Tabriz Journal of Electrical Engineering (TJEE), vol. 53, no. 4, Winter 2023 Serial no. 106
DOI: 10.22034/tjee.2023.54721.4569

Content-adaptive Gradient-based Error Concealment Scheme for H.264/AVC

Ali Radmehr, Ali Aghagolzadeh*, Seyyed Mehdi Hosseini Andargoli

Faculty of Electrical and Computer Engineering, Noshirvani University of Technology, Babol, Iran
a.radmehr.ir@ieee.org, aghagol@nit.ac.ir*, smh_andargoli@nit.ac.ir
*Corresponding author

Received: 03/01/2023, Revised: 30/03/2023, Accepted: 14/10/2023.

Abstract
A challenging aspect of video error concealment (VEC) is reducing blockiness around the missing region. In this paper, we have devised a novel method for calculating the boundary matching distortion based on the edge direction surrounding the lost area to deal with this problem. After detecting the corresponding slice's erroneous Macroblock (MB), it is divided into four 8×8 pixels sub-blocks. Then, the gradient of outer boundary pixels is derived and determines each side's smoothness, which is further used for drawing the hypothetical line. Furthermore, this paper proposes a novel optimization-based model that can accurately measure boundary matching error for loss recovery. We observe better results for our proposed technique than other related VEC algorithms in terms of PSNR and SSIM. The proposed algorithm improves the average PSNR by 1.09 dB and increases the average SSIM by 0.0135 at the packet loss rate of 10%. In addition, for a PLR of 20%, the PSNR is increased by 1.28 dB, and the SSIM enhancement is 0.019. This algorithm is slightly more computationally complex than the compared methods, but it is still acceptable. Thus, by adding a few computations to the video decoder, the proposed method maintains the quality of the video, especially in the rough, damaged regions of a decoded frame.

Keywords
Video error concealment, H.264/AVC video coding, gradient-based error concealment, boundary matching algorithm.

1. Introduction

Multimedia video communication has become an integral part of daily human life. Therefore, it is a challenging research topic in this field to transmit video data over error-prone networks [1], [2]. H.264/AVC is a video compression codec widely used in consumer video technology [3]–[5]. Also, one bit of loss could severely degrade the video quality due to the high compression ratio [6], [7]. Many error resilience techniques are presented in the literature to solve this issue. Some of them increase redundancy at the encoder, which decreases the coding efficiency [8]. On the other hand, error concealment (EC) techniques utilize the correctly received information at the decoder side [9]. It is mainly classified into three categories, spatial, temporal, and hybrid methods [10].

In the literature, methods for error resilience (ER) and EC have been proposed to mitigate error propagation [7]. Error resilience (ER) and error correction (EC) methods are employed to improve the robustness of the video bitstream against channel errors. Forward error correction (FEC) is used to encode the video bitstream, while automatic repeat request (ARQ) procedures are utilized in case of link failure or network congestion during transmission. However, FEC increases video data bandwidth, and ARQ is not suitable for interactive communications[11]–[12].

Video EC (VEC) is a crucial technique to minimize the impact of transmission errors on the received video signal, particularly in scenarios such as broadcasting and multicasting. VEC is especially important in multicasting because transmission errors can propagate to many receivers, causing a significant decline in the received signal quality. To enhance network throughput, researchers have introduced novel methods for multicasting, such as using the random distribution of popularity among users to increase the probability of sending the same content to multiple users simultaneously. [13].

This paper proposes a hybrid method that utilizes spatial and temporal information from intra-frame and inter-frame redundancies. We develop a novel method to refine motion data based on the gradient information around the corrupted region. Furthermore, we present an optimization-based gradient algorithm for formulating the boundary matching distortion. Many algorithms have been proposed in the literature to measure the boundary matching error. The proposed approach addresses several challenging issues:

1) Evaluating the smoothness of the whole corrupted block is a traditional EC method. It is not always the case since each block's different regions can be smooth or rough. In this paper, we consider a novel boundary distortion evaluation metric for each side of the corrupted block, which enables the proposed motion vector (MV) refinement procedure to be more accurate.

2) Various approaches are developed to determine the lost motion information using the boundary matching algorithm (BMA) or one of its modifications. Although it can recover lost blocks, the surrounding boundary edges cannot be preserved efficiently. In this paper, we develop an optimization-based EC algorithm to determine the spatial distortion of the boundary pixels accurately.

3) Many previous directional EC approaches are computationally intensive, as they apply the directional matching function to the entire boundary pixels. However, the proposed approach classifies different situations and omits unnecessary calculations.

4) Previous studies calculate the block's boundary edges by comparing the block's mutual sides. It may be noisy

DOI: 10.22034/tjee.2023.54721.4569

since two boundary pixels are often far apart. This paper presents an algorithm that exploits more immediate information to find reliable boundary edges.

5) Many directional EC methods only utilize gradient intensity. This paper developed a technique that exploits the gradient phase and reliability factor to improve the accuracy of the results.

This paper proposes an edge-preserving content-adaptive EC algorithm that selects a proper MV from the MV candidates of the missing block. We also present a low-complexity method to alleviate edge discontinuity around the damaged area.

First, the corresponding slice's missing Macroblock (MB) is detected and divided into its 8×8 pixels' sub-blocks. Then, the gradient of outer boundary pixels for each side of the sub-blocks is calculated, and the largest ones are selected. These pixels are potential candidates for finding the dominant edges. Furthermore, we select the boundary edge which goes through the proper directions, resulting in a drastic reduction in computational complexity.

Second, each side of the sub-block is categorized as smooth or rough. Also, we proposed a set of MV candidates for each sub-block. In addition, we developed a novel equation for measuring the boundary matching distortion to find the best MV candidate. Thus, this paper exploits each candidate block's inner boundary pixels to improve the accuracy of edge recovery.

The paper is organized as follows. Section II is devoted to an overview of the latest related work on EC algorithms. Section III presents the problem statement for motion recovery. In section IV, we describe the proposed algorithm in detail. Then, experimental results are presented in section V. Section VI concludes the paper with a summary of our algorithm. Table 1 summarizes the main symbols utilized in our proposed method.

2. Related work

2.1. Spatial methods

Spatial error concealment (SEC) algorithms utilize the correctly received intra-frame pixel information to recover the missing blocks. Many EC methods in the literature mainly focus on SEC. In [14], authors introduced a novel method for multi-directional interpolation SEC to improve the correctness of boundary edge pixels. Also in [15], the authors present a new method which involves 3 steps, including edge estimation, pixel approximation based on predicted directions, and computing weighted average. The proposed method outperforms the previous state-of-the-art methods based on both subjective and objective measures.

