14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

AN IMPROVED ERROR CONCEALMENT STRATEGY DRIVEN BY SCENE MOTION PROPERTIES FOR H.264/AVC DECODERS

Susanna Spinsante, Ennio Gambi, Franco Chiaraluce

Dipartimento di Elettronica, Intelligenza artificiale e Telecomunicazioni, Universitá Politecnica delle Marche
via Brecce Bianche, I-60131, Ancona, Italy
phone: + (39) 0712204894, fax: + (39) 0712204835, email: s.spinsante, e.gambi, f.chiaraluce @univpm.it
web: www.deit.univpm.it

ABSTRACT

This paper deals with the possibility of improving the concealment effectiveness of an H.264 decoder, by means of the integration of a scene change detector. This way, the selected recovering strategy is driven by the detection of a change in the scene, rather than by the coding features of each frame. The scene detection algorithm under evaluation has been chosen from the technical literature, but a deep analysis of its performance, over a wide range of video sequences having different motion properties, has allowed the suggestion of simple but effective modifications, which provide better results in terms of final perceived video quality.

1. INTRODUCTION

Compressed video information, as H.264/AVC encoded video, is in general very sensitive to channel errors which may appear during transmission. The adoption of Variable Length Coding (VLC) at the encoder side, together with more complex techniques, like Motion Compensation, can lead to error propagation. Error correction methods for compressed videos consist in increasing the system robustness: i) by adding extra parameters or more synchronization points, or ii) by applying error concealment schemes to reconstruct the lost or erred blocks. In general, solution i) needs to change the encoding scheme and in some situations this could be not compatible with current systems; for this reason, solution ii) is usually preferred in most practical applications. Error concealment has two basic approaches: spatial interpolation and motion compensation. The former, used in the case of Intra frames for which previous reference frames do not exist, exploits only the information deriving from surrounding macroblocks to regenerate the missing pixels. Methods based on spatial interpolation [1] can hardly achieve high resolution for the lost blocks; consequently, they are not well suited to detailed video sequences. On the other hand, the temporal prediction method with motion compensation exploits the correlation between Inter frames to recover the damaged blocks [2]. It is possible to compensate a lost block by using Motion Vectors estimated from the temporal reference frame, but this approach is not effective in high motion areas, or when scene changes occur, since the temporal correlation becomes low. Errors at scene changes are obviously less frequent than errors during the scene; however, errors at scene cuts could be catastrophic for the decoding mechanism, so they should be avoided or compensated somehow.

In this paper we present an error concealment strategy which is basically driven by an evaluation of the sequence dynamism, in terms of scene change events and panning effects, rather than being based on the coding features of each frame. In Section 2 we summarize the features of the Reference Software JM9.7 concealment algorithm and the scene change detector, that are the starting point of our study. Then, in Section 3, after having motivated the flaws of the current solutions, we introduce some modifications to the scene change detection algorithm, that permit its integration with the reference concealment strategy. Explicative examples, developed on standard sequences, are also provided. Finally, Section 4 concludes the paper.

2. ERROR CONCEALMENT AND SCENE CHANGE DETECTOR

2.1 Reference Software JM9.7 Concealment Algorithm

The concealment algorithm implemented within the Reference Software version JM9.7 [3] is a hybrid scheme, which applies different recovering strategies according with the nature of the frame, i.e. Intra or Inter. If the block to be concealed belongs to an Intra frame, the missing pixels are obtained as a weighted interpolation of the pixels located in the available adjacent macroblocks. In the case of a lost macroblock belonging to an Inter frame, the algorithm adopts two different strategies, depending on the motion degree of the video content: low motion macroblocks are simply replaced by copying the corresponding macroblocks from the last reference frame, whereas high motion macroblocks require further computation. More specifically, among the Motion Vectors of the available adjacent macroblocks, the one that points to the area in the frame that minimizes the pixel difference over the edges is selected. This way, the concealed version of the frame is quite uniform, being the motion of adjacent areas highly correlated. Both the strategies use the so called *status map* of the macroblocks, in which every macroblock is marked as *correctly received*, *lost* or *concealed*. In order to recover a current macroblock marked as *lost*, not only the available *correctly received* macroblocks but also the previously *concealed* ones can be used. In regard to the concealment of Intra frames provided by the JM9.7 Reference Software, first of all the macroblocks available for bilinear interpolation of the missing pixels are located. In the case of the luminance component, restoration is performed on a macroblock basis, whereas for the chrominance components it is carried out on a block basis. Pixels of the missing block are restored row by row, by using for computation four adjacent blocks, in a cross fashion. If the interpolation process cannot be carried out because of the unavailability of the adjacent macroblocks, the predicted DC coefficient is used instead of the current pixel. Once the interpolation is completed, the macroblock is marked as *concealed* in the status map, and can be consequently used for restoration of other macroblocks.

