Video Error-Resilience Research Based on Error-Resilient Screen Content

Zhe Liu [1,*], He Chen [1] and Songlin Sun [2]

[1]   School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China;
      beili2020@126.com
[2]   School of Information and Communication Engineering, Beijing University of Posts and
      Telecommunications, Beijing 100876, China; slsun@bupt.edu.cn
*   Correspondence: liuz2020bit@163.com

check for
updates

Abstract: In order to make video transmission more stable, various error-resilient mechanisms are proposed on video coding in the literature. However, the redundancy mechanism behind classical redundant coding algorithms is relatively simple and is not suitable for the network environment and video content in the context of screen content sequence with multiple abrupt frames and still frames. Motivated by this, a frame-level coding selection mechanism is proposed in this paper for the error-resilience transmission of screen content, where additional code stream or redundant information is considered to improve error-resilient performance with redundant coding and acceptable video quality is obtained in the case of frame transmission error. In addition, selective allocation redundancy is conducted to take the importance of the video frame ROI (region of interest) area into account in the co-encoding process. As a result, the redundancy insertion efficiency and the reliability are improved in return. The corresponding experiments validate the effectiveness of the schemes proposed in this paper.

Keywords: screen content coding; error-resilience; redundant coding; HEVC

1. Introduction

With the rapid development of multimedia technology, video resolution has been constantly improving. The storage and transmission of high-definition and ultra-high-definition video bring great challenges to video coding technology. Under the background of emerging video applications and continuous innovation of network and computer technology, high-efficiency video coding (HEVC) has further optimized the H.264 coding standard according to the characteristics of high definition (HD) video. Compared with H.264/AVC, 50% bit-rate can be saved for HEVC in terms of high profile on the premise of the same video coding quality. Consequently, HEVC is widely accepted and applied [1].

Screen content image is generally composed of natural and computer-generated images—image [2]. The encoding method of screen content image can adapt to the content of both images at the same time. At present, this coding method is widely used in remote desktop sharing, video conference, radio and television, and other video applications [3]. In modern coding standards, time-domain prediction is used to eliminate the redundancy, and video data are highly dependent on each other. Packet loss will cause serious distortion after subsequent decoding. Fault-tolerant coding algorithms can reduce the errors in the process of video transmission [4]. In order to make better use of the characteristics of screen content video and improve the encoding efficiency of screen content video, VCEG and MPEG jointly issued a proposal solicitation announcement of screen content video encoding in January 2014, and officially started the work of formulating HEVC-SCC standard [5]. After a two-year proposal collection, experimental verification, performance evaluation, and other work, the video

coding standard for screen content was developed and completed in February 2016 [6]. Nevertheless, the distortion will still appear in a loss of packets for HEVC-SCC due to the bad network transmission environment, which will undermine the quality of video and affect the user experience. Therefore, the video fault tolerance mechanism is discussed to improve video quality based on HEVC-SCC in this paper.

In the literature, various error-resilient mechanisms are discussed. Carreira et al. [7] have proposed a two-stage method to improve the error robustness of HEVC by reducing the time error propagation in the case of frame loss. In the coding stage, the number of prediction units that depend on a single reference is reduced to distribute the use of reference pictures; in the streaming stage, the MV prediction mismatch at the decoder is reduced by the priority sorting algorithm of motion vectors. In another scheme [8], frame loss simulation and corresponding error concealment are proposed to find the most effective method where the decoder is used to recover lost frames. Ferre et al. [9] have proposed a redundant coding method with macroblock rate and potential distortion information. The proposed algorithm is based on the given redundancy rate constraint and can achieve error resistant wireless video transmission without relying on retransmission. Xu [10] put forward a method of macroblock classification based on texture characteristics, motion characteristics, and other information, and then redundantly encoded important macroblocks. A joint rate-distortion optimization algorithm for primary and redundant images is proposed and this method [11] traverses all coding modes and coding parameter combinations of redundant macroblocks and selects the best combination of rate-distortion as the final coding mode. In [12], two macroblock coding modes are added to enhance the transmission robustness of the coded bit stream. Tillo [13] utilized the main frame and redundant frame to generate different descriptions of MDC and analyzed the relationship between redundant frame and quantization coefficient of main frame. Because the source distortion is not related to the pixel value, and the total end-to-end distortion can be represented by the sum of single lost distortion, and the overall distortion optimization process is simplified as a separate optimization process for each frame.

