



Error concealment with parallelogram partitioning of the lost area

Sahar Hojati¹ · Mohammad Kazemi¹  · Payman Moallem¹

Received: 29 January 2019 / Revised: 29 October 2019 / Accepted: 26 November 2019/

Published online: 23 December 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

In this paper, we proposed a video error concealment algorithm using Motion Vector (MV) recovery for parallelogram partitions in the lost area. Error concealment is inevitable when some video packets are lost during transmission and correction or retransmission is not feasible. In conventional methods, MVs are recovered for the square shaped blocks which are then used for motion compensated temporal replacement. But in our proposed method, by parallelogram partitioning of the lost area, the MVs are found for more general shaped blocks. The parallelograms with various sizes and angles are examined, and then the best combination (size and angle) is selected with the assist of a border matching algorithm and a blind quality assessment method. Experimental results show that our method outperforms the other error concealment algorithms, both subjectively and objectively.

Keywords Error concealment · Dynamic programming · Parallelogram partitioning

1 Introduction

Due to the tremendous growth of the video consumer systems in different fields, such as multimedia broadcasting, video conferences and video telephony applications, video transmission and networking over wired or wireless channels are currently important topics and active subjects in research and industry communities. It is known that video transmission traffic will occupy the highest percentage of the internet consumer's traffic. The present network infrastructure and broadcasting technologies do not have the required capacity to transmit the large amount of raw video data, thus the necessity of video compression is more pronounced. However, compressing the video data leads to a high sensitivity of the bitstream to data loss. Data loss is experienced usually in communication channels, due to the bit error rate in the wireless channels or the congestion in the wired channels.

✉ Mohammad Kazemi
m.kazemi@eng.ui.ac.ir

¹ Electrical Engineering Department, University of Isfahan, Isfahan 8174673441, Iran

The newest video coding standard, High Efficiency Video Coding (HEVC), provides the best compression ratio compared to the H.264/AVC and the other preceding standardized codecs [29]. Even though HEVC was introduced for high resolution videos, i.e. HD, UHD and beyond, our experiments showed that, even for low resolution videos such as Common Intermediate Format (CIF), it is also better than H.264/AVC but with a more computational cost. For this reason, HEVC is the underlying codec in this paper, even though it is less error resilient than H.264/AVC as shown in [21].

Error Concealment (EC) is a post-processing technique that is used to cover up the packet losses or bit errors. In some applications the packets are rarely lost; in these cases if quickly retransmission of the lost data is feasible, error concealment is not needed. However, there are many applications that there is no feedback channel and correctly delivering the packets are not guaranteed; such as broadcasting the video (terrestrial, satellite or cable TV) over lossy channels. In these applications, some packets are certainly lost and they must be concealed if we want to preserve the video display. In the communications with the back channel, if Round Trip Time (RTT) is too long, the responses to the feedbacks of the lost packets arrive too late; therefore, error concealing the unreceived packets is inevitable.

Two loss scenarios can be assumed. In the first one, the packets are randomly dropped due to the congestions. In that scenario, all information encapsulated in the lost packet is not available. Each packet contains an integer number of slices and each slice composed of one or several Coding Tree Units (CTUs). This means that each missing packet leads to missing several CTUs. Bit error is another loss scenario which occurs in wireless communications. In that case, the erroneous bits make the slice completely or partly undecodable, due to the context adaptive entropy coding of the symbols and the predictive coding nature used in HEVC. Therefore, with the packet drop or bit error, one or several CTUs in a frame are not available at the decoder and must be concealed. This is the actual problem error concealment techniques deal with.

With the EC techniques, the decoder tries to estimate the lost video content by means of the known correlation between the available and the missed data [28, 30, 31]. This correlation exists between the spatially adjacent pixels of a frame, and the techniques which exploit this correlation to conceal the lost area are known as Spatial EC (SEC) techniques. On the other hand, the temporal adjacent frames are also correlated and this similarly is beneficial for error concealment. The error concealment algorithms that use temporal information are categorized as Temporal EC (TEC) techniques. There are also some hybrid algorithms that use both SEC and TEC techniques concurrently.

SEC methods reconstruct the lost area blocks using the information of surrounding pixels in the same frame. Clearly, these algorithms are applicable for the images and for the videos if the lost area and the available (intactly received) data are sufficiently close to each other [5]. However, it is not the case for relatively large loss areas in HEVC streams. As already mentioned, TEC methods restore the erroneous blocks using the correlation between the video frames; this correlation exists and is helpful even for large losses. Therefore, TEC methods are generally more efficient for HEVC error concealment.

In order to figure out the best block for replacement, many MV recovery works have been developed. However, they have used square blocks as concealment unit; i.e., MVs for 16×16 , 8×8 or 4×4 pixel square blocks are recovered and used for motion compensated replacement. In this paper, we have proposed to use the parallelogram blocks as the concealment unit; it is a generalized form of the rectangular blocks. The relatively large lost area is partitioned into several parallelograms, which are now more general than the squares and provide a higher

potential for more accurate error concealment. For each partition, an MV is obtained which is then used for lost area replenishment. This procedure is carried out for various angles and sizes of the partitions, then the angle and size that leads to the best quality are selected as the optimal concealment unit attributes. The best quality output is selected by a boundary smoothness measure and a no-reference (blind) Image Quality Assessment (IQA) method.

The paper is structured as follows: The related works are reviewed in Section 2. In Section 3, we explain our proposed method step-by-step; the formulation and the procedure to examine the various partitioning and then how to select the best one is described in this section. Our method is evaluated and compared to other concealment techniques in Section 4; and finally, the paper conclusion is provided in Section 5.

2 Related work

There are some techniques used for concealing the whole frame loss. One simple and efficient method is to use MV of the collocated block in the prior frame for motion compensated replacement. This method is known as Motion Copy. This approach needs to know the MVs of the previous frame. As another approach, a method for temporal error concealment is Decoder Motion Vector Estimation (DMVE) [34], in which the boundary pixels of the lost MacroBlock (MB) is used for motion estimation. In that method, a full search within the prior frame is carried out to find an MB which its boundaries have the least difference with the up, down, and left boundaries of the lost MB. One can use either the inner or outer boundary for matching computation. The inner Boundary Matching Algorithm (BMA) and testing several candidate MVs was firstly presented in [13]. The algorithm presented in [12] examines both the outer and the inner BMA, and select one which its output provides a smoother border between the concealed and loss-free area. A particle filtering can be used for denoising the MVs recovered by the BMA [23]. In [33], the authors take advantage of the BMA and object shape preservation. In that method, the objects are detected in the reference frame and the appropriate motion vector is found by applying the motion estimation on the detected objects. With the aim of BMA, error concealment is performed for preserving the objects' shapes. However, as already explained, the losses are sometimes too large so that the object and its boundary may be completely unavailable. In [35], MVs of the erroneous MBs can be obtained by the BMA and the lost area is filled in the first round. Then an auto-regressive model is fitted to the available spatial-temporal neighboring pixels which is then used to refine the recovered pixels. A similar solution but using the sparse optimization approach is presented in [17]. These methods are not actually appropriate for large lost areas, due to special error propagation caused by the model dependency on the spatial information. There are also some other works propose to use sparse recovery for error concealment, the examples are [2, 19] where the former does not work for standard bitstream and the latter is suitable for spatial error concealment. In the method presented in [9], a homography-based registration is performed between the available pixels around the lost area and their counterparts in the reference frame, it is then used for mapping the matching points at the lost area. The authors propose to find this registration according to the whole frame as well as a patch around the lost area and then apply which one provides smaller distortion. This procedure is applied in both forward and backward directions, and the remaining points are then filled by the SEC. However, the success of this method depends on the correlation of the lost area with the available surrounding pixels, which is not necessarily provided specially in CTU losses in HEVC. With the same reason, the

sapio-temporal works such as the one presented in [26] are not appropriate for HEVC loss concealment. In another method, the lost block surrounding pixels are related to the reference pixels with a tensor model, which is then used for pixel recovery [36]. In [7], the MVs of the collocated blocks in the prior frame are used for loss concealment, but if the residual of the collocated block is relatively high or it is intra coded, its MV is unreliable. These blocks are merged together; then the MV of the merged block is set as the average of the surrounding blocks' MVs. Interpolation of MVs with the help of a plane fitted to the surrounding MVs is presented in [14].

