A Low Complexity Concealment Algorithm for H.264 Encoded Video over DVB-S2

Susanna Spinsante[1], Claudio Sacchi[2], and Ennio Gambi[1]

[1] DIBET, Universitá Politecnica delle Marche
Via Brecce Bianche 12, I-60131 Ancona, Italy
{s.spinsante,e.gambi}@univpm.it
[2] DISI, Universitá degli Studi di Trento
Via Sommarive 14, I-38123 Trento, Italy
sacchi@disi.unitn.it

Abstract. Due to the high level of compression, H.264/AVC video information is very sensitive to channel errors, and in particular the Variable Length Coding applied at the encoder side, together with Motion Compensation, can amplify the effects of transmission errors during the decoding phase. If, in addition, we consider that in real time applications lost or damaged data cannot be retransmitted, the importance to introduce concealment techniques in order to recover lost or corrupted video data emerges, through the exploitation of spatial and/or temporal correlation among the residual available information. In this paper we present a technique for polygonal edge interpolation, to improve the performance of Intra concealment in the case of video frames rich in details. Experimental results show that the proposed low complexity scheme may provide substantial improvements to the perceived video quality, confirmed by an increase in PSNR values.

Keywords: Satellite video broadcasting, video coding, video transmission, error concealment.

1 Introduction

The amazing spreading of Mobile TV, video-enabled cellular phones, and portable video terminals favored the broad adoption of the H.264/AVC video coding standard [1], mainly thanks to its increased compression performance (up to 60%) with respect to all the previous video compression techniques. For the same reason, the high compression efficiency attainable by H.264/AVC made it possible to deliver High Definition (HD) video over satellite, for example through DVB-S2 [2,3]. Considering the improvements in channel coding (LDPC) and modulation techniques (e.g. QPSK, 8PSK, 16APSK and 32APSK), DVB-S2 standard achieves a capacity increase of 30% compared to its predecessor DVB-S, under the same transmission conditions. The inclusion of H.264/AVC in DVB-S2 specification brings an additional gain of a factor 2 in bit rate at the same perceived video quality, when compared to the previously adopted MPEG-2 video compression. DVB-S2

G. Giambene and C. Sacchi (Eds.): PSATS 2011, LNICST 71, pp. 94–103, 2011.

already provides advanced features and characteristics targeted not only to consumer but also to professional applications, such as the increased number of modulation modes available, the Adaptive Coding and Modulation (ACM) tool, very powerful FEC techniques, new roll-off factor values for a tighter bandwidth shaping. The joint adoption of H.264/MPEG-4 coding further extends the range of supported services, targeted to both average viewers and professionals.

However, as well known, the transmission of compressed video streams over error-prone networks and channels (such as the RF broadcasting channels) is vulnerable to bit errors and packet losses that lead to data corruption and serious video quality deterioration. As a consequence, error concealment algorithms may be implemented at the receiver to enhance the quality of the reconstructed video. Error concealment has the attractive advantage of being encoder-independent: to compensate the effects of transmission errors, and to approximate the missing data, only the residual available correlated information in the received video signal is exploited, without the need of any specific customization at the encoder side. A rich technical literature exists in the field of error concealment techniques (just to mention a few examples, see [4,5,6]); many proposals relate to spatial only, or temporal only solutions, better known as Intra and Inter concealment, respectively, or introduce a mix of different strategies. Other solutions involve the adoption of image processing operators, to be applied to each single frame, usually borrowed from still image compression techniques.

