Refined Boundary Matching Algorithm
for Temporal Error Concealment

Tao Chen
Sarnoff Corporation, Princeton, NJ 08543, USA
tchen@sarnoff.com

Abstract — In this paper, an adaptive technique is presented for temporal error concealment based on the boundary matching algorithm. Existing approaches usually estimate only one motion vector for a damaged macroblock by minimizing the boundary difference. This may generate edge fragmentation, object deformation, and other noticeable artifacts, especially in damaged areas involving non-homogeneous motion. We propose a framework of using various motion vectors for different regions within a damaged macroblock. The damaged macroblock is divided into small blocks, and each of the blocks is then concealed with an estimated motion vector. The estimation of motion vectors is independent, but imposed with a set of constraints in order to make it robust for concealment in different areas. The advantage of the new algorithm has been shown by simulations.

1 Introduction

In the transmission of packetized video through unreliable communication links, channel noise or congestion often leads to packet loss. This can dramatically degrade the visual quality of the decoded sequence. In hybrid motion compensated coding, such as the MPEG (e.g., [1]) series and H.26x (e.g., [2]) video compression, objectionable reconstruction artifacts observed in a frame may propagate to subsequent frames. This results in even worse visual quality. To combat this, signal reconstruction and error concealment have been proposed to obtain a close approximation of the original video, or to make the output signal at the receiver less objectionable to human eyes [3].

Error concealment is viable because a video bitstream still possesses certain degree of statistical redundancy, due to various constraints such as coding delay, implementation complexity and availability of perfect source modeling. Though tremendous research effort has been devoted to achieving the highest possible compression efficiency. On the other hand, the source encoder may employ some error-resilient mechanisms to intentionally introduce redundancy in the bitstreams and to facilitate the recovery of lost information. Additionally, human eyes can tolerate a limited degree of signal distortion. All the factors make the error concealment worth exploiting at the receiver, without any dependence upon the sender. A simplified block diagram is shown in Figure 1 to depict the process of video decoding that embeds error concealment.

There are in general two key components in error concealment: (1) an appropriate transport format to assist in identifying picture areas that correspond to the missing or damaged data, and (2) a spatial- and/or temporal- interpolation techniques to fill the damaged regions. This work is focused on the second aspect. The proposed scheme assumes that the locations of damaged areas
are available at this stage. Readers interested in the detection of damaged areas are referred to [3] for detailed discussions on the topic.

![Simplified block diagram of a video decoder incorporating error concealment.](image)

Figure 1: Simplified block diagram of a video decoder incorporating error concealment.

In particular, error concealment in video communications is intended to recover the loss of information, with the use of available image data. It is therefore to ameliorate the negative influence brought in by the data loss and to present the reconstructed video more pleasing visually. In block-based hybrid video compression, three types of information at the macroblock (MB) level may be damaged and need to be reconstructed [4]. They are (i) textural information such as the pixel or DCT coefficient values for either an intracoded MB or the prediction residue of an intercoded MB, (ii) the motion vectors of an MB coded in either P- or B-mode, and (iii) lastly the coding mode of the MB. A number of methods have been developed to conceal errors in video communications by adaptive interpolation in the spatial, temporal, and frequency domains. Extensive coverage of these schemes can be found in [3, 4, 5]. In this paper, we concentrate on techniques particularly devised for the estimation of damaged motion vectors in error concealment.

The estimation of lost MVs may employ several simple strategies [6, 7]: (i) assuming the lost MVs to be zeros, which works well for video sequences contain small motion; (ii) using the MVs of the corresponding macroblocks in the previous frame; (iii) using the average of the MVs from spatially adjacent macroblocks; (iv) taking the median of MVs from the spatially adjacent macroblocks; and (v) re-estimating the lost MVs using the boundary matching algorithm (BMA) [7]. It has been found that the last two operations yield the best results [8, 4]. A scheme similar to the BMA was proposed in [9] where the side match criterion measured the difference between the external boundary of the damaged MB and the internal boundary of the replacing MB in the reference frame. Alternatively, in the BMA, the difference measure for best matching takes on the external surrounding boundaries of both the damaged MB and the replacing MB. In this way, the overlapping boundaries can be multiple pixels wide [10]. On the other hand, instead of using the mean or median of the candidate MVs, one can also use other approaches for the interpolation of lost MVs. In [11], vector rational interpolation was applied in motion field estimation for error concealment.