Motivated by the work mentioned in [8–13], a video error-resilient scheme is proposed for error-resilient transmission of screen content coding (SCC) in this paper. In the proposed scheme, additional code stream or redundant information is considered to improve error-resilient performance with redundant coding and acceptable video quality is obtained in the case of frame transmission error. In addition, selective allocation redundancy is conducted to take the importance of the video frame ROI area into account in the co-encoding process. Consequently, the redundancy insertion efficiency and reliability are improved in return. The corresponding experiments validate the effectiveness of the schemes proposed in this paper. The main contribution of this paper is concluded as follows.

1. This paper provides an adaptive error prevention coding mechanism for frame-level-based video transmission security, where the total end-to-end rate-distortion cost can be optimized by adjusting the number of coded redundant frames.
2. To further optimize the bit rate and redundancy allocation, the abrupt frames are distinguished according to the characteristics of SCC, and the region of interest is considered as a redundant allocation.

The rest of this paper is organized as follows. Section 2 discusses the related work. Section 3 provides the main contribution of this paper. Simulation experiments and result analyses are accomplished in Section 4. Finally, Section 5 concludes this paper.

2. Related Work

In the video error recovery technology, fault-tolerant coding belongs to the forward technology. The forward technology refers to introducing redundant error correction bits to the transmission stream at the encoding end. Under certain conditions, the transmission error can be automatically corrected in the decoding process, and the BER (bit error rate) of the received stream can be reduced. According to

the standard coding scheme, fault-tolerant video code stream is generated to solve the error problem in the video transmission process. Through the change of coding mode, the impact of error on the decoded video quality is suppressed as much as possible [14]. Error concealment, as a post-processing technology, repairs and conceals the data in lossy transmission according to the normal received stream information, so that the decoded video can be as close as possible to the original encoded image [15]. In the feedback-based coding transmission framework, when the receiver finds data loss, it can choose the coding unit correctly received in the reference image sequence to realize error recovery. For the coding transmission system without feedback, the redundant slice coding method can add one or several redundant representations of a coding unit in the code stream [16]. As the decoder decodes the received stream, the ordinary information is first utilized. If the received stream is lost, the added redundant slice is considered to decode and reconstruct.

When the transmission error occurs in the main frame, the redundant frame is decoded so that the decoder can obtain acceptable video quality compared with the original video. In the process of video transmission, redundancy coding can be realized in video source coding and channel coding, respectively. The additional redundant content can be either the redundant image encoded for the whole image, or the redundant slice encoded for a specific region or some macroblocks [17].

In the coding process, redundant frames often use different coding parameters from the main frame to improve the rate-distortion performance. Zhu et al. [18] propose a redundant image coding method based on reference frame selection with motion vector and error diffusion distortion to determine whether the current frame encodes redundant frames, where reference frame selection is just utilized in redundant frames. The probability of using redundant frames is only 1-p, where p denotes the probability of independent packet losses for image frames. As a result, the loss reduction with reference frame selection mechanism in redundant frames may not offset the increase of rate-distortion cost caused by the increase of code rate.

Redundant slice embedding is equivalent to the repeated encoding of pictures, which consumes more encoding bits [19]. Therefore, it is preferable to add redundancy to important areas rather than to protect all areas. For error-resilient coding with redundant frames, redundant frames are encoded to improve the robustness of I-frames, so as to further enhance the image transmission quality of video [9]. The method based on adjusting the coding parameters of the redundant frame and main frame does not consider the influence of different scene features. The adaptive redundant frame coding algorithm only uses a fixed reference frame selection mode for redundant frames and does not use other error-resilient methods for main frames, which makes the generation mechanisms of redundant images (redundant slices) simple and limits its adaptability to different network conditions and video scenes.