There are also some attempts for error concealment with smaller blocks; i.e. 8×8 and 4×4 pixel blocks. Some neighboring MVs have a stronger correlation to the MV of the lost MB than the others, and hence they should have more impact in MV recovery, as presented in [11]. In that paper, with processing the MVs of the collocated and correctly available MB in the previous frame, a tendency of the MV in each 4×4 block is estimated. This tendency actually shows that the MV of the current block is closest to which two 4×4 blocks' MVs. A pixel-wise MV can be obtained as the weighted average of the available nearest pixels' MVs around the lost MB as presented in [16]. The weights are inversely proportional to the distance of the lost pixel and the surrounding available pixel. This way, an MV for each pixel is recovered which can be used for pixel replacement from the earlier frame.

To combat the burst loss of MBs, there are some algorithms which presented for H.264/AVC but are applicable for HEVC as well. The challenge is this fact that the concealing an MB will affect the concealment of the other MBs in the lost area. A solution is presented in [22] where the concealment order of MBs in the lost area is prioritized. The MB which has neighboring (available or concealed) MBs with higher texture has higher priority for concealment. The MV is recovered by DMVE algorithm [34] and it is used for error concealment of this MB. As the next step, the next highest priority MB is found again with this approach; the information of already concealed MBs is also exploited. Another approach for MV recovery in the case of missing several successive MBs is presented in [15]. In [15], Lie et al. have developed a method that estimates MV of each block considering the possible replacements for the boundary pixels of spatially adjacent blocks. In other words, the MVs of the lost MBs are found by considering the MV recovery of the next MBs as well. Since the effects of concealing one MB propagate to the next MBs, the problem is formulated as a Dynamic Programming (DP) problem. Considering the possible MVs for the neighboring MB, boundary mismatch associated with each candidate MV is calculated and accumulated for all lost MBs in one row. Minimizing this cost function gives a set of MVs such that the boundary mismatch error over MBs in one row is minimized accordingly. A temporal error concealment exploiting Principle Component Analysis (PCA) is presented in [8]. With a trained PCA, the lost parts are error concealed with the information of previous frames. Due to the scene change, the previous frames may not be always suitable; therefore, the authors propose to update the buffered frames in the case of scene changes.

Recently, there are some works propose and show how to use Deep Neural Networks (DNNs) for error concealment. In [25], with a combination of convolutional Long Short Term Memory (LSTM) layers and simple convolutional layers, a network for predicting the optical flow in the lost area is presented. In [32], a Generative Adversarial Network (GAN) is trained for video error concealment. The main issue in DNNs is a relatively large needed database and the complexity of the training phase.

As reviewed above, all error concealment techniques use rectangular blocks with predefined sizes as error concealment unit. However, the concealment unit in our proposed

method is not limited to square shaped blocks; we select it adaptively from parallelogram blocks with various sizes and angles. For this purpose, we need to have the up and down boundaries of the partitions and to use an algorithm for partitions' MV recovery; these two steps are carried out exploiting the algorithms presented in [15, 22], respectively. We propose to find the best angle and size of the parallelograms with two levels of filtering, a boundary smoothness measure and a blind IQA method.

3 Proposed method

We are dealing with a relatively large loss area; therefore, we decided to apply two rounds of concealment. The first round is for preliminary filling of the lost area. We then apply a parallelogram partitioning; afterwards, the best MV for each parallelogram partition is obtained with the approach of DP. With these MVs, the second round of the concealment is performed; but this step is repeated for several sizes and angles of parallelograms. The final step is to find which partitioning will provide the best concealment quality. These steps are explained in the following subsections.

3.1 Using AECOD for the preliminary filling of the lost area

The lost area in HEVC is relatively large; therefore, similar to the algorithm presented in [15], we partitioned it into several rows of blocks. In order to have the up and down boundaries of these rows, the lost area is filled as the first round of the error concealment. Based on our experiments, we found that using AECOD [22] as the first version concealment gives the best results. As already explained in Section 2, the lost region is portioned into several MBs in AECOD, and then the order of MBs for error concealment is determined based on the neighboring MBs textures. In that method, the matching of external boundaries is used as the measure to find the best MVs.

3.2 MV recovery with the DP approach for parallelogram partitioned lost area

From the early era of image and video coding standards, the coding unit was in square shape. The main reason is the rectangular form of the images and display screens. There exist some works discussing mesh-based partitioning for motion estimation [10], but these schemes were not included in any standard. However, for error concealment, there is no standard limitation. Here, the effort is to replace the not-received data with the some data from reference frame. The criterion for finding the best MV is BMA. Clearly, the concealment unit and its boundary form can significantly affect the recovered MVs. We propose to use parallelogram partitioning of the lost area, with various vertex angles and sizes; and these parallelograms are used as the concealment unit. The partitioning procedure is shown in Fig. 1; one can see there that the square is a special form of the parallelograms.

In our experiments, θ has values of 30° , 60° , 90° , 120° and 150° , W and H could be 64, 32, and 16 pixels.

How to assign the MVs to these parallelograms is explained in the remaining part of this subsection. In most of the concealment methods, MV recovery has been done for MBs, individually and by using temporal/spatial BMA. BMA method is usually in conjunction with Flexible Macroblock Ordering (FMO). But in the case of not using FMO (in H.264/AVC) or

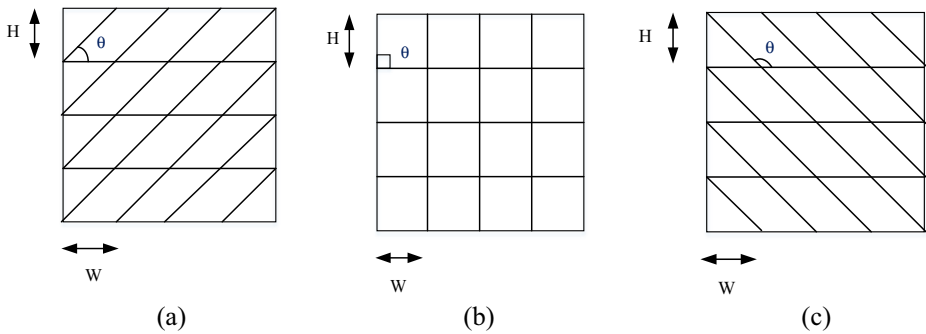


Fig. 1 Partitioning using parallelograms with different angles: **a** $\theta = 45^\circ$, **b** $\theta = 90^\circ$, **c** $\theta = 135^\circ$

HEVC losses (FMO is not supported in HEVC), the losses include a series of connected 16×16 blocks. As already mentioned, the concealment of one block affects the concealment of the next blocks. Therefore, it is reasonable to take into account the concealment of all connected blocks, concurrently.

As already mentioned, one solution to this problem is to use DP, which is an efficient method for the problems that can be split into several sub-problems [4]. In other words, the goal of DP is to find the best path leading to the least cost, where the original problem is divided into sub-problems, and every sub-problem is solved considering its effects on the other sub-problems. In our application, the concealment of the first and second block in the lost area are directly related to each other, but this is not the case for example for the first and the fourth block. Therefore, the problem of concealing the whole lost area is divided into many sub-problems where their goal is minimizing the cost function of concealment of the consecutive blocks.

DP formulation for parallelograms is as follows. Let $X_{(M,N)}(i,j)$ denotes the pixel data of the (M,N) th parallelogram block, and $X_{(M,N)}^v(i,j)$ denotes the prediction of $X_{(M,N)}(i,j)$ from the reference frame associated with the candidate motion vector v (see Fig. 2a). If we move the cartesian coordinate axes on the top-left sides of the parallelograms, the spatial matching of the boundaries for a candidate $v = v_1$ is defined as:

$$D_{M,N}^{bd}(v_1) = D_{M,N}^{Up}(v_1) + D_{M,N}^{Down}(v_1) \quad (1)$$

$$D_{(M,N)}^{Up}(v_1) = \sum_{j=0}^{W-1} \left| X_{(M-1,N)}(H-1,j) - X_{(M-1,N)}^{v_1}(H-1,j) \right| \quad (2)$$

$$D_{(M,N)}^{Down}(v_1) = \sum_{j=0}^{W-1} \left| X_{(M+1,N)}(0,j) - X_{(M+1,N)}^{v_1}(0,j) \right| \quad (3)$$

where $D_{(M,N)}^{Down}(v_1)$ and $D_{(M,N)}^{Up}(v_1)$ are the SAD between external boundaries of the block and its down and up neighboring blocks' boundaries, respectively. Parameters W and H are width and height of the parallelograms as shown in Fig. 1. With the preliminary filling of the lost area described in the previous subsection, the concealed up and down boundaries of the parallelograms are now available and used in the above formulations. Note that Eqs. (2) and (3) compute the matching in outer strips of the parallelograms.

