APPLICATION OF VECTOR RATIONAL INTERPOLATION TO ERRONEOUS MOTION FIELD ESTIMATION FOR 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
A study on the use of vector rational interpolation for the estimation of lost or erroneously received motion vectors of MPEG-2 coded video bitstreams for error concealment purposes is reported. Four different nonlinear interpolation schemes are examined. They are capable of adapting their behaviour to neighbouring motion information. Simulation results prove the success of their application to concealment of predictively coded frames. Such concealment is fast and thus adequate for real-time applications.

1. INTRODUCTION
Transmission of compressed video through physical communication channels may lead to information loss due to channel noise or congestion. Such loss further results in visual quality degradation of the decoded video sequence. When motion compensated coding is employed (MPEG-2 codec [1]), error propagation is inevitable leading to even worse visual results. Error concealment (EC) methods have emerged to deal with such problems at the decoder [2, 3, 4, 5]. Many of these attempt to estimate lost motion information of predictively coded frames and conceal lost blocks by motion compensated temporal replacement [2, 3, 4, 5]. Such methods are simple and fast and attain satisfactory concealment results.

Among the motion field estimation EC methods, one can distinguish the zero motion (ZM) EC, the motion compensated (MC) EC, the boundary matching algorithm (BMA) EC [2], the motion vector estimation by boundary optimizing (MVE-BO) EC [5] and the forward-backward block matching (F-B BM) EC [4]. ZM EC sets lost motion vectors to zero. MC EC estimates lost motion vectors from available adjacent ones by finding their mean (MC-AV) or vector median (MC-VM). It performs well when the assumption of smooth motion is valid among all neighbours or when adjacent blocks are not intra-coded. The remaining methods attempt to deal with the latter cases by selecting the best temporal neighbour that leads to a smooth spatial transition between the concealed block and its spatial neighbours. BMA EC defines a number of motion vector candidates and selects the one that leads to the minimum boundary matching error. MVE-BO EC and F-B BM EC define search regions in previously decoded frame(s). The former centers the search region around the block pointed at by the vector median of adjacent motion vectors and locates the optimal candidate by boundary matching error minimization. The latter locates the “best match” of adjacent blocks by MAD minimization. Concealment involves copying the block of the previously decoded frame, pointed at by the estimated motion vector, to the lost region in the current frame.

A study on the use of vector rational interpolation for the estimation of lost or erroneously received motion vectors of MPEG-2 coded video bitstreams for error concealment purposes is undertaken. Rational functions, i.e. the ratio of two polynomials, have been extensively used for image filtering and restoration [6, 7], enhancement [8, 9] and interpolation [10, 11] because of their desirable properties. Since rational function operators are able to adapt their behaviour with respect to the local source content, they exhibit remarkable performance in all previously mentioned applications. For this reason, their use in the estimation process of erroneous motion fields for error concealment purposes has been studied. Such EC method proves to be fast for real-time applications.

2. VECTOR RATIONAL INTERPOLATION
Vector rational interpolation has been introduced in [10, 11] for color image interpolation applications, where every pixel is considered as a 3-component vector in the considered color space.

3. MOTION VECTOR RATIONAL INTERPOLATION EC METHOD
Vector rational interpolation is employed in the estimation of erroneous motion fields of predictively coded frames of an MPEG-2 video bitstream (MVRI EC). Transmission errors lead to loss of the decoder synchronization and consequently to partial or entire slice information loss. Information loss is translated to missing motion vectors, coding modes and prediction errors of the respective macroblocks. Thus, only top or bottom correctly received adjacent block information is available for the estimation of the missing one. The neighbourhood employed in the motion vector rational interpolation has the structure shown in Figure 1. In a first approach, four different interpolation schemes have been considered [12] and are described in the sequel:

- 2-stage 1 – D Case (Figure 1a)
  Row Interpolation. Initially, estimates \( \mathbf{v}_T \) and \( \mathbf{v}_B \) of the top and bottom adjacent motion information are obtained by applying the 1 – D vector rational interpolation function:

\[
\mathbf{v}_T = \frac{w_{ab}(a + \frac{1}{2}b) + w_{bc}(c + \frac{1}{2}b)}{w_{ab} + w_{bc}}
\]  