3. Proposed Method

Traditional error-resilient coding algorithms are based on a single error suppression mechanism and do not adaptively consider different error-resilient coding methods according to the application environment. For this reason, this section proposes an algorithm based on a frame-level error-resilient coding selection mechanism for the transmission security of screen content sequences. If the complexity of the current frame is greater than three times the average encoded frames, the current frame defined as a mutation (abrupt) frame. Due to the lack of correlation with the encoded frames, the mutation frame is often encoded by intra prediction, which consumes a lot of bits. By judging the mutation frame, the abrupt frame is regarded as a more important frame and more redundancy is assigned to it to achieve adaptive frame-level video error-resilient coding algorithm.

Note that the total end-to-end rate-distortion is optimized by virtue of number adjustment of coded redundant frames in this paper. In consequence, error prevention coding varies adaptively in the process of video transmission. In addition, the abrupt frames are determined based on SCC characteristics and the region of interest is considered as a redundant allocation for the proposed scheme. The flowchart of the proposed scheme is described in Figure 1. According to Figure 1, the detailed operations are concluded as follows.

```mermaid
flowchart
```
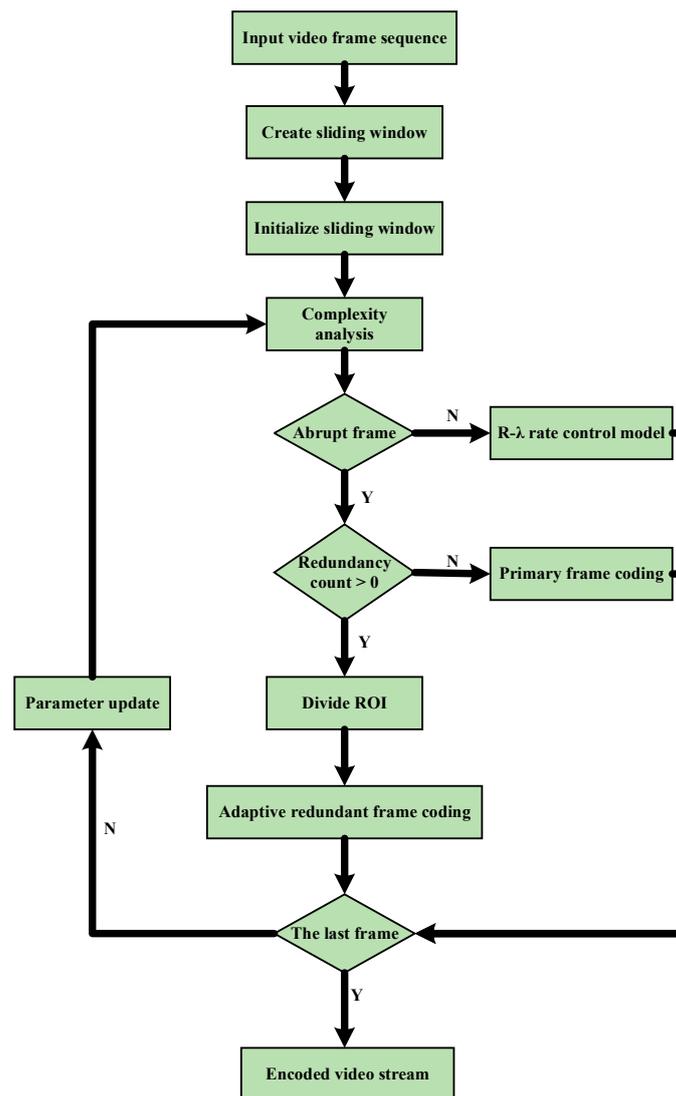
Figure 1. The flowchart of the proposed scheme.

1. The proposed algorithm first creates and initializes a sliding window with size 5.

2. Then the complexity of the current frame is analyzed and compared with the threshold to determine whether the current frame is an abrupt frame.

3. If the current frame is not an abrupt frame, the original model of HEVC is used to control the bit rate.

4. If the current frame is an abrupt frame, the algorithm will judge whether the redundancy (frame) count in the slice header is greater than 0. Note that the information of the redundant frame count is in the slice header in the process of error-resilient encoding. A slice contains part or all of the data of a frame image [17]. In other words, a frame of a video image can be encoded into one or more slices. Each slice is composed of two parts, where one part is the slice header and is used to save the overall information of the slice. During the encoding process, the information of the redundant frame count is included in the slice header.