For the right boundary, we should know the boundary of the right block; therefore, it is temporarily filled with motion vector v_2 as a candidate. Then by considering a pair of MV candidates (v_1, v_2) , the SAD of inner/outer motion compensated boundary strips (see Fig. 2b), with cartesian coordinate axes on the top-left sides of the parallelograms, is computed as follows:

$$D_{(M,N)(M,N+1)}^{smh}(v_1, v_2) = D_{(M,N)(M,N+1)}^{smhe}(v_1, v_2) + D_{(M,N)(M,N+1)}^{smho}(v_1, v_2) \tag{4}$$

$$D_{(M,N)(M,N+1)}^{smhe}(v_1, v_2) = \sum_{i=0}^H |\hat{X}_{(M,N+1)}^{v_2}(i, 0) - \hat{X}_{(M,N)}^{v_1}(i, H-1)| \tag{5}$$

$$D_{(M,N)(M,N+1)}^{smho}(v_1, v_2) = \sum_{i=0}^H |\hat{X}_{(M,N)}^{v_2}(i, H-1) - \hat{X}_{(M,N+1)}^{v_1}(i, 0)| \tag{6}$$

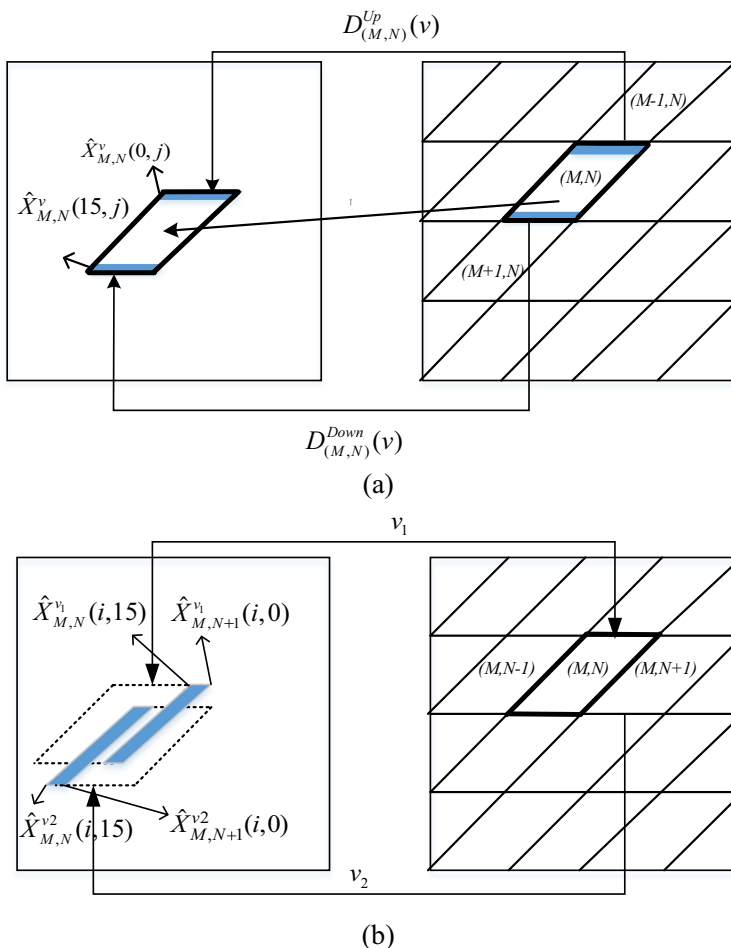


Fig. 2 Boundary matching of a up/down strips and b left/right strips for H = 16, W = 16

where $D_{(M,N)(M,N+1)}^{smhe}(v_1, v_2)$ and $D_{(M,N)(M,N+1)}^{smho}(v_1, v_2)$ measure the matching of the inner/outer boundary strips by using (v_1, v_2) and (v_2, v_1) for the $(M, N)^{\text{th}}$ and $(M, N + 1)^{\text{th}}$ blocks, respectively. Therefore, the final cost function is as given by (7):

$$D_{total}(\mathbf{v}_M) = D_{(M,0)}^{bd}(v_{M,0}) + \dots + D_{(M,N-1)}^{bd}(v_{M,N-1}) + D_{(M,0)(M,1)}^{smh}(v_{M,0}, v_{M,1}) + D_{(M,1)(M,2)}^{smh}(v_{M,1}, v_{M,2}) + \dots + D_{(M,N-2)(M,N-1)}^{smh}(v_{M,N-2}, v_{M,N-1}) \quad (7)$$

where \mathbf{v}_M is the set of MVs for all blocks in one row; i.e., $\mathbf{v}_M = (v_{M,0}, v_{M,1}, \dots, v_{M,N-1})$. The vector \mathbf{v}_M which minimizes the above cost function is used for concealment of one row of parallelogram blocks. This procedure is continued for the next rows of parallelograms as long as the lost area is completely covered. The whole process is repeated for several angles (30, 60, 90, 120 and 150 degrees) and various sizes (64, 32, and 16 pixels); where in total we have 45 combinations of angles and sizes. Which one is finally selected as the best partitioning for getting the best concealment output is discussed in the next subsection.

3.3 Selecting the best angle and size for parallelograms

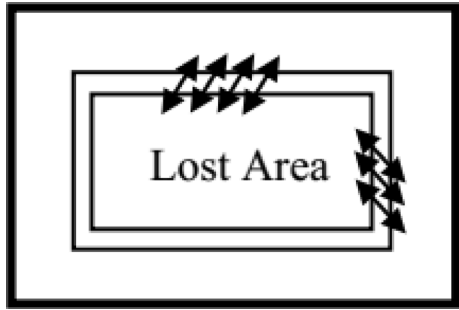
As already mentioned, a fixed size and shape block is used as processing unit in video encoding standards; where with the assist of smaller blocks for inter/intra prediction, the encoder can adaptively handle the various video contents. But the shapes are still rectangles. However, we examined a more general parallelogram shaped blocks as the concealment unit. Clearly, the best angle and size of the parallelograms are content dependent. In the previous step, the various angles and sizes are examined, but how does one select the best one? The problem is here that we have no reference to know which partitioning leads to the best concealment quality.

With a similar approach presented in [26], one solution is to use directional matching (as shown in Fig. 3) between the external and internal boundaries of the concealed lost area. First, we calculate the dominant edge direction in the surrounding pixels of the loss-free area. The Canny edge detector is applied and the dominant edge is found in each 2×4 window of the pixels. We denote this criterion as Border Mismatch Along Dominant Edge Direction (BMADED). From the “Border” here, we mean the rectangle around the lost area as shown in Fig. 3.

With MV recovery for the parallelograms with different angles and sizes, 45 concealed versions of the lossy frame have been obtained which are now checked for minimum BMADED. The smaller BMADED could be a criterion for predicting the best quality; however, the results show that this criterion is not usually enough. Therefore, we apply a next level of filtering as explained below.

After choosing the top 25% versions of the concealed frame with the smallest BMADED, we propose to use a no-reference IQA method for final selection. IQA methods are able to measure the quality of the images without any reference. Different approaches have been developed for no-reference IQA (the source codes of several IQAs are publicly available in [3]). We evaluated different metrics such as BLIINDS-II [24], Blind Image Quality Index (BIQI) [20], and Spatial-Spectral Entropy-based Quality (SSEQ) [18]; and we found that SSEQ is more appropriate for this work according to the visual qualities of the selected outputs. SSEQ is a general-purpose method for assessment of video frames’ quality, which supports different distortion types such as noise, compression, blurriness, and fading. SSEQ measures spatial and spectral entropies which are favorably sensitive to concealment distortion.