(1)
This interpolation scheme estimates the lost motion vector $\mathbf{v}_{\text{lost}}$ by employing the expression:

$$
\mathbf{v}_{\text{lost}} = \frac{w_{\text{ab}}(\mathbf{a} + \mathbf{b}) + w_{\text{be}}(\mathbf{b} + \mathbf{e}) + w_{\text{ef}}(\mathbf{c} + \mathbf{f}) + w_{\text{ed}}(\mathbf{e} + \mathbf{d}) + w_{\text{af}}(\mathbf{a} + \mathbf{f}) + w_{\text{cf}}(\mathbf{c} + \mathbf{d})}{w_{\text{ed}} + w_{\text{ef}}}
$$

(8)

where $w_{\text{mn}}$ is given by (3).

All interpolation schemes attempt to estimate lost motion information in such a way that smoothness of motion is attained in smooth motion areas, whereas irregular motion of adjacent blocks does not result in high estimation errors. Intra-coded neighbours are simply considered as having zero motion vectors. No coding mode information of adjacent blocks is exploited (all other motion field estimation methods use such kind of information). In smooth motion areas, where Euclidean distances between adjacent vectors are small, the weights $w_{\text{mn}}$ are close to 1.0, thus leading to an averaging interpolator. When Euclidean distances increase, the respective weights decrease limiting thus the contribution of the respective candidate neighbouring pair in the final motion vector estimate.

In a second approach, the boundary matching criterion has been employed in order to locate that interpolation scheme, i.e. the "optimal" motion vector estimate, of the above mentioned four that leads to the best concealment with respect to the minimum boundary matching error:

$$
\mathbf{v}_{\text{opt}} = \arg \min_{\mathbf{v}} \sum_{i=0}^{N-1} \sum_{x=x_0}^{x=N-1} \left[ f(x + dx_i, y_0 + dy_i) - f(x, y_0 - 1) \right]^2
$$

(9)

$N \times N$ denotes the size of the block, $(x_0, y_0)$ the spatial coordinates of the top-left pixel of the lost block, $f$ the reference frame (forward or backward), $f_c$ the currently considered frame and $(dx_i, dy_i) = \mathbf{v}_i$ the estimated motion vector for the lost block by each one of the four interpolation schemes.

In a third approach, interpolation has been attempted in the direction of minimum change while preserving the transition in the direction of maximal change in an inverse manner than in the methods of [8, 9] in which enhancement of images is attempted using rational control functions of the rate of change. In order to find the rate and direction of change (maximal or minimal) an approach similar to the one proposed in [13] has been adopted. Let $\mathbf{x}, \mathbf{y} : \mathbb{R}^2 \to \mathbb{R}^2$ be a two-valued two-dimensional function denoting the estimated motion field of a frame, i.e. $\mathbf{MF}(x, y) = [Dx(x, y), Dy(x, y)]^T$, where $Dx$ and $Dy$ represent the functions of the horizontal and vertical displacements composing the motion vectors at points $(x, y)$. When the Euclidean distance of two points $(x_0, y_0)$ and $(x_1, y_1)$ tends to zero, then the difference of the values of $\mathbf{MF}(x, y)$ at those points, $\Delta \mathbf{MF} = \mathbf{MF}(x_0, y_0) - \mathbf{MF}(x_1, y_1)$, becomes the arc element:

$$
d\mathbf{MF} = \frac{\partial \mathbf{MF}}{\partial x} dx + \frac{\partial \mathbf{MF}}{\partial y} dy
$$

(10)

Its squared norm, known as the first fundamental form, is given by:

$$
d\mathbf{MF}^2 = \begin{bmatrix} dx \\ dy \end{bmatrix}^T \begin{bmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix}
$$

(11)

where $(g_{xy} = g_{yx})$:

$$
g_{xx} = \frac{\partial^2 \mathbf{MF}}{\partial x^2}, \quad g_{xy} = \frac{\partial^2 \mathbf{MF}}{\partial x \partial y}, \quad g_{yy} = \frac{\partial^2 \mathbf{MF}}{\partial y^2}
$$

(12)

Figure 1: MVRI Schemes: (a) 2-stage 1 - $D$ case, (b) 2 - $D$ case, (c) 2-stage combined 1 - $D$ and 2 - $D$ case and (d) 2 - $D$ case considering all directions.

