

# Motion characteristic differentiated error concealment

Heng Li · YanChun Zhong

Published online: 20 May 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** In this paper, we propose a brand-new motion characteristic differentiated error concealment (MDEC) method based on motion field transfer. Firstly, the FMO checkerboard pattern is used at the encoder, so as to prevent MBs of a large area getting lost. Then at the decoder, Greedy Spread Motion Region Extraction (GSMRE) method is used to distinguish low-motion region from high-motion region in each frame based on different motion characteristic, and apply different strategies to recover regions with different characteristics respectively. Simulation results show that the proposed algorithm reconstructs the lost frame with both a better visual quality and a higher PSNR, comparing to error concealment methods including Joint Model, boundary matching, inpainting, and block motion vector extrapolation as well. For example, the PSNR gain of our approach over boundary matching algorithm reached about 0.6 dB, and 1.4 dB when the packet loss rate is 3 and 7%, respectively, which demonstrates that our method has an good application within a wide scope of packet loss rate.

**Keywords** Error concealment · Motion field transfer · Low motion · High motion

## 1 Introduction

Video communication over the internet and wireless network is becoming ubiquitous and popular in academic and industrial fields. While the bit error and packet loss is a well-known enemy of the visual quality. Researchers have proposed various

---

H. Li (✉) · Y. Zhong  
State Key Lab of Software Development Environment, New Main Building,  
BeiHang University, Beijing, Haidian 100191, China  
e-mail: liheng@nlsde.buaa.edu.cn

Y. Zhong  
e-mail: zhongyc@nlsde.buaa.edu.cn

solutions to combat the transmission error. However, reliable transmission protocol and Automatic Retransmission Request (ARQ) may introduce high delay. Adding redundant information into the bit stream will decrease both the coding efficiency of the encoder and the available bandwidth of the channel. So when transmission errors occur, error concealment at the decoder is the method of more effective. Lots of error concealment algorithms have been proposed to recover pictures in terms of MB or pixel [32]. Existing error concealment methods at the decoder typically rely on the motion vector smoothness, block boundary smoothness or pixel smoothness. The recovery performance depends on the number of correctly received MBs around the lost MB, the more correct neighboring MBs, the more information available for recovery, and the better visual quality obtained accordingly.

Other concealment methods adopt the same concealment method to recover MBs or pixels with different motion characteristics. Other methods such as recover the MVs by splitting a whole object, still or moving, into different patches may lost integrity and cause blocky artifacts comparing with recovery on pixel level. To solve these problems, the decoder should firstly know the motion characteristics of the lost regions, and then apply different error concealment strategies for each of them. However, the best effort network can't guarantee that lost patches always belong to the same motion characteristic region. Besides, recognizing the motion characteristics of lost region is a problem of high complexity within a video sequence carrying errors. These problems require the concealment method has the following three features:

1. Using certain optimization strategies to make as much available supporting information as possible information for recovery in the case of packet loss. FMO technique is used in this article to isolate the MBs with their neighboring MBs and ensure that the packet loss during transmission will not make a large area missing. By comparing several FMO patterns, we use checkerboard pattern in this paper.
2. Automatically recognize the motion characteristic of the lost region when packet loss occurs to the video sequence. It can be do through the correctly received slice group number packetized in the slice header and find out which groups of slices lost. Then, for the missing MBs, the decoder needs to determine the motion characteristics of the lost MBs and classify them dynamically. Thus to apply different recovery methods.
3. Motion characteristic differentiated error concealment method should be applied. since the motion characteristic of the lost slice group is known at the decoder, we propose a motion characteristic differentiated error concealment method. The best feature of it is that applying different concealment methods to regions with different motion characteristics.

The main contributions of this paper can be summarized under the following points:

1. Proposed an error control and recovery framework with co-working between the encoder and the decoder. At the encoder, FMO technology is used to isolate the losing MBs and obtains more additional supporting information. At the decoder, lost MBs are classified in accordance with the motion characteristics, and different recovery methods are applied for the MBs of each characteristic.
2. Proposed an approach to identify high-motion and low-motion regions. It based on the assumption that the interested regions of human eyes are of high motion,

- while not interested regions are of low motion. For the MBs of high motion regions, the motion characteristics need to be reconstructed. While for MBs of low motion regions, only the time domain-based copy and paste methods will be used for recovery, which the human eyes not concerned much about.
3. Proposing a motion differential error concealment strategy for the cases of a large number stream data loss, such as almost half of the whole frame. The proposed error concealment method expresses the motions of each correctly received pixel using optical flow, and then through the movement of adjacent pixels, the movement trends of loss pixels, also called optical flow, can be predicted using motion field transfer method. Finally, the color of each missing pixel will be further assigned using color propagation method according to the reconstructed optical flow.

The rest of this paper is organized as follows. In Section 2, state-of-art will be introduced. Section 3 gives some preliminaries and an overview of our approach. Our concealment method is described in detail in Section 4. In Section 5, the performance and the computational cost of our method will be analyzed by both subjective and objective results. Finally, we conclude our paper in Section 6.

## 2 State-of-art

Recover the Motion Vector (MV) of each lost MB is often a useful method. Boundary matching algorithm (BMA) is a classical method to recover the motion vector of lost MB [17]. It works with an additional consideration of the smoothness of the concealed MB along its boundaries. BMA is also extended to compare the edge characteristics instead of pixel color value to smooth the edge of the lost MB [10]. Although the two methods proposed based on boundary matching is very practical to conceal a few of MBs lost situation, the blocky artifacts is the biggest problem for the visual quality. To estimate the lost MV with more accuracy, Chen et al. proposed a spatial-temporal boundary matching algorithm (STBMA) based on BMA, and use a novel partial differential equation (PDE) based algorithm to refine the reconstruction instead of directly copying the reference MB to perfect matching between the recovered MB and the surrounding boundaries [6]. Decoder motion vector estimation (DMVE) algorithm is similar to BMA except that it uses the temporal correlation instead of the spatial correlation [35]. The best matching is searched in the previous frame by minimizing assume of absolute differences (SAD) between boundary pixels of the corrupted MB in the current frame and their corresponding pixels in the previous frame. In [11], Hwang et al. choose one of two results obtained by BMA and DMVE based on the normalized SAD values for each lost MB.

Motion vector extrapolation first extrapolates motion vectors of MBs to lost frame; and then, estimated the overlapped areas between damaged block and the motion extrapolation MBs, and the best estimation motion vector is obtained, the lost block is concealed with the block in previous frame [24]. Chen et al. proposed a modified pixel-wised MV extrapolation (PMVE) method to conceal the whole frame [5]. In addition, recover pixels is also a popular error concealment method. Ranjbar et al. first split the erroneous blocks to different patches based on a context-dependent exemplar-based segmentation method, and then replace each patch of lost

pixels by another patch of the image that contains correct pixels [25]. Chen et al.'s method extrapolates two motion vectors from the motion vector of the previous reconstructed frame and the next frame [5]. Yan et al. proposed a new hybrid motion vector extrapolation (HMVE) algorithm which can adaptively select the possible MV and discard impossible MV [33]. HMVE use extrapolation technology to produce the candidate MV. Mengyao et al. proposed an edge-directed error concealment (EDEC) algorithm, to recover lost slices. Based the strong edges estimation in a lost frame, EDEC can adaptively recover the pixels referencing the pixels neighboring frames and the received area of the current frame [22].

In [2], the authors implement error concealment by spatial and temporal interpolation. Also a gradient based boundary matching mode selection method is proposed to choose if spatial, temporal, or spatial and temporal interpolation could be used to recover the lost MBs. It seems to be very efficient, especially for intra-coded frames, it is computationally more intensive compared with the one originally presented in the test model [12, 30].

In addition to mentioned [2] above, there are also some other methods recover the lost MBs based on classifying MBs into different types. In [34], the authors embed extra data into the encoded bit stream to convey the necessary high level information in a standard compliant way to help improve the decoder choose different error concealment strategies for different MBs. As the weakness of error control in the encoder, the extra information not only adds the complexity of encoding but also decreases the available bandwidth for the actual encoded bit stream. Similarly in [13], all lost blocks are first classified into foreground blocks or background blocks based on neighboring pixels. If a lost block belongs to background region, the corresponded block in the previous frame will be used to replace the lost block. If a lost block belongs to the foreground block, the candidates of most similar blocks in each of the multiple reference frames are searched and either selected or averaged. These methods haven't taken into account that high motion also is possible to happen in the background. Moreover, block deformation and distortion often occurs in foreground, simply copying the data from the reference frames or using weighted average lead to poor recovery.

Moreover, the predictive coding and variable length coding adopted in most of video coding standards makes the code streams more sensitive to the errors. Several error concealment algorithms have been proposed for MB losses based on FMO resilience tool of H.264, such as spatial interpolation based on edge direction [9] and motion compensated temporal prediction based on boundary matching [28] in order to prevent error propagation. Image inpainting techniques have also been used in recent EC algorithms [3, 7, 29]. In [7], an exemplar based image inpainting approach is proposed. The authors demonstrate that the quality of the synthesized image is highly influenced by the order in which the filling process proceeds. So they propose to give higher synthesis priority to those regions lying on the continuation of image structures.