5. When a slice with a redundancy count is greater than 0, the slice is identified as a redundant slice, otherwise, it will be identified as a primary slice. In the encoding process, the encoding of the current frame is divided into two phases: primary frame encoding and redundant frame encoding.

6. The primary slice will be encoded by the primary frame encoding. The ROI of the redundant slice will be divided, then the redundant slice will be encoded with the adaptive frame-level video error-resilient coding algorithm based on ROI.

7. If the current frame is not the last frame, repeat step 2 to step 6, otherwise, the algorithm will terminate.

3.1. Region Division

The HEVC-SCC standard is developed on the basis of the HEVC coding framework. The reliability can be ensured by judging the abrupt frames in the sequence of screen content coding and implementing the redundant coding. The judgment is made by utilizing the proportion of the complexity of the current frame in the entire sliding window. The complexity of the current frame refers to the SAD (sum of absolute differences) value of the current frame and the reference frame. $\frac{3*C_{pre}}{N_{pre}}$ represents three times the average SAD value for all previous coding frames, which is used as the judgment threshold to determine whether the current frame is a mutation frame. C_{pre} represents the sum of the complexity of all encoded frames before the current frame, and N_{pre} denotes the number of currently encoded frames. Note that $\frac{3*C_{pre}}{N_{pre}}$ is set due to numerical experience.

Since the mean value can better reflect the complexity of the already encoded frame, and test, the decision threshold can accurately identify the abrupt frame. When the complexity of the current frame is greater than three times the average value of all previously encoded frames C_{pre}, the current frame is an abrupt frame, and vice versa. If the current frame is an abrupt frame, the current frame and subsequent frames (four frames) are subjected to redundant coding; otherwise, the code is controlled according to the original code rate control model.

Figure 2 exhibits the flowchart of region division. Note that region division aims at determining whether the current frame is the abrupt frame. According to the decision results, different actions will be taken just as mentioned above.
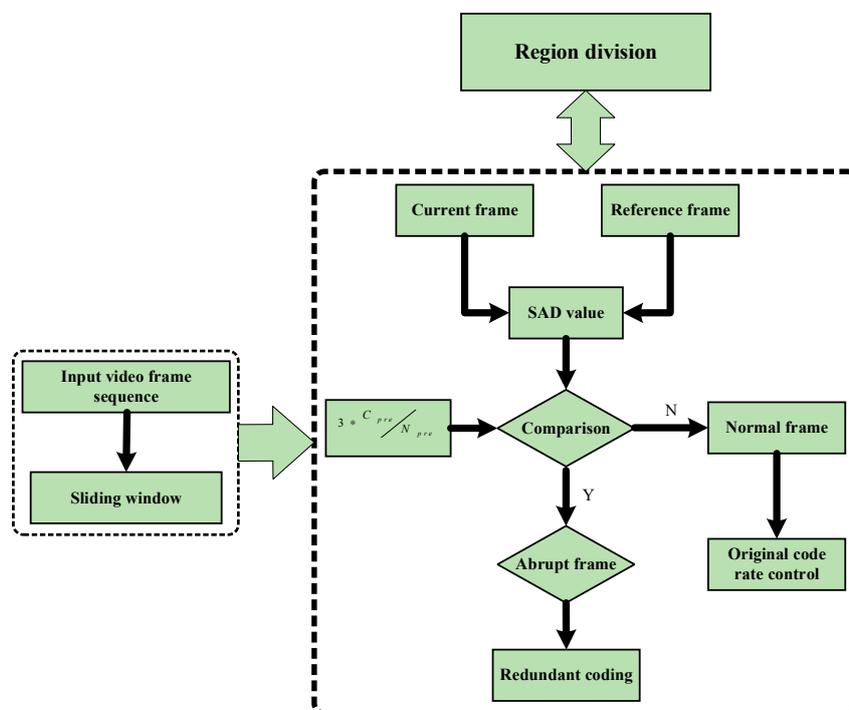
Figure 2. The flowchart of region division.