Fig. 3 Directional matching in BMADED



SSEQ has the following four steps. In the first step, the distorted image is decomposed into three scales: low, middle and high which enable us for multi-scale analysis. As the next step, the images in each scale are divided into 8×8 size blocks and then the spatial and spectral entropy of each block is computed. These two features will be sorted in an ascending order during step three and the mean and the skew of 60% central values of the entropies are selected as the image features. This selection is done for both spatial and spectral domains and for each scale. With these features, a framework is trained as step four of the algorithm; it then can be used for predicting the quality of images. As already mentioned, the sensitivity of the SSEQ algorithm to the spatial and spectral entropies makes it relatively successful in our application, since the misalignments caused by non-ideal concealment can be detected by spatial and/or spectral entropies.

3.4 Complexity issues

We should mention that the proposed algorithm is rather complex; firstly, due to DP used for motion recovery and; secondly, due to the checking numerous sizes and angles of parallelograms. Algorithms of AECOD used as the preliminary filling phase, and BMADED and SSEQ are not so complex compared to the DP. However, the rather high complexity of the proposed method does not limit its application even for battery powered devices, due to the following reasons:

- (a) Error concealment is needed occasionally

Note that most of the times, the packets are error-free and it is not needed to run error concealment. For example, for a loss rate of about 5%, the redundant complexity of the concealment algorithm is added to the decoding process only about 5% of the times.

- (b) The error concealment complexity is less than motion estimation (ME) complexity at the encoder

BMA is checked for each candidate MV in the search window. If the number of candidate MVs is k , then the number of SAD computations becomes $O(k^2)$ in DP; since BMA must be computed for all possible combinations of MVs of the two adjacent blocks. However, the number of candidate MVs for ME is $O(k)$, but the encoder does Block Matching (BM) for ME which needs more SAD computations compared to BMA. The reason is that BMA is applied

to the pixels on the block perimeter, while BM is applied to the pixels on the block area. With the assumption of $W \times H$ pixel blocks, N_{SAD} (number of SAD computations) for BMA and BM is:

$$\begin{aligned} N_{W \times H}^{BMA} &= k^2 (2(W + H)) \left(\frac{LostArea}{W \times H} \right) \\ N_{W \times H}^{BM} &= k(W \times H) \left(\frac{FrameSize}{W \times H} \right) = k (FrameSize) \end{aligned} \quad (8)$$

Therefore, for 5 values of θ and 3 values of W and H in our proposed algorithm, the number of SAD computations required for error concealment, N_{con}^{BMA} , becomes:

$$\begin{aligned} N_{con}^{BMA} &= \sum_{\theta=30, 60, 90, 120, 150} \sum_{W=16, 32, 64} \sum_{H=16, 32, 64} N_{W \times H}^{BMA} = 5 k^2 (LostArea) \sum_{W=16, 32, 64} \sum_{H=16, 32, 64} \left(\frac{2(W + H)}{W \times H} \right) \\ &= k^2 (LostArea)(6.5625) \end{aligned} \quad (9)$$

As equation (8) shows, $N_{W \times H}^{BM}$ does not depend on the partition size; therefore we can simply obtain the number of SAD computation required for ME of a frame, N_{ME}^{BM} , as:

$$N_{ME}^{BM} = \sum_{Block_{W \times H}} N_{W \times H}^{BM} = 13 N_{W \times H}^{BM} = 13k (FrameSize) \quad (10)$$

where $Block_{W \times H}$ in (10) is the allowed block sizes for inter prediction units. By excluding asymmetric motion partitioning (explained in [29]), with CTU size of 64×64 pixels and depth of 3, $Block_{W \times H}$ consists of 4×4 , 4×8 , 8×4 , 8×8 , 16×8 , ..., 64×64 modes, which becomes 13 different sizes.

Finally, we can write that:

$$\frac{N_{con}^{BMA}}{N_{ME}^{BM}} = \frac{k^2 (LostArea)(6.5625)}{13k (FrameSize)} \cong \frac{k (LostArea)}{2 (FrameSize)} \quad (11)$$

It means that for a sufficiently small lost area or a limited number of MV candidates, the complexity of the proposed concealment algorithm is less than that of ME of a frame. That is, if a consumer device is able to encode an HEVC stream with the acceptable quality, it has enough processing power to run our method for error concealment. Note that the encoder has even more complexity; e.g. for intra prediction or mode decision, but we found a minimum for that is usually done. The aim of the above discussion is to show that, even though our method is complex compared to the other methods, its complexity is not beyond the processing power of the video communication devices. If this device is used for receiving a live video from a lossy network, the device is also capable of executing the proposed algorithm without a problem. Furthermore, if the complexity of the algorithm is handled, it has no latency since all information needed for the error concealment is available when the frame packets are received.

It should be noted that the parallelogram partitioning is more challenging for the hardware compared to the rectangular blocks. One possible solution is to save the frame pixels in a tilt order at a temporary memory, such that the processor fetches the pixels just as the rectangular blocks. The memory does not need to save the whole frame; just the region under processing is saved. Therefore, with an additional memory and a simple algorithm to apply the tilt ordering of the pixels, the issue of parallelograms can be solved in the hardware.

4 Performance comparison

To evaluate our proposed method against the other methods, we implemented our algorithm as well as the competitors in HEVC reference software (HM-16.15) [1]. Our proposed method and five other algorithms were implemented and tested on *Ice*, *Soccer* and *Football* CIF sequences, and *Speed-bag*, *Pedestrian-area*, and *Tractor* full-HD sequences. The Group of Picture (GOP) structure is IPPPPP. The Quantization Parameter (QP) value in this paper is set to 20, 22, and 27. With Packet Loss Rate (PLRs) of 4% and 8%, some CTUs of the frames are intentionally lost and concealed. The lost CTUs are spread around the frames covering various contents of the videos.

First of all, we show the performance of BMADED+SSEQ for the best quality concealed frame selection. Our experiments show that the best visual/subjective quality concealed frame may not be the same as selected with BMADED+SSEQ, but they are so close. In Fig. 4, the concealed frames selected by BMADED+SSEQ and those selected subjectively are compared for some instances. This figure verifies the performance of our selection strategy.

We now compare our proposed method with algorithms proposed in [7, 12, 14, 15], and AECOD [22]. The concealment unit is 4×4 in [14] and 16×16 pixel square shaped blocks in the other methods. Figures 5, 6, 7, 8, 9, 10, 11 and 12 show STRRED [27] curves for frames of the CIF and full-HD test sequences. Note that the smaller STRRED means higher quality. This figure shows that our proposed method gives smaller STRRED in most of the frames of full-HD sequences, but it is not always true for CIF sequences. For full-HD sequences, there are a few cases for which the proposed method is not the best. It is common that a concealment

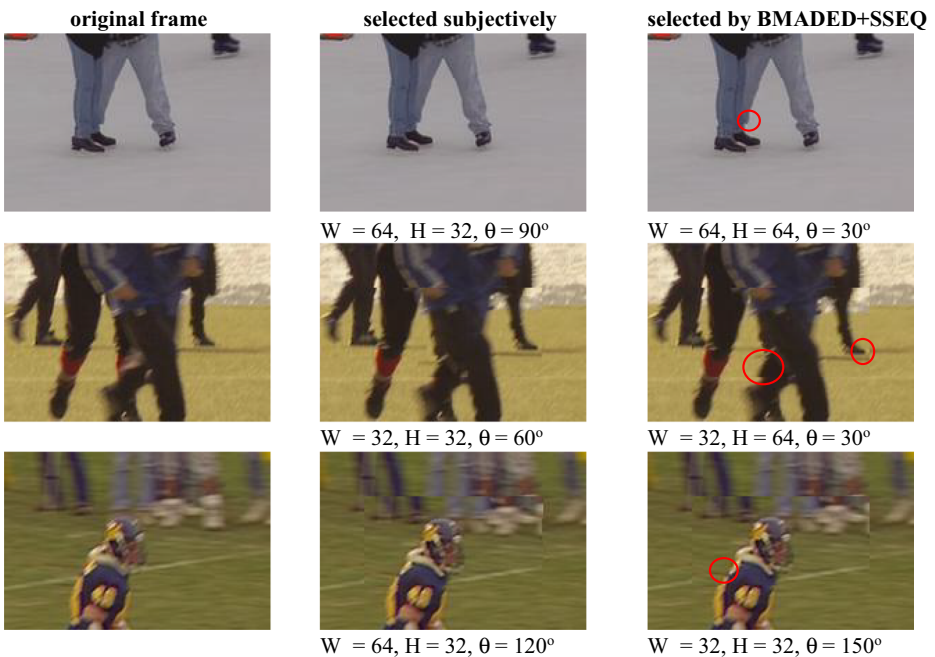


Fig. 4 Some instances for comparison of the selected concealed frame among 45 outputs, left column: original frame, middle column: selected subjectively, and right column: selected by BMADED+SSEQ. The parameters θ , W and H are the selected angle, width, and height, respectively