Some research has been performed using inpainting methods. In [16], Hwang et al. devised an exemplar-based image inpainting SEC method to derive the best matching patch. Patch regions are calculated using a weighted sum of squared differences (SSD). In [17], The authors introduce a new scalable spatial error concealment algorithm that employs K-MMSE estimator to attain high-quality reconstructions. The technique is comprised several hierarchically stacked layers, with a dynamic layer management mechanism based on visual complexity. It outperforms other algorithms and matches K-MMSE quality. In [18], the authors presented a spiral-like pixel reconstruction (SPR) framework for H.264/AVC. A corrupted MB boundary is matched using the edge matching method, and the dominant edge group is determined based on its edge ratio. This paper shows a higher visual quality than the traditional directional SEC methods. Also, several shape preservation methods for loss concealment attempt to restore the corrupted object's shape.

Edges are points in an image where brightness changes rapidly and are often detected using mathematical techniques known as edge, contour, or boundary detection techniques. Directional edge analysis is used to locate edges in missing areas and the weighted average of two corresponding edges is used to conceal the missing areas. This approach results in a more number of specific computations but is well-suited for images with significant edges and details [9].

In [19] a new algorithm called edge-directed error concealment (EDEC) recovers lost slices in videos by estimating edges in a corrupted frame using neighboring frames and received area. Lost regions along these edges are recovered using spatial and temporal neighboring pixels. EDEC outperforms existing approaches in visual quality and decoder peak signal-to-noise ratio (PSNR). The authors in [20] propose a new spatial error concealment algorithm for image or video recovery, which uses an adaptive linear predictor to reconstruct missing blocks of pixels. The algorithm determines the predictor's order and support shape using Bayesian information criterion and proposes a novel scan order based on pixel uncertainty to reduce error propagation. The proposed method outperforms existing EC algorithms in objective and subjective evaluations.

In [21], Hsia et al. provided a scheme for categorizing the blocks into four classes using different strategies. If the blocks are inside the object or in the background, they have not been utilized for shape preservation. Contrarily, if one of the blocks is on the boundary of the object's shape, it is utilized for error concealment by utilizing the adjacent edges. The method relies on finding the shapes' boundaries, which is unreliable. The SEC algorithms suffer from high computational complexity despite the better subjective and objective quality.

Table 1. The description of key symbols

Symbol	Description
$\boldsymbol{f}(.)$	Objective function
$\mathbb{R}^2$	2D vector space
$\mathbb{R}$	Real number
X	Motion vector candidate
$\boldsymbol{S_{mv}}$	The set of MV candidates
$\boldsymbol{\tau}$	Smoothness parameter
$\boldsymbol{s}$	Scaling factor
r	Penalty flag
$\boldsymbol{\psi}(.)$	Matching function
$\boldsymbol{\phi}(.)$	Penalty function

DOI: 10.22034/tjee.2023.54721.4569

N_s	Total number of the available block's sides
G_x, G_y	Gradient matrix (kernel)
δ	The MV's resolution
$I_c(X)$	Boundary pixels of candidate blocks
I_n	Boundary pixels of neighboring blocks
$g(.)$	SAD function
$h(.)$	Boundary distortion measurement function for rough side of the block
θ	Gradient direction
(x_0, y_0)	Position of the dominant edge
(x, y)	Candidate block's boundary pixel position
(x_i, y_i)	Candidate block's boundary pixels
$d_j(.)$	Distance measurement function for the j^{th}side of the block
$a_j(.)$	Difference of gradients' direction of the neighboring and candidate blocks
$g_j(.)$	Difference of gradients' magnitude of the neighboring and candidate blocks' edges
g_{v_n}	Gradient magnitude of vertical boundary edges of neighboring block
g_{h_n}	Gradient magnitude of horizontal boundary edges of neighboring block
g_{v_c}	Gradient magnitude of vertical boundary edges of candidate block
g_{h_c}	Gradient magnitude of horizontal boundary edges of candidate block
M	Highest video signal value
N	Number of candidate blocks
j	Index of side number for each missing block
i	Index of missing block
k	Index of candidate block
p	Index of block's sides
E_{total}	Total error for a block
$\|G\|_{i,j,k}$	Gradient magnitude for j^{th} side of k^{th} candidate block for i^{th} missing block
$\|G\|_{i,j}$	Gradient magnitude for thej^{th} side of the i^{th} missing block

2.2. Temporal methods

Temporal error concealment (TEC) methods attempt to recover corrupted regions of a video frame utilizing the correlation among consecutive frames of a video sequence. Many TEC algorithms are devised to find a selected set of MVs for P-pictures or a relation between them for B-pictures. Also, there are several methods attempting to estimate the optical flow. These approaches could be useful in predicting missing video data due to transmission errors or packet loss. By predicting the missing information, the method could be used to conceal the errors and improve the visual quality of the reconstructed video. For example in [22] the authors suggest using a deep neural network with fuzzy relationship and optical flow to predict video interactions in advance. They extract two fuzzy images from each video frame using gradient and optical flow, then extract features using a convolutional neural network. The final prediction is made by combining the two outputs of the network. The method demonstrates encouraging outcomes on two interaction datasets.

In [23], the principal component analysis (PCA) is trained with a buffered previous frame to conceal the lossy parts of a decoded frame. Besides, the problem of the scene change is addressed using the updated buffer algorithm. However, it is inefficient in the large, lossy area since this method is highly correlated to the surrounding pixels.

Some EC schemes combine several TEC methods to increase video quality. In [24], Kim et al. proposed the loss concealment procedure, which is applied in two steps. First, if the corrupted block is in an intra-frame, it is replaced using the co-located block in the previous frame. Second, the other MVs are recovered using DMVE. Despite relatively good results for large missing areas, the high dependence of MVs on the first stage results in poorer performance for more dynamic frames.

Furthermore, Li et al. devised a new MV modeling by a plane fitting in [25]. The neighboring MVs are linearly interpolated to derive the MV candidates of the missing MB. Finally, the selected MV is determined based on the OBMA algorithm. Filter-based TEC schemes have also been developed for MV refinement. In [26], particle filtering is exploited for MV recovery. The results are superior in many cases due to the non-stationary nature of MVs in a video frame.

2.3. Hybrid methods

Hybrid EC (HEC) methods exploit spatial and temporal information to recover the missing blocks better. The important issue with HEC is to recover lost motion information while preserving boundary edges. In [27], the authors proposed a sparse recovery method where the encoder transmits side information to help the decoder cope with packet loss. However, this algorithm increases the bitrate, and thus it has a limited application.

Moreover, there are spatio-temporal gradient based VEC methods which are designed to enhance the overall visual quality of videos by smoothing out any errors that may occur during transmission or processing. While these techniques may not always result in a significant increase in the objective quality metrics such as PSNR, they can be successful in recovering edges and details that are lost due to video compression or transmission errors [28], [29].

