MOTION FIELD ESTIMATION BY COMBINED VECTOR RATIONAL AND BILINEAR INTERPOLATION FOR MPEG-2 ERROR CONCEALMENT

S. Tsekeridou†, F. Alaya Cheikh‡, M. Gabbouj‡, I. Pitas†

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

‡Signal Processing Laboratory
Tampere University of Technology
P.O. Box 553, FIN-33 101 Tampere, Finland
e-mail: {faouzi,moncef}@cs.tut.fi

ABSTRACT

The combined use of bilinear and vector rational interpolation for motion field estimation of erroneously received MPEG-2 video bitstreams is reported. Motion vector rational interpolation is capable of adapting its behaviour with respect to neighbouring motion information. Bilinear interpolation, operating on a finer interpolation grid, takes neighbouring spatial correlations into account. Thus, the combined method is expected to lead to improved motion estimates of the lost motion data and thus satisfactory concealment. It additionally proves to be adequately fast for real-time applications (MPEG-2 decoder). Simulation results prove the efficiency of the method compared to similar approaches.

1 INTRODUCTION

Transmission errors to highly compressed bitstreams (e.g. MPEG-2 video bitstreams) lead to observable deterioration of the decoded sequence quality. In the case of MPEG-2 coding, this is due to the use of VLC and differential coding. When an error occurs, the decoder is able to resynchronize only at the next resynchronization codeword in the received bitstream. Thus, all information in-between (prediction errors, coding modes, motion data) becomes "useless". Furthermore, due to the use of motion compensated prediction, temporal error propagation leads to even worse results.

Error concealment (EC) methods at the decoder side present a way to deal with such problems. A class of such methods are the motion field estimation ones. Among these, one can distinguish the motion compensated EC which calculates the average or vector median of adjacent motion vectors (MC-VM EC), the boundary matching algorithm EC (BMA EC) which selects the optimal motion vector with respect to the minimal boundary matching error [1], the motion vector estimation by boundary optimizing EC (MVE-BO EC) which defines search regions in reference frames and reestimates motion information [2], the bilinear motion field interpolation EC (BMFI EC) which estimates one motion vector per pixel instead of block using bilinear interpolation among adjacent motion vectors [3] and the combined BMFI EC (Comb. BMFI EC), an extension of the BMFI EC, which, additionally, exploits the outcome of the BMA EC [3].

The use of vector rational interpolation [4] in the estimation of erroneously received motion fields has proven to lead to remarkable concealment results and real-time implementation capabilities [5]. Motivated by the approach of [3], the motion vector rational interpolation scheme (MVRI EC) [5] has been extended to include bilinear interpolation principles in order to further exploit spatial correlations. Furthermore, a finer interpolation grid has been defined according to which one motion vector per 4×4 block inside the lost 16×16 MPEG-2 MB is estimated. Thus, 16 motion vectors (instead of 1) per lost MB are estimated. Smoother concealment, especially at the MB borders, is attained. The method is called motion vector rational-bilinear interpolation EC (MVRI-BMFI EC). Both spatial correlations (bilinear approach) and neighbouring motion similarities (rational approach) lead to better performance than the BMFI EC method. Furthermore, only few computations are added to the MVRI EC ones due to the identical spatial neighbourhood structure for all lost MBs. Spatial weights are only calculated once.

The organization of the paper is as follows. In section 2, a description of the vector rational-bilinear interpolation is given and its application to MPEG-2 error concealment is discussed. Simulation results are reported in section 3. Conclusions are drawn and possible extensions are discussed in section 4.