These methods have a common drawback that is represented by the blocking artifacts introduced in the recovered areas. In [12], the overlapped block motion compensation (OBMC) [13, 2] has been combined with the side match criterion [9] in order to minimize the blocking artifacts. Since this reduces the differences gradually across the block boundaries, it produces better visual quality. Improvement over the BMA has also been obtained by using multi-frame boundary matching scheme [14]. It exploits the boundary smoothness property in the current frame as well as in the succeeding intercoded frames that use the current damaged frame as a reference. However, this unavoidably introduces larger latency in addition to requiring extra memory for decoding. Its practical usage is
lessened especially in the applications that require low delay, e.g., video conferencing.

The above schemes are simple, but erroneous estimate could lead to noticeable reconstruction artifacts. In addition, most of the approaches usually estimate only one MV for a damaged macroblock. This works well in the areas involving translational motion, where pixels in the MB have almost identical motion. However, it fails when the damaged macroblock is in the regions containing rotational movements, zooming, and deformation of objects, or incoherent motion of different regions within the MB. In these cases, a simple compensation based on block translation will produce the effect of edge fragmentation, object deformation, and other artifacts. These could be ameliorated to certain degree by using different MVs for different regions within a damaged MB. One solution is to refine the BMA by dividing a damaged MB into a number of sub-blocks and estimating an MV for each of the small sub-blocks. This paper proposes a framework of refined boundary matching algorithm (RBMA) and examines its performance of dividing the missing MB into four blocks.

The paper is organized as follows. The boundary matching algorithm is first briefly reviewed in Section 2. Section 3 presents in detail the proposed refined boundary matching algorithm. Experimental results are given in Section 4. Finally, Section 5 concludes the work.

2 Boundary Matching Algorithm

The BMA exploits the fact that adjacent pixels in a video picture exhibit high spatial correlation. A boundary matching (BM) measure is defined in Eq. (1) as the squared difference between the external boundary of the damaged MB and the internal boundary of the replacing MB in the reference frame.

$$E_{BM} = \sum_{x=-x_0}^{x_0+N-1} \left[ \left( F_{x,y_0} - F^r_{x+v_x,y_0+v_y} \right)^2 + \left( F_{x,y_0+N} - F^r_{x+v_x,y_0+N+v_y} \right)^2 \right]$$

$$= \sum_{y=y_0}^{y_0+N-1} \left[ \left( F_{x_0,y} - F^r_{x_0+v_x,y+v_y} \right)^2 + \left( F_{x_0+N,y} - F^r_{x_0+N+v_x,y+v_y} \right)^2 \right],$$

(1)

where \((x_0,y_0)\) is the coordinate of the top-left pixel in the damaged MB, \((v_x,v_y)\) is the candidate MV under consideration, and \(F_{x,y}\) and \(F^r_{x,y}\) denote the pixels of the current and reference frames, respectively. Here, \(N = 16\) is the macroblock size. Instead of using full search, the motion vector of damaged MB is selected from a set of candidates by minimizing the BM measure [7]. The candidates are usually the MVs of the macroblocks spatially adjacent to the damaged MB. The selected replacing MB is summed with the prediction residue of the damaged MB. In the case that the residue is also lost, the motion compensated residue of a spatially adjacent MB is used instead.

Similar solutions have been proposed to find the motion vector that yields the closest or best overlapping between the external boundaries of both the lost MB and the replacing MB in the reference frame. The error function of overlapping boundary matching (OBM) is defined as:

$$E_{OBM} = \sum_{x=-x_0}^{x_0+N-1} \left[ \left( F_{x,y_0} - F^r_{x+v_x,y_0+1+v_y} \right)^2 + \left( F_{x,y_0+N} - F^r_{x+v_x,y_0+N+v_y} \right)^2 \right]$$

$$+ \sum_{y=y_0}^{y_0+N-1} \left[ \left( F_{x_0,y} - F^r_{x_0-1+v_x,y+v_y} \right)^2 + \left( F_{x_0+N,y} - F^r_{x_0+N+v_x,y+v_y} \right)^2 \right].$$

(2)
3 Refined Boundary Matching Algorithm

The framework of the proposed RBMA is naturally hierarchical, and based on the results obtained for $M \times M$ blocks, one can readily extend the algorithm by further dividing the block into smaller sub-blocks ($\frac{M}{2} \times \frac{M}{2}$). The multi-level refinement could be adaptively controlled according to some rules, such as thresholding a cost function of boundary error. In other words, smaller blocks are used if the boundary matching differences at the current level exceed a predefined value.