Furthermore, the authors in [10] proposed a method using the surrounding pixel information to find the most correlated MV for each 8×8 block. The final MVs are selected based on the four candidates. MV is chosen if the motion trajectory in the previous frame is sufficiently close. The statistically significant MV with the highest correlations is selected, and the weighted average of the other MVs is computed. The recovered MV is calculated by averaging these MVs to find the missing MB's displacement. Chung et al. in [30] devised an HEC to recover all corrupted regions of a frame. Firstly, they perform TEC and propagate source information in the forward direction and backwardly for the remaining missing area in a group-of-frames (GOF). Moreover, they propose an adaptive homography-based registration approach for lost blocks' recovery. This method shows comparably good performance for relatively large unknown regions.

In [31], the authors propose a regression-based method for fixing errors in video frames using nearby pixels. The method combines regression coefficients from spatial and temporal pixels to fill in corrupted blocks, improving visual quality compared to other methods. However, it

DOI: 10.22034/tjee.2023.54721.4569

only works well when there is a sufficient correlation between the missing area and the surrounding pixels. Also, deriving an appropriate matching point is not always applicable to various video sequences.

More recently, a generative adversarial network (GAN) has been applied for HEC [32]. They trained completion and discriminator networks exploiting the spatially and temporally neighboring pixels. Next, the missing regions are recovered using the trained generator to conceal the unknown area regarding the whole frame. Although this approach can recover the details in a corrupted region of a video frame, it is very time-consuming due to training the GAN network. The paper [33] suggests a network that predicts optical flow of lost areas using convolutional LSTM and a simple convolutional layer. It uses two rows of MBs above and below of each slice for three frames. However, this approach may not be enough for some content, and training the network for potential loss locations may not be practical in real-life situations.

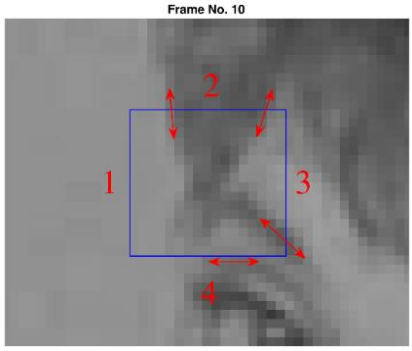

Fig. 1 Demonstration of different types of the block's sides in the 10th frame of the "Mother and daughter" sequence. A blue rectangle represents the 6th row and 13th column of the corresponding block. The numbers 1, 2, 3, and 4 are the neighboring regions of this block

3. Problem statement

3.1. Motivation of the proposed method

An important objective of EC algorithms is to preserve details and edges around the damaged areas. Therefore, most of these methods attempt to estimate boundary distortion for the corrupted blocks. Usually, the damaged blocks are classified into three categories: flat block, edge block, and texture block [34]. However, this classification is not always valid. There are many cases where a block cannot be explicitly assigned to one of these three classes. This is shown in Fig. 1 for the 10th frame of the "Mother and daughter" sequence.

In this figure, four neighboring regions of the considered block are illustrated, and the blue rectangle area shows the 6th row and 13th column of this frame. Region 1 is flat, and there is no edge near the boundary. Regions 2 and 3 have dominant edges. Even though the outer boundary pixels of region 4 show a dominant edge, this edge does not continue within the block. Hence, the replacing block for the considered block should not include a continuing edge at region 4.

Although the MVs are derived in rate-distortion optimization (RDO) process, it is a process that occurs in video encoder. It is an operation to compromise between the calculated distortion and the data transfer rate. In other words, in RDO process, different MVs are evaluated and those are selected that create the best compromise between the visual quality and the bit rate. Additionally, in RDO process, the compressed frame is compared with the original frame and the amount of calculated distortion. However, in EC problem, due to the existence of errors, there is not enough information to accurately measure the amount of distortion for MVs. Therefore, the available neighboring information of the damaged areas of the video frame is used to measure the amount of distortion for candidate MVs. Hence, two issues need to be addressed: 1) The matching distortion criteria should be calculated for each side of the corrupted block, not only for the whole block. 2) The surrounding edges of the corrupted block should be measured by analyzing the inner and outer boundary pixels rather than relying solely on the information of the outer boundary pixels.

To address these two problems, we develop a novel edge-preserving EC scheme. The proposed method suggests a new boundary distortion measurement algorithm to enhance video quality. Furthermore, the proposed method optimizes the number of computations for the boundary matching formula by exploiting the lost block's neighboring information.

3.2. Problem statement

This paper studies the problem of finding the proper replacements for the corrupted blocks among all feasible solutions. We characterize this issue as an optimization problem; the goal is to find the optimal solution to reconstruct the missing areas in the received video frames using the neighboring information. To achieve this, we propose a modeling strategy for MV field to find the most appropriate MV among MV candidates to minimize the boundary matching difference between the candidate and spatially neighboring blocks, while satisfying the following certain assumptions:

1) MVs are generated for each erroneous block within the missing 16×16 pixels (i.e., macroblock) region.

2) We make assumptions about each side of the missing block: i) Smooth: no dominant edge is present or no edge go through the missing block. ii) Rough: a dominant edge is present, which passes through the missing block.

3) The recovering MV is assumed to derive the block containing all necessary samples for the missing block.

4) The proposed algorithm assumes that the frames are from a natural frame scene, meaning that the information is highly correlated and similar blocks are available in the previous frame. This can be achieved by having close contexts at the missing block's boundaries.

5) We assume that the higher the amount of context to the boundary pixels, the lower the total error.

6) Better MVs for EC are assumed to provide a higher amount of good context for motion estimation and higher similarity with missing blocks.

7) Samples are assumed to come from an existing block in the reference frame and not constructed.

8) We assume that two regions on different sides of a missing block are uncorrelated.

9) Straight lines passing through the missing area are assumed to represent the missing edges.

To formulate this problem, consider the function $f(.)\colon \mathbb{R}^2 \to \mathbb{R}$ to be the objective function that calculates

DOI: 10.22034/tjee.2023.54721.4569

the overall boundary matching distortion among the replacement candidates and neighboring blocks of the lost MB:

$$f(\boldsymbol{X},\tau,s,r)=\sum_{j=1}^{N_s}\big(\psi(\boldsymbol{X},\tau)+\phi(\boldsymbol{X},s,r)\big) \quad (1)$$

$$\min_{\boldsymbol{X}} f(\boldsymbol{X},\tau,s,r) \quad (2)$$
$$s.t.\,\boldsymbol{X}\in S_{\boldsymbol{mv}}$$
$$\mathrm{r}\in\{0,1\}$$
$$\mathrm{s}\in\mathbb{R}$$

where $S_{\boldsymbol{mv}}$ is the set of MV candidates and N_s is the total number of the available sides for each block. Also, $\psi(\boldsymbol{X},\tau)$ is the matching function and τ is the smoothness parameter for each side. Besides, $\phi(\boldsymbol{X},s,r)$ is the penalty function where s and r are the scaling factor and penalty flag, respectively. Equation (2) aims to minimize the artificial discontinuities that damage the dominant boundary edges of the corrupted block. These discontinuities can greatly impact the subjective quality of the decoded video frame [35].