These methods recover the image in high quality based on the spatial smoothness assumption, but not effective for recovery in time domain. Meanwhile, Image inpainting method highly depends on the area selection, and sometimes it is necessary for human to label the areas that need inpainting. A common drawback of inpainting methods is that the lack of the edge texture information in spatial domain of image will lead to fuzzy results.

Another recovery method which takes Human Visual System (HVS) into consideration is edge based spatial error concealment (EBSEC) method, it focus on edge information recovery of each frame by using the weighted average method for each pixel in the edge region and non-edge regions (known as structured region and unstructured region) respectively [1, 14–16, 26]. This type of approaches assumes that the lost edges are straight lines, the edge directions and the main edge line are firstly estimated, and the area around the main edge line by a certain distance is selected as the structured region. Then the pixels in the two regions are reconstructed using the weighted average, the weight values depend on the distance from the loss pixel to the edge of current MB.

Most of the above methods need correctly received spatial information. If large area of a frame is lost, these methods may not work effectively. Although EDEC [22] uses checkbox FMO can combat slice-level packet loss, it can't reflect different motion characteristic of different region in each frame. As we know, motion characteristic between frames is the biggest difference between video and image. It is also what the HVS most sensitive.

Therefore, in this paper, we combine the advantages of spatial and temporal error concealment (STEC) and edge based spatial error concealment (EBSEC) above, and then a new error control and concealment framework is proposed. We use FMO resilience tool of H.264 in [22] to avoid the error propagation between two slices in one frame. Inspired by [2, 13, 34], lost MBs are divided into two types: low inter-frame motion with no significant change compared with reference frames, and high inter-frame motion. The low motion lost MBs are reconstructed by copying in the time domain, and for high motion lost MBs, we use the error concealment method based on motion field transfer to recover each lost pixel.

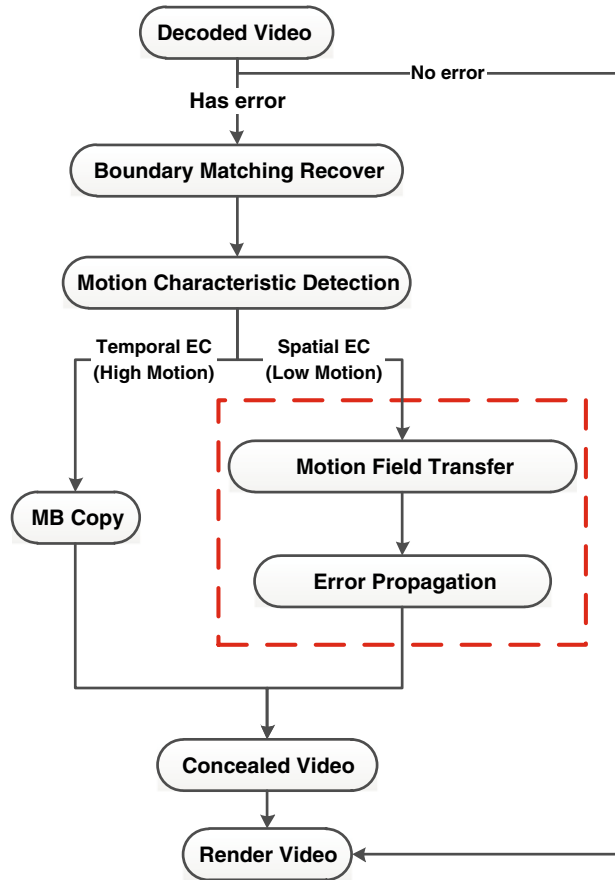
### 3 Approach overview

Our method can be divided into the four steps including boundary matching recover, motion characteristic detection, high-motion region error concealment, and low-motion region error concealment. The flowchart of our approach is shown as Fig. 1.

*Boundary matching recover* Applying boundary matching for all the lost MBs. Different with traditional BMA methods, the lost MBs use the external boundaries for matching in this paper. At the same time, pixels are directly compared to limiting the search depth, instead of using pre-defined motion vectors as candidates. And for the MBs not properly matched, which have distortions of high degree, motion field transfer approach will be further used for recovery.

*Motion characteristic detection* In this step, the high motion and low motion MBs will be detected and applying different error concealment method respectively. We will use the above step's recovered picture as the input picture, which indicates that all the MBs of this frame will be used to estimate the motion characteristic. The estimated MAD (Mean Absolute Difference) value of each MB is used to determine whether the MB belongs to high-motion region. And then the high-motion region and low-motion region will be partitioned into two independent slices for transmission.

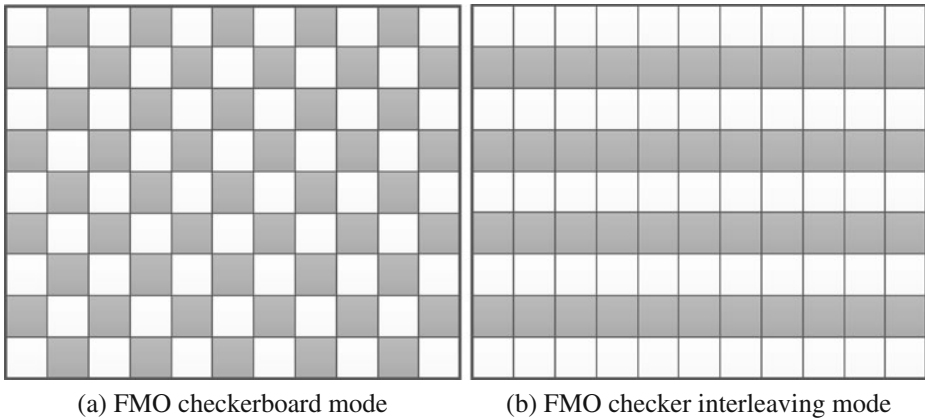
**Fig. 1** Flow chart of out approach



*High-motion region error concealment* The high-motion region is corresponding to the sensitive region of HVS. As the previous section described, optical flow is able to represent the pixel movement accurately. Therefore, optical flow is usually used to represent the motion of each pixel. Furthermore, we use the motion field transfer technology to represent the trend of motion, called as optical flow, and transfer the motion from previous frame to the current frame. After recovering the lost motion, we will use the color propagation technology to assign color for the lost pixels based on reconstructed optical flow.

*Low-motion region error concealment* Since the HVS is not sensitive to the low-motion region, a way of not very precise is used to conceal it. In the initialized boundary matching step, the matching standard is fit for the less change MBs and the low motion MBs. Therefore, the result obtained by boundary matching can be directly used, and the experimental results prove that this is feasible.

Described above is the error control and recovery framework. It provides a universal solution in error control and concealment for video communication which doesn't depend on the video compression standard. Comparing with former methods, our method works well and achieves good visual quality both at low and high packet loss rate. Since our method classify the lost MBs based on their different



**Fig. 2** FMO checkerboard mode and interleaving mode

movement characteristics, it is called as **Motion Characteristic Differentiated Error Concealment (MDEC)**.

From Fig. 1, it can be seen that our method recovers the lost pixels directly, instead of recovering motion vectors and then decoding video stream again, thus saves a lot of time overhead. To avoid the error propagation between two slices in one frame and attain more available neighboring MBs for concealment. Firstly, we use FMO resilience tool of H.264 like [21, 28] to encode the video stream. There among, checkerboard pattern and interleaved pattern, as Fig. 2 shows, are most suitable for EC at the decoder. In the Fig. 2, slice 1 contains all the white MBs and slice 2 contains all the grey MBs. Because in these two patterns, the lost MBs of one slice distribute in the spatial evenly. Moreover, suppose one of the two slices gets lost in a frame, for each lost MB, there are two and four correctly neighboring MBs available in interleaved pattern and checkerboard pattern respectively.

Thus, the checkerboard pattern is selected to be used in our paper. Suppose based on the checkerboard FMO mode in Fig. 2a, in which Then these two slices are packed and transmitted separately over the Internet. Suppose slice 1 of frame  $t$  is lost during the transmission, as predictions between neighboring MBs are abandoned in checkerboard FMO, slice 2 of frame  $t$  can be correctly received and decoded. We will use both the decoded slice 1 and the previous frame  $t - 1$  to conceal the lost MBs in slice 1 in our proposal. What follows are different parts of our method, including four main algorithms: Boundary Matching Recover, Motion Characteristic Detection, Motion Field Transfer, and Color Propagation. We will describe them in detail separately in the following of this section.

#### 4 MDEC

This section describes our proposed algorithm Boundary Matching Recover, Motion Characteristic Detection, High-motion MB Error Concealment and Low-motion MB Error Concealment.

