MOTION FIELD ESTIMATION BY VECTOR RATIONAL INTERPOLATION FOR ERROR CONCEALMENT PURPOSES

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

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

‡Signal Processing Laboratory
Tampere University of Technology
P.O. Box 553, FIN-33 101 Tampere, Finland

ABSTRACT

A study on the use of vector rational interpolation for the estimation of erroneously received motion fields of an MPEG-2 coded video bitstream has been performed. Four different motion vector interpolation schemes have been examined using motion information from available top and bottom adjacent blocks since left or right neighbours are usually lost. The presented interpolation schemes are capable of adapting their behaviour according to neighbouring motion information. Simulation results prove the satisfactory performance of the novel nonlinear interpolation schemes and the success of their application to the concealment of predictively coded frames. The motion vector rational interpolation concealment method proves to be a fast method, 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 possible loss of visual quality of the decoded video sequence. When motion compensated coding has been employed, which is the case with the MPEG-2 compression algorithm [1], errors observed in a frame propagate to subsequent in coding order frames leading, thus, to even worse visual quality. Error concealment methods have emerged to deal with such problems at the decoder [2, 3, 4, 5]. Many of these methods attempt to estimate lost motion information for predictively coded frames (P- or B-frames) and conceal lost blocks by motion compensated temporal replacement [2, 3, 4, 5]. Such methods prove to be fast and attain satisfactory concealment results.

Among the motion field estimation error concealment (EC) methods, one can distinguish the *zero motion EC* (ZM EC), the *motion compensated EC* (MC EC), the *boundary matching algorithm EC* (BMA EC) [2], the *motion vector estimation by boundary optimizing EC* (MVE-BO EC) [5] and the *forward-backward block matching EC* (F-B BM EC) [4]. The *ZM EC* sets the lost motion vectors to zero. The *MC EC* estimates the lost motion information from adjacent existing information by either averaging (MC-AV) or finding the vector median (MC-VM). It performs well when the assumption of smooth motion among all neighbours is valid 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. The *BMA EC* defines a number of motion vector candidates for the lost one and selects the one that leads to the minimum boundary matching error. The *MVE-BO EC* and *F-B BM EC* define search regions in the previously decoded frame. The first method centers the search region around the block pointed at by the vector median of the neighbouring motion vectors and locates the optimal candidate with respect to the minimum boundary matching error. The second method locates the "best match" of the adjacent blocks (block matching) by MAD minimization. For all methods, concealment is performed by 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 erroneously received motion fields of an MPEG-2 coded video bitstream for error concealment purposes is undertaken in this paper. Rational functions, i.e. the ratio of two polynomials, have been extensively used for image filtering and restoration, image enhancement as well as image interpolation because they are universal approximators, good extrapolators, able to be trained using a linear algorithm and requiring lower degree terms than Volterra expansions [6, 7]. Since rational function operators are able to adapt their behaviour with respect to the local source content, they exhibit remarkable performance in the previously mentioned applications. For this reason, their use in the estimation process of erroneous motion fields by vector rational interpolation for error concealment has been investigated. It proves to be fast enough for real-time applications and produces visually satisfactory concealment without the use of available coding mode information [6, 7].

2. VECTOR RATIONAL INTERPOLATION

Vector rational interpolation has been introduced in [6, 7] for color image interpolation applications, where every pixel is considered as a 3-component vector in the considered color space. The input/output relationship of the vector rational function is defined by:

$$\mathbf{y} = \frac{\alpha_0 \mathbf{1} + \sum_{j=1}^{m} \alpha_{1j}\mathbf{x}_j + \sum_{j=1}^{m}\sum_{k=1}^{m} \alpha_{2jk}\mathbf{x}_j\mathbf{x}_k + \cdots}{\beta_0 \mathbf{1} + \sum_{j=1}^{m} \beta_{1j}\mathbf{x}_j + \sum_{j=1}^{m}\sum_{k=1}^{m} \beta_{2jk}\mathbf{x}_j\mathbf{x}_k + \cdots} \quad (1)$$

where $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m$ are l-component input vectors and α_0, β_0, α_{ij}, β_{ij} are the coefficients defining the rational function. If $l = 1$ then equation (1) reduces to the scalar rational function input/output relationship. In (1), $\alpha\mathbf{x} = [\alpha x_1, \cdots, \alpha x_l]^T$, $\mathbf{xz} = [x_1 z_1, \cdots, x_l z_l]^T$, $\mathbf{x} + \mathbf{z} = [x_1 + z_1, \cdots, x_l + z_l]^T$ and $\frac{\mathbf{x}}{\mathbf{z}} = [\frac{x_1}{z_1}, \cdots, \frac{x_l}{z_l}]^T$.