Additionally, several studies in the literature have utilized the opposite sides of a lost MB to estimate dominant edges [18]. However, two major issues need to be considered: 1) The boundary pixels of the opposite sides of the lost MB are far from each other. Therefore, they are uncorrelated, and matching them leads to false results. 2) The boundary edge of the lost block often continues into the lost block without any matching edge being found on the other side.

4. The proposed method

In our proposed approach, we use only one side of the block to predict the corresponding edge for the lost block recovery. In addition, we exploit the neighboring motion information, which is included in the set $S_{\boldsymbol{mv}}$. This set consists of the neighboring MVs and their average, the zero MV, and the co-located one. Also, we suggest the resolution limitation of the MVs to reduce computational complexity. Accordingly, the set $S_{\boldsymbol{mv}}$ contains the MVs with limited resolution, denoted by δ.

In the present manuscript, we use the Sobel operator for edge detection, which can accurately determine the gradient of the boundary pixels. The Sobel operator is applied to the Y component of a video signal using G_x and G_y:

$$G_x=\begin{bmatrix}-1&0&1\\-2&0&2\\-1&0&1\end{bmatrix}, G_y=\begin{bmatrix}1&2&1\\0&0&0\\-1&-2&-1\end{bmatrix} \quad (3)$$

Furthermore, if the gradient magnitude of the boundary pixels' Y component is higher than a fixed threshold, it is regarded as a dominant edge. The Sobel operator calculates the magnitude and direction of the edge pixels. Also, it is widely used in video processing because it has low computational complexity [36].

4.1. The block side without any boundary edge

The proposed method exploits the Sum of Absolute Difference (SAD) method while the block's side has no dominant boundary edge. It is depicted as the first side of the block in Fig. 1. It is a smooth boundary, and the boundary pixels are almost the same in this case. Thus, $\psi(\boldsymbol{X},\tau)$ in equation (1) is formulated as follows:

$$\psi(\boldsymbol{X},\tau)=\tau g(I_c(\boldsymbol{X}),I_n)+(1-\tau)h(I_c(\boldsymbol{X}),I_n) \quad (4)$$

where τ is one. Also, $I_c(\boldsymbol{X})$ and I_n are boundary pixels of the candidate and neighboring blocks, respectively. Also, $g(.)$ is the SAD function for measuring the boundary distortion of each 8×8 block and $h(.)$ is used when the boundary is rough. Thus, $g(.)$ can be formulated as equation (5) for the smooth block side:

$$g(I_c(\boldsymbol{X}),I_n)=\frac{1}{8}\sum_{i=0}^{i=7}|I_c(\boldsymbol{X})-I_n| \quad (5)$$

where $I_c(\boldsymbol{X})$ denotes the boundary pixels of the candidate block, and I_n represents the boundary pixels of the neighboring block which is illustrated in Fig. 2.

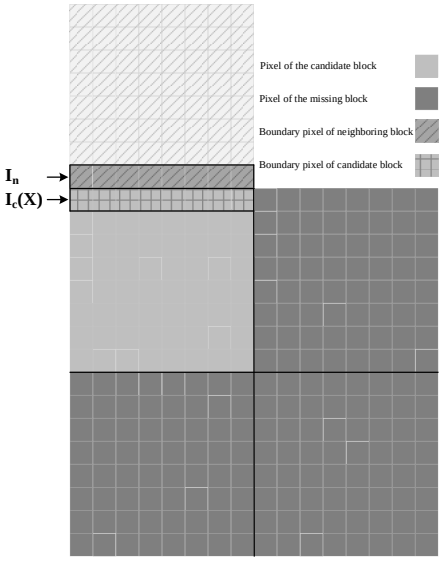

Fig. 2 Illustration of the pixels of the missing and candidate blocks. Also, $I_c(\boldsymbol{X})$ and I_n are the boundary pixels of the candidate and neighboring blocks, respectively

4.2. The block side with the boundary edge

A rough block side can challenge EC methods when one or more dominant edges are on the block boundary. The recovery of the blocks with rough sides is considerably more complicated because of the greater level of detail in such blocks. Moreover, the traditional boundary matching algorithm is not applicable since the SAD calculation can only preserve smooth regions. Therefore, this paper presents a new EC procedure to maintain the continuity of the edges on the block sides. The proposed method uses the dominant edge to draw a hypothetical line that goes through the candidate block, which is illustrated in Fig. 3. Since the Sobel operator is a 3×3 matrix, the three immediate neighboring rows of pixels are exploited for edge detection. An edge is considered as dominant if its gradient magnitude is more than a fixed threshold. Also, the gradient is always in the direction of maximum variation.

Consider θ to be the gradient direction. Thus, its degree with the horizontal line is $\theta-\frac{\pi}{2}$. Also, consider the pixel position of the dominant edge to be (x_0,y_0), then the

DOI: 10.22034/tjee.2023.54721.4569

equation of the hypothetical line can be formulated as follows:

$$y - y_0 = -\cot(\theta).(x - x_0) \quad (6)$$

where (x, y) is the boundary pixel position in the candidate block and $-\cot(\theta)$ is the hypothetical line's slope. The candidate block should have the nearest edge position to the hypothetical line, which leads to the least discontinuity at the boundaries.

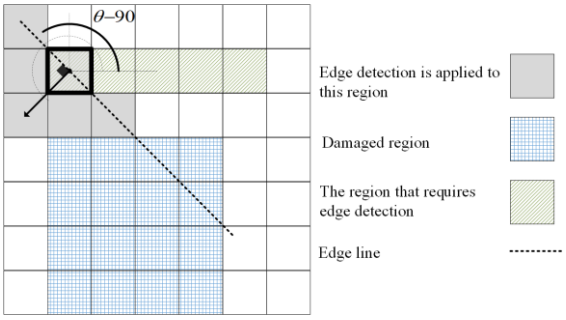

Fig. 3 Illustration of dominant edge and hypothetical line. The dominant edge is calculated, which is exploited for drawing a hypothetical line. This line goes through the corresponding missing block

In this paper, candidate blocks can have two types of sides, i.e., smooth and rough. In another case, the edge enters the block and does not continue to the other side of the block. This situation is studied in [18]. The given method in [18] does not consider this situation when no matching edge pixels can be found for the corresponding block. However, it contrasts with the assumption of preserving the edge integrity at the boundaries of the corrupted blocks. Thus, in order to maintain the edge integrity at the rough sides of a block, the proposed method measures three following parameters:

1) The distance of the candidate block's boundary edge to the hypothetical line is calculated as follows:

$$d_j(\theta, x_i, y_i) = |\sin(\theta)| \times |\cot(\theta).x_i - x_0 + (y_i - y_0)| \quad (7)$$

where (x_i, y_i) is the coordinates of boundary pixels of the candidate block for the j^{th} side of the block.