This paper focuses on the development of an interpolation technique for Intra concealment, optimized for the case of polygonal edges, in order to overcome some limitations found in classical Intra approaches ([7,8]). Common techniques, such as bilinear and bi-cubic interpolations, are widely used for their low computational complexity; however they often cause blurring artifacts around the edges, which are annoying to visual perception. The proposed technique improves linear and directional interpolation, to efficiently and accurately recover lost macroblocks (MBs) containing a single, but polygonal, edge. A two-step approach is applied, by which a one-dimensional interpolation follows an edge detection phase. By means of the directional reconstruction, typical smoothing effects due to pixel interpolation are avoided, and image details preserved. Further extensions may include multi directional edges in a single MB that require a more complex analysis of the image information, to select the proper interpolation strategy. A major constraint in the design of such a technique for the recovery of polygonal edges is the limited complexity and computational load required, that are necessary in order to ensure real time applicability in video decoders. Many interpolation algorithms proposed in the literature for image recovery suffer from high complexity and computational requirements, and are not suitable in real time video applications, though able to provide very good performance in image reconstruction. Intra error concealment performance are especially efficient in the following cases: 1) high motion frame areas, for which the temporal recovery is not satisfactory, due to the very low correlation among subsequent scenes; 2) when data losses occur at scene changes in the video sequence; 3) when image variations are present at a local scale (MB level), and the most suitable interpolation algorithm is selected on the basis of the available information. Most of the Intra concealment solutions presented in the literature recovers the missing pixels of a MB by applying interpolation algorithms (such as linear or bilinear ones) on the pixels of the available neighboring and correctly received MBs. This is motivated by the fact that adjacent MBs in the same frame usually show minimal variations in their pixel values. An improvement to interpolation based approaches may be represented by a polygonal technique, which exploits the presence of two dominant directions in image areas rich in details, which usually feature geometric patterns, such as in the case of objects' edges.

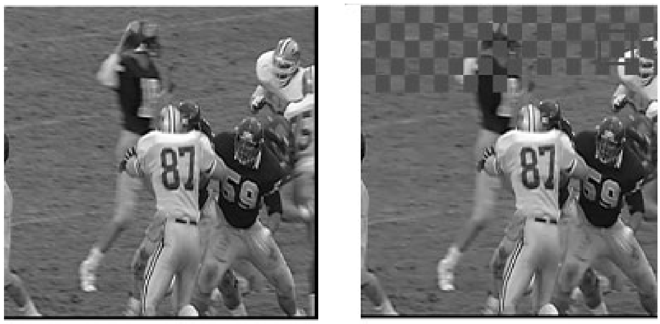
The paper is structured as follows: Section 2 will describe the proposed concealment technique; Section 3 will show some simulation results. Finally paper conclusions will be drawn in Section 4.

2 The Proposed Concealment Technique

The human visual system is particularly sensitive to edges' distortions, as those due to blurring or blockiness artifacts that may happen in a coded video frame. As a consequence, a polygonal interpolation process could substantially improve the reconstruction results provided by a linear or bilinear interpolation, by preserving the real edge directions in the image, and avoiding the generation of artificial edges. To such an aim, before moving to the true interpolation operation, it is necessary to perform a local, pixel-based analysis of the stream to decode, in order to preserve the original image structure and obtain an as much faithful as possible reconstruction of the video content. Obviously, a constraint on the computational effort required is always to be considered, especially for real-time applications. Available pixels located at the edges of the missing MB allow recovering information about the lost geometric pattern, and are used to reconstruct the missing pixels by a polygonal interpolation. The solution proposed in the following cannot be applied in any case, but only when a single, polygonal edge in the missing MB is to be recovered. In this sense, the linear interpolation scheme may be seen as a specific case of application of the polygonal interpolation solution. The edge direction detection step necessary to perform polygonal interpolation is executed by applying the well known Sobel operators, with a simple but effective modification with respect to their classical adoption. The Sobel operators S_x and S_y applied to pixels $P(x, y)$ of an image I provide the gradient components G_x and G_y to compute the gradient vector amplitude and direction:

$$
\begin{aligned}
|G| &= \sqrt{G_x^2 + G_y^2} \; . \\
\Theta &= tan^{-1}\left(G_y/G_x\right) \; . \qquad (1)
\end{aligned}
$$

If the edge located by the Sobel operators along the boundaries of the missing MB has a direction included in [-5°,+5°] horizontally, and [-85°,+85°] vertically, the edge is discarded, as it represents a direction parallel to the MB boundaries, i.e. an edge that does not intercept the missing MB itself. This edge does not

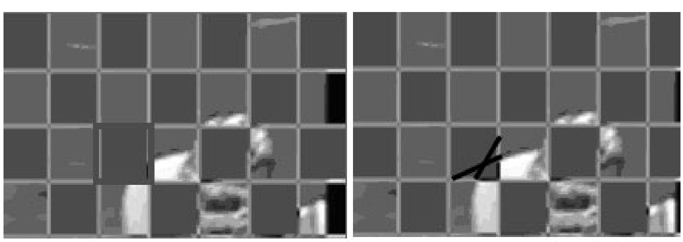


a) b)