Inter frame concealment is carried out on a block basis, according with the motion properties of the damaged areas. Once one or more missing macroblocks are found in the status map, three concealing strategies are available: from top to bottom, from bottom to top or bidirectional. The restoring process starts from the macroblocks at the edge of the frame and proceeds towards the center, column by column. The recovering algorithm for each missing macroblock is selected according with the frame motion degree, which is evaluated by comparing the average amplitude of its Motion Vectors to a threshold. When the threshold is overcome, a suited subroutine selects the Motion Vectors of the macroblocks adjacent to the missing one, to get back to their corresponding predictor macroblocks. The predictor for which the difference among the pixels located on its edges and the pixels located on the edges of the missing macroblock is minimal is chosen to replace the missing block in the current frame. When the threshold is not exceeded, another sub-routine is executed, which recovers the missing macroblock by copying the area of the reference frame located in the same position. At the end of this process, the information stored within the *status map* is updated.

2.2 Scene Change Detector

The details of the scene change detector we refer to in this paper can be found in [4]; in the following, we will provide motivations for some modifications, to obtain improved performance.

As well known, in the case of Inter frames the encoder exploits their temporal correlation to find a good predictor for each macroblock, using Motion Compensation based techniques. When a scene change occurs, however, the correlation gets lost and the majority of the blocks in the frame is encoded as Intra. According with this consideration, if the number of 8x8 blocks encoded as Intra within an Inter frame gets higher than a threshold, we can assume a scene change has occurred. In this case, even in presence of an Inter frame, it is better to perform a spatial concealment, by pixel interpolation from adjacent macroblocks. Following [4], the threshold has been experimentally set to 30%; actually, we will point out that this value is not appropriate in some specific situations related to panning effects of the camera. It is also important to set a minimal threshold to the number of correctly received macroblocks, because in presence of wide damaged areas, copying the missing macroblocks from the previous frame is better than restoring them by interpolation, at least from the point of view of subjective perception.

As no information on temporal correlation is available in the case of Intra frames, scene change detection is performed by properly "sampling" four distinct macroblocks, one within each quadrant. Once the sample macroblocks are selected, the co-located blocks in the most recent reference image are used to compute the SAD (Sum of Absolute Difference) between their corresponding pixels. By applying a suited algorithm, one out of the four SAD values is chosen and compared to a threshold experimentally set to 10000; if higher, a scene change is revealed. Fig.1 shows the scanning process performed in order to locate the four sample macroblocks within an Intra CIF (Common Intermediate Format) frame.

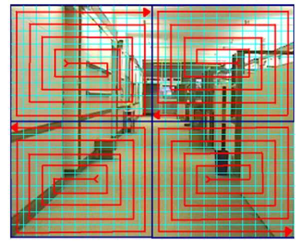

Figure 1: Scanning order within each quadrant of an Intra CIF frame

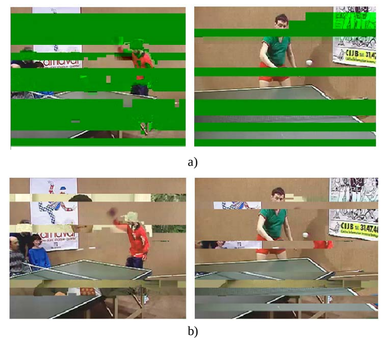

Figure 2: Scene change between frames 189 and 190 of TableTennis CIF sequence, when decoded a) without concealment, b) by applying the JM9.7 concealment function for Inter frames (packet loss = 13.66%)