3.2. Error-Resilient Coding

To improve the robustness of video transmission, different error-resilient mechanisms are considered. The error-resilient code stream can be improved by providing additional code streams or information. During the encoding process, the information of the redundant frame count is included in the slice header. When a slice with a redundancy count is greater than 0, the slice is identified as a redundant slice, and a slice with a redundant frame count equal to 0 is identified as a primary slice. In the encoding process, the encoding of the current frame is divided into two phases: primary frame encoding and redundant frame encoding.

By analyzing the code rate and distortion characteristics of the current frame, an adaptive redundant coding algorithm is implemented to optimize the reference end-to-end rate-distortion cost. At the same time, the redundant coding method is adaptively selected based on the ROI region, thus completing the security-based research on video error prevention mechanisms.

The reconstructed value of the redundant pixels is made equal to the reconstructed value of the corresponding pixel in the original picture, which can reduce the error concealment bias. The embedding of redundancy can effectively enhance the error-resilient ability of the bit stream, but it also brings an increase in bit consumption. Therefore, it is necessary to strike a balance between coding efficiency and error prevention capability.

Different from classical error-resilient coding algorithms, such as [20–23], the proposed scheme in this paper provides an adaptive error prevention coding mechanism for frame level-based video transmission security. The total end-to-end rate-distortion cost can be optimized by adjusting the number of coded redundant frames. In order to further optimize the bit rate and redundancy allocation, the abrupt frames can be distinguished according to the characteristics of the screen content coding, and the region of interest can be considered as a redundant allocation. On the premise of secure transmission, it can better adapt to different network conditions and scenarios and ensure the rated performance.

In the process of image processing, the region of interest (ROI) is an image region selected from the image. This region is the focus of image analysis and is circled for further processing. Circling the target with ROI can reduce the processing time and increase the precision. In addition, a video error-resilient scheme is proposed for error-resilient transmission of screen content coding (SCC) in this paper. More specifically, the primary slice will be encoded by the primary frame encoding in the proposed scheme. The ROI of the redundant slice will be divided, then the redundant slice will be encoded with the adaptive frame-level video error-resilient coding algorithm based on ROI. Taking the importance of the video frame ROI area into account in the co-encoding process contributes to realizing the security-based research on video error prevention mechanisms. Consequently, the redundancy insertion efficiency and the reliability are improved.

3.3. Analysis of Distortion

In this subsection, the distortion situation is discussed to evaluate the performance of the error-resilience coding scheme proposed in this paper. The lower the distortion, the higher the error-resilience capacity.

In the video transmission process, the overall end-to-end distortion D_{total} mainly comes from three parts, namely coding distortion, error concealment distortion, and potential error diffusion distortion from the reference block. D_{total} could be described as:

$$D_{\text{total}} = E\left\{\left(f_i - \tilde{f}_l\right)^2\right\} = (1-p)\cdot D_Q + (1-p)\cdot D_r + p\cdot(D_{ec} + D_{ec\_r}) \tag{1}$$

where D_Q denotes the quantization distortion; D_r represents the distortion of potential error diffusion caused by the reference frame; D_{ec} denotes the distortion between the content of the error concealed video macroblock and the correct video macroblock; D_{ec} depends on the current coding parameter and

error concealment; D_{ec_r} represents the potential distortion caused by hiding the wrong macroblock. In addition, f_i signifies the original video frame, $\tilde{f}_i$ denotes the video frame with redundancy encoding, $E\{\cdot\}$ denotes the expectation operator and p denotes the probability of independent packet losses for image frames. Note that errors may occur due to different color ranges in the process of image color conversion, and the error diffusion algorithm [24] reduces the visual errors caused by transmitting the errors to the surrounding pixels.

The distortion D_{pri} when the redundant frame is not currently embedded can be written as:

$$D_{pri} = (1-p) \cdot D_r + p \cdot (D_{ec} + D_{ec\_r}) \tag{2}$$

When the redundant frame is embedded, the distortion could be formulated as [21]:

$$D_{pri_{red}} = (1-p) \cdot D_r + p \cdot (1-p) \cdot \left(D_{dif} + D_{red\_r}\right) + p \cdot (D_{ec} + D_{ec\_r}) \tag{3}$$

where D_{dif} denotes the distortion caused by the difference between the main frame and the redundant frame, and D_{red_r} signifies the distortion generated by the potential error diffusion of the redundant frame. If packet loss happens for the current frame with probability p, the redundant frame will be used. As a result, the redundancy of the current frame is not lost. Then the remaining frames are normal with probability $p \cdot (1-p)$.