An H.264/AVC decoder is capable to detect the frame loss through parsing the slice headers. The standard specifies a procedure to verify the start of a new “primary

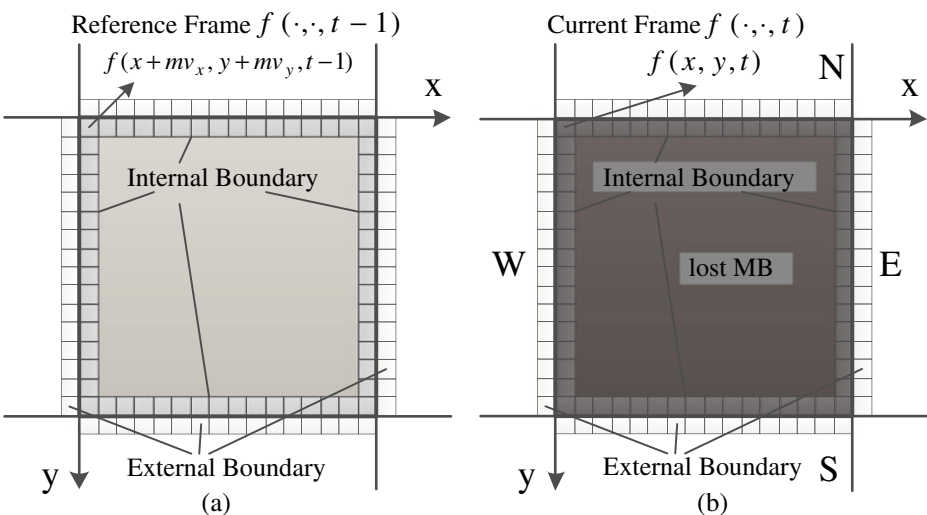
coded picture” by checking the coherence of some values in the slice header that are expected to remain constant for all the slices of a picture. In particular, the frame number is a field that increases by one unit for each frame sent by the encoder. Therefore, it is possible to define a procedure that checks this field for every slice and activate the concealment when it detects an unexpected gap.

Only detect which frame is lost is not enough, we should know which slice in the frame is lost, which can be figured out through the field of slice group number, and then the corresponding position of lost MBs could be known.

### 4.1 Boundary matching recover

In this paper, we use the external boundary of each lost MB to measure the MB distortion based on the spatial smoothness consideration as shown in Fig. 3. The classic boundary matching algorithm (BMA) is utilized to recover the lost MVs from the candidate MV set, which minimizes the side match distortion between the internal boundary of reference MB and the external boundary of the reconstructed MB. However, there may be a big change between the internal boundary and external boundary of a MB, while BMA assumes they are smooth enough. Thus, the MB boundary matching standard of lost MBs in our paper is defined as the difference between both the external boundary of the corresponding MB in the reference frame and the MB in the current frame.

H.264/AVC has introduced multiple reference frames, which helps predicting the motion vectors of the block more accurately. And we gather statistics about the probabilities of which reference frame is selected in actual motion estimation using Lagrange rate distortion optimal search strategy. Several standard test video sequences are used in this experiment. Based on the experimental results we know that the tree motion prediction selecting the previous frame as the reference frame achieves a large percentage up to 90%. Thus, boundary matching recover step only



**Fig. 3** Illustration of the boundary matching



find the most match MB by calculating external boundaries difference comparing with the previous frame.

Most of methods including the Joint Model Reference [12] and the BMA use the luminance component of each pixel as a pixel matching criteria, leading to inaccurate matches when there are small changes in the brightness of the circumstances. Euclidean distance is used to measure the difference between the two pixels in our paper:

$$d(p, q) = \sqrt{(Y_p - Y_q)^2 + (U_p - U_q)^2 + (V_p - V_q)^2} \tag{1}$$

where  $Y_i, U_i, V_i$  denote the luminance component and two chrominance components respectively.

For the lost MB  $P_t$  in the current frame  $t$ , we use its external boundary for matching as Fig. 3 shows and the matching distortion from its reference MB in the previous frame  $t - 1$  can be measured by the following formula:

$$\begin{aligned} D(P_t) &= \frac{1}{N} (D_{\text{top}} + D_{\text{bottom}} + D_{\text{left}} + D_{\text{right}}) \\ &= \frac{1}{N} \left( \sum_{i=0}^{M-1} d(f(x_t + i, y_t - 1, t), f(x_t + mv_x + i, y_t + mv_y - 1, t - 1)) \right. \\ &\quad + \sum_{i=0}^{M-1} d(f(x_t + i, y_t + M, t), f(x_t + mv_x + i, y_t + mv_y + M, t - 1)) \\ &\quad + \sum_{i=0}^{M-1} d(f(x_t - 1, y_t + i, t), f(x_t + mv_x - 1, y_t + mv_y + i, t - 1)) \\ &\quad \left. + \sum_{i=0}^{M-1} d(f(x_t + M, y_t + i, t), f(x_t + mv_x + M, y_t + mv_y + i, t - 1)) \right) \tag{2} \end{aligned}$$

where  $N$  is the number of external boundary points, and  $N \leq 4M$ ,  $N$  get maximum value  $4M$  when none of the boundaries in the four directions is out of the range of the frame,  $M$  is the size of MB (e.g.,  $M = 16$  in H.264),  $D_{\text{top}}, D_{\text{bottom}}, D_{\text{left}}$  and  $D_{\text{right}}$  is the distortion of top, bottom, left and right boundary respectively.  $x_t$  and  $y_t$  represent the top and left coordinate of  $P_t$ ,  $mv_x$  and  $mv_y$  represent the motion vector of  $P_t$ .  $f(., ., t)$  stands for current frame and  $f(., ., t - 1)$  is the previous frame.

The minimum distortion ( $MD$ ) is defined as the minimum average difference between the external boundaries of the candidate block in the previous frame and the lost block in current frame as the following:

$$MD = \min (D(P_t)) \tag{3}$$

where  $-\omega \leq mv_x \leq \omega, -\omega \leq mv_y \leq \omega$ ,  $\omega$  is the search range corresponding to the current MB, in our experiment we set  $\omega = 16$ .

Correspondingly, the best matching motion vector is:

$$mv_{P_t} = \arg \min (D(P_t)) \tag{4}$$

## 4.2 Motion characteristic detection

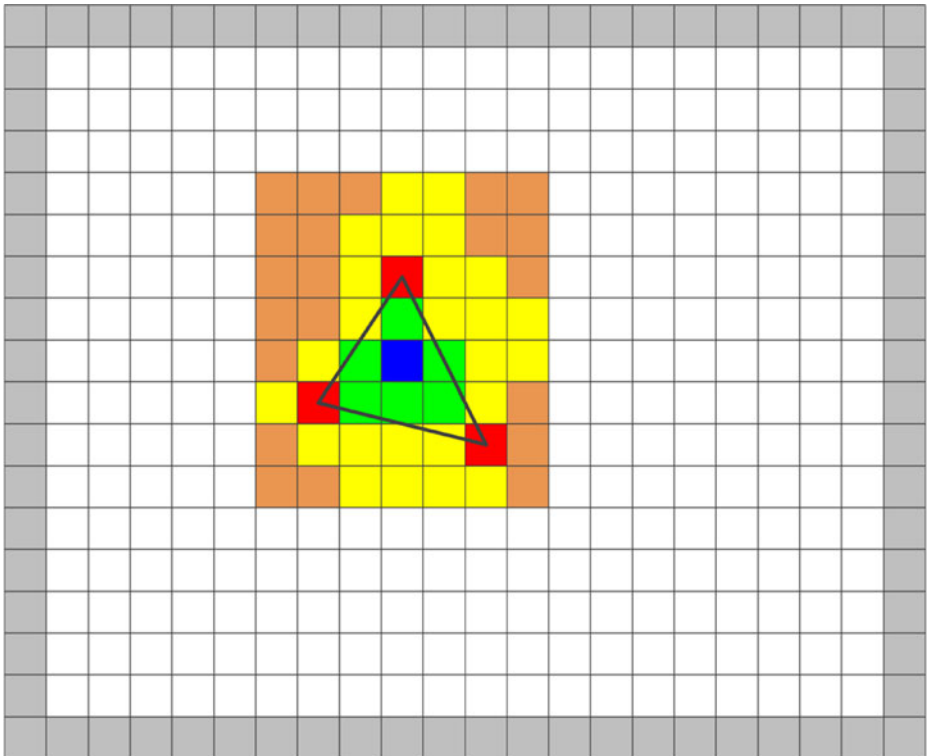
Since the decoder can't know the actually MAD values of lost MBs, in our approach, we use direct MAD ( $MAD_{\text{direct}}$ ) [19] to solve this problem, which could be estimated directly as its name shows. The  $MAD_{\text{direct}}$  value is used to measure the motion of the MB approximately.

However, the MAD not always represents the high-motion MBs such as in the case of isolated point. Our approach can't directly use the MAD value to estimate the high-motion MBs. We propose a Greedy Spread Motion Region Extraction (GSMRE) algorithm as the following steps to extract high motion regions effectively.

In Fig. 4, each box represents a MB, Red MBs are the seed MBs, Blue and green MBs is the spread MBs in the first call.