3. THE PROPOSED SOLUTION

The original concealment algorithm provided by the Reference Software JM9.7 shows some drawbacks which are basically related to the following considerations:

- the restoring strategy applied to Intra frames is not effective if the damaged areas are wide or the image is rich in details. Under these conditions, spatial interpolation determines a shading effect and a loss of details;
- the restoring strategy conceived for Inter frames is effective in the case of null or very low motion video sequences, whereas it determines annoying artifacts when scene change events occur, as shown in Fig.2, for the *TableTennis* CIF sequence.

With the aim of improving the behavior of the JM9.7 concealment scheme, we propose its integration with a modified scene change detector. Modifications are motivated by experimental evidences of the inefficiency of the suggested thresholds under specific conditions. The new scheme performs concealment of the damaged frames on the basis of the detected scene change events, instead of referring to the coding features of the frames. A simplified block diagram of the modified scheme is plotted in Fig.3. As shown in the diagram, either in the case of an Intra or Inter frame, the scene change detector is applied. If no scene change is revealed in the case of an Intra frame, then a plain temporal replacement strategy [5] is adopted to restore the damaged blocks, i.e. the missing blocks are simply copied from the previous frame; otherwise the concealment algorithm for Intra frames provided by the JM9.7 Reference Software is invoked. The JM9.7 concealment scheme for Intra frames is activated also in the case of a scene change detected within an

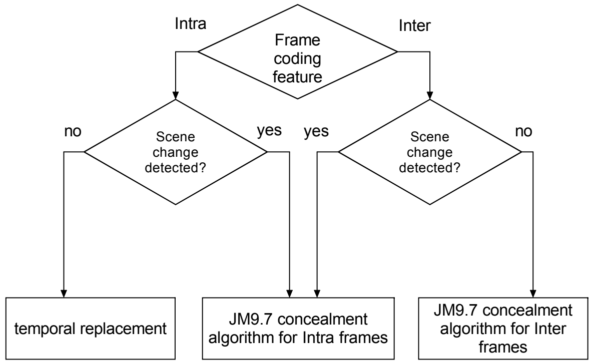

Figure 3: Block diagram of the modified error concealment scheme based on scene change detection

Inter frame, whereas the JM9.7 concealment solution specifically conceived for Inter frames is applied, in the case of no scene changes detected. Before integrating the scene change detector into the JM9.7 Reference Software, we tested its performances when applied to "ad hoc" video sequences, named *ensemble* sequences. They consist of a concatenation of standard sub-sequences, each one with its own features (i.e. panning effects, high/low motion, scene details...). At the same time, the ensemble sequence shows a number of scene changes adequate for testing the efficiency of the algorithm. An example of ensemble sequence is reported in Table 1.

3.1 Scene change detector improvements for Inter frames

The performance of the scene change detector has been tested on a CIF sequence characterized by the so called panning effect, a movement of the camera that provides gradual transition of the scene in sequential frames. More in details, we considered the CIF ensemble video sequence reported in Table 1; this is composed by a total amount of 300 frames, with sub-sequences switching every 20 frames. An RTP (Real Time Protocol) packetization has been simulated, with one slice (22 macroblocks) in each packet and a packet loss equal to 13.66%, i.e. a very hostile transmission context. The first frame of the sequence is encoded as Intra, whilst all the following ones are predicted, i.e. Inter frames; the *intraperiod* parameter is zero, as we want to test the efficiency of the scene change detector for Inter frames. The Quantization Parameter has been set to 28, with a 30 Hz frame rate.