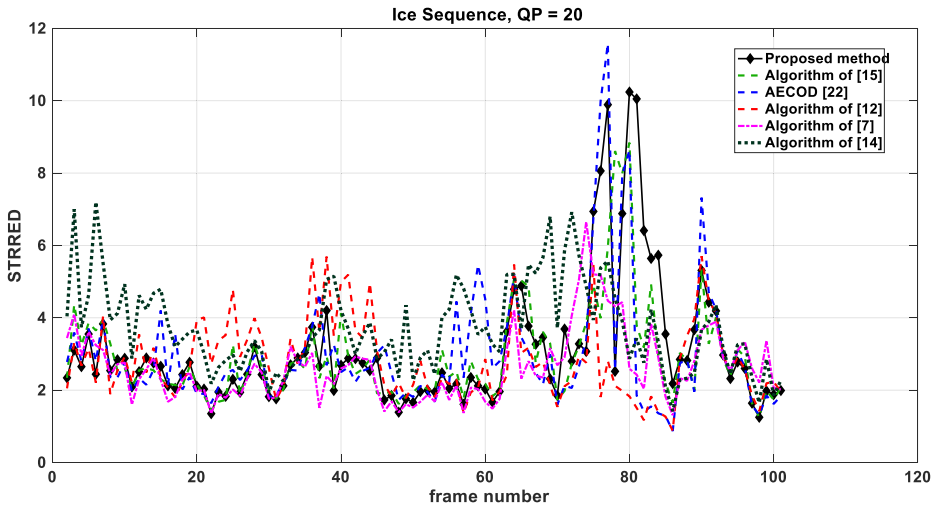


Fig. 5 STRRED curves for the proposed method in comparison to the other methods, for *Ice* CIF sequence, QP = 20, PLR = 8%

technique might not be optimum for all cases. For getting an overall good performance, several concealment techniques can be examined at the encoder and the best one is signaled to the receiver side [6]. The overhead of such information signaling is negligible (maybe several bytes); therefore, it can be easily transmitted many times so that they are correctly being delivered, even over the channels with very high loss rates.

Table 1 provides the average STRRED gains of the proposed method over the other EC methods, for CIF and full-HD sequences. PLR is 8% and QP is 20 and 22 for CIF and full-HD sequences, respectively. Note that the larger negative values in Table 1 mean higher visual quality. It can be seen that for full-HD sequences, similar to the curves in Figs. 8, 9, 10, 11 and 12, the numbers in Table 1 confirm the superiority of the proposed method. However, it can be

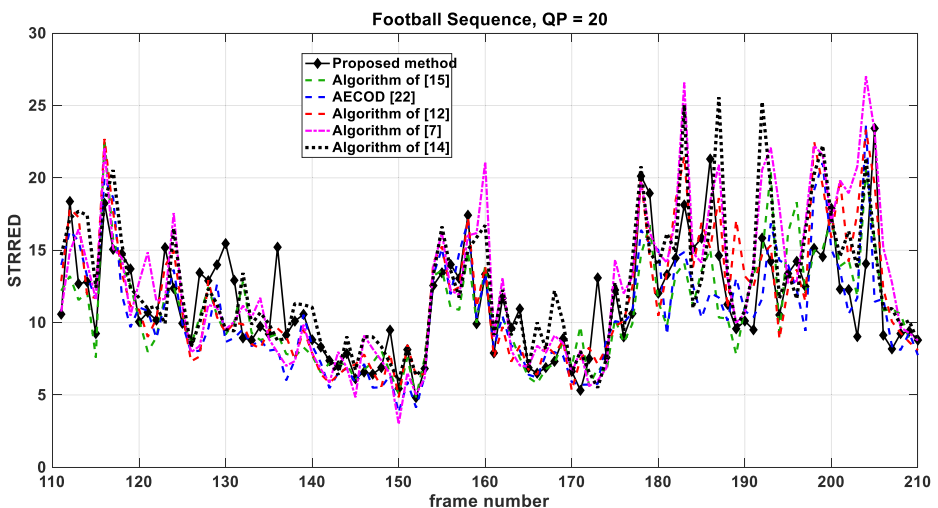


Fig. 6 STRRED curves for the proposed method in comparison to the other methods, for *Football* CIF sequence, QP = 20, PLR = 8%

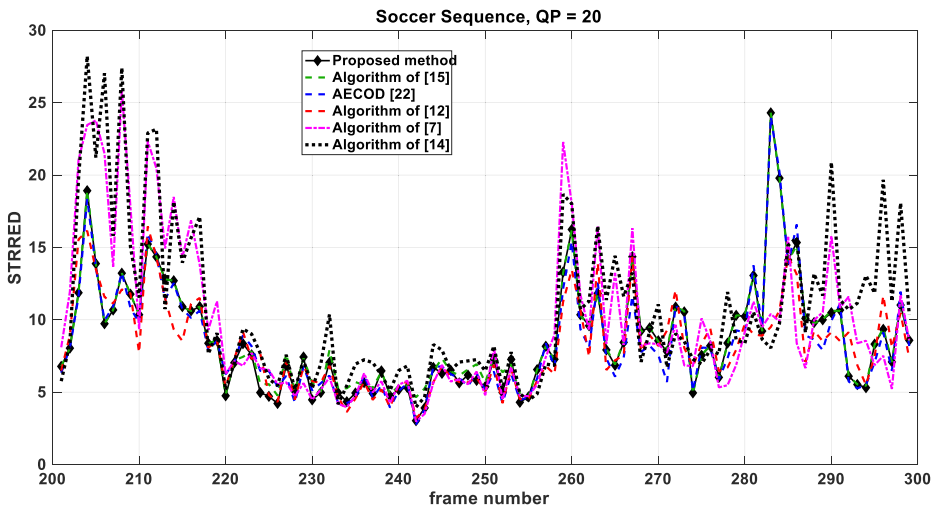


Fig. 7 STRRED curves for the proposed method in comparison to the other methods, for Soccer CIF sequence, QP = 20, PLR = 8%

seen that our proposed method does not outperform the others for CIF sequences in STRRED metric.

We also provided the average of PSNR gains achieved by the proposed method in Table 2. The positive and larger values mean higher quality here. The numbers associated with full-HD sequences are again in favor of the proposed method. However, there exist some conflicts between PSNR and STRRED for CIF sequences, especially for Football and Ice. PSNRs are very good for the proposed method in Table 2, but it is not the case for STRRED scores in Table 1. Our subjective tests also verify this in general for these two sequences, i.e. the visual qualities of our method outputs are better than the others. It should be also noted that, in some cases, sharp changes in the scores affect the averaging. From Fig. 5, one can see that for Ice

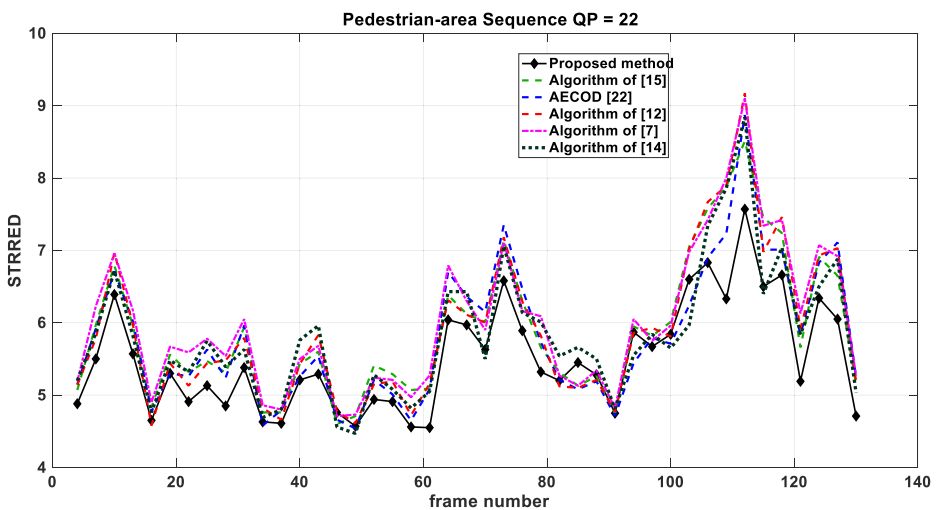


Fig. 8 STRRED curves for the proposed method in comparison to the other methods, for Pedestrian-area full-HD sequence, QP = 22, PLR = 8%