### 4.2.1 Initial the high-motion region

With the assumption proposed by Lee et al., we know that the HVS focuses on the center of the picture, and the high-motion regions always concentrate in the center of the picture [18]. Therefore, the MBs in the upper, left and right columns are excluded outside the high-motion regions as in Fig. 4. The gray MBs are the excluded MBs. Another reason for initialing the high-motion region is that the boundary MBs sometimes have large MAD values, which have bad effects to the seed MBs selection.



**Fig. 4** Greedy Spread Motion Region Extraction

#### 4.2.2 Select the MB seeds

In the initial region, our approach will select three MBs which have the largest  $MAD_{\text{direct}}$  value as in Fig. 4. The red MBs are the MB seeds. We add the three MB seed seeds into the high-motion region set  $S_h$ .

Next, link the three MB seeds into a triangle, and add the MBs included in or intersect the triangle into  $S_h$ . The strong orange MBs are all added into  $S_h$ . Initialize the non-high-motion region set  $S_n$  to empty. Calculate the average  $MAD_{\text{direct}}$  of the three MB seeds as initial  $MAD_{\text{direct}}$ .

#### 4.2.3 Rectangular the MBs region

Choose the smallest rectangle  $R_{\text{tem}}$  which can include all the MBs in  $S_h$ . Let  $x_l$ ,  $x_r$  represent the leftmost and rightmost MB's x-coordinate in  $S_h$ , and let  $y_t$ ,  $y_b$  represent the topmost and bottommost MB's y-coordinate in  $S_h$ . Let  $R_H$  represents the result of the high-motion region, then the top left point of  $R_H$  is the MB at  $(x_l, y_t)$  and the bottom right point of  $R_H$  is the MB at  $(x_r, y_b)$ . If the area of  $R_{\text{tem}}$  is larger than quarter of the frame, the detection process terminates, if  $R_H$  have no value, then  $R_H$  can be initialized by  $R_{\text{tem}}$ .

#### 4.2.4 Greedy spread the MBs region

For each MB in  $S_h$ , add its four-neighbor MBs into the check set  $R_c$  if it not in  $R_n$ . If  $R_c$  is empty, the detection process terminates. Otherwise, sort  $R_c$  ascending according to MB's  $MAD_{\text{direct}}$  value. Pop the largest MB in  $R_c$  into  $S_h$  until the average  $MAD_{\text{direct}}$  value in  $S_h$  is smaller than  $0.90 * MAD_{\text{direct}}$ . And then, pop all MBs in  $R_c$  and the last MB in  $S_h$  into  $R_n$ , loop the "Rectangular the MBs region" step.

The final detected high-motion region is  $R_H$  (colored MB in Fig. 4) and the left MBs in the frame belong to the low-motion region  $R_L$ . The result of this step is that the high-motion object is been included into the high-motion region and the low-motion region include stillness object.

### 4.3 High-motion MBs error concealment

The high-motion region is the HVS's sensitive region and has motion smoothness in temporal domain. Motion field transfer is effective to recover the slice lost and frame lost separately in accordance with different motion characteristic of video [20, 27]. Here, we also adopt a modified motion field transfer method to estimate the lost motion characteristic which can be used to propagate the pixel color.

In most cases, the inner patches and the high-motion region of the same frame belong to the same object, and have similar move characteristic. There is the biggest difference between high-motion extraction and foreground subtraction. The high-motion region extraction aims at obtaining the MBs with distinctive movement while the foreground subtraction aims at obtaining the interest object. Since surrounding patches of the high-motion of each frame belong to low-motion and the inner patches belong to high-motion, both of the high-motion and low-motion region can use the spatial information from each other to recover them. Although we can use the following motion field transfer method to recover the lost motion field as a whole region, it may cause imprecise estimation since the region is always too large. In

our method, we don't recover each MB as the minimum unit. We split each high-motion MB into four 8\*8 patches and recover each patch's motion field separately. We recover only one pixel of one color each time.

Former motion field transfer [20, 27] recover the motion field of each MB along with the raster sequence and the previous recovered motion field is the base of the following ready recover motion field. Such a strategy may cause error spread and the color may wrong more further. In the paper, using the feature of chessboard FMO that most lost MBs have four neighboring correctly decoded MB, which can avoid this error spread well. In the following subsection, we will describe the details.

#### 4.3.1 Local motion estimation and dissimilarity measure

The first step is to computer the local motion field. To do this, we use the two-frame G. Lucas-Kanade optical flow computation method [8, 23]. We represent the estimated the motion vector at point  $P = \{x, y, t\}^T$  in the video sequence by  $(u(P), v(P))^T$ . These 2D optical flow can be viewed as 3D vector in spatia-temporal domain with the constant temporal element being  $t$ , the 3D vector  $m$  is defined as  $m \equiv (u_t, v_t, t)^T$ . Then we can define the distance between of two motion vectors using angular difference (in 3D space) as following:

$$d_m(m, \tilde{m}) = 1 - \frac{m \cdot \tilde{m}}{|m| * |\tilde{m}|} = 1 - \cos \theta \quad (5)$$

Where  $\theta$  is the angle between two motion vectors  $m$  and  $\tilde{m}$ . This angular error measure accounts for differences in both direction and magnitude, since measurements are in homogeneous coordinates.

#### 4.3.2 Motion field transfer

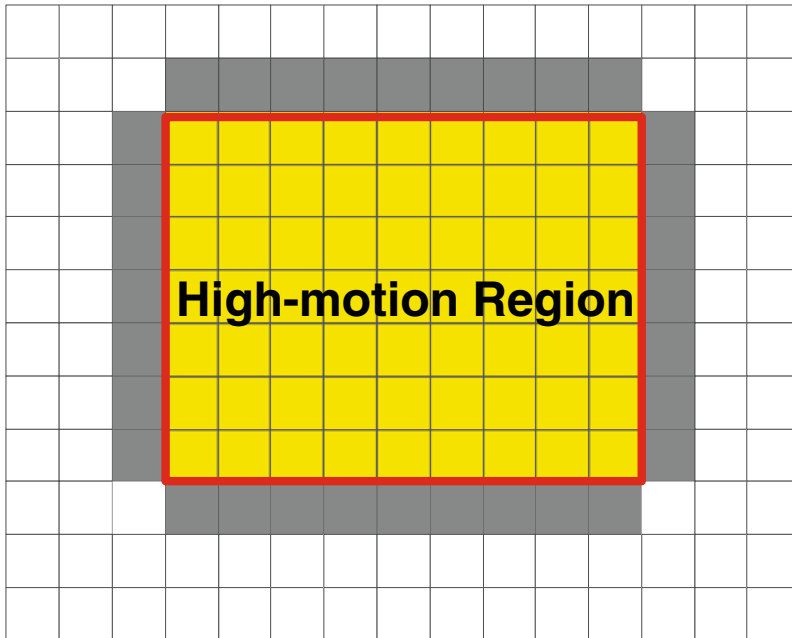
Using the dissimilarity measure defined in formula (5), we can search the most similar patch with the lost patch in the previous and follow-on frames. As the do, we also define the aggregate distance between the source and target patches as:

$$d(P_s, P_t) = \sum_{m_s \in P_s, m_t \in P_t} d_m(m_s, m_t) \quad (6)$$

Where  $m_s$  is a pixel in source patch,  $m_t$  is the corresponding pixel with  $m_s$ , they have the same relative position with its parent patch. The number of corrected received pixels in  $P_s$  is a fixed number. For reduce the computation complex, we don't calculate the average dissimilarity. If our method searched in one neighboring frame  $m$ , the optimal object patch is obtained by minimizes the following formula as:

$$\widehat{P}_m^t = \arg \min d(P_s, P_t) \quad (7)$$

It is important to point out that the each frame's search region in the following sub-section is not the whole frame. We also don't use the high-motion region as our search region because its acquiring the high-motion region obtained by the GSMRE is restricted by many issues, for example the high-motion region area constraint, etc. For getting the search region more precisely, we enlarge boundary of the high-motion by one MB as Fig. 5 shows. The grey patch is the extended patches. One of advantages for using this method is reducing the search region in each frame comparing with the full frame search region, resulting in reducing the computational complex.