2 MOTION VECTOR RATIONAL-BILINEAR INTERPOLATION EC METHOD

Transmission errors to MPEG-2 coded bitstreams cause partial or entire slice information loss. Thus, when slices are a successive row of MBs, only correctly received top or bottom adjacent block data becomes available for the estimatiing lost information. The aim is to estimate the lost motion vector(s) in the best possible way using information from adjacent ones. The neighbourhood employed in the motion vector rational-bilinear interpolation approach is shown in Figure 1. It is identical with the neighbourhood considered by the MVRI EC method. The finer interpolation grid defined on the lost area is additionally illustrated in the same Figure. 16 motion vectors $\mathbf{v}_{ij}$ are estimated per lost MB, one per 4×4 block inside the lost region. Consequently, the fi-

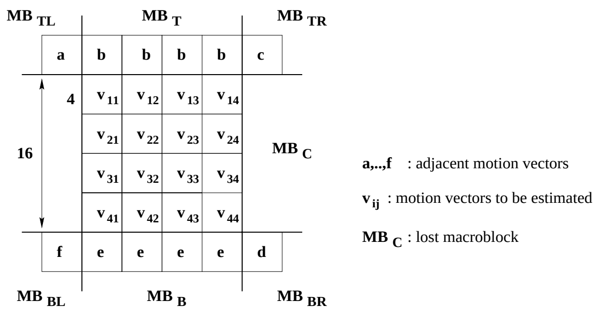

Figure 1: Graphical presentation of the estimation basis in the MVRI-BMFI method.

nal motion estimate $\mathbf{v}$ is a 4×4 matrix $\mathbf{v} = [\mathbf{v}_{ij}]$. The estimation scheme, a combination of vector rational and bilinear interpolation, involves the computation of each $\mathbf{v}_{ij}$:

$$\mathbf{v}_{ij} = \frac{\sum_{(\mathbf{u},\mathbf{v}) \in \{(\mathbf{a},\mathbf{d}),(\mathbf{b},\mathbf{e}),(\mathbf{c},\mathbf{f})\}} w_{\mathbf{uv}} (w_{ij\mathbf{u}} \mathbf{u} + w_{ij\mathbf{v}} \mathbf{v})}{\sum_{(\mathbf{u},\mathbf{v}) \in \{(\mathbf{a},\mathbf{d}),(\mathbf{b},\mathbf{e}),(\mathbf{c},\mathbf{f})\}} w_{\mathbf{uv}} (w_{ij\mathbf{u}} + w_{ij\mathbf{v}})} \tag{1}$$

for $i, j = 1, .., 4$. In the above equation, $\mathbf{a}, .., \mathbf{f}$ represent available neighbouring motion vectors as shown in Figure 1. Eq. (1) is an extension of the $2d$ vector rational interpolation case of [5]. $w_{\mathbf{uv}}$ are rational interpolation weights given by:

$$w_{\mathbf{uv}} = \frac{1}{1 + k||\mathbf{u} - \mathbf{v}||} \tag{2}$$

where $||.||$ denotes the L_2 vector norm and k is a positive constant that controls the degree of nonlinearity of the rational filter. $w_{\mathbf{uv}}$ are calculated only once per lost MB since they are independent of spatial locations. $w_{ij\mathbf{u}}$ represent weights based on spatial locations. For each $\mathbf{v}_{ij}$, a 2×3 matrix $W_{ij} = [w_{ij\mathbf{u}}]$, $\mathbf{u} \in \{\mathbf{a}, .., \mathbf{f}\}$ is defined that contains the spatial location weights. The structure of W_{ij} is proportional to the spatial localization of the 4×4 block motion vector with respect to the motion vector neighbourhood in Figure 1. Obviously, 16 such matrices have to be calculated, one per each $\mathbf{v}_{ij}$. They are defined as spatial location look-up tables because they need to be estimated only once due to the identical spatial structure of each $\mathbf{v}_{ij}$ with respect to adjacent motion vectors for different MBs. This results in computation savings. The spatial weight calculation is done in an identical manner as in bilinear interpolation. Distances from the estimated motion vector to its neighbour estimator are measured in the motion vector coordinate system. These are converted to integers for faster computations and less memory requirements.