The proposed redundant coding algorithm assumes that when the slice of the primary frame is lost, the decoding operation of the redundant frame on the decoding side replaces the content of the lost frame. As a result, the use of redundant frames and redundant slices reduces error propagation and improves video reliability.

Rate-distortion performance can be seen from the change rate of distortion (for a given source distribution and distortion metric, the minimum expected distortion that can be achieved at a specific bit rate) and the change of code rate:

$$\Delta D = p \cdot (1-p) \cdot \left[\sum_k D_{dif}(k) - \sum_k D_{ec}(k) \right] \tag{4}$$

$$\Delta R = R_{cur} - R_{pri} = R_{red} \tag{5}$$

where R_{cur} denotes the bit consumption common to the current primary frame and the redundant frame in the case of embedded redundancy, R_{pri} represents the bit consumption of the current frame. ΔR, the amount of change in the code rate is also the code rate of the redundant frame. k signifies the order number of the pixels belonging to the rectangular ROI part. When a redundant frame is used, the probability that the redundant frame and the main frame are simultaneously lost is p^2, it is obvious that the change rate of the distortion change and the code rate are both related to the packet loss rate p. When the packet loss rate p is small, a single redundant frame can be used to implement error-resilience. When the packet loss rate p is large, different redundant frame allocation strategies can be adopted according to the type of frame. For key frames in the screen content encoding, more redundant frames are allocated for encoding in order to adapt to different characteristics of the SCC and other scenarios, thereby achieving reliability improvement under secure transmission.

Then we evaluate the rate-distortion (R-D) performance with a redundant slice, which is the ratio of the changing distortion and the changing bit consumption (the distortion change per unit bit consumption) after embedding a redundant slice:

$$\delta = \frac{\Delta D}{\Delta R} \tag{6}$$

After obtaining the Lagrangian coefficient λ from the code rate, the quantization parameter is obtained based on the relationship between QP and λ.

$$\lambda = \alpha \cdot R^{\beta} = -\frac{\partial D}{\partial R} \tag{7}$$

$$\lambda = (1-p) \cdot p^{m} \cdot \frac{\partial D_{dif}}{\partial R_{red}} \tag{8}$$

$$QP = 4.2005 \ln \lambda + 13.7122 \tag{9}$$

In general, natural images can be divided into the following categories:

(1)　A completely static scene, that is, the image content of the current frame and the previous frame are exactly the same.

(2)　A continuously changing scene, that is, there is continuity between adjacent frames in the video. The moving object or the angle of view of the shot has moved.

However, for the sequence of screen contents, in addition to the above two types of images, there are images with scene mutations appearing. In the proposed solution, several cases based on the coding of the screen content are considered. Most of the intra prediction in a picture is likely to mean scene conversion. If the motion vector of all macroblocks is large, this indicates that the features in the picture may be more complicated; and if the ROI area in a picture is large, it means that a larger protected area is needed. After the current picture is encoded, the encoding parameters are updated based on the $R - \lambda$ model [25].

3.4. R-λ Model

In this section, the update of the encoding parameters is discussed based on the $R - \lambda$ model [25]. Figure 3 exhibits the $R - \lambda$ rate control model. Generally, the $R - \lambda$ model contains two steps, bit allocation and coding parameter updating.
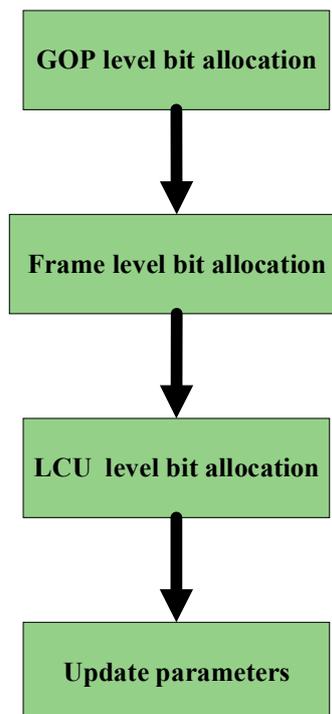
Figure 3. The flowchart of the $R - \lambda$ model.