3. MOTION VECTOR RATIONAL INTERPOLATION CONCEALMENT METHOD

Vector rational interpolation is employed in the estimation of erroneously received 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 modes of the respective macroblocks. Thus, only top or bottom correctly received neighbouring block information is available for estimating the missing one. The neighbourhood employed in the motion vector rational interpolation has the structure shown in Figure 1. Four different interpolation schemes have been considered and are

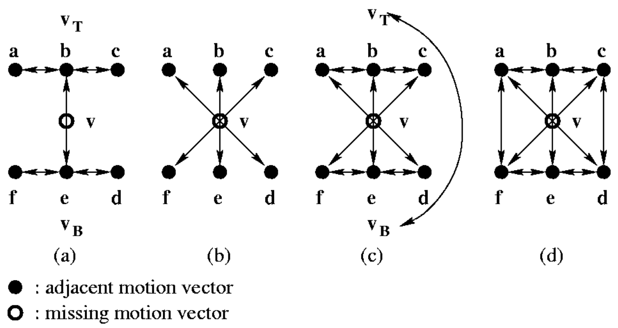


Figure 1: MVRI Schemes: (a) 2-stage $1-D$ case, (b) $2-D$ case, (c) 2-stage combined $1-D$ and $2-D$ case and (d) $2-D$ case considering all directions.

described in the sequel:

◇ **2-stage $1-D$ Case (Figure 1a)**

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

$$\mathbf{v}_T = \frac{w_{\mathbf{ab}}(\mathbf{a}+\frac{1}{2}\mathbf{b}) + w_{\mathbf{bc}}(\mathbf{c}+\frac{1}{2}\mathbf{b})}{\frac{3}{2}(w_{\mathbf{ab}}+w_{\mathbf{bc}})}$$

$$\mathbf{v}_B = \frac{w_{\mathbf{de}}(\mathbf{d}+\frac{1}{2}\mathbf{e}) + w_{\mathbf{ef}}(\mathbf{f}+\frac{1}{2}\mathbf{e})}{\frac{3}{2}(w_{\mathbf{de}}+w_{\mathbf{ef}})} \qquad (2)$$

where the coefficients $w_{\mathbf{uw}}$ are defined by ($\mathbf{u}, \mathbf{w}$, $\mathbf{u} \neq \mathbf{w}$ take values from $\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}\}$):

$$w_{\mathbf{uw}} = \frac{1}{1+k\|\mathbf{u}-\mathbf{w}\|} \qquad (3)$$

In equation (3), $\|.\|$ denotes the Euclidean distance and k is a positive constant that controls the degree of nonlinearity of the rational filter.

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

$$\mathbf{v} = \frac{\mathbf{v}_T + \mathbf{v}_B}{2} \qquad (4)$$

◇ **$2-D$ Case (Figure 1b)**

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

$$\mathbf{v} = \frac{w_{\mathbf{ad}}(\mathbf{a}+\mathbf{d}) + w_{\mathbf{be}}(\mathbf{b}+\mathbf{e}) + w_{\mathbf{cf}}(\mathbf{c}+\mathbf{f})}{2(w_{\mathbf{ad}}+w_{\mathbf{be}}+w_{\mathbf{cf}})} \qquad (5)$$

where $w_{\mathbf{uw}}$ 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:

$$\mathbf{v} = \frac{w_{\mathbf{ad}}(\mathbf{a}+\mathbf{d}) + w_{\mathbf{be}}(\mathbf{b}+\mathbf{e}) + w_{\mathbf{cf}}(\mathbf{c}+\mathbf{f}) + w_{\mathbf{v}_T\mathbf{v}_B}(\mathbf{v}_T+\mathbf{v}_B)}{2(w_{\mathbf{ad}}+w_{\mathbf{be}}+w_{\mathbf{cf}}+w_{\mathbf{v}_T\mathbf{v}_B})} \qquad (6)$$