2) The gradient direction difference between the edge of the candidate and the adjacent block is calculated using equation (8). The smaller the difference, the greater the integrity of the boundary edge. Consider $a_j(g_{v_n}, g_{h_n}, g_{v_c}, g_{h_c})$ to be the difference between the directions of the gradients of the neighboring and candidate block. It can be calculated as follows:

$$a_j(g_{v_n}, g_{h_n}, g_{v_c}, g_{h_c}) = |\tan^{-1}\frac{(g_{v_n} \times g_{h_c} - g_{v_c} \times g_{h_n})}{(g_{h_n} \times g_{h_c} - g_{v_n} \times g_{v_c})}| \quad (8)$$

where $g_{v_n}, g_{h_n}, g_{v_c}$, and g_{h_c} are, respectively, the gradient magnitude of vertical and horizontal boundary edges of neighboring and candidate blocks.

3) The gradient magnitude difference between the candidate and the adjacent block's boundary edge is calculated using equation (9). The closer this value is to zero, the more similar the two edges are. Consider $g_j(.)$ to be the magnitude difference function. It can be formulated as follows:

$$g_j(g_{v_n}, g_{h_n}, g_{v_c}, g_{h_c}) = |\sqrt{g_{h_n}^2 + g_{v_n}^2} - \sqrt{g_{h_c}^2 + g_{v_c}^2}| \quad (9)$$

Furthermore, τ equals 0 in equation (4) when one or more dominant boundary edges are detected and $\psi(\boldsymbol{X}, \tau)$ equals $h(I_c(\boldsymbol{X}), I_n)$, which can be formulated as follows:

$$h(\boldsymbol{X}_c, \boldsymbol{X}_n) = d_j(\theta, x_i, y_i) + a_j(g_{v_n}, g_{h_n}, g_{v_c}, g_{h_c}) + g_j(g_{v_n}, g_{h_n}, g_{v_c}, g_{h_c}) \quad (10)$$

On the other hand, there is a case where the edge does not go through the damaged block. It is demonstrated in Fig. 1 for the fourth side of the corresponding block. In other words, the edge is detected on the boundary pixels, but no edge is discovered for the candidate block. In this case, the gradient direction is not between the minimum and maximum possible degree, which is shown in Fig. 4, and equation (4) measures the distortion of the boundary pixels.

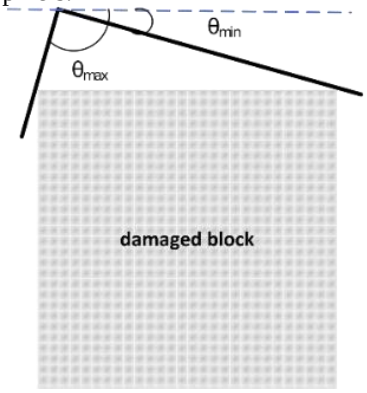

Fig. 4 Maximum and minimum possible gradient direction

Additionally, suppose there is no boundary edge for the candidate block, but the neighboring block contains a dominant edge, and its direction is between the minimum and maximum possible degree. In this case, the penalty function, i.e., $\phi(\boldsymbol{X}, s, r)$ is applied using equation (11) and increases the total error, which is formulated as follows:

$$\phi(\boldsymbol{X}, s, r) = s.r.g(I_c(\boldsymbol{X}), I_n) \quad (11)$$

where $\phi(\boldsymbol{X}, s, r)$ is the penalty function. Also, the penalty flag r equals 1 in this case, and the scaling factor s is determined empirically. Moreover, $g(I_c(\boldsymbol{X}), I_n)$ is the boundary matching distortion for the boundary pixels of the candidate and neighboring blocks. In this case, the

DOI: 10.22034/tjee.2023.54721.4569

candidate block which causes the edge discontinuity is penalized in our proposed method.

Fig. 5 illustrates the diagram of the proposed method. First, the gradient of the boundary pixels of the neighboring block is calculated for the j^{th} side of the i^{th} block. Second, if the gradient magnitude is less than a fixed threshold, the matching error is calculated using equation (5). Otherwise, if the gradient direction is between the minimum and maximum possible value, the gradient magnitude is calculated for the k^{th} block for all N candidate blocks. Then, the boundary matching error is calculated using the penalty function, which is formulated in equation (11). Otherwise, the boundary distortion is measured using equation (10). In this paper, a threshold level of 0.1 is considered for gradient magnitude to find the edges. First, the operation of finding the dominant edge for the boundary pixels is performed. Next, we select the edge with the highest gradient value as the dominant edge, based on the threshold level of 0.1 for the gradient magnitude. It is used to find the dominant edge of the boundary pixels.

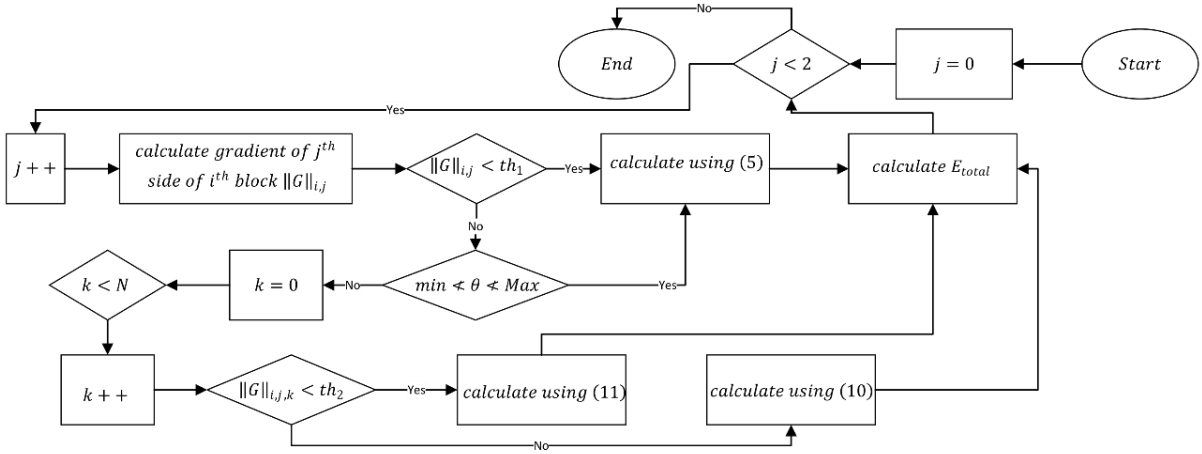

Fig. 5 Illustration of the proposed algorithm in detail

5. Experimental results

In this section, the performance and computational complexity of the proposed method are compared with other related EC algorithms, such as the traditional boundary matching Algorithm (BMA), Particle filtering based method (PF) [26], Lin [37] approach, and Nam [10] method. Furthermore, the MV resolution, i.e., δ, is 0.25 and 0.5 for various experiments. The video coding structure is "IPPP," consisting of only one intra-frame (I-frame) after 50 inter-frames (P-frame). The peak-signal-to-noise ratio (PSNR) is evaluated for the performance comparison of the decoded video sequences, which is formulated as follows:

$$PSNR = 10 \times log \frac{M^2}{MSE} \qquad (12)$$

where M denotes the highest video signal value, 255 in the experiments, and MSE is the Mean Square Error difference between two frames [29]. Besides, this paper exploits Structural Similarity Index Measurement (SSIM) to find the perception-based information from the experimented video sequences [11],[38],[39]. All the simulations are tested on Dual-Core, 2.6 GHz MacBook laptop. The Quantization Parameter (QP)s were 24 and 28. Also, the packet loss ratio (PLR) is 10% and 20%. The experiments were conducted on the H.264/AVC platform using the MATLAB software for standard CIF (352×288 pixels) test sequences encoded with 30 Frames per Second (fps).

Table **2** and Table 3 compare the performance of the proposed method with other EC algorithms in terms of PSNR and SSIM. Moreover, the PLR is set to 10% and 20% for Table **2** and Table 3. This experiment is carried out for 100 frames of the "Foreman" and "Coastguard" video sequences. The proposed method outperforms the Lin algorithm on average by 1.09 dB and 1.28 dB in terms of PSNR for PLR of 10% and 20%, respectively. Moreover, the proposed approach increases the SSIM by 0.0135 and 0.019 on average for PLR of 10% and 20%, respectively. The improved video quality is consistent with the proposed approach compared to the other algorithms, and the superiority of the proposed method is confirmed for various EC techniques. Moreover, the proposed method is evaluated with MV resolutions of 0.25 and 0.5. In addition to a slight performance improvement, the complexity increases for $\delta = 0.25$.

Fig. 6 depicts the subjective comparison of the proposed method for the 8th frame of the "Coastguard" test sequence. The erroneous frame is illustrated in Fig. 6(a), where the PLR is set to 20%. The visual quality of the concealed frame is enhanced for the proposed method, and the patterns are better preserved. Moreover, the edges are recovered, and the motion information is better estimated, which is evident for the moving boat, while the blockiness is visible for other techniques.

We attribute the performance enhancement to three main factors: 1) While most EC algorithms consider the smoothness of the entire missing block, the proposed method considers the boundary edges for each side separately. Thus, the proposed technique can preserve the details more accurately. 2) Edge detection only occurs when the hypothetical line passes through the candidate block, which drastically reduces computational complexity. 3) Unlike other EC techniques, our proposed method finds the boundary edge derived using the same

DOI: 10.22034/tjee.2023.54721.4569

side of the corrupted block. It increases the accuracy of motion recovery since it minimizes boundary matching distortion.

5.1. Computational complexity analysis

The required decoding time is compared in
Table **2** and Table 3 for different EC algorithms. EC methods increase the overall decoding time since they are post-processing techniques. In the present paper, the proposed method optimizes the computations needed for edge recovery by exploiting the available information. In other words, the proposed approach depends on the region's smoothness surrounding the corrupted area. The smoother the region, the fewer computations are required. The proposed method's computational complexity is higher than fast EC methods, such as Nam [10] algorithm but still acceptable. The proposed technique provides a balance between video quality and computational complexity.

6. Conclusion

This paper proposed a novel gradient-based boundary matching method for lossy video frame recovery. The proposed approach addressed the issue of edge discontinuity and presented a practical edge-preserving EC approach. It categorizes each side of the lost block as rough and smooth, which is further processed separately to calculate the boundary matching distortion. Furthermore, this paper proposed novel criteria for calculating boundary differences for the rough regions of a video frame. This method exploited boundary pixels' gradient information and devised a hypothetical line to estimate the candidate block's similarity with neighboring ones accurately. Also, other approaches can be studied to find the more proper hypothetical line in the future. The proposed method can restore complicated edges and textures of the damaged regions of a video frame even when the edge does not cross opposite sides of a corrupted block.

Moreover, this paper determined the proposed method's computational complexity and the required decoding time for the proposed method and compared it against the other EC algorithms. Results indicate that the proposed method achieves superior results over the other tested EC algorithms while maintaining a reasonable trade-off between video quality and computation complexity.

Acknowledgments

This work was supported by Babol Noshirvani University of Technology through grant program No. BNUT/389059/1401.

Table 2. PERFORMANCE COMPARISON OF H.264/AVC IN TERMS OF PSNR, SSIM, AND RUN TIME FOR PACKET LOSS RATE 10%

QP	Video		BMA	Lin [38]	Nam [10]	Proposed($\delta = 0.5$)	Proposed($\delta = 0.25$)
24	Foreman	PSNR (dB)	29.66	29.80	31.23	32.73	32.74
		SSIM	0.9088	0.9119	0.9120	0.9296	0.9297
		Time (s)	0.3366	0.0912	0.3312	0.8441	0.8667
	Coastguard	PSNR (dB)	31.13	29.58	31.36	31.71	31.73
		SSIM	0.9152	0.8984	0.9011	0.9227	0.9229
		Time (s)	0.3346	0.0928	0.3291	0.6002	0.6206
28	Foreman	PSNR (dB)	29.06	29.21	30.47	31.57	31.58
		SSIM	0.8932	0.8956	0.9056	0.9110	0.9116
		Time (s)	0.3349	0.0901	0.3252	0.6940	0.7144
	Coastguard	PSNR (dB)	29.98	28.73	29.09	30.42	30.45
		SSIM	0.8905	0.8769	0.8920	0.9001	0.9003
		Time (s)	0.3324	0.0903	0.3088	0.5912	0.6001
Average		PSNR (dB)	29.96	29.33	30.54	31.61	31.63
		SSIM	0.9019	0.8957	0.9026	0.9158	0.9161
		Time (s)	0.3346	0.0911	0.3235	0.6824	0.7004