By applying the scene change detector to the ensemble sequence, we found that all the true scene change events are correctly revealed, whereas false detections take place in frames 41 - 60 occupied by the *Stefan* sub-sequence. No scene change takes place, but a gradual change of the scene due to the movement of the camera. This panning effect causes false detections, because the percent amount of Intra 8x8 blocks within these frames gets higher than 30%, as evidenced in Fig.4. Taking into account the experimental results, we introduce a second parameter, named *delta*. This parameter is calculated by evaluating the number $N1$ of Intra 8x8 blocks in the frame for which a potential scene change event has been detected, and the corresponding number $N2$

Table 1: Ensemble sequence

Frame	Sequence
1 - 20, 241 - 260	Akiyo
21 - 40	Foreman
41 - 60	Stefan
61 - 80	MobileCalendar
81 - 100	Children
101 - 120	Container
121 - 140	GoldenGate
141 - 160	HallMonitor
161 - 180, 281 - 300	Sean
181 - 200	TableTennis
201 - 220, 261 - 280	News
221 - 240	Weather

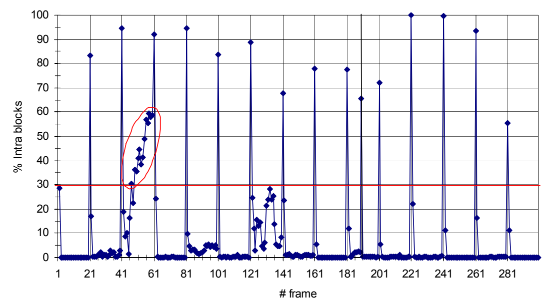

Figure 4: Percent amount of blocks coded as Intra within the Inter frames of the ensemble sequence

in the frame immediately preceding. Then, $delta = N1 - N2$; this value is compared against a second threshold also fixed to 30%, as experimentally optimized.

Thanks to this modification, the scene change detection algorithm does not provide false detections. More precisely:

- in the case of true scene change events, not only the number of Intra 8x8 blocks is higher than 30%, but also its increment;
- in the case of panning phenomena, no false detections take place, because even if the number of Intra 8x8 blocks is higher than 30%, its increment is lower.

3.2 Scene change detector improvements for Intra frames

The ensemble video sequence described in the previous subsection has been used to test the performance of the scene change detector, when applied to Intra frames. To this aim, for sequence encoding, the *intraperiod* parameter has been set to 1, in such a way as to force the Intra coding of each frame, with a Quantization Parameter equal to 28 and a frame rate of 30 Hz. Tests have been performed considering three different packet loss values: 13.52%, 16.10% and 10.59% respectively, and a fixed value of 10000 for the threshold

Table 2: Percent amount of scene change events correctly detected, in the case of Intra frames, for different values of the SAD threshold and the number of sample macroblocks

SAD Threshold	4 MBs	2 MBs	9 MBs
10000	90.47	90.47	90.47
9000	92.85	90.47	97.61
8500	97.61	90.47	100
7000	100	100	100
5000	100	100	100

against which the SAD value obtained by the four sample macroblocks is to be compared, in order to detect a scene change event. Different packet loss profiles imply the selection of different macroblocks in the four quadrants of the frame, to compute the SAD value required by the detector. This way, it is possible that the same scene change event is correctly revealed or not, in each situation under evaluation, because the SAD value that is compared to the threshold can vary from test to test.

Scene change events are always correctly detected in the first situation, when the packet loss equals 13.52%; in the second case, for a packet loss of 16.10%, a true scene change event is not detected at frames 95 - 96, corresponding to the *Table Tennis* sub-sequence, because the SAD value gets lower than 10000. The same happens in the third case, for a packet loss of 10.59% and for the same frames. In order to solve the problem of undetected scene change events, we tried to modify the threshold and the number of sample macroblocks for the SAD computation (not only 4, but also 2 and 9), obtaining the results collected in Table 2, where the percent amount of change events correctly revealed is reported, for each configuration of the detector. As a further verification, among the configurations giving a 100% amount of scene change events correctly revealed, we have selected that corresponding to a SAD threshold equal to 5000 and 4 sample macroblocks, and tested these new detector settings on a sequence comprising also some frames from *Gladiator*, which has a very high motion degree and is very rich in details. Again, tests have been performed by considering three different packet loss values, 5.57%, 11.26%, and 17.97%: in any case, all the scene change events have been correctly detected.