Fig. 1. a) Original video frame, b) Corrupted video frame (the sample missing MB of interest is evidenced)

need to be reconstructed by the interpolation step. In order to better illustrate the proposed scheme, let us consider Figs. 1 a) and b), where a video frame and a corrupted version of it, featuring several missing MBs in a chessboard pattern, are shown. Suppose that the missing MB evidenced in Fig. 1 b) is to be recovered. As shown in the enlarged version of the image detail provided in Fig. 2 a), a correct reconstruction of the missing pixels may be accomplished only by a polygonal interpolation of the lost edge, the geometry of which is determined by the intersection of the two straight lines found by the edge detection algorithm, and the direction of which can be provided by an edge direction detection step. An example is shown in Fig. 2 b).

Given the quantities P_{1a}, Θ_{1a}, P_{1b} and Θ_{1b}, as represented in Fig. 3 a), it is possible to derive the analytical expression of the two lines r_1 and r_2 that divide the missing MB into four areas, as shown in Fig. 3 b). The proposed algorithm associates the missing pixel to one out of the four areas in the MB, according to the intersection points P_{r1}, P_{r2}, P_{r3}, and P_{r4} previously located (Fig. 3 c)). Considering, as an example, the case of a missing pixel in the area #1, only P_{r1}



a) b)

Fig. 2. a) Detail of the missing MB to recover, b) Edge direction detection on the missing MB

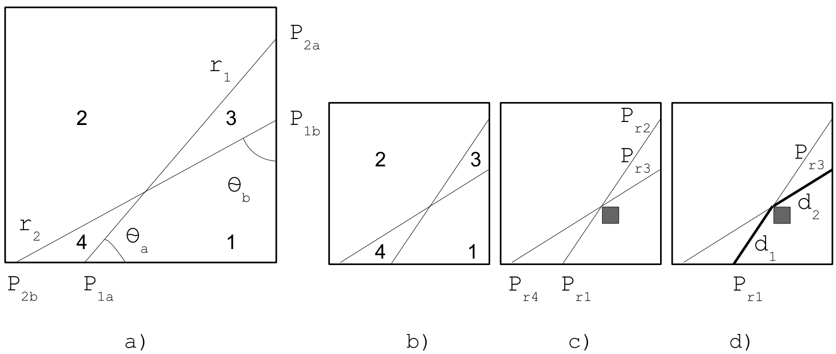


Fig. 3. a) Geometric representation of the polygonal interpolation scheme, b) The four areas located in the missing MB, c) Areas and pixel association, d) Geometric quantities for the polygonal pixel interpolation

and P_{r3} contribute the value of the interpolated pixel, according to their relative distances d_1 and d_2 (Fig. 3 d)) and the general formula (2), that is valid when n different points are used in the reconstruction:

$$p_0 = \frac{\sum_{i=1}^{n} d_i \cdot P_{ri}}{\sum_{i=1}^{n} d_i} . \tag{2}$$

Similar computations are applied when the missing pixels to recover are located in the other areas of the MB subjected to concealment.

3 Simulation Results

The interpolation technique proposed in this paper is integrated in a more structured concealment strategy, according to which a preliminary edge detection and classification step establishes the features of the missing MB to recover. Based on the outcome of this process, the proper Intra concealment technique is selected (bidirectional or polygonal). If the number of edges revealed inside the correctly received MBs adjacent to the missing MB is greater than 2, in order to search for a possible dominant direction to recover, a refinement step is applied. Such a refinement relies on the application of a Directional Entropy (DE) calculation block, to decide if the number of revealed directions (angles) denote a single dominant direction, a limited number of dominant directions, or so many directions that they cannot be accounted for. The DE block implements the following formula:

$$DE = -\sum_{i=1}^{n} p(d_x) \cdot log_2(p(d_x)) \tag{3}$$

where $p(d_x)$ is the probability density function of direction d_x, estimated by enumerating all the edges located by the edge detection step, and falling into 8 groups of directions, from -$\pi/2$ to $\pi/2$, in steps of $\pi/8$. A control parameter β is also defined as:

$$\beta = \frac{-\sum_{i=1}^{n} p(d_x) \cdot log_2(p(d_x))}{Max(DE)} \quad (4)$$

and compared against a set of thresholds, to differentiate the possible cases. In our experiments we empirically set threshold values at 0.3 and 0.75: by this choice, when $0 < \beta \leq 0.3$, a single directional edge area is located; when $0.3 < \beta < 0.75$, a textural multidirectional edge area is found; when $\beta \geq 0.75$, a complex multidirectional edge area is identified. In the case of complex multidirectional edge areas, the bilinear interpolation could be used, motivated by the multitude of different edges and directions featured by the image. In the textural multidirectional case, quite complicated to cope with, advanced algorithms with increased computational complexity should be conceived.

In order to evaluate the performance of the proposed concealment technique, the operations discussed have been integrated in the JM Reference Software v.14 decoder [9]. Simulations have been performed by encoding a synthetic video sequence obtained by collecting frames extracted by well known test sequences, showing different features and dynamic effects, and then applying a packet loss simulator in order to delete the 1%, 2%, 4%, and 10% amount of the original encoded data. Such values are reasonable, in the case of link degradation due to rain or other atmospheric impairments (like e.g. scintillations). Rain attenuation in high frequency band like Ka-band is a major factor for lowering the link capacity in satellite broadcasting services. The interesting analysis of packet-error-rate (PER) performance of DVB-S2 contained in [10] evidenced that PER curves achieved for all the 28 coded modulation modes of DVB-S2 have the typical waterfall pattern. This means that small SNR variations may result in strongly increased PERs. Each mode is characterized by a SNR threshold guaranteeing error-free transmission (See Tab.1 of [10]). However, the SNR margin allowed to have quasi error-free is very narrow (0.3 dB). A decrease of the SNR of 1 dB below the assigned threshold may involve PERs value larger than 10^{-2}. This is the reason why, error concealment is necessary to guarantee satisfactory QoS (Quality of Service) also in the presence of relevant packet loss. The application of a concealment technique at the decoder may enforce the performance provided by DVB-S2 through its ACM tool, that allows the transmission parameters to be changed on a frame by frame basis, depending on the particular conditions of the delivery path for each individual user. The H.264 video sequence is encoded in the Baseline profile, with a CIF 4:2:0 format, a Search Range of 16 pixels, 1 reference frame, Intra Period = 15, and a dispersed slice group map type. The behavior of the proposed concealment solution is compared to the performance obtainable by applying the default concealment approach included in the original Reference Software, that combines a bilinear interpolation scheme for the spatial concealment of all the type-I frames, and an enhanced boundary matching scheme for the Inter concealment of type-P frames. The applied scheme is selected on the basis of a motion index associated to the motion vectors computed during decoding. The global performances over the entire video sequence are evaluated with respect to the quality of the decoded sequence, described by the Peak Signal to Noise Ratio (PSNR) of the Y component, and the Structural Similarly Information quality metric (SSIM). The former is typically used to estimate the video quality in an objective way; the latter is a numerical tool to estimate subjective quality perceived by the user. A more precise definition for the PSNR would be "similarity to original", as it calculates a quality difference, by comparing each pixel in a distorted (reconstructed) frame to each pixel in the corresponding original frame. As the difference between frames becomes smaller, the PSNR index gets bigger. The PSNR index was calculated by the following formula:

$$PSNR = 10 \cdot log_{10} \frac{255^2 \cdot mn}{\sum_{i,j=1}^{m,n} (x_{ij} - y_{ij})^2} \tag{5}$$

where x and y represent the original and reconstructed frames, respectively, m and n are frame dimensions, i and j denote the pixel position in the frame. Only the luminance component of each pixel was considered for PSNR computation.