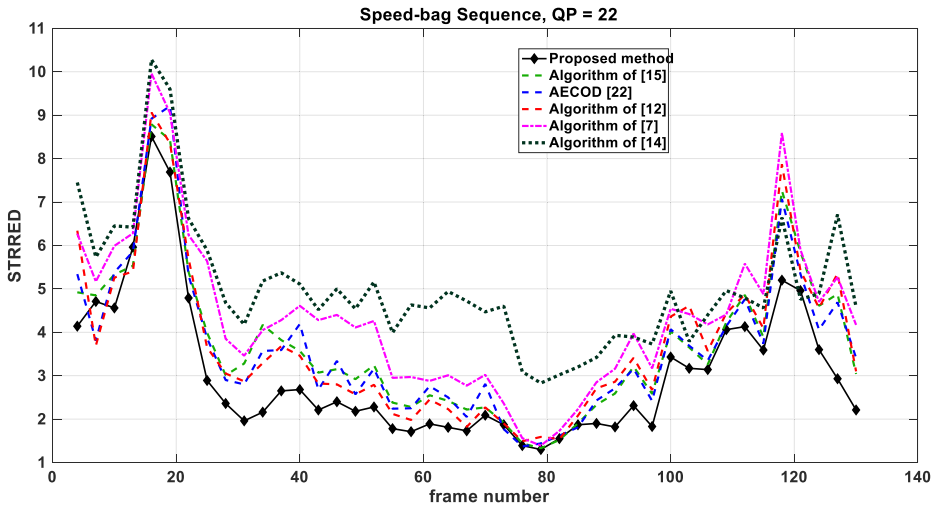


Fig. 9 STRRED curves for the proposed method in comparison to the other methods, for *Speed-bag* full-HD sequence, QP = 22, PLR = 8%

sequence, the algorithm of [7] is the best, while the STRRED curve of the proposed method is most of times close to this algorithm. However, due to the very high STRREDs for the frames around 80, the averages in Table 1 are mostly positive showing that all other methods are better than ours, the judgment that it is not true as Fig. 5 shows. Furthermore, we also checked these high STRRED (low quality) frames, their visual qualities were not such disappointing. It is due to the sometimes wrong sensitivity of the full reference (e.g. PSNR) and reduced reference (e.g. STRRED) metrics to the concealment distortion. Concealment distortion is a type of pixel misalignment and is much different from the other distortion types such as quantization, blurriness or environment noise. The misfunctionality of the quality assessment metrics has been always a challenge. However, even though STRRED and PSNR are not ideal, but are

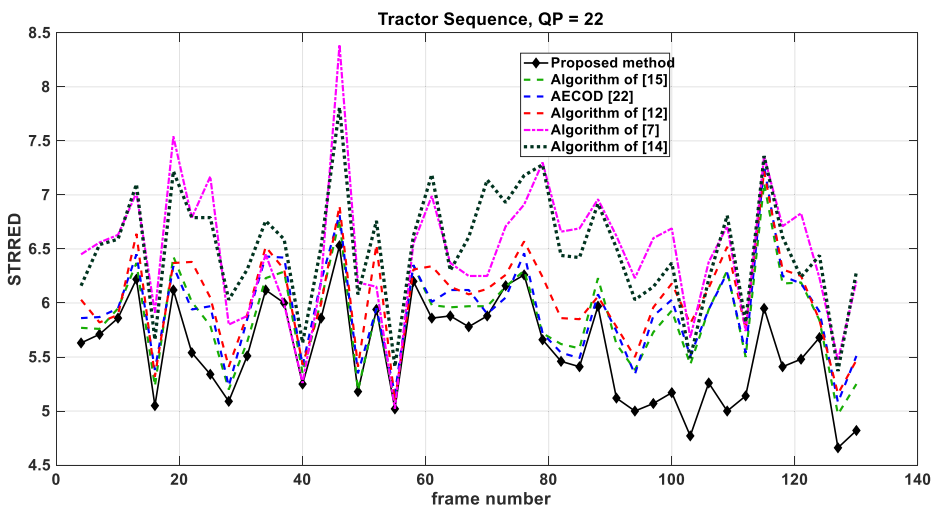


Fig. 10 STRRED curves for the proposed method in comparison to the other methods, for *Tractor* full-HD sequence, QP = 22, PLR = 4%

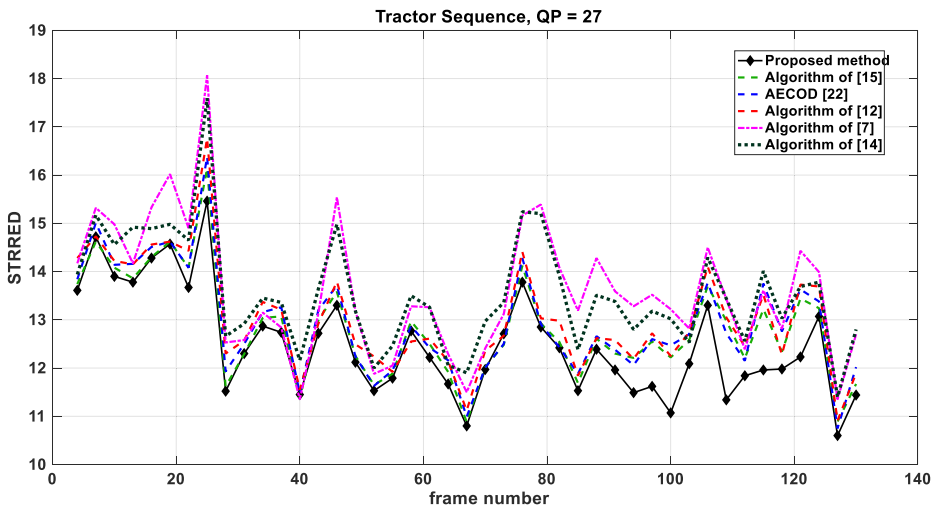


Fig. 11 STRED curves for the proposed method in comparison to the other methods, for *Tractor* full-HD sequence, QP = 27, PLR = 4%

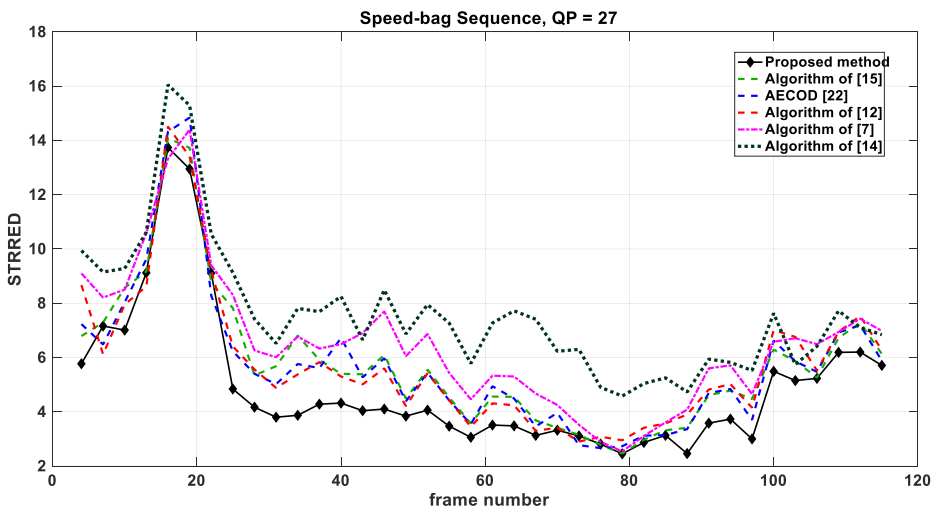


Fig. 12 STRED curves for the proposed method in comparison to the other methods, for *Speed-bag* full-HD sequence, QP = 27, PLR = 8%