Column Interpolation of Row Estimates. The lost motion vector is finally estimated by averaging the row estimates $\mathbf{v}_T$ and $\mathbf{v}_B$:

$$
\mathbf{v} = \frac{\mathbf{v}_T + \mathbf{v}_B}{2}
$$

(4)

○ 2 - $D$ Case (Figure 1b)

This interpolation scheme estimates the lost motion vector $\mathbf{v}$ by employing the expression:

$$
\mathbf{v} = \frac{w_{\text{ab}}(\mathbf{a} + \mathbf{d}) + w_{\text{be}}(\mathbf{b} + \mathbf{e}) + w_{\text{ef}}(\mathbf{c} + \mathbf{f}) + w_{\text{ed}}(\mathbf{e} + \mathbf{d}) + w_{\text{af}}(\mathbf{a} + \mathbf{f}) + w_{\text{cf}}(\mathbf{c} + \mathbf{d})}{w_{\text{ed}} + w_{\text{ef}} + w_{\text{af}} + w_{\text{cf}} + w_{\text{be}}}
$$

(5)

where $w_{\text{mn}}$ is given by (3).

○ 2-stage Combined 1 - $D$ and 2 - $D$ Case (Figure 1c)

Row Interpolation. It is performed in the same way as in the first stage of the 1 - $D$ case.

Combined Interpolation. The row estimates of the previous stage are used as input vectors in combination with the input vectors of the 2 - $D$ case to estimate the lost motion vector:

$$
\mathbf{v} = \frac{w_{\text{ab}}(\mathbf{a} + \mathbf{d}) + w_{\text{be}}(\mathbf{b} + \mathbf{e}) + w_{\text{ef}}(\mathbf{c} + \mathbf{f}) + w_{\text{ed}}(\mathbf{e} + \mathbf{d}) + w_{\text{af}}(\mathbf{a} + \mathbf{f}) + w_{\text{cf}}(\mathbf{c} + \mathbf{d})}{w_{\text{ed}} + w_{\text{ef}} + w_{\text{af}} + w_{\text{cf}} + w_{\text{be}}}
$$

(6)

where $w_{\text{mn}}$ is again given by (3).

○ 2 - $D$ Case of All Directions (Figure 1d)

This interpolation scheme is an extension of the 2 - $D$ case in the sense that almost all directions between neighbours are considered in the final estimation:

$$
\mathbf{v} = \frac{w_{\text{ab}}(\mathbf{a} + \mathbf{d}) + w_{\text{be}}(\mathbf{b} + \mathbf{e}) + w_{\text{ef}}(\mathbf{c} + \mathbf{f}) + w_{\text{ed}}(\mathbf{e} + \mathbf{d}) + w_{\text{af}}(\mathbf{a} + \mathbf{f}) + w_{\text{cf}}(\mathbf{c} + \mathbf{d})}{w_{\text{ed}} + w_{\text{ef}} + w_{\text{af}} + w_{\text{cf}} + w_{\text{be}}}
$$

(7)
According to [13], $dM^2$ is a measure of the rate of change in a prespecified direction. The extrema of (11) are obtained in the direction of the eigenvectors of the matrix $\begin{bmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{bmatrix}$ and the values attained there are the corresponding eigenvalues. Thus, the eigenvectors provide the direction of maximal/minimal rate of change ($\lambda_+$ and $\lambda_-$ respectively). In order to $\lambda_+$ and $\lambda_-$ in different directions, six different "edge-sensing" masks have been employed for the evaluation of $g_{xx}$, $g_{xy}$ and $g_{yy}$. The selection of these masks was done bearing in mind the singularities of the lost motion information problem (e.g. no horizontal neighbouring motion vector information).