This interpolation scheme attempts to estimate lost motion information in such a way that smoothness of motion is attained in smooth motion areas (linear operation performs best in such a case), whereas irregular motion of adjacent blocks does not result in high estimation errors (spatial correlations perform well in such cases). Thus, the method achieves to attain good motion field characteristics: motion smoothness and motion discontinuity, as long as adjacent motion data is sufficient for correct estimation. It should be noted at this point that intra-coded neighbours are simply considered as having zero motion vectors and that no coding mode information of adjacent blocks is actually exploited in the estimation process (all the other methods, except BMFI, use such kind of information).

After the lost motion vector matrix has been estimated, concealment of predictively coded frames is performed by copying the displaced, with respect to the estimated motion vector, block (size 4×4) of the previously decoded frame, to the respective lost one. This is done for all 4×4 blocks inside the lost MB. In the case of B-frames, where two motion fields are available (forward and backward motion fields), estimation is performed in both and the one that leads to the minimum boundary estimation error is selected for concealment. Intra-coded frames are concealed by the forward-backward block matching method [6].

3 SIMULATION RESULTS

In order to evaluate the performance of the motion vector rational-bilinear interpolation method (MVRI-BMFI EC), three different CCIR 601 sequences at 4:2:0 chroma sampling format have been used, namely the Flower Garden (125 frames), the Mobile & Calendar (40 frames) and the Football (50 frames) sequence. These have been coded at 5Mbps at 25 fps (PAL) using slice sizes equal to an entire row of macroblocks. A PER value of 2% has been considered. The locations of errors are assumed known. Objective performance evaluation of the presented concealment method is based on average $PSNR$ values whereas subjective evaluation is achieved by observing the visual quality of the concealed sequence. In order to assess the performance of the different motion field estimation processes, the *Motion Field Estimation Error*, MFE, is used [5]:

$$MFE = \frac{1}{b_x \times b_y} \sum_{x=1}^{b_x} \sum_{y=1}^{b_y} ||\mathbf{v}(x,y) - \mathbf{v}_{or}(x,y)|| \tag{3}$$

$b_x \times b_y$ represent the total number of block motion vectors in a frame. $\mathbf{v}_{or}$ is the original motion vector of the lost block whereas $\mathbf{v}$ is the estimated one. Finally, since fast concealment is a requirement, execution times have been calculated for the different concealment simulated methods.

Table 1 illustrates the average $PSNR$ values of the luminance component evaluated on the concealed test sequences by the MC-VM EC, the BMA EC, the MVE-BO EC, the BMFI EC, the Combined BMFI EC and the MVRI-BMFI EC method. These methods have been chosen to be compared because they lead to satisfactory and sometimes fast concealment. It can be seen that in almost all cases the MVRI-BMFI EC method attains the best result. The satisfactory performance of the MVRI-BMFI EC method can be further established by observing the achieved visual quality of part of the concealed frame of the Flower Garden sequence in Figure 2e compared with the visual quality of the concealed ones by the BMFI EC and the Combined BMFI EC methods shown in Figures 2c and 2d, respectively. It can be seen that the MVRI-BMFI EC behaves better than the BMFI EC method and leads to smoother concealment in areas where irregular motion is observed. Extension of the method to include boundary matching minimization might lead to better performance. Furthermore, the MVRI-BMFI EC method behaves better than the Comb BMFI EC especially in cases without entirely prevailent irregular motion as in Flower Garden and Mobile sequences. When the motion fields are mostly discontinuous, both MVRI-BMFI and Comb BMFI ECs behave almost similarly as in the case of the Football sequence. The Comb BMFI EC accounts for irregularities due to the additional use of BMA EC which may be thought to use zero order motion field interpolation when the method finds it as most appropriate. Zero order interpolation is a nice approach for handling and preserving discontinuities.

Table 1: Average $PSNR$ values measured on the three test sequences (Y component) after their concealment by the methods under study for a PER value equal to 2%.

