ERROR CONCEALMENT OF INTRA CODED VIDEO FRAMES

M. Petrani\textsuperscript{2}, R. Bernardini\textsuperscript{2}, L. Celetto\textsuperscript{1}, G. Gennari\textsuperscript{1}, R. Rinaldo\textsuperscript{2}

\textsuperscript{1} STMicroelectronics - AST - Agrate Brianza (MI) - Italy
\textsuperscript{2} University of Udine, DIEGM, Udine, Italy

ABSTRACT
In this paper, we propose a procedure to recover lost slices in intra-coded H.264/AVC video frames. The procedure recovers missing pixels by exploiting only the video frame spatial redundancy, with no reference to past frames. This situation is interesting in correspondence to scene changes. The proposed scheme determines the presence of a dominant edge by analyzing the pixels in a neighbor of the lost slice. Missing pixels are reconstructed by interpolating in the direction of the dominant edge, if it is present, or by smooth interpolation. Experimental results show that the proposed scheme compares favorably with existing state-of-the-art methods, both objectively and subjectively.

1. INTRODUCTION
The recent H.264/AVC standard for video coding provides several techniques to mitigate the effects of packet losses due to transmission errors in unreliable channels \cite{1}. In particular, slice partitioning allows to confine the effects of a packet loss to a spatially limited region of a video frame. Appropriate error concealment techniques have to be adopted at the decoder to recover the missing pixels corresponding to the lost macroblocks in the slice \cite{2}. Excellent results can be obtained by using techniques that take advantage of the temporal correlation between the current and the past frames, and that use motion information to recover missing blocks in the current frame \cite{2,3}. In many cases, e.g., in correspondence to scene changes, the use of temporal information is however not recommended, and recovery has to be accomplished by using only spatial redundancy in the current frame. A similar problem arises for the transmission of block-based compressed images \cite{4}.

In this paper, we focus on the recovery of missing blocks by exploiting only the spatial redundancy of the video frame. For video decoding, our approach can be used together with a scene change detection procedure \cite{5}. In this application, the quality of the reconstructed frame is crucial, since any error propagates to the following frames due to video differential coding. Moreover, the proposed technique can be used also to correct missing blocks in still coded images.

A number of different approaches have been proposed to perform error concealment at the decoder. In particular, in \cite{6,7,8} blocks are reconstructed with the purpose to maximize the smoothness of
the image. Recovery using POCS (Projection Onto Convex Set) is considered in [9, 10], while [11] uses fuzzy logic techniques.

Edge detection can be used to help the reconstruction procedure. In [12], the Authors propose the use of edge-based interpolation. In particular, four possible edge directions (horizontal vertical and two diagonal) are considered for each missing block. The edge direction is estimated by using a differential operator and a thresholding process in the blocks surrounding the missing one. If one or more edges are found, the DCT coefficients of the missing block are computed as a weighted average of the those in the neighbor blocks.

In [4], the Authors estimate edges by computing local differences between adjacent pixels in the single-pixel border surrounding the missing block. The largest difference in each row/column is considered and compared to a threshold. Differences above the threshold are supposed to identify the initial and ending points of an edge crossing the missing block. If two edges are estimated for each block. Finally, interpolation is performed along the edge directions by a weighted average of the border pixels.

In this paper, we propose a technique that is similar in spirit to the one described in [4], with the important difference that we do not consider local differences, but estimate a dominant edge, if any, and interpolate along the edge direction. The dominant edge is estimated by means of the Sobel operator applied in a region of pixels surrounding the missing block. The procedure uses gradient information from all the pixels in this region, so that the edge detection procedure becomes reliable. In case no edge is detected, the missing region is smoothly interpolated from the pixel borders. The fact that edge-based interpolation is applied only if an edge is actually present, allows for improved consistency of the recovery process, as demonstrated by experimental results. In case a missing block intersects multiple edges, no dominant edge is detected, and the block is smoothly interpolated from the pixel borders.

This paper is organized as follows. Section 2 describes the proposed technique in some detail. Section 3 presents the experimental results, while Section 4 draws the conclusions.