Table 1 Average STRED of our proposed method over the other methods, the more negative means the higher quality

Sequences	Algorithm of [15]	AECOD [22]	Algorithm of [12]	Algorithm of [7]	Algorithm of [14]
Ice CIF	0.1478	0.0838	0.2683	0.4514	-0.6519
Soccer CIF	-0.1516	0.2789	0.5284	-0.9097	-1.9015
Football CIF	0.7081	0.8330	-0.1616	-1.0206	-1.0031
Speed-bag full-HD	-0.6640	-0.6221	-0.6684	-1.3293	-1.9728
Tractor full-HD	-0.3074	-0.3693	-0.4819	-0.9177	-0.9412
Pedestrian-area full-HD	-0.3719	-0.2926	-0.3716	-0.4770	-0.3107

Table 2 Average PSNR gain (dB) of our proposed method over the other methods

Sequences	Algorithm of [15]	AECOD [22]	Algorithm of [12]	Algorithm of [7]	Algorithm of [14]
Ice CIF	0.9828	0.9715	1.0785	-1.1355	4.5144
Soccer CIF	0.1494	-0.5232	-0.7049	-0.1778	1.6046
Football CIF	0.3404	0.3287	0.8919	1.1142	1.7580
Speed-bag full-HD	1.1174	1.0451	1.1395	2.0319	3.8163
Tractor full-HD	0.8779	1.0616	1.4895	1.5540	2.7314
Pedestrian-area full-HD	0.6228	0.6519	0.7619	0.9174	1.1265

sufficient for objectively evaluation of the proposed method. In the next part of this section, some subjective tests are also provided.

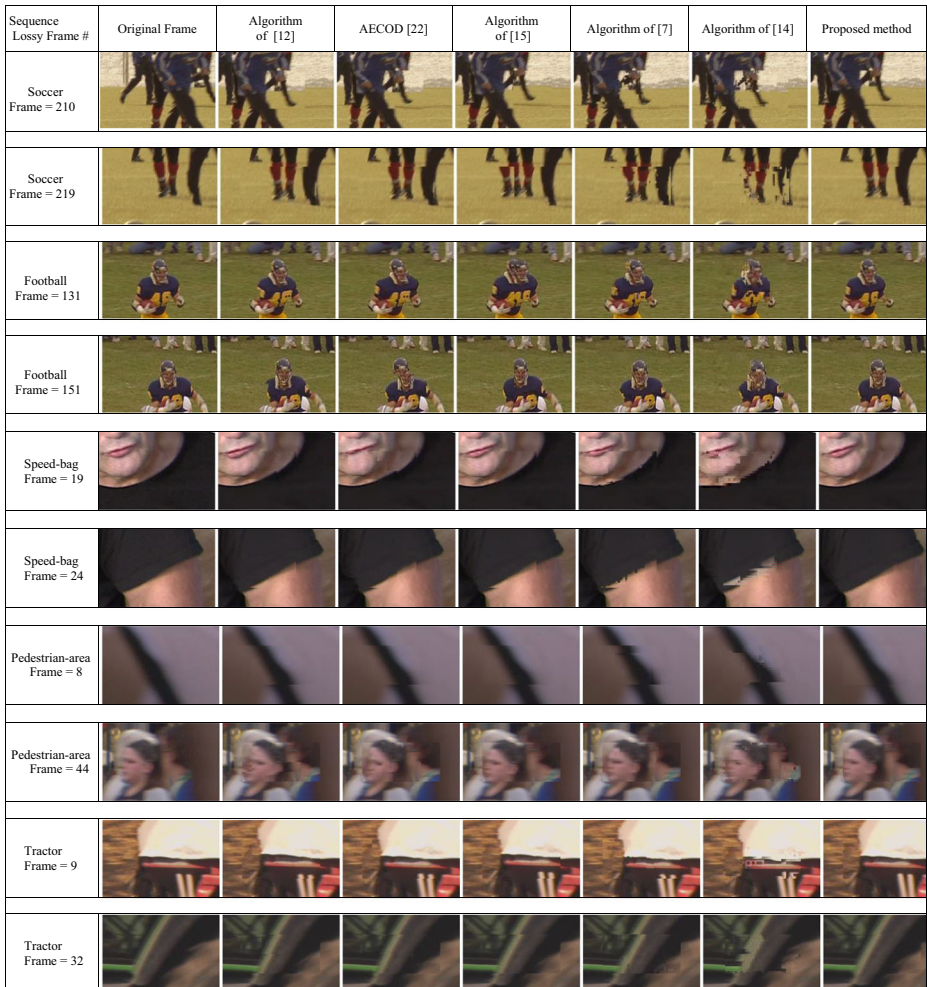


Fig. 13 Some sample of the subjective comparison of the error concealed outputs

Sequence Lossy Frame #	Original Frame	Algorithm of [12]	AECOD [22]	Algorithm of [15]	Algorithm of [7]	Algorithm of [14]	Proposed method
Soccer Frame = 21							
Soccer Frame = 20							
Football Frame = 132							
Football Frame = 152							
Speed-bag Frame = 20							
Speed-bag Frame = 25							
Pedestrian-area Frame = 9							
Pedestrian-area Frame = 45							
Tractor Frame = 10							
Tractor Frame = 33							

Fig. 14 Error propagation in various methods, the next frame after the lossy frame

Figure 13 shows the visual quality after applying different concealment algorithms for some frames of the test sequences. Images in columns from left to right are the original frames, the outputs of Algorithm [12], AECOD [22], and Algorithms of [15, 7, 14] and our proposed method, respectively. This figure shows that our proposed method is more successful in recovering the various parts of the objects.

The actual scenario for evaluating the concealment performance is the quality of the video when it is displaying. That is the effect of concealment distortion propagated to the next frames is also or even more important. Therefore, the visual quality of the next frame after the lossy frame is provided in Fig. 14. It is worthy to mention that this frame part is error free, and the distortion in this part is due to the non-ideal concealment in the lossy frame which has propagated to this frame due to inter prediction. This figure also verifies the performance of the proposed method against the other methods. However, for frame 20 of *Speed-bag* sequence, even though the edges are correctly preserved in the proposed method, but some blockiness artifacts are observable. They are caused by the residual signals added to an

erroneous reference (error concealed frame); therefore, this part of the picture is reconstructed erroneously and maybe with blockiness-like artifacts. The relatively homogeneous regions which have not exactly uniform luminance if are error concealed based on edge matching approach, such as algorithm of [15] and our method, have higher potential for these artifacts. However, this phenomenon occurs in a few cases as we checked.

5 Conclusion

The innovations in this work are as follows: (a) our paper proposes using various block shapes as the concealment unit, where it has been limited to rectangular blocks in the existing error concealment works. Partitions with various heights, widths, and angles are examined as the concealment units. Each combination of sizes/angles leads to a version of error concealed frame. The second innovation in this paper is to use a no-reference IQA method for selecting the best one among these versions.

The proposed method consists of three steps. First, AECOD algorithm is applied as a preliminary EC. Second, the corrupted region is split into parallelogram partitions and then, the DP optimization is applied to find the optimal MV associated with each partition, this MV is used for error concealment of this partition. This step is carried out for several values of sizes and angles of parallelograms. For choosing the best partitioning, the top 25% outputs with the smallest BMADED are selected as the first level of filtering; then an IQA method (SSEQ) is applied for final selection. Simulation results show that with this strategy, most of the times, the best visual quality concealed frame is selected; therefore, the performance of BMADED + SSEQ is verified.

In comparison with the other works used for error concealing the videos, our method provides significant improvement for some frames, subjectively and objectively, especially for full-HD videos. It is also true with taking into account the error propagation. For some frames, our method will not offer any improvement or even degrades the quality; however, the average gains in the objective metrics are in favor of the proposed algorithm.