where $w_{\mathbf{uw}}$ 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:

$$\mathbf{v} = \frac{w_{\mathbf{ad}}(\mathbf{a}+\mathbf{d}) + w_{\mathbf{be}}(\mathbf{b}+\mathbf{e}) + w_{\mathbf{cf}}(\mathbf{c}+\mathbf{f}) + w_{\mathbf{ab}}(\mathbf{a}+\mathbf{b}) + w_{\mathbf{bc}}(\mathbf{b}+\mathbf{c}) + w_{\mathbf{fe}}(\mathbf{f}+\mathbf{e}) + w_{\mathbf{ed}}(\mathbf{e}+\mathbf{d}) + w_{\mathbf{af}}(\mathbf{a}+\mathbf{f}) + w_{\mathbf{cd}}(\mathbf{c}+\mathbf{d})}{2(w_{\mathbf{ad}}+w_{\mathbf{be}}+w_{\mathbf{cf}} + w_{\mathbf{ab}}+w_{\mathbf{bc}}+w_{\mathbf{fe}} + w_{\mathbf{ed}}+w_{\mathbf{af}}+w_{\mathbf{cd}})} \qquad (7)$$

where $w_{\mathbf{uw}}$ is given by (3).

All the above described interpolation schemes attempt 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. 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 other motion field estimation methods use such kind of information). By observing the interpolation expressions outlined above, one can conclude that the previously mentioned desirable estimation properties are achieved. In smooth motion areas, where the Euclidean distance between adjacent vectors is rather small, the weights $w_{\mathbf{uw}}$ are close to 1.0, thus leading to an averaging interpolator. When Euclidean distances increase significantly, the respective weights are becoming rather small limiting significantly the contribution of the corresponding candidate neighbouring pair in the final motion vector estimate.

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 forward-backward block matching method [3, 4]. Motion vector rational intepolation could be also employed in the estimation of concealment motion vectors that the MPEG-2 standard defines for transmission along 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:

$$MFE = \frac{1}{b_x \times b_y} \sum_{x=1}^{b_x} \sum_{y=1}^{b_y} ||\mathbf{v}(x,y) - \mathbf{v}_{or}(x,y)|| \qquad (8)$$

In (8), $b_x \times b_y$ 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.

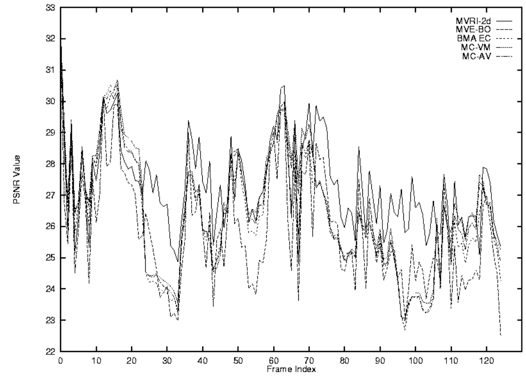
Table 1 illustrates the average $PSNR$ values of the Y component evaluated on the concealed test sequences by the EC methods under study. It can be seen that, in almost all cases, the MVRI EC method attains the second best result. To observe this outcome better, the $PSNR$ versus frame index plot is shown in Figure 2a for the Flower Garden sequence. 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 3 for the same sequence. Noticeable shifts are avoided when lost motion info is reconstructable by adjacent data and concealment is performed smoothly.

Table 1: Average $PSNR$ values ($PER = 2\%$).