The masks are shown in Table 1. For every type, a minimum and a maximum rate of change, $\lambda_{+1}(i)$ and $\lambda_{-1}(i)$, $i = 1, ..., V$, respectively, are estimated for the specific direction defined by the employed mask. The differences $\lambda_{+1}(i) - \lambda_{-1}(i)$ are a good measure for detecting transitions in the motion field. Thus, they have been chosen to control the estimation of the weights employed in the motion vector rational interpolation among the candidate motion vector estimates defined for every type as the average of the considered neighbouring motion vectors. For example, for type I, the candidate motion vector estimate, $v(x,y)$, is given by $v(x,y) = (a + b + c + d + e + f)/6$. In case of motion uniformity among neighbours, their average presents a good estimate of the lost motion vector. The final motion vector estimate is given by:

\[
\mathbf{v} = \frac{\sum_{i=1}^{V} w(i) \mathbf{v}(i)}{\sum_{i=1}^{V} w(i)}
\]

(13)

where the weights $w(i)$ are estimated by:

\[
w(i) = \frac{1}{1 + k(\lambda_{+1}(i) - \lambda_{-1}(i))}
\]

(14)

After the lost motion vector has been estimated, concealment of predictively coded frames is performed by copying the displaced, with respect to the estimated motion vector, block of the previously decoded frame to the current lost one. In the case of B-frames, where two motion fields are available (forward and backward motion fields), estimation is accomplished in both and the one that leads to the minimum boundary estimation error is selected for concealment. Intra-coded frames are concealed by the F-B BM EC [3, 4]. Motion vector rational interpolation can also be employed for recovering lost concealment motion vectors of I-frames.

4. SIMULATION RESULTS

In order to evaluate the performance of the MVRI EC method, 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) sequences. 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 error locations are assumed known. Objective performance evaluation 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 incorporated in the concealment methods, the Motion Field Estimation Error (MFE) is introduced:

\[
MFE = \frac{1}{b_x \times b_y} \sum_{x=1}^{b_x} \sum_{y=1}^{b_y} ||v(x,y) - v_{cor}(x,y)||
\]

(15)

In (15), $b_x \times b_y$ represent the total number of block motion vectors in a frame. $v_{cor}$ is the original motion vector of the lost block (generated by the codec) whereas $v$ is the estimated one.

Table 2 illustrates the average PSNR values of the Y component evaluated on the concealed test sequences by the EC methods under study. It can be seen that, in almost all cases, the MVRI

<table>
<thead>
<tr>
<th>EC Method</th>
<th>Flower</th>
<th>Mobile</th>
<th>Football</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Free</td>
<td>29.751</td>
<td>35.504</td>
<td>32.398</td>
</tr>
<tr>
<td>ZM EC</td>
<td>24.104</td>
<td>31.120</td>
<td>26.955</td>
</tr>
<tr>
<td>MC-AV EC</td>
<td>26.476</td>
<td>33.133</td>
<td>27.884</td>
</tr>
<tr>
<td>MC-VM EC</td>
<td>26.646</td>
<td>33.588</td>
<td>27.918</td>
</tr>
<tr>
<td>BMA EC</td>
<td>26.333</td>
<td>33.614</td>
<td>28.096</td>
</tr>
<tr>
<td>MVE-BO</td>
<td>25.865</td>
<td>32.626</td>
<td>28.135</td>
</tr>
<tr>
<td>MVRI-1-D</td>
<td>27.215</td>
<td>33.553</td>
<td>28.063</td>
</tr>
<tr>
<td>MVRI-2-D</td>
<td>27.309</td>
<td>33.852</td>
<td>28.021</td>
</tr>
<tr>
<td>MVRI-Combined</td>
<td>27.276</td>
<td>33.817</td>
<td>28.015</td>
</tr>
<tr>
<td>MVRI-2-D-All</td>
<td>27.237</td>
<td>33.787</td>
<td>28.007</td>
</tr>
<tr>
<td>MVRI-BM</td>
<td>27.261</td>
<td>33.894</td>
<td>28.121</td>
</tr>
<tr>
<td>MVRI-RoC</td>
<td>27.294</td>
<td>33.843</td>
<td>28.030</td>
</tr>
<tr>
<td>F-B BM EC</td>
<td>27.736</td>
<td>33.849</td>
<td>28.396</td>
</tr>
<tr>
<td>Erroneous</td>
<td>13.878</td>
<td>20.140</td>
<td>17.524</td>
</tr>
</tbody>
</table>
EC method (of either approach, i.e. four interpolation schemes or selection of best based on boundary matching (MVRI-BM) or optimal weighting based on minimum rate of change (MVRI-RoC)) attains the second best result. The satisfactory performance of the novel MVRI EC method can be further established by observing the achieved visual quality of the concealed frames in Figure 2 for the same sequence. Noticeable shifts are avoided when lost motion information is reconstructible by adjacent data and concealment is performed smoothly.