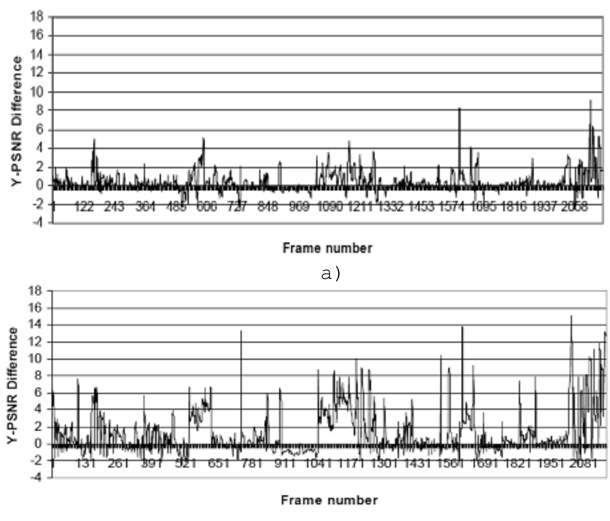


Fig. 4. Y-PSNR gain due to the proposed concealment framework, with respect to the original JM decoder, for a packet loss rate of a) 1% and b) 10%

Fig. 4 shows the performance of the modified decoder, with respect to the original one, expressed as the PSNR gain due to the proposed technique, averaged over 5 decoding iterations, for packet loss ratios of 1% (Fig. 4 a)) and 10% (Fig. 4 b)). On average, the proposed solution is able to improve the quality performance of the decoder in a remarkable way, and the improvement is more evident when the packet loss rate increases. On the other hand, the negative values for the PSNR shown in the graphs correspond to frames for which the proposed technique provides lower PSNR values than the ones obtained by

Table 1. Average Y-PSNR performance for different PER values: comparison among the concealment solutions

% PER	Avg. Y-PSNR (dB) JM Ref. Software	Avg. Y-PSNR (dB) Proposed
1%	36	36.39
2%	33.93	34.75
4%	31.31	32.42
10%	27.35	26.68

Table 2. Average Y-SSIM of the video sequence: comparison among the concealment solutions

Concealment technique	Y-SSIM (dB)
JM Reference	0.974
Proposed	**0.985**
Bilinear	0.975

the application of the original JM Reference Software concealment. A frame-by-frame analysis of the results, together with the visual inspection of these specific frames, reveal that the lower PSNR values are basically due to the Inter concealment strategy, which is not optimized (proper threshold setting is missing), and to noise effects that in some cases distort the edge detection performance. It is obvious that the concealment strategy included within the JM Reference Software is complete either in the Intra and Inter sections; the solution herein tested, on the contrary, has been optimized in its Intra component but not in the Inter one, which is the object of research activities currently ongoing. Table 1 shows how the quality gain provided by the proposed solution, in terms of average Y-PSNR over the whole video sequence, increases as the % PER affecting transmission increases, thus confirming the effectiveness of the concealment technique described in the paper, with respect to the solution included in the JM Reference Software. Results about the SSIM estimation confirm the better perceived video quality of the sequence concealed by means of the proposed technique, with respect to the original JM concealment, as reported in Table 2. A detailed analysis, performed frame by frame over the decoded sequence, shows that the most relevant quality gain is obtained for high motion sub-sequences, and frame areas rich in details, which confirms that the goals of the proposed concealment approach have been fulfilled. In such conditions the concealment process performed through polygonal interpolation provides very satisfactory results, because it is able to recover and preserve the details in the frame that most affect the video quality perceived by the user. An example is shown in Fig. 5, where a detail of the MB recovered by the proposed approach is presented, together with the whole concealed frame. In Fig. 5 a), the evidenced MB shows the edge of the player's shoulder reconstructed by the polygonal interpolation

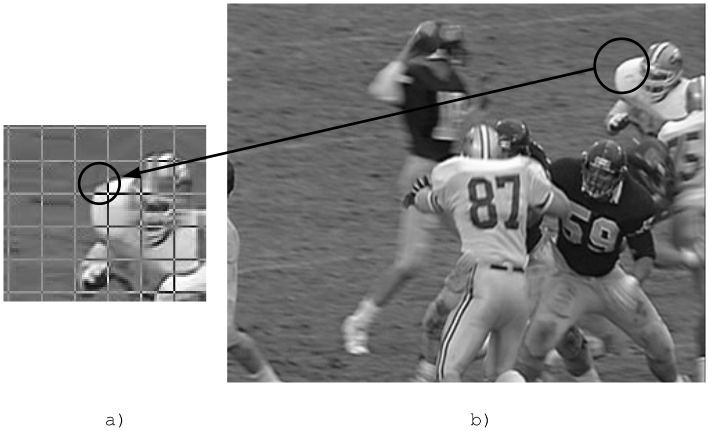


Fig. 5. Polygonal interpolation based concealment: a) Detail of the reconstructed MB, b) The concealed video frame