Table 3. PERFORMANCE COMPARISON OF H.264/AVC IN TERMS OF PSNR, SSIM, AND RUN TIME FOR PACKET LOSS RATE 20%

QP	Video		BMA	Lin [38]	Nam [10]	Proposed($\delta = 0.5$)	Proposed($\delta = 0.25$)
24	Foreman	PSNR (dB)	27.46	27.54	28.48	30.14	30.16
		SSIM	0.8565	0.8591	0.8701	0.8961	0.8967
		Time (s)	0.6722	0.0963	0.6581	1.2031	1.4035
	Coastguard	PSNR (dB)	28.73	27.32	29.48	31.15	31.16
		SSIM	0.8536	0.8213	0.8581	0.8688	0.8697
		Time (s)	0.6935	0.0961	0.6500	0.8861	0.1063
28	Foreman	PSNR (dB)	26.84	26.99	29.43	29.47	29.50
		SSIM	0.8421	0.8460	0.8692	0.8791	0.8801
		Time(s)	0.6621	0.0950	0.6510	1.071	1.2810
	Coastguard	PSNR (dB)	27.80	26.57	27.07	28.74	28.75
		SSIM	0.8287	0.8020	0.8211	0.8478	0.8479
		Time (s)	0.6771	0.0941	0.6495	0.8716	1.0141
Average		PSNR (dB)	27.70	27.10	28.61	29.87	29.89

	SSIM	0.8452	0.8321	0.8546	0.8729	0.8736
	Time (s)	0.6762	0.0953	0.6521	1.0079	1.1904

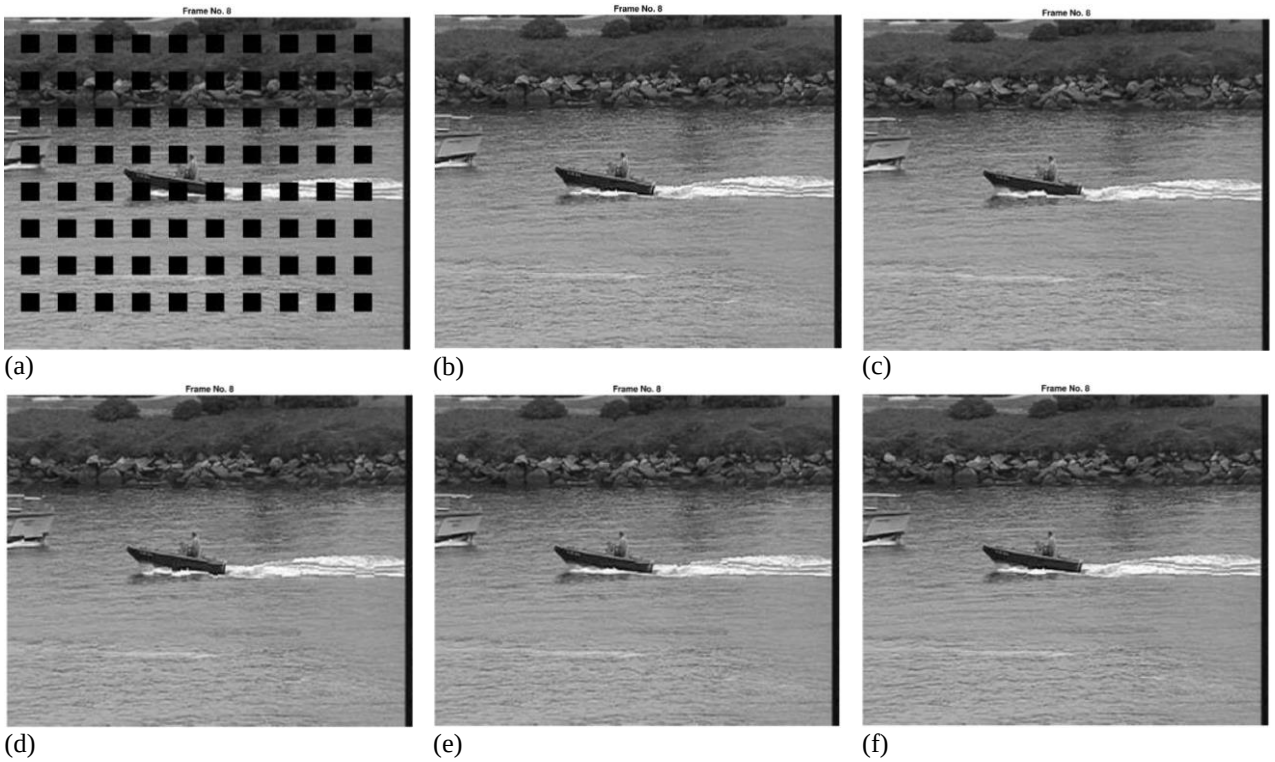

Fig. 6 Visual comparison of the reconstructed 8th frame of "Coastguard" video sequence (PLR=20%).(a) corrupted (b) BMA (c) Lin method [38] (d) Nam method [10] (e)the proposed method ($\delta = 0.5$) (f)the proposed method ($\delta = 0.25$)

7. References

[1]A. J. Hussain and Z. Ahmed, "A survey on video compression fast block matching algorithms," Neurocomputing, vol. 335, pp. 215–237, 2019.

[2]V. Tiwari and C. Bhatnagar, "A survey of recent work on video summarization: approaches and techniques," Multimed Tools Appl, vol. 80, no. 18, pp. 27187–27221, 2021.

[3]International Standard ISO/IEC, "Recommendation ITU-T H.264," 2017.

[4]Cisco, "Cisco Visual Networking Index: Forecast and Methodology," Cisco White Paper, 2017. [Online]. Available:https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html. [Accessed Dec 1, 2022]

[5]J. He, Y. Xu, W. Luo, S. Tang, and J. Huang, "A novel selective encryption scheme for H.264/AVC video with improved visual security," Signal Process Image Commun, vol. 89, p. 115994, 2020.

[6]R. Ma, T. Li, D. Bo, Q. Wu, and P. An, "Error sensitivity model based on spatial and temporal features," Multimed Tools Appl, vol. 79, no. 43–44, pp. 31913–31930, 2020.

[7]R. Chellappa and S. Theodoridis, Academic Press Library in Signal Processing, Volume 6: Image and Video Processing and Analysis and Computer Vision. Academic Press, 2017.

[8]K. Go, M. Kim, S. Kang, and Y. Yoon, "A systematic reallocation and prioritization scheme for error-resilient transmission of video packets," Multimed Tools Appl, vol. 76, no. 5, pp. 6755–6783, 2017.

[9]M. Usman, X. He, M. Xu, and K. M. Lam, "Survey of error concealment techniques: Research directions and open issues," in 31st Picture Coding Symposium (PCS), pp. 233–238, 2015.