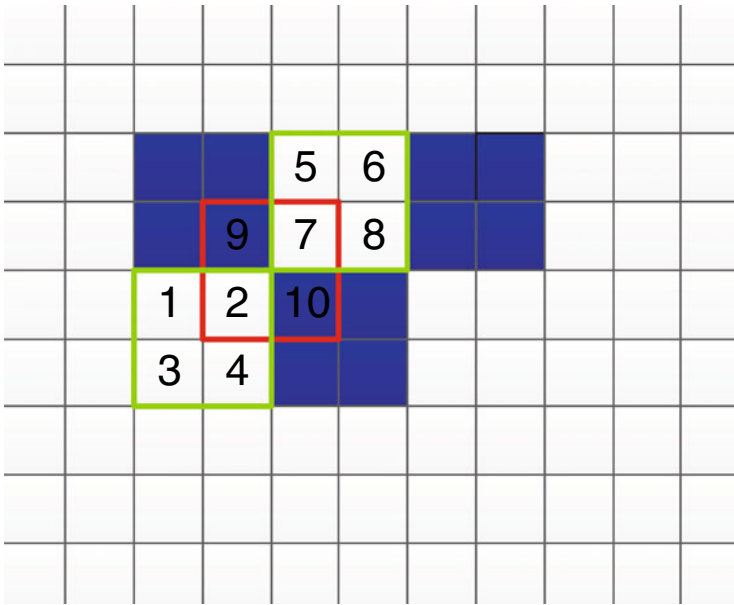
**Fig. 5** Motion field transfer

The transfer order is not very important for the transfer step. In this paper, we start from the top most left boundary of the high-motion region, progressively towards the bottommost right MB in a raster sequence. As Fig. 6 shows, for each MB to be recovered, it will be divided into four  $8 \times 8$  sub-patches, and at least half of the four patches is used to calculate the dissimilarity value, that means up to half of the patches is the lost pixels. For example, the deep blue box in Fig. 5 is the correctly received MBs, and the green box MBs is the lost MBs whose motion field is unknown. Patches 1–8 are the  $8 \times 8$  sub-patches which divide the green box into eight patches. We choose the red box MB as the motion field transfer MB, then Patches 9 and 10 are the correctly received patches and Patches 2 and 7 are the patches which should be recovered. Patches 9 and 10 are the known motion field patch which used to calculate the dissimilarity value of the MB and then we can obtain the motion field of Patches 2 and 7. According to these methods, each lost MB will be recovered.

Computational complexity is also an important issue. Slow search speed may prevent it to be applied in real-time environment. We adopt parallel computation technology to search the previous and next frames in parallel, then we can complete this work in about half time in dual-CPU. Additionally, we use a fast search block based strategy proposed by Venkatachalapathy et al. [31] to reduce the count of calculation. The detail analysis will be proposed in the Computational Cost Analytic subsection.

#### 4.3.3 Color propagation

In Shiratori et al.'s method, each pixel's propagative pixels are the connected pixels in the previous frame. It is under the situation that the connected pixels in the previous frame may move to the current pixel. In high-motion video sequence, this



**Fig. 6** Motion field transfer

assumption can't hold since high-motion causes the previous frame's pixels move to some faraway position of the current frame. So, in the initiative step, we should locate the propagative region for color propagation.

For each patch, the patch which has the most similarity with it in previous frame will be searched, the dissimilarity measure method use the formula defined in formula (6). All the above patches can form a graph as show in Fig. 7. The motion vector motion can be treated as undirected edges which represent pixel correspondences among frames.

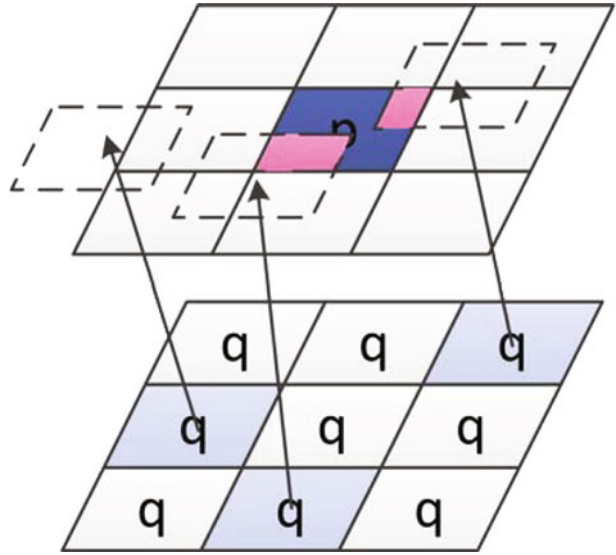
Suppose we are to assign a color value to a pixel  $p$  using the connected pixels  $q \in \pi$ ,  $\pi$  is initialized as the corresponding pixel and its eight-connected pixels in the previous frame. A point  $q$  in the previous frame may be connected to a fractional location of pixel  $p$ . We use the sizes of overlapped areas  $s(p, q)$  as weight factors to determine the contribution of neighboring pixels  $q$  to pixel  $p$ . We also use the reliability of the edge  $r(p, q)$ , which is measured by the inverse of dissimilarity measure defined in formula (6). The contribution from the neighboring pixel  $q$  to the pixel  $p$  is given by the product of  $r$  and  $s$  as:

$$w(p, q) = s(p, q) * r(p, q) \tag{8}$$

Thus, the color  $c(p)$  at pixel  $p$  is a weighted average of colors at the neighboring pixels  $q$ :

$$c(p) = \frac{\sum_{q \in \pi} w(p, q) * c(q)}{\sum_{q \in \pi} w(p, q)} \tag{9}$$

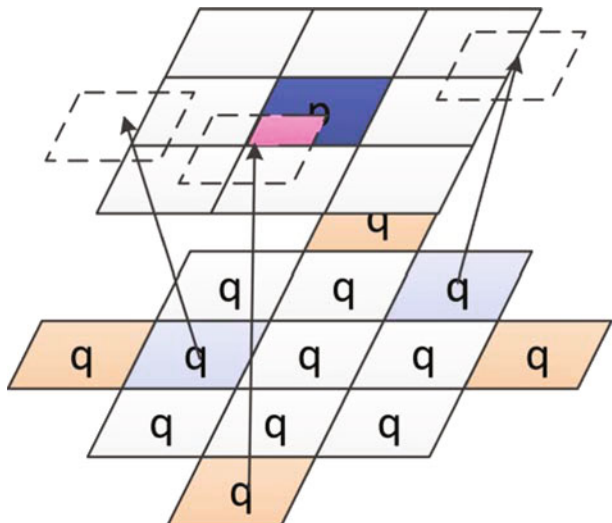
**Fig. 7** Eight-connected region color propagation



In the above formula, the pixels set  $\pi$  should change along with the calculate of  $w(p, q)$ .

1. If any eight-connected pixels'  $w(p, q)$  is less than the threshold value  $T_w$ , it will be removed from  $\pi$ .
2. If  $\pi$  is empty, extend the propagative region to the pixels whose Manhattan Distance with the current pixel is less than 2 pixels as shown in Fig. 8, and check the pixels which belong to this region instead of eight-connected region using the method as 1) does. Else fix the pixel set  $\pi$  and calculate the color using formula (9).

**Fig. 8** Extended region color propagation



3. If the final  $\pi$  is also empty, then the previous propagative region is not enough for the current pixel, and the color of current pixel is empty. The maximum  $w(p, q)$  and its corresponding  $c(q)$  are saved, if already have maximum  $w(p, q)$ , then replace it with the bigger one. Else fix the pixel set  $\pi$  and calculate the color using formula (9).

In the above method, there may exist some pixels that do not overlap with any pixels, i.e. the previous frame is not enough for propagation. If this happens, our method will search the best matching region in the next frame, reversed lengthen the edge of each pixel to the current frame. Then we can use the same method above to the next frame to propagate the left pixels. If the follow-on is not buffered, then we can ignore this step.

After the above two color propagation steps, there may also exist pixels have empty color. The number of these pixels is very small. If any these pixel have maximum  $w(p, q)$  we can use formula (9) to recover its color, otherwise inherit the color of the corresponding pixel in the previous frame since the pixel has no movement.

#### 4.4 Low-motion region error concealment

The low-motion and high-motion recognition in the encoder is different from the high-motion detection. Background is the actually litter move pixels in a frame, while the low-motion region is the litter move MB in a frame. Mean-while, limited by the area of high-motion region, there are many pixels in low-motion region while belong to high-motion pixels. These pixels is the clue for concealing the low-motion region.

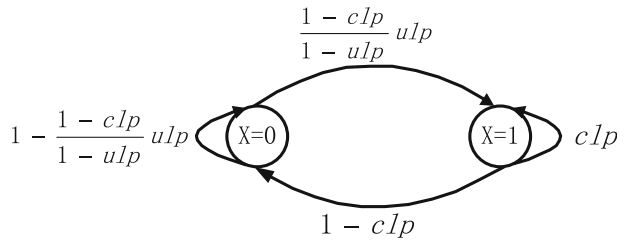
## 5 Experiment and performance

Our experiments compare the proposed MDEC algorithm with the other four algorithms, including the boundary matching algorithm (BMA) error concealment algorithm, block-based motion vector extrapolation (BMVE) approach, error concealment based on image inpainting (IP), the method proposed in Joint Model (JMEC). The search mode in BMA is the classic mode and the boundary width for the motion estimation of lost MBs is set to 1. We use the joint video team reference software JM 17.2 (baseline profile) in the experiment, and test video sequences are 100 frames of Foreman\_CIF, Sign\_Irene\_CIF, Akiyo\_CIF and Silent\_CIF, compressed at 30 frames per second. For the whole frame concealment, the previous frame is directly copied to the current frame. The default fast full search is used for motion estimation. In the experiment, each frame is encoded into two slice groups, using the chessboard FMO pattern. The Group of Picture (GOP) size is also adjusted, there are two types of GOP size in our experiment, first of which is set to be 15, and the other one is only the first frame is encoded as an INTRA-frame, with all the subsequent ones INTER-frames. The maximum number of reference frames is 5. The initialized Quantization Parameter (QP) is adjusted to achieve different bit rates during the transmission.

