to the lost parts and because each of the algorithm's four steps is executed only when considered necessary.

Figure 4: Frame 0: Error Free, $PSNR = 32.2193$ (top-left), Erroneous, $PSNR = 15.7172$ (top-right), processed by method [1], $PSNR = 26.1087$, (bottom-left) and by Spatial S/M EC method (using anisotropic diffusion), $PSNR = 26.4613$ (bottom-right).

The performance of the proposed forward-backward temporal block matching concealment method (F/B BM EC) is compared with that achieved by the adaptive EC of [4] and by the temporal ECs of [1] referred to as motion compensated temporal EC and zero motion EC. The average values of $PSNR$ measured on the first 125 frames of the test sequence are given in Table 2. The observation of these results

Table 2: Average $PSNR$ measured on the 125 Frames (Y Component) processed by the F/B BM EC, the adaptive EC of [3] and the zero motion and motion compensated temporal ECs of [1].

Method	Average $PSNR$
Error Free Seq.	28.155256
F/B BM EC	26.4991
Adaptive EC	25.2962
Motion Comp. EC	25.0903
Zero Motion EC	23.9627
Erroneous Seq.	15.9498

leads to the conclusion that the proposed method performs significantly better than the others. The visual quality of the entire sequence also is proved to be satisfactorily improved with respect to that achieved by the other methods. Figure 5 shows the achieved visual quality of a B-frame. This significant visual and $PSNR$ improvement results from the fact that concealment in I-frames is much better performed when

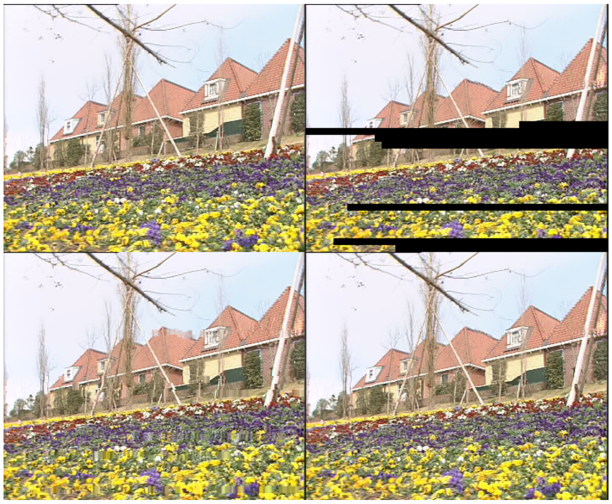

Figure 5: Frame 92: Error Free, $PSNR = 27.1917$ (top-left), Erroneous, $PSNR = 11.0338$ (top-right), processed by method [4], $PSNR = 21.1499$ (bottom-left) and by Temporal F/B BM EC method, $PSNR = 24.5583$ (bottom-right).

taking into account temporal information. Their best possible reconstruction results in a remarkable improvement of the visual quality of the decoded sequence. The MPEG-2 standard has, of course, the possibility to transmit concealment motion vectors for I-frames in the bitstream. However, bitstream errors might lead to their loss as well.

Although the concealed decoded video sequence is of satisfactory visual quality, some defects and propagated concealment errors cannot be avoided. Such kind of reconstruction errors are caused in occluded regions, where mere past time information proves rather useless. Defects might appear in the borders of the reconstructed frame, where new image content enters. Furthermore, in cases where the assumption of uniform motion does not apply, the method will not perform satisfactorily. However, the visual artifacts caused in such cases are not easily perceived, considering that, e.g. for PAL quality, 25 frames are displayed per second. Another drawback might prove to be the processing time. This is true when a great number of MBs in a frame are lost. In order to reduce the processing time, the search region in the reference frame may be chosen smaller if the frames are nearer in time and larger when they are far apart.

6. CONCLUSIONS

In conclusion, the Spatial Split-Match EC method achieves line, edge and smooth texture continuation with high probability. Its processing time is kept as small as possible. The method fails only in regions which cannot be reconstructed by mere spatial information. A drawback of the algorithm is its dependence on thresholds. Through the use of anisotropic diffusion instead of copying, postprocessing is not required. The Forward-Backward Temporal Block Matching EC method achieves significant improvement of both $PSNR$ and visual quality of the concealed video sequence. Its processing time depends on the loss percentage per frame. Defects due to occlusions or new image content entering from the frame borders as well as propagated imperfect concealment errors cannot be prevented. However, these visual artifacts are hardly obvious in most cases.

7. REFERENCES

[1] S. Aign and K. Fazel, “Temporal & Spatial Error Concealment Technique for Hierarchical MPEG-2 Video Codec”, *in Proc of the Int. Conf. on Communications, ICC’95*.

[2] B. Ramamurthi and Al. Gersho, “Nonlinear Space-Variant Postprocessing of Block Coded Images”, *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-34, no. 5, pp. 1258-1267, October 1986.

[3] P. Perona and J. Malik, “Scale-Space and Edge Detection Using Anisotropic Diffusion”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, July 1990.

[4] H. Sun and J. Zdepski, “Adaptive Error Concealment Algorithm for MPEG Compressed Video”, *SPIE Conf. on Visual Communications and Image Processing*, vol 1818, pp. 814-824, 1992.

[5] S. Tsekeridou, I. Pitas and C. Le Buhan, “An Error Concealment Scheme for MPEG-2 coded video sequences”, *ISCAS’97*, vol. II, pp. 1289-1292, Hong Kong, 9-12 June 1997.