[10]C. Nam, C. Chu, T. Kim, and S. Han, "A novel motion recovery using temporal and spatial correlation for a fast temporal error concealment over H.264 video sequences," Multimed Tools Appl, vol. 79, no. 1, pp. 1221–1240, 2020.

[11]H. R. Wu and K. R. Rao, Digital video image quality and perceptual coding. CRC press, 2017.

[12]D. R. Bull, Communicating pictures: A course in image and video coding. Academic Press, 2014.

[۱۳]محمود دی پیر و فرزاد گویا، "به کارگیری ارتباطات چندکاربره ی دستگاه به دستگاه برای تحویل بیسیم محتوای ویدئویی در شبکه های سلولی،" *مجله مهندسی برق، دانشگاه تبریز*، دوره پنجاهم شماره چهارم، صفحات ۱۵۸۱-۱۵۹۱، سال ۱۳۹۹.

[14]H. Byongsu, J. Jonghyon, and R. Cholsu, "An improved multi-directional interpolation for spatial error concealment," *Multimed Tools Appl*, vol. 78, no. 2, pp. 2587–2598, 2019.

[15]H. Asheri, H. Rabiee, N. Pourdamghani, and M. Ghanbari, "Multi-directional spatial error concealment using adaptive edge thresholding," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 880–885, 2012

[16]H. Byongsu, J. Jonghyon, and R. Cholsu, "An improved exemplar-based image inpainting algorithm for error concealment," *ICTACT Journal on Image and Video Processing (IJIVP)*, vol. 8, no. 1, pp. 1583–1587, 2017.

[17]J. Koloda, J. Seiler, A. M. Peinado, and A. Kaup, "Scalable Kernel-Based Minimum Mean Square Error Estimator for Accelerated Image Error Concealment,"

IEEE Transactions on Broadcasting, vol. 63, no. 1, pp. 59–70, 2017.
[18]H. C. Shih, C. T. Wang, and C.-L. Huang, "Spiral-Like Pixel Reconstruction Algorithm for Spatiotemporal Video Error Concealment," *IEEE Access*, vol. 6, pp. 6370–6381, 2018.
[19]Mengyao Ma, O. C. Au, S.-H. G. Chan, and Ming-Ting Sun, "Edge-Directed Error Concealment," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 3, pp. 382–395, 2010.
[20]Jing Liu, Guangtao Zhai, Xiaokang Yang, Bing Yang, and Li Chen, "Spatial Error Concealment With an Adaptive Linear Predictor," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 3, pp. 353–366, 2015.
[21]S.-C. Hsia and C. H. Hsiao, "Fast-efficient shape error concealment technique based on block classification," *IET Image Process*, vol. 10, no. 10, pp. 693–700, 2016.
[۲۲] مه لقا افراسیابی،حسن ختن لو و محرم منصوری زاده، "شبکه عصبی عمیق برای پیشبینی تعامل انسان در ویدئو با استفاده از روابط فازی و شار نوری"، *مجله مهندسی برق، دانشگاه تبریز*، دوره پنجاهم، شماره سوم، صفحات ۱۰۳۵-۱۰۴۶، سال ۱۳۹۹.
[23]G. Choe, C. Nam, and C. Chu, "An effective temporal error concealment in H.264 video sequences based on scene change detection-PCA model," *Multimed Tools Appl*, vol. 77, no. 24, pp. 31953–31967, 2018.
[24]D. Kim, Y. Kwon, and K. Choi, "Motion-Vector Refinement for Video Error Concealment Using Downhill Simplex Approach," *ETRI Journal*, vol. 40, no. 2, pp. 266–274, 2018.
[25]Y. Li and R. Chen, "Motion vector recovery for video error concealment based on the plane fitting," *Multimed Tools Appl*, vol. 76, no. 13, pp. 14993–15006, Jul. 2017.
[26]A. Radmehr and A. Ghasemi, "Error concealment via particle filter by Gaussian mixture modeling of motion vectors for H.264/AVC," *Signal Image Video Process*, vol. 10, no. 2, pp. 311-318, 2016.
[27]T. L. Lin, T. L. Ding, C. Y. Fan, and W. C. Chen, "Error concealment algorithm based on sparse optimization," *Multimed Tools Appl*, vol. 76, no. 1, pp. 397–413, 2017.
[28]M. Kazemi, M. Ghanbari, and S. Shirmohammadi, "A review of temporal video error concealment techniques and their suitability for HEVC and VVC," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 12685–12730, 2021.
[29]M. Kazemi, M. Ghanbari, and S. Shirmohammadi, "The Performance of Quality Metrics in Assessing Error-Concealed Video Quality," *IEEE Transactions on Image Processing*, vol. 29, pp. 5937–5952, 2020.
[30]B. Chung and C. Yim, "Bi-Sequential Video Error Concealment Method Using Adaptive Homography-Based Registration," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 6, pp. 1535–1549, 2020.
[31]Y. Zhang, X. Xiang, D. Zhao, S. Ma, and W. Gao, "Packet video error concealment with auto regressive model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 1, pp. 12–27, 2012.
[32]C. Xiang, J. Xu, C. Yan, Q. Peng, and X. Wu, "Generative adversarial networks based error concealment for low resolution video," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1827–1831, 2019.
[33]A. Sankisa, A. Punjabi, and A. K. Katsaggelos, "Video Error Concealment Using Deep Neural Networks," in *25th IEEE International Conference on Image Processing (ICIP)*, pp. 380–384, 2018.
[34]S. Chapaneri, S. Mistry, and S. Dixit, "Performance Evaluation of Edge-based Video Error Concealment using H. 264 Flexible Macroblock Ordering," *International Journal of Computer Science & Engineering Technology (IJCSET)*, vol. 3, no. 12, pp. 605–615, 2015.
[35]J. Korhonen, "Study of the subjective visibility of packet loss artifacts in decoded video sequences," *IEEE Transactions on Broadcasting*, vol. 64, no. 2, pp. 354–366, 2018.
[36]T.-H. Tsai, S.-S. Su, and T.-Y. Lee, "Fast mode decision method based on edge feature for HEVC inter-prediction," *IET Image Process*, vol. 12, no. 5, pp. 644–651, 2018.
[37]T.-L. Lin, X. Wei, X. Wei, T.-H. Su, and Y.-L. Chiang, "Novel pixel recovery method based on motion vector disparity and compensation difference," *IEEE Access*, vol. 6, pp. 44362-44375, 2018.
[38] U. Sara, M. Akter, and M. S. Uddin, "Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study," Journal of Computer and Communications, vol. 07, no. 03, pp. 8–18, 2019.
[39] F. Tommasi, V. de Luca, and C. Melle, "Packet losses and objective video quality metrics in H.264 video streaming," J Vis Commun Image Represent, vol. 27, pp. 7–27, 2015.