In Table 3, the average values of the $MFE$ Errors are shown. It is seen that the smallest errors are achieved by the MVRI schemes justifying our previous observation about their good adaptive behaviour with respect to local motion content. Figure 3 depicts the estimated motion fields. In the erroneous field the horizontal continuous lines denote the locations of lost motion information. It should be noted that the MVRI EC method performs recursively.

For the Flower Garden sequence, which exhibits large uniform motion of the tree and smaller uniform motion of the background, estimation is very well performed. Small irregular motion in the Mobile & Calendar sequence is also well estimated by the MVRI EC method. The rather irregular motion of the Football sequence can hardly be well estimated by any motion field estimation process but the MVRI method does not introduce large estimation errors and performs a smooth transition between differently moving objects.

The last aspect that has been examined is the processing time of the EC methods under study. Table 4 illustrates the execution times in secs required for the total concealment of the test sequences. Simulations were executed under an Ultra-1 Sun Sparc Workstation at 143MHz. The processing time calculation has been performed on the concealment of the predictively coded frames. It is seen that methods using a search region (MVE-BO EC, F-B BM EC) are actually rather time consuming whereas the MVRI EC methods attain a remarkably fast concealment suitable for real-time applications.

<table>
<thead>
<tr>
<th>EC Method</th>
<th>Flower</th>
<th>Mobile</th>
<th>Football</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZM EC</td>
<td>8.661</td>
<td>2.604</td>
<td>8.006</td>
</tr>
<tr>
<td>MC-A V EC</td>
<td>3.559</td>
<td>2.144</td>
<td>7.724</td>
</tr>
<tr>
<td>MC-VM EC</td>
<td>3.934</td>
<td>2.345</td>
<td>9.362</td>
</tr>
<tr>
<td>BMA EC</td>
<td>4.542</td>
<td>2.516</td>
<td>8.717</td>
</tr>
<tr>
<td>MVE-BO</td>
<td>3.586</td>
<td>2.456</td>
<td>8.727</td>
</tr>
<tr>
<td>MVRI-1 – D</td>
<td>2.808</td>
<td>1.932</td>
<td>6.059</td>
</tr>
<tr>
<td>MVRI-2 – D</td>
<td>2.995</td>
<td>1.945</td>
<td>6.342</td>
</tr>
<tr>
<td>MVRI-Combined</td>
<td>2967</td>
<td>1.930</td>
<td>6.277</td>
</tr>
<tr>
<td>MVRI-2 – D-All</td>
<td>3.102</td>
<td>1.996</td>
<td>6.597</td>
</tr>
<tr>
<td>MVRI-RoC</td>
<td>2.976</td>
<td>1.944</td>
<td>6.267</td>
</tr>
<tr>
<td>Erroneous</td>
<td>8.661</td>
<td>2.604</td>
<td>8.006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EC Method</th>
<th>Flower</th>
<th>Mobile</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZM EC</td>
<td>0.84</td>
<td>0.23</td>
</tr>
<tr>
<td>MC-AV EC</td>
<td>4.05</td>
<td>1.31</td>
</tr>
<tr>
<td>MC-VM EC</td>
<td>4.29</td>
<td>1.49</td>
</tr>
<tr>
<td>BMA EC</td>
<td>8.13</td>
<td>2.50</td>
</tr>
<tr>
<td>MVE-BO</td>
<td>289.32</td>
<td>89.27</td>
</tr>
<tr>
<td>MVRI-1 – D</td>
<td>2.84</td>
<td>0.93</td>
</tr>
<tr>
<td>MVRI-2 – D</td>
<td>3.10</td>
<td>0.90</td>
</tr>
</tbody>
</table>
5. CONCLUSIONS

Motion field estimation by vector rational interpolation has been investigated for error concealment purposes. In a first approach, four different interpolation schemes have been examined. In a second approach, the optimal estimate is searched for based on boundary matching minimization. In a third approach, the minimum rate of change and the respective direction have been incorporated in the interpolation scheme. The motion vector rational interpolation error concealment method has been found to perform well and be fast. The interpolator adapts its behaviour according to the local motion information, thus leading to a well estimated motion field.

6. REFERENCES