In the simplest case of the refined boundary matching algorithm, the damaged macroblock is divided into four blocks, $B_i$ ($i = 0, \ldots, 3$). Accordingly, the external boundary of a damaged MB is also divided into four vectors. As shown in Figure 2, the vectors are denoted $b_i$ ($i = 0, \ldots, 3$), corresponding to the top-left, top-right, bottom-left, and bottom-right blocks in the MB, respectively. The motion vectors of the four blocks are estimated independently, but with some constraints in order to ensure the spatial coherence of estimated motion vectors in homogeneous regions.

![Figure 2: Boundary pixels used in the RBMA are marked as the shadowed areas, where the overlapping boundary is considered and it can be multiple pixels wide.](image)

The MV of $B_i$, for $i = 0, \ldots, 3$, is estimated by matching $b_i$. In the case of one pixel wide boundary, the motion vector $(v_x, v_y)$ that minimizes the following error is selected for block $B_0$:

$$
E_{RBMA} = \left\| b_0 - b_0' \right\|^2
$$

$$
= \sum_{x=x_0}^{x_0+n-1} \left( F_{x,y_0-1} - F_{x+v_x,y_0-1} \right)^2 + \sum_{y=y_0}^{y_0+n-1} \left( F_{x_0-1,y} - F_{x_0-1,y+v_y} \right)^2
$$

$$
+ \left( F_{x_0-1,y_0-1} - F_{x_0-1,y_0-1} \right)^2,
$$

where $b_0'$ denotes the corresponding external boundary of a replacing $n \times n$ block in the reference frame, $x_r = x_0 + v_x$, $y_r = y_0 + v_y$, and block size $n = 8$. In the above equation,

$$
b_0 = \begin{bmatrix} F_{x_0,y_0-1}, F_{x_0+1,y_0-1}, \ldots, F_{x_0+n-1,y_0-1}, \\
F_{x_0-1,y_0}, F_{x_0-1,y_0+1}, \ldots, F_{x_0-1,y_0+n-1}, F_{x_0-1,y_0-1} \end{bmatrix}^T,
$$

and

$$
b_0' = \begin{bmatrix} F_{x-r,y_r-1}, F_{x_r+1,y_r-1}, \ldots, F_{x_r+n-1,y_r-1}, \\
F_{x_r-1,y_r}, F_{x_r-1,y_r+1}, \ldots, F_{x_r-1,y_r+n-1}, F_{x_r-1,y_r-1} \end{bmatrix}^T.
$$
Similarly, the boundary vectors and error functions of the other three blocks can be constructed. Note that the corner pixels are also included in the calculation (see Figure 2).

So far, the overlapping boundary matching is described, and the non-overlapping case can be derived straightforwardly. In this work, we use the overlapping boundary matching technique to recover the damaged MVs. This approach has the advantage that edge or signal continuity between the lost MB and its neighbors can be preserved. The re-estimation of motion vector does not require any differential information from the adjacent MBs.

In the proposed RBMA, the motion vector re-estimation is not confined with selecting one of the MVs of adjacent macroblocks. Rather it searches for the best MV based on candidate motion vectors that are sifted with some constraints. The detailed operations of the RBMA are laid out in the following subsections. The configuration of isolated MB loss is considered for simplicity, and is used to demonstrate the advantage of employing multiple motion vectors for a damaged MB in temporal error concealment. The algorithm can be readily modified and extended to handle the case of consecutive missing MBs in a general scenario.

3.1 Temporal Activity

The framework of the proposed RBMA is imposed with several constraints in estimating the MVs of the missing blocks. The motion vector re-estimation of the RBMA requires higher computation, and therefore it is performed only when needed. In addition, the standard BMA works well in regions with no motion, or involving translational motion only where the motion of blocks in the lost MB is usually homogeneous. Therefore, the following criterion is devised to switch the concealment operations between the BMA and RBMA.