And a network model is used to simulate the packet loss. The two states Markov chain model in [4] could simulation the internet packet loss well. As shown in Fig. 9, the state “1” is the loss state and the state “0” is the non-loss state, which



**Fig. 9** Markov model of packet loss

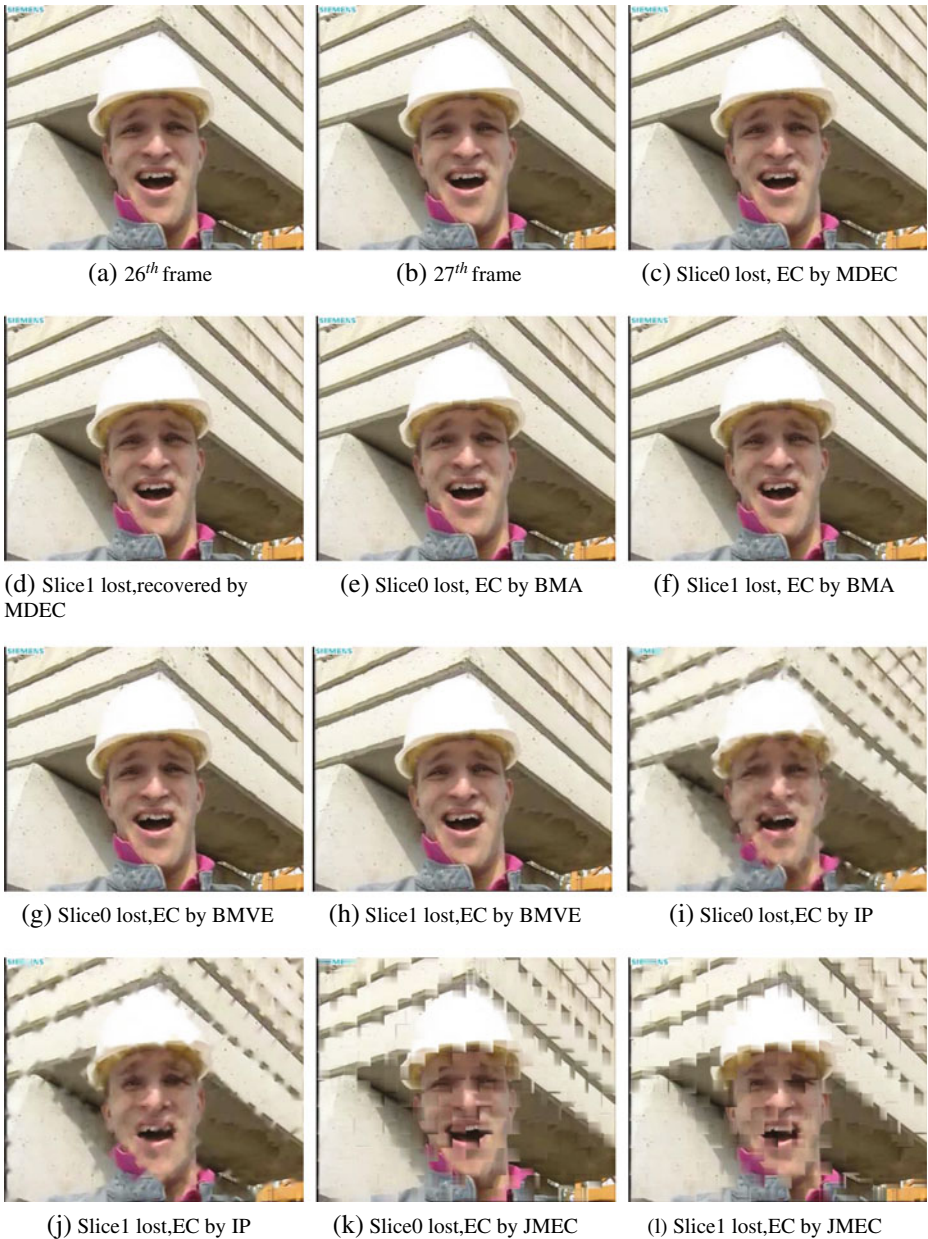


represent the packet is lost and the packet is received correctly respectively. This model allows for correlated losses and one of the parameters of the model is the so called conditional loss probability, *clp*, which is the probability that a packet is lost (loss state), given that the previous packet was lost (loss state). The higher this probability, the more correlated the losses are. The other parameter of the model is the unconditional loss probability, *ulp*, which is the overall probability that a packet is lost. Figure 10 illustrates the visual quality after using the above three error concealment method on Foreman(CIF). The first two pictures are the original frame. Comparing with BMA, BMVE, IP and JMEC, MDEC can transfer the motion characteristic better. The concealed frame of BMA is very similar to its previous frame, i.e. the BMA lost some detailed motion characteristic and take along block artifacts. BMVE concealed frame reflects a litter motion characteristics while pull-in block artifacts. We can also observe from the picture that the concealed frame by MDEC looks much better than it of the other two methods, and have no block artifacts. Figure 13 in the end of this paper shows more details about visual quality obtained by MDEC.

In our experiment, the loss rates are  $P = ulp = 3, 5, \text{ and } 7\%$ . Decoder PSNR is used as the objective measurement, which is computed using the original uncompressed video as reference. Given a packet loss rate  $P$ , the video sequence is transmitted 100 times, and the average PSNR of the 100 transmissions is calculated at decoder. Figure 11 illustrates the PSNR after applying different error concealment algorithms on four video sequences. The rate-control of Joint Model is enabled, and the bit rate is limited to 256, 384, 512, 640 and 768 kbps, respectively.

Figure 11 shows the Rate-distortion curves for different sequences and different error concealment method when GOP size is 15. With MDEC, the PSNR could increase nearly 1.5 dB than BMVE when packet loss rate is 7%. From the figure we can see that the RD curves of MDEC are much higher than those of BMVE, BMA, IP and JMEC, i.e., Akiyo\_CIF sequence, when  $P = 3\%$ , 1.6 dB higher than BMVE and 1.8dB higher on average; when  $P = 7\%$ , 1.2 dB higher than BMVE and 0.5 dB than BMA on average. Specifically, for a sequence with high movement, such as Sign\_irene, MDEC performs especially well. For example, when the loss rate  $P$  is 3%, the proposed MDEC 0.7 dB PSNR higher than BMA on average. In the whole experiment, IP and JMEC performed bad.

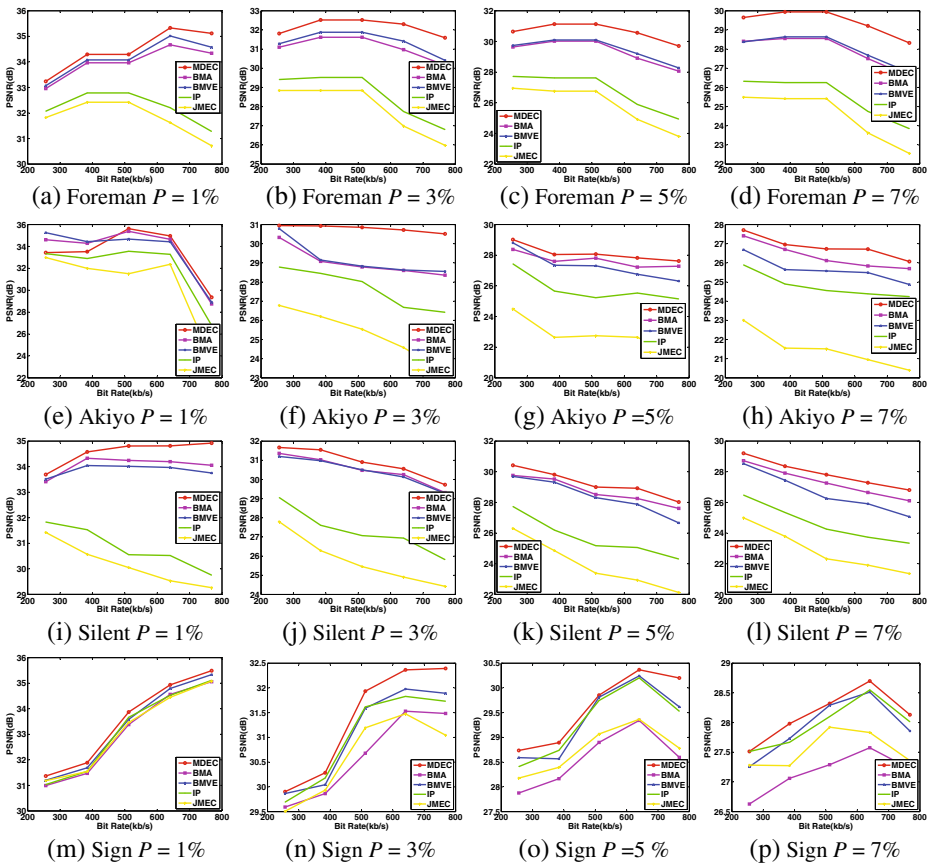
We can see from Fig. 11 that the PSNR decreases when bit rate increases. That is because of each slice is partitioned into more packets when at high bit rate than that at low bit rate, then if one of its packets get lost, the slice may not be decoded correctly, making it more likely to get damaged. Meanwhile, frame copy leads to error accumulated. In the next research, this problem will be further researched.