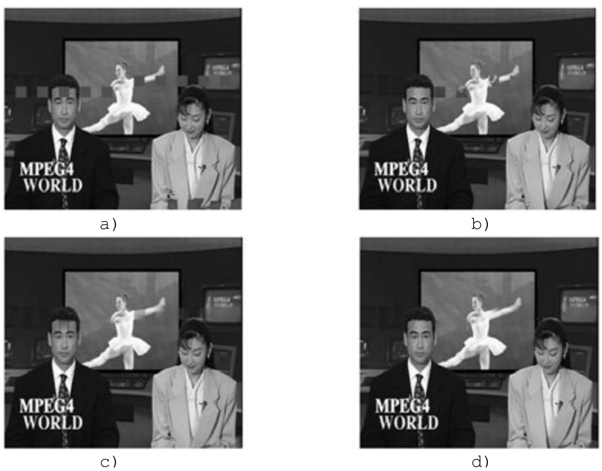


Fig. 6. Detail of the proposed scheme performance: a) Corrupted frame, b) JM concealment, c) Bilinear approach, d) Proposed concealment

scheme. The quality of the reconstruction makes the effects of video data losses not perceivable by the user.

A further example of the effective results provided by the proposed concealment, even at the expense of a limited complexity, is provided in Fig. 6, where a sample frame is shown, in which the suggested scheme applies different recovering strategies, according to the different properties exhibited by the frame sub areas at the decoder.

4 Conclusion

In this paper, a novel error concealment algorithm has been proposed for H.264 coded videos in the framework of DVB-S2 satellite TV broadcasting. The proposed algorithm relies on polygonal edge interpolation in order to improve Intra-frame concealment performance. Visual improvements in performance in terms of increased video quality are evident, also in the case of high packet error rates that are possible in satellite broadcasting, due to atmospheric and rain attenuations. It is shown that the gain in quality increases as the effects of channel impairments increase, manifested by higher percent values of the packet error rate.

References

1. ITU-T Recommendation H.264 (11/2009), Advanced video coding for generic audiovisual services — ISO/IEC 14496-10:2008, Information technology - Coding of audio-visual objects (MPEG-4) - Part 10: Advanced video coding, Version 13 (November 2009)
2. Digital Video Broadcasting (DVB): Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications, ETSI draft Specification, EN302 307, V1.1.1 (June 2004)
3. Morello, A., Mignone, V.: DVB-S2: The Second Generation Standard for Satellite Broadband Services. Proc. IEEE 94(1), 210–227 (2006)
4. Yi, J-W., Cheng, E., Yuan, F.: An Improved Spatial Error Concealment Algorithm Based on H.264. In: 3rd International Symposium on Intelligent Information Technology Application, pp. 455–458. (2009)
5. Beesley, S., Armstrong, A., Grecos, C.: An Edge Preserving Spatial Error Concealment Technique for the H.264 Video Coding Standard. In: Ph. D. Research in Microelectronics and Electronics, pp. 113–116 (2006)
6. Nemethova, O., Al-Moghrabi, A., Rupp, M.: Flexible Error Concealment for H.264 Based on Directional Interpolation. In: International Conference on Wireless Networks, Communications and Mobile Computing, pp. 1255–1260 (2005)
7. Weibin, Z., Zhiheng, Z., Zhiliang, X.: Error concealment for video sequences based on Sobel edge detector. In: International Conference on Communications, Circuits and Systems, pp. 723–726 (2008)
8. Senel, H.G.: Gradient Estimation Using Wide Support Operators. IEEE Trans. Image Proc. 18(4), 867–878 (2009)
9. JM Reference Software v.14.0, `http://iphome.hhi.de/suehring/tml/`
10. Chi, W.S., Seo, K.-D., Lee, I.K., Chang, D.-I.: Joint Source-Channel Coding Scheme for SVC-based DVB-S2 Satellite Broadcasting System. In: Proc. of 2010 Int. Conf. on Consumer Electronics (ICCE 2010), pp. 1–2 (2010)