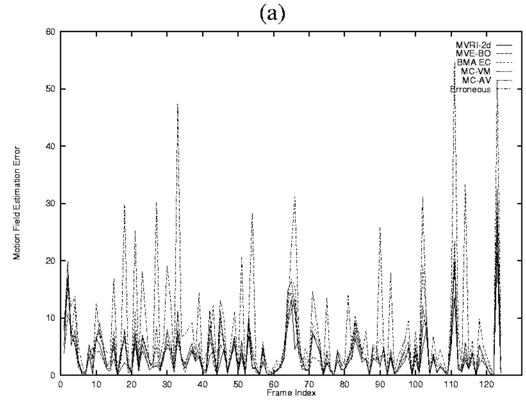
EC Method	Flower	Mobile	Football
Error Free	29.751	35.504	32.398
ZM EC	24.104	31.120	26.955
MC-AV EC	26.476	33.133	27.884
MC-VM EC	26.646	33.588	27.918
BMA EC	26.333	33.614	28.096
MVE-BO	25.865	32.626	**28.135**
MVRI-1 $- D$	27.215	33.553	28.063
MVRI-2 $- D$	**27.309**	**33.852**	28.021
MVRI-Combined	27.276	33.817	28.015
MVRI-2 $- D$-All	27.237	33.787	28.007
F-B BM EC	**27.736**	**33.849**	**28.396**
Erroneous	13.878	20.140	17.524

Table 2: Average MFE Errors.

EC Method	Flower	Mobile	Football
ZM EC	8.661	2.604	8.006
MC-AV EC	3.559	2.144	7.724
MC-VM EC	3.934	2.345	9.362
BMA EC	4.542	2.516	8.717
MVE-BO	3.586	2.456	8.727
MVRI-1 $- D$	**2.808**	1.932	**6.059**
MVRI-2 $- D$	2.995	1.945	6.342
MVRI-Combined	2.967	**1.930**	6.277
MVRI-2 $- D$-All	3.102	1.996	6.597
Erroneous	8.661	2.604	8.006



(a)



(b)

Figure 2: (a) $PSNR$ values of the concealed Flower Garden sequence by the MC-AV EC, the MC-VM EC, the BMA EC, the MVE-BO EC and the MVRI-$2 - D$ EC. (b) MFE Errors evaluated on the predictively coded frames of the same sequence after concealment.

In Table 2, 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. The same conclusions are reached also by observing the MFE Error plots versus frame index in Figure 2b. Figure 4 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 3 illustrates the execution times in secs required for the total concealment of the test sequences. Simulations were run under an Ultra-1 Sun Sparc Workstation. The processing time calculation has been performed on the concealment of the predictively coded frames. It is seen that meth-

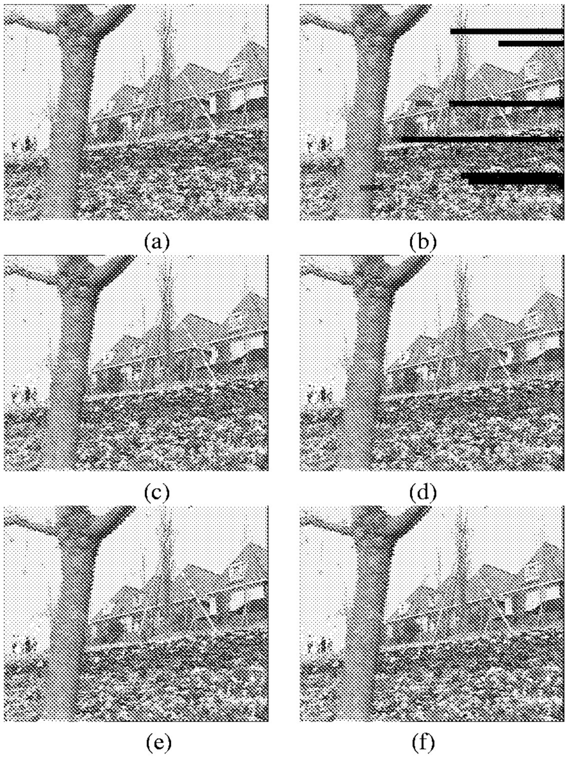


Figure 3: (a) Frame 25 (B-frame) of the Flower Garden Sequence, (b) Erroneous, $\boldsymbol{PER} = \mathbf{0.02}$, Concealed by: (c) the MC VM EC, (d) the BMA EC, (e) the MVE-BO EC and (f) the MVRI-$\mathbf{2} - \boldsymbol{D}$ EC.

ods 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. The postscript files of the figures in the paper can be accessed at: http://poseidon.csd.auth.gr/∼sofia/ICASSP99/ for the review process. In future, they will be available upon request.