2. THE PROPOSED EDGE-BASED INTERPOLATION PROCEDURE

We first describe the case of a single missing 16 × 16 macroblock (MB). We will explain later how we deal with the loss of entire slices. We consider a region \( R \) of pixels surrounding the missing block (see Fig. 1). We compute the gradient by convolving the pixels in \( R \) with the Sobel operators

\[
G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, \quad G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix},
\]

and obtain, for each pixel, the gradient \( \nabla I = [g_x, g_y] \), its magnitude \( G = \sqrt{g_x^2 + g_y^2} \) and direction \( \Theta = \arctan \left( \frac{g_y}{g_x} \right) \). The magnitude of the gradient is compared to a threshold \( \alpha \) and only the values greater than \( \alpha \) are considered for further processing. The direction of the surviving gradients is quantized in steps corresponding to a 10 degree angle. The procedure then computes the frequency of the quantized direction values, with the idea that a dominant edge can be detected only if one direction is more frequent than the others. In the experiments, a dominant edge is detected if the difference between the
two largest frequencies is greater than 10% of the largest frequency value.

\[
p = \frac{D_2}{D_1 + D_2} p_1 + \frac{D_1}{D_1 + D_2} p_2. \tag{1}
\]

When a dominant edge is not detected (smooth block), the area in the missing block is smoothly interpolated from the border. In particular, we compare in the experiments two possible strategies. The first one uses simple bilinear interpolation of the four pixels on the border surrounding the missing block, along the horizontal and vertical directions (see Fig. 1). More precisely, if we denote by \(p_{1h}\) and \(p_{2h}\) the closest available pixels in the horizontal direction, and by \(p_{1v}\) and \(p_{2v}\) the ones in the vertical direction, we set \(p = (p_h + p_v)/2\), where

\[
p_h = \frac{D_{2h}}{D_{1h} + D_{2h}} p_{1h} + \frac{D_{1h}}{D_{1h} + D_{2h}} p_{2h}, \tag{2}
\]

\[
p_v = \frac{D_{2v}}{D_{1v} + D_{2v}} p_{1v} + \frac{D_{1v}}{D_{1v} + D_{2v}} p_{2v}. \tag{3}
\]

The second interpolation strategy reconstructs missing pixels in such a way that the interpolated values are optimally “smooth”. The method is to solve the combinatorial Laplace equation with boundary conditions given by the known values [8].

In our experiments, we use the MATLAB® function roifill to this purpose.

When several contiguous blocks are missing, as it happens when a slice is lost in a video frame, we adapt the procedure outlined above by interpolating missing macroblocks one at a time, starting from the leftmost and rightmost slice macroblocks and converging to the center. In this way, we use reconstructed blocks to conceal the remaining ones. The procedure is simply modified by including in region \(\mathcal{R}\) only the available pixels. In this situation, however, it may happen that, for some dominant edge directions \(\theta\), one of the closest available pixels \(p_1\) and \(p_2\) in (1) is indeed too far away for meaningful interpolation (see Fig. 1). Thus, when \(D_i, i = 1, 2,\) is greater than a maximum value \(D_M\), we set \(D_i = +\infty\) in (1). The same procedure is adopted for interpolation via (2) and (3).

3. EXPERIMENTAL RESULTS

In all the experiments, the region \(\mathcal{R}\) consists of a border surrounding the missing block, with a thickness equal to 6 pixels. The maximum distance \(D_M\) is set to two times the block side, while the threshold \(\alpha\) for the gradient magnitude is set to 90. The procedure is not very sensitive to the values of these parameters.

The first set of experiments is relative to 512 × 512 grayscale images. For comparison purposes,
Table 1. PSNR comparison of concealment techniques.