**Fig. 10** Visual results of applying different error concealment algorithms on “Foreman\_CIF” for one slice loss (encoded with checkerboard pattern, 27th frame,  $QP = 28$ )

### 5.1 Computational cost analytic

The biggest difference between our method and former methods is that our method doesn’t need to decode the recovered data stream, for our method obtains the final color of each lost pixel, which saving a lot of decoding time.



**Fig. 11** RD curves of different EC algorithms, GOP size is 15,  $QP = 28$

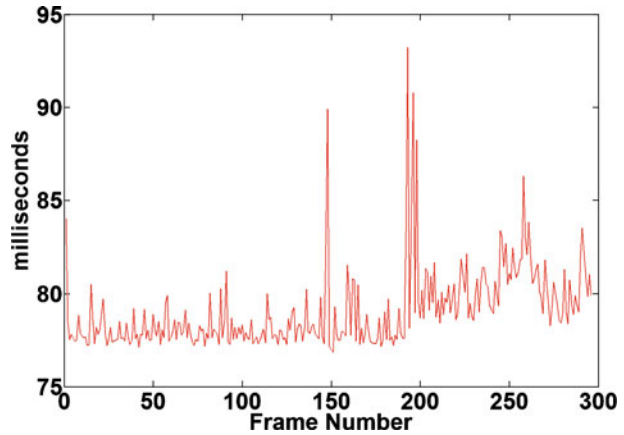
In addition, the motion field transfer is the most time-consuming part of our method. There are three small steps in the motion field transfer:

1. Calculate the Lucas-Kanade optical flow.
2. Search the similar patch in the adjacent frames.
3. recover color of each lost pixel.

Recover color of each lost pixel is always complete in 5 ms, without further optimizations. We only optimize the first two computation steps. Under statistics, the area of high-motion region is less than a quarter of the whole frame. It means that a CIF video sequence’s high-motion region sequence’s frame size is about QCIF. The motion field transfer only need to know the moving trend, but not need to know the precise optical flow. Thus, little iteration is needed in calculating the Lucas-Kanade optical flow. Five iterations cost about 30 ms to calculate each QCIF frame’s Lucas-Kanade optical flow.

For searching similar patches, since they are searched in different frames independently, it can be assigned in different CPUs in a multi-core CPU using

**Fig. 12** Motion field transfer time, sequence: Foremain\_CIF,  $QP = 28$



parallel computing technology. If searching in two frames, the compute work can be completed in about half time compare with the serial execution. In real experiments, the parallel computing technology accelerates the search work about 0.48 times.

The experiment result shows that, the fast search strategy in [31] complete the searching work in 45 ms in average, compare with 15,220 ms of full search, fast search algorithm can accelerates the search speed about 338 times. The actual calculate time is as the above figures. Figure 12 shows that most of the motion field transfer time is between 75 and 80 ms. For real-time, this computing time is acceptable in a similar configuration of computer hardware.

## 6 Conclusion

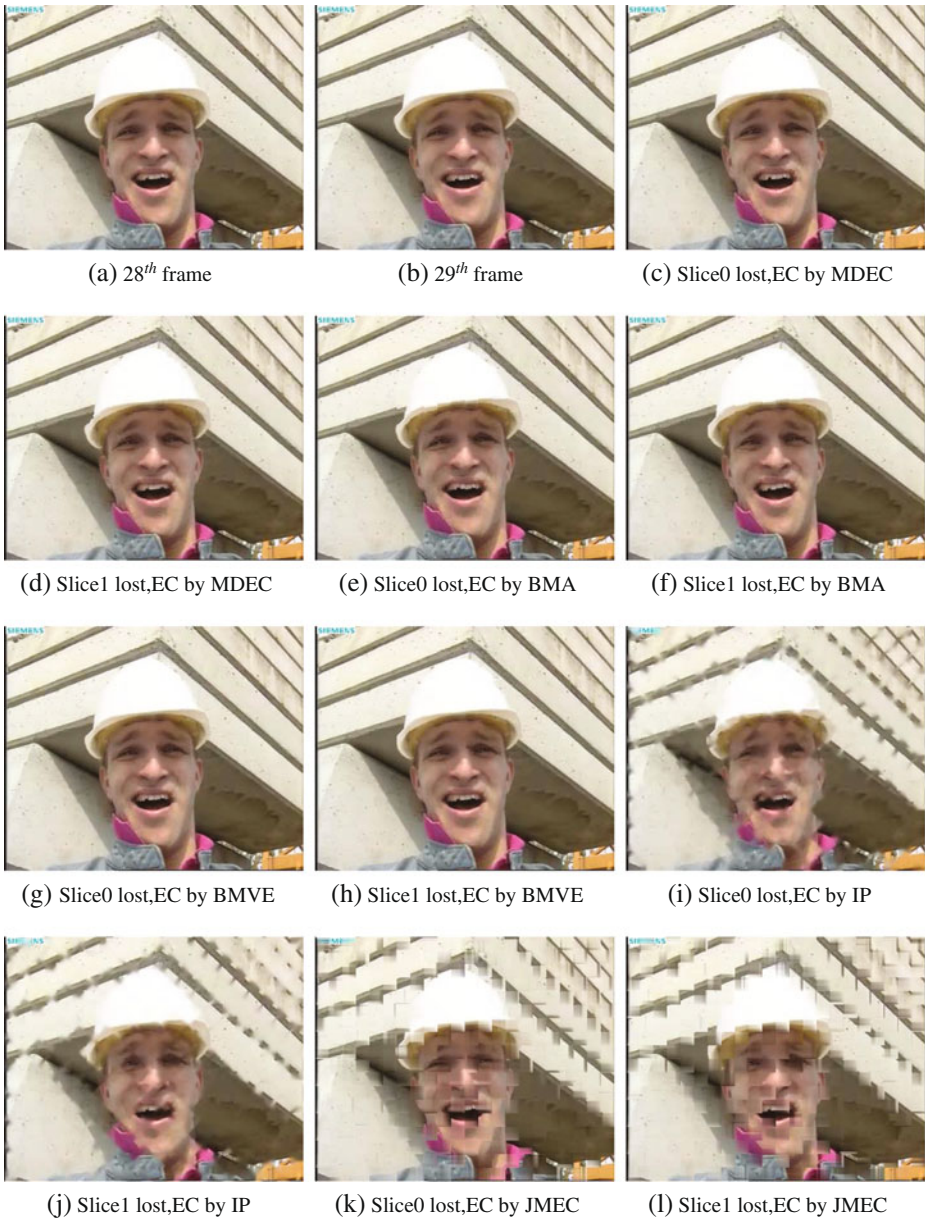
In this paper, we proposed a novel motion characteristic differentiated error concealment (MDEC) method based on motion field transfer. It represents the motion change trend by motion fields. Then different conceal methods are applied for different regions according to their motion characteristics. Meanwhile, the FMO checkerboard pattern is used at the encoder, so as to prevent MBs of a large area getting lost. At the decoder, the proposed GSMRE method is used to distinguish low-motion region from high-motion region in each frame based on different motion characteristic. We adopt parallel computer technology and fast block search strategy to reduce the computer time of motion field transfer operation.

Simulation results show that the proposed MDEC can reconstruct the corrupted frame with a higher PSNR value than BMVE, BMA, IP and JMVC. Besides, MDEC achieves a better subjective video quality than the other four methods. the PSNR gain of our approach over boundary matching algorithm reached about 0.6 dB, and 1.4dB when the packet loss rate is 3% and 7% respectively, which demonstrates that our method has an good application within a wide scope of packet loss rate. The experiments shows that frame copy may cause badly error spread when the whole frame is lost. So, our further work will focus on the entire frame recovery.