The complexity of the proposed algorithm is high as mentioned before. However, as discussed in Section 3.4, the number of computations is less than ME computations performed when encoding. Also, it is worthy to mention that the complexity of our method is mostly due to the relatively high computations inherently needed in the DP approach.

References

1. [Online]. Available: <https://hevc.hhi.fraunhofer.de/trac/hevc/browser>
2. Akbari A, Trocan M, Granado B (2017) Sparse recovery-based error concealment. *IEEE Trans Multimed* 19(6):1339–1350
3. Available: <http://live.ece.utexas.edu/research/Quality/noreference.htm>.
4. Bellman R (1954) The theory of dynamic programming. *Bull Am Math Soc*
5. Byongsu H, Jonghyon J, Cholsu R (2017) An improved multi-directional interpolation for spatial error concealment. *Multimed Tools Appl* 76(1):397–413
6. Carreira JFM, Assunção PA, de-Faria SM, Ekmekcioglu E, Kondoza A (2018) Error concealment-aware encoding for robust video transmission. *IEEE Trans Broadcast* 65(2):282–293
7. Chang YL, Reznick YA, Chen Z, Cosman PC (2013) Motion compensated error concealment for HEVC based on block-merging and residual energy. In: *International packet video workshop (PV)*

8. Choe G, Nam C, Chu C (2018) An effective temporal error concealment in H.264 video sequences based on scene change detection-PCA model. *Springer Multimed Tools Appl* 77(24):31953–31967
9. Chung B, Yim C (2019) Bi-sequential video error concealment method using adaptive homography-based registration. *IEEE Trans Circ Syst Video Technol*. <https://doi.org/10.1109/TCSVT.2019.2909564>
10. Dufaux F, Moscheni F (1995) Motion estimation techniques for digital TV: a review and a new contribution. *Proc IEEE* 83:858–876
11. Huang Z, Cai Q (2018) A video data recovery algorithm in wireless communication networks. In: *IEEE 18th international conference on Communication Technology*
12. Hwang MC, Kim JH, Duong DT, Ko SJ (2008) Hybrid temporal error concealment methods for block-based compressed video transmission. *IEEE Trans Broadcast* 54(2):198–207
13. Lam WM, Reibman AR, Liu B (1993) Recovery of lost or erroneously received motion vectors. In: *Proceedings of IEEE international conference on Acoust. Speech Signal Process*, pp 417–420
14. Li Y, Chen R (2017) Motion vector recovery for video error concealment based on the plane fitting. *Springer Multimed Tools Appl* 76(13):14993–15006
15. Lie WN, Lee CM, Yeh CH, Gao ZW (2014) Motion vector recovery for video error concealment by using iterative dynamic-programming optimization. *IEEE Trans Mulimed* 16(1):216–227
16. Lin TL, Ding TL, Yang NC, Wu PY, Tung KH, Lai CK, Chang TE (2016) Video motion vector recovery method using decoding partition information. *J Disp Technol* 12(11):1451–1463
17. Lin TL, Ding TL, Fan CY, Chen WC (2017) Error concealment algorithm based on sparse optimization. *Multimed Tools Appl* 76(1):397–413
18. Liu L, Liu B, Huang H, Bovik AC (2014) No-reference image quality assessment based on spatial and spectral entropies. *Signal Process Image Commun* 29(8):856–863
19. Liu X, Zhai D, Zhou J, Wang S, Zhao D, Guo H (2016) Sparsitybased image error concealment via adaptive dual dictionary learning and regularization. *IEEE Trans Image Process* 26(2):782–796
20. Moorthy AK, Bovik AC (2010) A two-step framework for constructing blind image quality indices. *IEEE Signal Process Lett* 17(5):513–516
21. Oztas B, Pourazad MT, Nasiopoulos P, and Leung VCM (2012) A study on the HEVC performance over Lossy Networks. In: *Proceedings of 19th IEEE international conference on Electron. Circuits Syst. (ICECS)*, pp 785–788
22. Qian X, Liu G, Wang H (2009) Recovering connected error region based on adaptive error concealment order determination. *IEEE Trans Multimed* 11(4):683–695
23. Radmehr A, Ghasemi A (2016) Error concealment via particle filter by Gaussian mixture modeling of motion vectors for H.264/AVC. *Springer Sig Image Video Process* 10(2):311–318
24. Saad MA, Bovik AC (2012) Blind image quality assessment: a natural scene statistics approach in the DCT domain. *IEEE Trans Image Process* 28(39):3339–3352
25. Sankisa A, Punjabi A, Katsaggelos AK (2018) Video error concealment using deep neural networks. In: *Proceedings of the IEEE international conference on Image Processing, Athens, Greece*, pp 380–384
26. Shih HC, Wang CT, Huang CL (2018) Spiral-like pixel reconstruction algorithm for spatiotemporal video error concealment. *IEEE Access* 16:6370–6381
27. Soundararajan R, Bovik AC (2013) Video quality assessment by reduced reference Spatio-temporal entropic differencing. *IEEE Trans Circ Syst Video Technol* 23(4):684–694
28. Suh JW, Ho YS (2002) Error concealment techniques for digital TV. *IEEE Trans Broadcast* 48(4):299–306
29. Sullivan GJ, Ohm JR, Han WJ, Wiegand T (2012) Overview of the High Efficiency Video Coding (HEVC) standard. *IEEE Trans Circ Syst Video Technol* 22(12):1648–1667
30. Usman M., He X, Xu M and Lam KM (2015) Survey of error concealment techniques: research directions and open issues. In: *Picture coding symposium (PCS), Cairns, QLD, Australia*
31. Wang Y, Zhu QF (1998) Error control and concealment for video communication: a review. *Proc IEEE* 86(5):974–999
32. Xiang C, Xu J, Yan C, Peng Q, (2019) Generative adversarial networks based error concealment for low resolution video. In: *IEEE international conference on Acoustics, Speech And Signal Processing (ICASSP)*
33. Yang SH, Chang CW, Chan CC (2015) An object-based error concealment technique for H.264 coded video. *Multimed Tools Appl* 74(23):10785–10800
34. Zhang J, Arnold JF, Frater MR (2000) A cell-loss concealment technique for MPEG-2 coded video. *IEEE Trans Circ Syst Video Technol* 10(4):659–665
35. Zhang Y, Xiang X, Zhao D, Ma S, Gao W (2011) Packet video error concealment with auto regressive model. *IEEE Trans Circ Syst Video Technol* 22(1):12–27
36. Zhou Z, Dai M, Zhao R, Li B, Zhong H, Wen Y (2016) Video error concealment scheme based on tensor model. *Springer Multimed Tools Appl* 76(14):16045–16061

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Sahar Hojati received the B.Sc. degree from the Shahid Chamran University of Ahvaz, Ahvaz, Iran, and the M.Sc. degree from the University of Isfahan, Isfahan, Iran, both in electrical and electronics engineering, in 2015 and 2019, respectively. Her current research interests include image and video processing, computer vision and machine learning.



Mohammad Kazemi received the B.Sc. degree from the Isfahan University of Technology, Isfahan, Iran, in 2003, and the M.Sc. and PhD degrees from the Sharif University of Technology (SUT), Tehran, Iran, both in electrical engineering, in 2005 and 2012, respectively. He currently works with Electrical Engineering Department, University of Isfahan, Iran, as an Assistant Professor. His current research interests include the areas of image/video processing, coding, and transmission, with emphasis on error resilient video coding, as well as digital systems design, specifically for video coding applications.



Payman Moallem was born in Tehran, Iran, in 1970. He received the B.Sc. and M.Sc. degrees in electronics engineering from Isfahan University of Technology, Isfahan, Iran, in 1992, and Amirkabir University of Technology, Tehran, Iran, in 1996, respectively, and the Ph.D. degree in electrical engineering from Amirkabir University of Technology, in 2003. From 1994 to 2002, he conducted research for the Iranian Research Organization Science and Technology (IROST) on the topics such as parallel processing, robot stereo vision, and DSP boards development. In 2003, he joined the University of Isfahan, Isfahan, Iran, as an Assistant Professor, and was promoted to Associate Professor and Full Professor in 2010 and 2015, respectively. He authored more than 300 papers published in peer-reviewed journals and conference proceedings, and five books. His research interests include real-time computer vision and image processing, neural networks, and pattern recognition.