3.3 Concealment strategy based on the scene change detector

Once the scene change detector has been optimized for both the cases of Inter and Intra frames, it has been integrated within the concealment strategy provided by the Reference Software JM9.7, as described in Fig.3. The resulting concealment scheme has been tested on a new ensemble CIF video sequence, different from that in Table 1, also including frames from *Gladiator*. The Quantization Parameter has been set to 28, with a frame rate of 30 Hz. Up to 5 reference frames are available for the JM9.7 concealment of Inter frames; the *intraperiod* equals 15. A packet loss of 13.98% has been simulated; scene change events due to subsequences switching take place every 10 frames. In Fig.5 c), it is clear that the introduction of the scene change detector

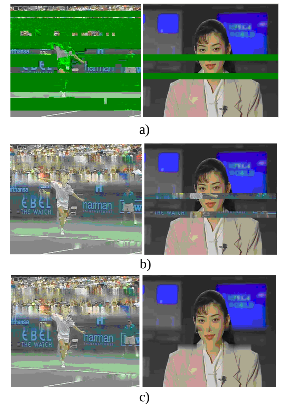

Figure 5: Adjacent frames decoded a) without concealment, b) by JM9.7 concealment, c) by JM9.7 concealment integrating the scene change detector

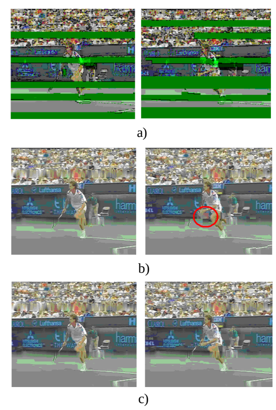

Figure 6: Adjacent frames decoded a) without concealment, b) by JM9.7 temporal concealment for Inter frames, c) by temporal replacement for Inter frames

improves the behavior of the original JM9.7 concealment algorithm, when a scene change takes place, avoiding the use of macrolocks that are not related to the video content, for the recovery of the current frame. Besides the detection of scene change events, a further improvement to the original JM9.7 concealment relies on the adoption of a temporal replacement strategy, when no scene changes take place, in the case of Inter frames. Fig.6 c) shows the results obtained by applying the pure temporal replacement; they are better than results in Fig.6 b), where the JM9.7 temporal concealment has been used. As a matter of fact, in the latter case the concealed blocks evidenced in the figure do not fit well the rest of the frame: they belong to *Foreman* that is rather far in the sequence, as permitted by the JM9.7 temporal concealment, that can use up to 5 previous reference frames. This way, however, it is not able to correctly manage the sequence dynamism. A similar situation happens in the case of the *Table Tennis* sub-sequence (frames 254 - 256), as shown in Fig.7. The introduction of the *delta* parameter for the correct detection of scene changes in panning frames allows to improve the concealment effect, by avoiding the loss of details due to pixel interpolation. However, the same parameter is not effective in the case of the *Gladiator* sub-sequence. No

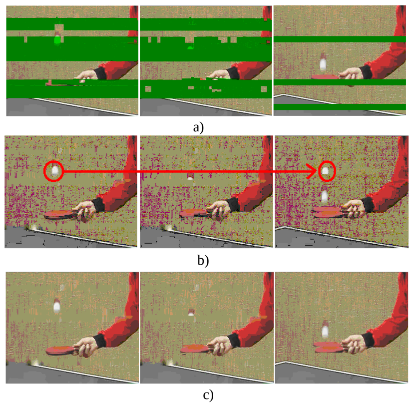

Figure 7: Consecutive frames decoded a) without concealment, b) by JM9.7 temporal concealment for Intra frames, c) by temporal replacement for Intra frames

scene change event takes place between frames 88 and 89, as *delta* is lower than 30%, and a temporal concealment is applied. However, the sequence high motion degree makes the correlation between consecutive frames very low: as a matter of fact, Fig.8 shows that better results are achieved by means of pixel interpolation (b) than by copying the missing macroblocks from previous frames (c). The need of improving the concealment strategy in the case of high motion sequences has suggested a further modification, which has been driven by the analysis of the evolution of the percent amount of Intra 8x8 blocks within the Inter frames of the sequence. The modification applies to the scene change detector for Inter frames, as described by the following pseudo-code:

```
if (%Intra 8x8 > 45%) then:
     JM9.7 spatial concealment
if (%Intra 8x8 < 45%)
   AND (%Intra 8x8 > 30%)
   AND (delta > 30%)  then:
     JM9.7 spatial concealment
if (%Intra 8x8 < 45%)
   AND (%Intra 8x8 > 30%)
   AND (delta < 30%)  then:
       JM9.7 temporal concealment
```