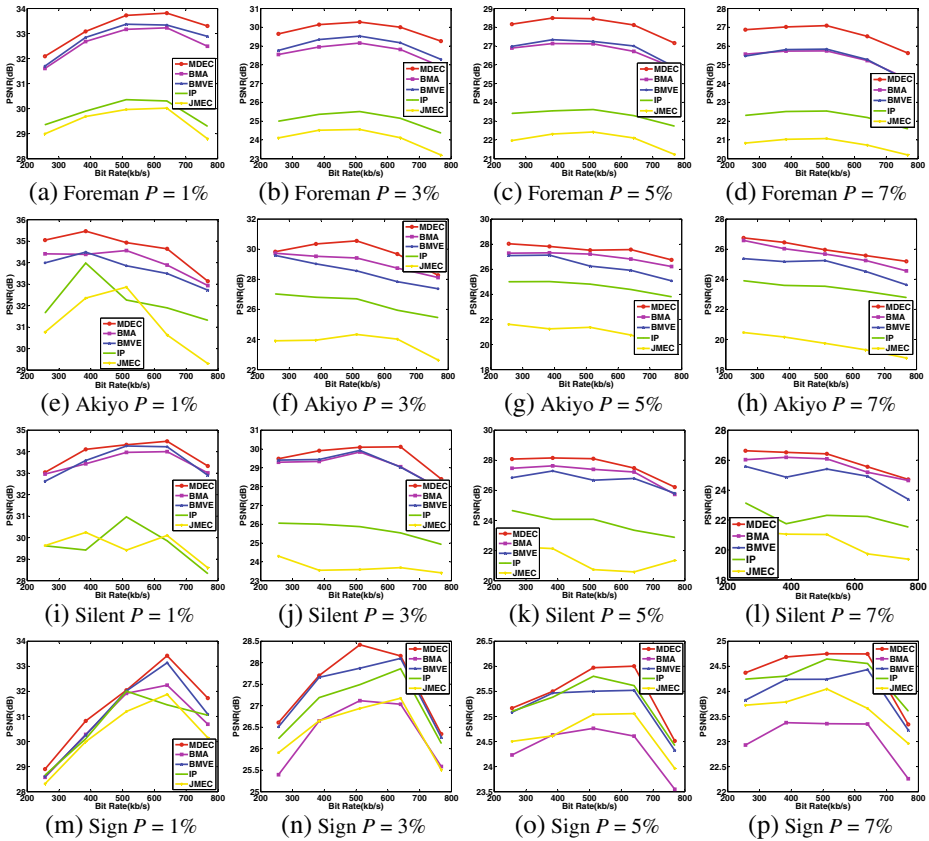


**Acknowledgement** The research was supported by National 973 Program of China (Grant No. 2005CB321901).

## Appendix



**Fig. 13** Visual results of applying different error concealment algorithms on “Foreman\_CIF” for one slice loss (encoded with checkerboard pattern, 27th frame,  $QP = 28$ )



**Fig. 14** RD curves of different EC algorithms, only first frame is INTRA-frame

**References**

1. Agrafiotis D, Bull DR, Canagarajah N (2006) Spatial error concealment with edge related perceptual considerations. *Signal Process, Image Commun* 21(2):130–142
2. Belfiore S, Granetto M, Magli E, Olmo G (2003) Spatio-temporal video error concealment with perceptually optimized mode selection. In: *Proc. IEEE int. conf. acoust., speech and signal processing*, vol 5, pp 750–753
3. Bertalmio M, Bertozzi AL, Sapiro G (2001) Navier–Stokes, fluid dynamics, and image and video inpainting. In: *Proc. conf. comp. vision pattern rec.*, vol I. Hawaii, pp 355–362
4. Bolot JC (1993) Characterizing end-to-end packet delay and loss in the internet. *J High-Speed Networks* 2:305–323
5. Chen Y, Yu K, Li J, Li S (2004) An error concealment algorithm for entire frame loss in video transmission. In: *Proceedings of the picture coding symposium*. San Francisco, California, USA, pp 389–392
6. Chen Y, Hu Y, Au O, Li H, Chen C (2008) Video error concealment using spatio-temporal boundary matching and partial differential equation. *IEEE Trans Multimedia* 10:2–15
7. Criminisi A, Prez P, Toyama K (2004) Region filling and object removal by exemplar-based image inpainting. *IEEE Trans Image Process* 13(9):1200–1212
8. Farneback G (2001) Very high accuracy velocity estimation using orientation tensors, parametric motion, and simultaneous segmentation of the motion field. In: *Proc. international conference on computer vision*, vol 1. Vancouver, Canada, pp 171–177

9. Hsia S-C, Cheng S-C, Chou S-W (2005) Efficient adaptive error concealment technique for video decoding system. *IEEE Trans Multimedia* 7(5):860–868
10. Huang Y-L, Lien H-Y (2005) An edge-based temporal error concealment for MPEG-coded video. In: *Proc. soc. photo-optic. instrum. eng. vis. commun. image process.*, pp 992–1000
11. Hwang M-C, Kim J-H, Duong DT, Ko S-J (2008) Hybrid temporal error concealment methods for block-based compressed video transmission. *IEEE Trans Broadcast* 54:198–207
12. Joint Model (2011) The JVT reference software for H.264/AVC. <http://iphome.hhi.de/suehring/tml/download>
13. Jung B, Jeon B, Kim M-D, Suh B, Choi S-I (2004) Selective temporal error concealment algorithm for H.264/AVC. In: *Proc. IEEE international conference on image processing*
14. Kim W, Koo J, Jeong J (2006) Fine directional interpolation for spatial error concealment. *IEEE Trans Consum Electron* 54(3):1052–1058
15. Kim D, Jeong S, Choi J, Jeon G, Kim S, Jeong J (2008) Error concealment using intra-mode information included in H.264/AVC-coded bitstream. *ETRI J* 30(4):526–535
16. Kim M, Lee H, Sull S (2008) Spatial error concealment for H.264 using sequential directional interpolation. *IEEE Trans Consum Electron* 56(4):1811–1818
17. Lam WM, Reibman AR, Liu B (1993) Recovery of lost or erroneously received motion vectors. In: *Proc. IEEE int. conf. acoust. speech signal process.*, pp 417–420
18. Lee S, Pattichis MS, Bovik AC (2002) Foveated video quality assessment. *IEEE Trans Multimedia* 4:129–132
19. Liu Y, Li ZG, Soh YC (2007) A novel rate control scheme for low delay video communication of H.264/AVC standard. *IEEE Trans Circuits Syst Video Technol* 17:68–78
20. Liu M, Chen S, Liu J, Tang X (2009) Video completion via motion guided spatial-temporal global optimization. In: *Proceedings of the ACM international conference on multimedia*, pp 537–540. Beijing, China
21. Luo Z, Song L, Zheng S, Xu Y, Yang X (2010) Improved error concealment of region of interest based on the H.264/AVC standard. *Opt Eng* 49:047003
22. Ma M, Au OC, Chan SHG, Sun MT (2010) Edge-directed error concealment. *IEEE Trans Circuits Syst Video Technol* 20:382–395
23. Open Computer Vision Library (OpenCV) (2011) <http://sourceforge.net/projects/opencvlibrary/>
24. Peng Q, Yang T, Zhu C (2002) Block-based temporal error concealment for video packet using motion vector extrapolation. In: *Proc. IEEE int. conf. communications, circuits and systems and West Sino expositions*, vol 1, pp 10–14
25. Ranjbar M, Kasaei S (2009) Spatial error concealment: a novel exemplar-based approach using segmentation. *Comput Electr Eng* 35:536–548
26. Ronfu Z, Yuanhua Z, Xiaodong H (2004) Content-adaptive spatial error concealment for video communication. *IEEE Trans Consum Electron* 52(1):335–341
27. Shiratori T, Matsushita Y, Kang SB, Tang X (2006) Video completion by motion field transfer. In: *Proceedings of IEEE conference on computer vision and pattern recognition*, pp 411–418
28. Su C-Y, Tsay S-H, Huang C-H (2006) Error concealment using direction-oriented candidate set and predicted boundary matching criteria. In: *Proc. IEEE int. conf. image process.*, pp 2221–2224
29. Telea A (2004) An image inpainting technique based on the fast marching method. *J Graphics Tools* 9(1):23–34
30. Tourapis A, Shring K, Sullivan G (2007) H.264/MPEG-4 AVC reference software manual. In: *24th JVT meeting, Doc. JVT-X072*, Geneva, Switzerland
31. Venkatachalapathy K, Krishnamoorthy R, Viswanath K (2004) A new adaptive search strategy for fast block based motion estimation algorithms. *J Vis Commun Image Represent* 15:203–213
32. Wang Y, Zhu QF (1998) Error control and concealment for video communication: a review. *Proc IEEE* 86:974–997
33. Yan B, Gharavi H (2010) A hybrid frame concealment algorithm for H.264/AVC. *IEEE Trans Image Process* 19:98–107
34. Zeng W (2005) Adaptive spatial temporal error concealment with embedded side information. *J Vis Commun Image Represent (Special Issue on Visual Communications in Ubiquitous Era)* 16(4–5):499–511
35. Zhang J, Arnold JF, Frater MR (2000) A cell-loss concealment technique for MPEG-2 coded video. *IEEE Trans Circuits Syst Video Technol* 10:659–665



**Heng Li** received the B.Sc. degree from School of Computer, Beihang University (BUAA), China, in 2003 and he was awarded as the best graduate student in Beihang University. In the same year, he was also selected to be a student in School of Advanced Engineering of Beihang University which is the best school in Beihang University. Only 1% of the first-year Ph.D. students can be selected into this school each year and receive the hardest mathematics and physics courses taught by the best experts in China. He is currently working toward the Ph.D. degree in the State Key Laboratory of Software Development Environment, School of Computing Science, Beihang University, Beijing, China. His research interests include image processing, video coding, content-based image and video retrieval.



**YanChun Zhong** received her B.Sc. degree from the School of Software Engineering, Zhejiang University of Technology, Zhejiang, China in 2008, and received her M.Sc. degree from the School of Computer Science, BeiHang University, Beijing, China in 2011. Her research interests include motion estimation algorithms, video error concealment, unequal error protection and ROI detection.