EC Method	Flower	Mobile	Football
Error Free	29.75	35.50	32.40
MC-VM EC	26.65	33.59	27.92
BMA EC	26.33	33.61	28.10
MVE-BO	25.87	32.63	28.14
MVRI-BMFI	**27.32**	**33.84**	28.03
BMFI	26.06	32.90	27.63
Comb BMFI	26.49	33.57	**28.22**
Erroneous	13.88	20.14	17.52

In Table 2, the MFE average values over the entire sequences are shown. It is concluded that rather small errors are achieved by the MVRI-BMFI approach justifying our previous observation about their good adaptive behaviour with respect to local motion content. In order to visually comprehend the performance of the motion field estimation methods, Figure 3 shows part of the original, the erroneous and the estimated motion fields by the BMFI and the MVRI-BMFI approaches. The erroneous or estimated part is included in a drawn rectangular. The motion fields are illustrated in a dense manner using 16 vectors per MB to better assess the estimation results. It is seen that the discontinuity in the motion field is better preserved by the MVRI-BMFI approach than the BMFI one.

Table 2: Average MFE values measured on the estimated motion fields of the predictively coded frames of the three test sequences.

EC Method	Flower	Mobile	Football
MC-VM EC	3.93	2.35	9.36
BMA EC	4.54	2.52	8.72
MVE-BO	3.59	2.46	8.73
MVRI-BMFI	2.99	**1.96**	6.29
BMFI	**2.87**	2.07	**5.99**
Erroneous	8.66	2.60	8.01

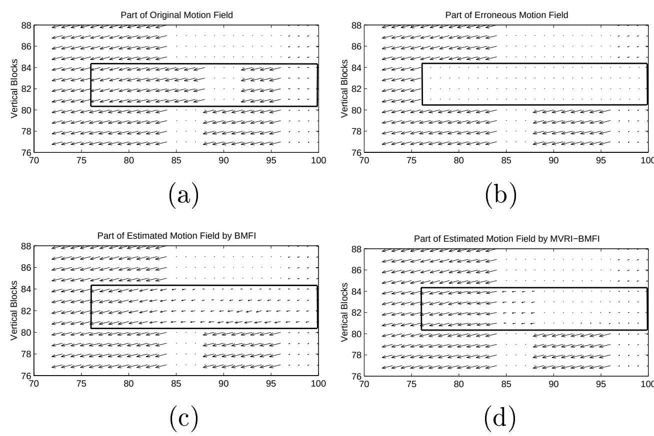

Figure 3: (a) Part of Backward Motion Field of Frame 8 (B-frame) of the Flower Garden Sequence, (b) Erroneous, $PER = 0.02$, Estimated by: (c) the BMFI EC and (d) the MVRI-BMFI EC.

Concerning processing time requirements, an evaluation has been performed for the different motion field estimation algorithms on an Ultra-1 Sun Sparc Workstation at 140 MHz. The results are tabulated in Table 3. It is seen that the MVRI-BMFI EC method is quite as fast as the BMFI EC method and remarkably faster than the Combined BMFI EC method, with which it otherwise exhibits similar performance. The MVE-BO EC requires too much processing time due to the search that it performs to re-estimate motion information.