A huge amount of Intra 8x8 blocks within an Inter frame, even if not related to a scene change event, denotes a low correlation among consecutive frames. In this case spatial interpolation of missing pixels performs better than temporal concealment. Condition expressed by the second if option accounts for this event. The last modification has been tested on the same ensemble CIF sequence, for a packet loss of 6.02%. Fig.9 shows the results obtained on frame 124 of *Gladiator*. Subjective comparison between frames b) and c) confirms the improvements obtained in b), that are supported also by the increase of the PSNR Y value, from 21.8 dB in c) to 24 dB in b). Finally, the computational time required at the decoder to perform the concealment process according with the final version of the algorithm is not increased with respect to the original JM9.7 scheme, whereas the resulting average PSNR Y value gets 1.2 times higher.

4. CONCLUSION

A scene detection algorithm recently proposed in the literature has been integrated in the JM9.7 concealment scheme, with the goal to improve the perceived video quality, without the need for complex processing at the decoder. The main contribution of this paper has been to introduce some important changes in the detection algorithm, that make concealment effective also in the hard-to-face case of high motion sequences.

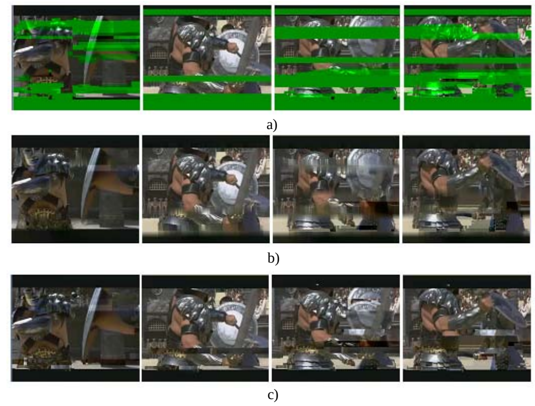

Figure 8: Adjacent frames decoded a) without concealment, b) by JM9.7 integrating the scene change detector without *delta* parameter, c) by JM9.7 integrating the scene change detector with *delta* parameter

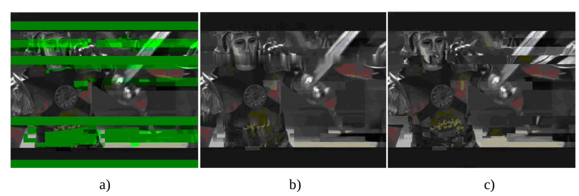

Figure 9: *Gladiator* sub-sequence decoded a) without concealment, b) by JM9.7 spatial concealment, c) by JM9.7 temporal concealment

REFERENCES

[1] G.S. Yu, M.M.K. Liu and M.W. Marcellin, "POCS-Based Error Concealment for Packet Video Using Multiframe Overlap Information," *IEEE Tran. Circuits Syst. Video Technol.*, vol. 8, n. 4, pp. 422–434, Aug. 1998.

[2] M.J. Chen, C.S. Chen and M.C. Chi, "Temporal Error Concealment Algorithm by Recursive Block-Matching Principle," *IEEE Tran. Circuits Syst. Video Technol.*, vol. 15, n. 11, pp. 1385–1393, Nov. 2005.

[3] http://iphome.hhi.de/suehring/tml/download/jm97.zip

[4] G. L. Gennari, G. A. Mian and L. Celetto, "A H.264 robust decoder with error concealment capabilities," *ST Journal of Research*, vol. 2, n. 1, pp. 67–82, Nov. 2005.

[5] F. Chiaraluce, L. Ciccarelli, E. Gambi and S. Spinsante, "Performance Evaluation of Error Concealment Techniques in H.264 Video Coding," *Proc. PCS04 Picture Coding Symposium*, December 15-17, 2004, San Francisco (CA), USA.