5. CONCLUSIONS

Motion field estimation by vector rational interpolation has been investigated for error concealment purposes. Four different interpolation schemes have been examined. The motion vector rational interpolation error concealment method has been found to perform rather well and be rather fast. The interpolator adapts its behaviour according to the local motion information, thus leading to a well estimated motion field.

Table 3: Execution times in secs.

EC Method	Flower	Mobile
ZM EC	0.84	0.23
MC-AV EC	4.05	1.31
MC-VM EC	4.29	1.49
BMA EC	8.13	2.50
MVE-BO	289.32	89.27
MVRI-$1 - \boldsymbol{D}$	2.84	0.93
MVRI-$\mathbf{2} - \boldsymbol{D}$	3.10	0.90

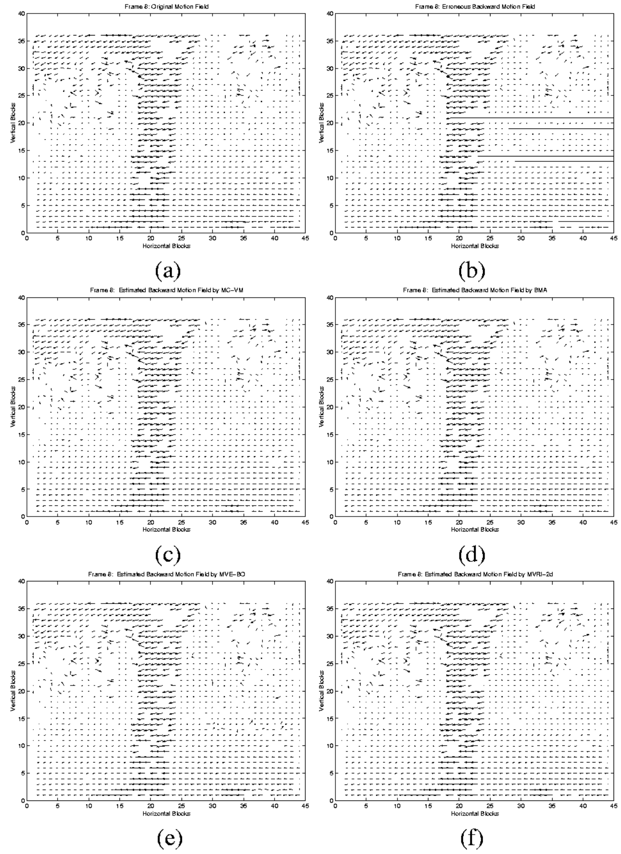


Figure 4: (a) Backward Motion Field of Frame 8 (B-frame) of the Flower Garden Sequence, (b) Erroneous, $\boldsymbol{PER} = \mathbf{0.02}$, Estimated by: (c) the MC VM EC, (d) the BMA EC, (e) the MVE-BO ECand (f) the MVRI-$\mathbf{2} - \boldsymbol{D}$ EC.

6. REFERENCES

[1] "Generic coding of moving pictures and associated audio information - part 2: Video", *ISO/IEC DIS 13818-2*, 1994.

[2] W.-M. Lam, A.R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors", in *Proc. of ICASSP'93*, Apr. 1993, vol. V, pp. 417–420.

[3] S. Tsekeridou, I. Pitas, and C. Le Buhan, "An error concealment scheme for mpeg-2 coded video sequences", in *Proc. of ISCAS'97*, 9-12 June 1997, vol. 2, pp. 1289–1292.

[4] S. Tsekeridou and I. Pitas, "Error concealment techniques in mpeg-2", in *CD-ROM of NSIP'97*, 7-11 September 1997.

[5] P. Salama, N.B. Shroff, and E.J. Delp, "Error concealment in encoded video", *submitted to IEEE Trans. IP*, 1998, http://dynamo.ecn.purdue.edu/∼ace/delp-pub.html.

[6] F. Alaya Cheikh, L. Khriji, M. Gabbouj, and G. Ramponi, "Color image interpolation using vector rational filters", in *Proc. of SPIE/EI Conf., Nonlinear Image Processing IX*, San Jose, CA, 24-30 Jan. 1998, vol. 3304.

[7] L. Khriji, F. Alaya Cheikh, and M. Gabbouj, "Multistage vector rational interpolation for color images", in *Proc. of CESA'98*, Hammamet, Tunisia, 1-4 April 1998.