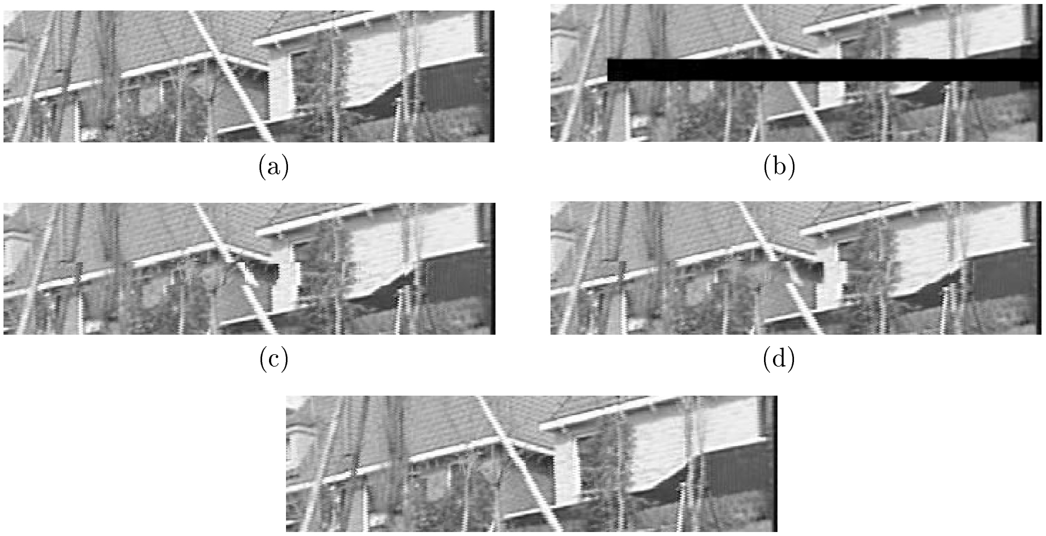

Figure 2: (a) Part of frame 25 (B-frame) of the Flower Garden Sequence, (b) Erroneous part, $PER = 0.02$, Concealed by: (c) the BMFI EC, (d) the Combined BMFI EC and (e) the MVRI-2d-BMFI EC.

Table 3: Execution times in seconds required for motion field estimation by the methods under study in order to achieve entire sequence concealment.

EC Method	Flower	Mobile
MC-VM EC	4.29	1.49
BMA EC	8.13	2.50
MVE-BO	289.32	89.27
MVRI-BMFI	3.76	1.05
BMFI	3.21	0.88
Comb BMFI	9.64	3.07

4 CONCLUSIONS

Motion field estimation by motion vector rational-bilinear interpolation has been investigated for error concealment purposes. Rational interpolation achieves to capture neighbouring motion smoothness or irregularities and adapt to the available information, whereas bilinear interpolation takes into account spatial correlations. Its performance is significantly better than the BMFI EC's performance. Although the performance of the Combined BMFI EC is similar to that of the MVRI-BMFI EC, the processing time requirements of the former are much higher than those of the latter method. Extensions of the MVRI-BMFI EC to incorporate different neighbourhood structures and boundary optimization might lead to even better field estimation. Additional use of other available neighbbouring information, such as prediction errors and coding modes, may result in better concealment.

Another application of the motion vector rational-bilinear interpolation, apart from concealment, could be the refinement of motion fields and their conversion from coarser to finer ones with very good estimation characteristics.

References

[1] W.-M. Lam, A.R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors", in *Proc. of 1993 Int. Conf. on Acoustics, Speech and Signal Processing*, Apr. 1993, vol. V, pp. 417–420.

[2] P. Salama, N.B. Shroff, and E.J. Delp, "Error concealment in encoded video", *submitted to IEEE Trans. on Image Processing*, 1998.

[3] M. Al-Mualla, N. Canagarajah, and D.R. Bull, "Temporal error concealment using motion field interpolation", *IEE Electronics Letters*, vol. 35, no. 3, pp. 215–217, Feb. 1999.

[4] L. Khriji, F. Alaya Cheikh, and M. Gabbouj, "High-resolution digital resampling using vector rational filters", *SPIE Optical Engineering*, vol. 38, no. 5, pp. 893–901, May 1999.

[5] S. Tsekeridou, F. Alaya Cheikh, M. Gabbouj, and I. Pitas, "Application of vector rational interpolation to erroneous motion field estimation for error concealment", in *Proc. of 1999 IEEE-EURASIP Wor. on Nonlinear Signal and Image Processing*, 20-23 June 1999.

[6] S. Tsekeridou and I. Pitas, "Mpeg-2 error concealment based on block matching principles", *IEEE Trans. on Circuits and Systems for Video Technology*, 1999, accepted for publication.