<table>
<thead>
<tr>
<th></th>
<th>Lenna</th>
<th>Boat</th>
<th>Peppers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyun [4]</td>
<td>33.24 dB</td>
<td>31.38 dB</td>
<td><strong>34.88 dB</strong></td>
</tr>
<tr>
<td>Park [10]</td>
<td>34.65 dB</td>
<td>30.78 dB</td>
<td>29.87 dB</td>
</tr>
<tr>
<td>Shirani [14]</td>
<td>31.69 dB</td>
<td>29.22 dB</td>
<td>31.72 dB</td>
</tr>
<tr>
<td>Proposed method (BL)</td>
<td><strong>34.82 dB</strong></td>
<td>32.28 dB</td>
<td>33.77 dB</td>
</tr>
<tr>
<td>Proposed method (RF)</td>
<td>34.77 dB</td>
<td><strong>32.30 dB</strong></td>
<td>33.77 dB</td>
</tr>
</tbody>
</table>

Fig. 2. Original image Boat.

we consider the same setting and test images of [4], where the results of several techniques are reported. In particular, the loss pattern consists of a checkerboard of $8 \times 8$ pixels blocks, covering $24\%$ of the image area. As an example, Fig. 2 shows the original image Boat, while Fig. 3.a shows the loss pattern. Fig. 3.b shows the reconstructed image using the outlined procedure. Note that the visual quality of the reconstruction is good, and that edges in the image are correctly preserved.

Fig. 4.a shows the concealed image Lenna for the same loss pattern considered before. In Fig. 4.b, we show a detail of Lenna’s hat, where many blocks in the textured region intersect multiple edges. It can be seen that the reconstruction can be poor for some blocks, which are interpolated smoothly from border pixels. However, the overall quality of the picture in Fig. 4.a is acceptable.

Table 1 shows the reported results of the various techniques considered in [4], compared with the proposed solution. For the proposed method, we report in the table the PSNR reconstruction values corresponding to the use of bilinear interpolation (BL) and roifill (RF) for smooth blocks. We see from the table that the proposed approach compares favorably against all the other techniques.

The proposed technique does not provide the best results only for the image Peppers, which consists of large homogeneous areas separated by a few edges.

The second set of experiments considers $288 \times 352$ CIF frames. For reproducibility, the loss pattern is applied to the original frames, and no coding artifacts are taken into account. The block size is $16 \times 16$ in this case, and each image is a $18 \times$
22 macroblock matrix. Using matrix notation, the loss pattern consists of four slices corresponding to macroblocks (3,2:9), (7,15:22), (10,9:17), (17,3:13), for a total of 36 lost macroblocks. Fig. 5 shows the results for one frame of the video sequence Foreman. Fig. 5.a shows the original frame, while the error pattern and the reconstructed frame are shown in Fig. 5.b and Fig. 5.c, respectively. In Fig. 5.c, the smooth blocks are reconstructed using bilinear interpolation. We note that the visual quality of the reconstructed frame is acceptable, and that edges are correctly preserved, especially in the wall and helmet areas. The reconstruction PSNR is equal to 35.15 dB.

Fig. 6 shows the results relative to the CIF frame Container for the same error pattern considered before. Note that the ship’s mast is correctly reconstructed, and that the overall visual quality is acceptable. However, the reconstruction quality is poor in correspondence to slices (3,2:9) and (17, 3:13), where smooth blocks are reconstructed using bilinear interpolation. The PSNR is 28.73 dB in this case.

4. CONCLUSIONS

In this paper, we proposed an error concealment technique for lost blocks in images and video frames. For video, the proposed technique does not use temporal information but only the inherent spatial redundancy of each frame. This setting is interesting in case of scene changes, which can be detected with simple algorithms. The main idea behind the proposed technique is that better interpolation results can be obtained if edges are reliably detected in the image, and if the interpolation is performed along the estimated dominant edge direction. For smooth blocks, when no dominant edge is detected,
bilinear interpolation or optimally smooth reconstruction can be adopted. Experimental results show that the proposed technique compares favorably with existing techniques, both from a visual and objective quality point of view.

5. REFERENCES


Fig. 5. *Foreman* CIF: (a) original; (b) lost slices; (c) concealed frame.

Fig. 6. *Container* CIF: (a) original; (b) lost slices; (c) concealed frame.