For a damaged macroblock, let $V_j$ ($j = 0, \ldots, 3$) denote the motion vectors of its four adjacent MBs, i.e., upper, lower, left, and right, respectively, to the damaged macroblock. The temporal activity, or motion coherence, of the damaged area is then quantified by

$$T_m = \frac{1}{6} \sum_{j=0}^{2} \sum_{k=j+1}^{3} \|V_j - V_k\|^2.$$  \hspace{1cm} (4)

In the case that $T_m$ does not exceed a threshold $t_{T_1}$, it is assumed that the surrounding area of the lost MB contains slow motion or translational motion. In this case, the standard BMA is used to estimate the MV of the lost macroblock. Otherwise, i.e., $T_m > t_{T_1}$, the error concealment algorithm steps into the search of the best motion vector for each block. It should be noted that the temporal activity $T_m$ will also be used to determine the search window size described in Section 3.3.

3.2 Unreliable Adjacent Motion Vectors

In the current video compression standards, block-based motion estimation is commonly employed. The reliability and accuracy of the motion vectors are limited by a number of factors, such as the search window size and block based structure causing motion uncertainty [15]. As a result, one may observe incoherent MVs of adjacent macroblocks even in regions where they are supposed to be well correlated. In order to reduce the probability of using an unreliable MV as a starting point of the re-estimation process, we introduce the following criterion to measure if each of the four motion
vectors of the adjacent MBs is an outlier, or irregular MV. The irregular motion vectors are then excluded from the candidate MV set in the subsequent estimation of the best MVs.

It consists of two steps to determine the existence of an irregular motion vector. First, for each \( c (c = 0, \ldots, 3) \), the variation of the motion vectors that exclude \( \vec{v}_c \) is defined as

\[
S_c = \frac{1}{3} \sum_{0 \leq j \leq 2, j \neq c} \sum_{0 \leq k \leq 3, k \neq c} \| \vec{v}_j - \vec{v}_k \|^2 .
\] (5)

If \( S_c \) is greater than a threshold \( t_S \), then it is regarded that \( \vec{v}_c \) is not an outlier. This is because the variation among the other three MVs is large, and in this case, the damaged MB may be located in the boundary area of different moving objects. Otherwise, i.e., \( S_c \leq t_S \), it goes to the second step.

Let \( \vec{v}_{BM} \) denote the motion vector that is estimated using the standard BMA, and the difference between \( \vec{v}_{BM} \) and \( \vec{v}_c \) is given by

\[
D_c = \| \vec{v}_{BM} - \vec{v}_c \|^2 .
\] (6)

If \( D_c > t_S \), then \( \vec{v}_c \) is determined as an irregular motion vector. Otherwise, \( \vec{v}_c \) is eligible to be a starting point in the following search of the best motion vectors.

### 3.3 Motion Vector Re-estimation

It is well known that block-based motion estimation may not be able to generate true motion vector and often cause motion uncertainty. In boundary matching algorithms, the motion vectors estimated with the use of only a few boundary pixels could have even higher unreliability. In order to ameliorate this problem, the search window size of motion vector re-estimation needs to be constrained in some way. In particular, for each of the four blocks \( B_i (i = 0, \ldots, 3) \), the starting points of motion vector search are selected as the reliable MVs of its two closest adjacent macroblocks. For example, the reliable MVs of the upper and left MBs are used for \( B_0 \). In addition, the zero motion vector \( \vec{0} \) is included in the candidate set. The search windows are then centered at these corresponding locations in the reference frame.

In the proposed RBMA, the size of the search window for boundary matching is determined as follows in regions with different motion characteristics.

\[
s = \begin{cases} 
2 & \text{if } T_m < T_2; \\
5 & \text{otherwise.}
\end{cases}
\] (7)

Here, \( T_m \) is the temporal activity index as defined in Eq.(4). The idea behind is to re-estimate the motion vector within a small search window if the MVs of adjacent MBs exhibit relatively high correlation, i.e., \( T_m \) is small. In the case that \( T_m \) is large, one can assume that the damaged MB is located in the region of moving object boundary, and a large window size is hence selected.

### 3.4 Block-Edge Effect Reduction

The final step of the proposed RBMA attempts to reduce the blocking artifacts generated by the temporal error concealment. This is important because the independent MV estimation for each of the blocks may cause noticeable discontinuity across the block boundaries. The situation is severer
in error concealment than that in video coding, since more than often the prediction residue of a damaged MB is also not available. The use of the prediction residue of an adjacent MB may not be able to compensate the discontinuity across the block boundaries, due to spatially low correlation between prediction residues of adjacent macroblocks.

Taking into account the computational complexity, a simple 3-tap low-pass filter, \([0.25, 0.5, 0.25]\), is employed to eliminate the block-edge effect across the block boundaries. Both sides of the sub-block boundaries inside the damaged MB and the boundaries of the damaged MB to its spatially adjacent macroblocks are filtered. Although simple, this filter improves the visual quality of reconstructed frames.

4 Simulation Results

The experimental results shown in this section demonstrate that improved quality has been achieved by the proposed error concealment technique with the use of various motion vectors for different regions in the damaged MB. For this purpose, the results for the configuration of isolated macroblock loss are reported. In addition, the case of one pixel wide boundaries is considered in the following experiments. The proposed algorithm can be readily extended to conceal the loss of multiple and consecutive macroblocks and use boundaries that are multiple pixels wide in the motion vector re-estimation. For performance comparison, the standard BMA and the BMA with OBMC scheme [12] are used. In addition, the zero motion vector replacement (ZMVR) is simulated, which refers to copying the macroblock at the same location from the reference frame.

A number of test sequences of CIF resolution \((352 \times 288)\) characterized by different spatial and temporal activities have been used in the performance evaluation. The concealment processes have been applied using a reference frame that is three frames ahead of the current frame. This corresponds to the case of recovering a P-frame with reference to an I- or P-frame in a common ‘IBBP’ GOP structure of motion compensated video coding. This is the largest temporal distance between a frame and its nearest reference in typical GOP configurations.

The damaged frame is generated by dropping about \(\%16\) of the macroblocks in the image, and the reference frame is assumed to be error free. In the experiments, this process is repeated for 50 frames of each sequence. In addition, we make the assumption that the prediction residue of a missing MB is also lost and is set to zero. Furthermore, the motion vectors of error-free macroblocks are estimated within a search range of \(\pm 16\). This is a typical choice in the MPEG video compression standards [1].

In the simulations, the parameters are set as follows: \(tT_1 = 1\) for deciding the temporal activity of the damaged area, \(tT_2 = 5\) for determining the window size in the subsequent motion vector re-estimation, and \(tS = 20\) for excluding possible outliers among the motion vectors of adjacent macroblocks, respectively.

Figure 3 presents the PSNR results of the error concealed frames from different sequences using the above approaches. In general, the proposed RBMA brings improvement on the overall performance over the competing algorithms. As summarized in Table 1, the RBMA achieves an average PSNR gain of up to 1.5dB over the scheme that combines the BMA with OBMC.

In addition to the improved quantitative quality, the visual quality of the images produced by the proposed RBMA are much better that those by the other techniques. Figure 4 shows the error
Figure 3: Comparative performance in terms of PSNR (dB) for different algorithms.
concealment results of sequence Foreman. From the reference and the current frames, Figures 4(a) and (b), one can observe non-translational motion in the face areas. The ZMVR scheme yields poor quality, and in fact, the motion involved is manifested in the results by the ZMVR. The BMA+OBMC scheme does not perform well either. The minimization of the boundary difference for all of the four sides of a damaged MB quite often cannot be satisfied simultaneously by the replacing macroblock found by the BMA. It is obvious that one motion vector for the entire MB is not enough in order to obtain a satisfactory result. In this case, the proposed RBMA produces the resultant frame more pleasing visually with less noticeable reconstruction artifacts. Better visual quality has also been observed in the other test video sequences concealed by the proposed technique.

## 5 Conclusions and Future Work

An adaptive technique based on the boundary matching algorithm has been proposed in this paper for temporal error concealment in video communications. To better satisfy the criterion of minimizing the boundary difference between the lost macroblock and a replacing macroblock in the reference frame, we propose to use different motion vectors for different regions within a damaged macroblock. For each of the blocks in a damaged macroblock, a motion vector is estimated by boundary matching. In order to make the RBMA robust for temporal concealment in different areas, the independent estimation of motion vectors is adaptively constrained. The proposed algorithm has shown its advantage of using multiple motion vectors to conceal different regions in a missing macroblock and therefore yielded better results than other competing methods.

More sophisticated constraints can be imposed to improve the robustness of the proposed technique. The framework of the RBMA is extendable to accommodate more levels of refinement using smaller sub-blocks. It can also be extended to handle the loss of multiple and consecutive macroblocks. Concealment of MPEG bitstreams using the new approach is under investigation, and the results will be presented in a further report.

## References


Figure 4: Results of Foreman: (a) reference frame, (b) original current frame, (c) damaged current frame, reconstructed frames by (d) ZMVR, (e) BMA+OBMC, and (f) RBMA, respectively.