The average bits of per frame is defined as:

$$R_{picAvg} = \frac{R_{tar}}{FR} \tag{10}$$

where R_{tar} is the target bit, and FR is the frame rate.

3.4.1. Bit Allocation

(1) Group of Picture (GOP) level bit allocation
The number of encoded frames is N_{codee}, and the number of bits used to encode these frames is R_{coded}. The number of frames in a GOP is N_{GOP}; and SW is the size of the sliding window for smoothing bit allocation, which is used to make bit consumption change and the quality of encoded pictures gentler. The bit allocation of GOP level is:

$$T_{GOP} = T_{AVgpic} * N_{GOP} \tag{11}$$

It is emphasized that the target code rate is achieved after SW. frames. If each frame in the SW. frames can exactly consume T_{AVgpic} bits, then the above equation can be written as follows:

$$T_{AVgpic} = T_{AVgpic} + \frac{R_{picAvg} * N_{coded} - R_{coded}}{SW} \tag{12}$$

(2) Frame-level bit allocation
Define Coded$_{GOP}$ as the consumed bits of the current *GOP*, and ω is the bit allocation weight of each frame, then the bits assigned to the current frame is defined as:

$$T_{Pic} = \frac{T_{GOP} * \text{Coded}_{GOP}}{\sum NotCodedPictures\omega_i} * \omega_{CurrPic} \tag{13}$$

(3) Largest Coding Unit (LCU) level bit allocation
Similar to the frame-level bit allocation strategy, the bits allocated to each LCU are calculated based on:

$$T_{LCU} = \frac{T_{CurrPic} - Bit_{header} - Coded_{pic}}{\sum NotCodedLCUs\omega_i} \tag{14}$$

where Bit_{header} is the estimated number of all header information bits, which is estimated from the actual number of header information bits of the previously encoded pictures. ω is the bit allocation weight of each LCU.

3.4.2. Coding Parameters Update

According to the relationship of λ and R [25]:

$$D(R) = CR^{-K} \tag{15}$$

where R denotes the bit rate; D is the MSE (mean squared error) distortion; and are model parameters related to the characteristics of the video source [25]; λ denotes the slope of the R-D curve, and α, β signifies the parameters related to the video content. That is, λ can be directly calculated by R through α and β However, α and β are parameters related to the content characteristics of the video, and the values of different contents are significantly different. The R–λ model solves this problem by introducing the following algorithm. Firstly, the following formula is used to calculate the λ of the frame and LCU [25]:

$$\lambda = \alpha bpp^{\beta} \tag{16}$$

where *bpp*. bit rate, α and β are different for each frame and each LCU. In order to realize content adaptation, α and β will be updated continuously in the encoding process. After completing the encoding of an LCU or a frame, α and β are updated as follows:

$$\lambda_{comp} = \alpha_{old} bpp_{real}^{\beta_{old}} \tag{17}$$

$$\alpha_{new} = \alpha_{old} + \delta_\alpha.(\ln\lambda_{real} - \ln\lambda_{comp}).\alpha_{old} \tag{18}$$

$$\beta_{new} = \beta_{old} + \delta_\beta.(\ln\lambda_{real} - \ln\lambda_{comp}).\ln bpp_{real} \tag{19}$$

where is the actual bpp_{real} is the actual *bpp* in the coding process; λ_{real} is the actual value of λ in the coding process; $\delta_\alpha = 0.1$ and $_\beta = 0.05$ [25].

Note that *bpp* is defined as:

$$bpp = \frac{R}{f.w.h} \tag{20}$$

where f is the frame rate, w and h are the width and height of the picture, respectively [25].

When λ is determined, *QP* is calculated according to [26].

$$QP = C_1 * \ln(\lambda) + C_2 \tag{21}$$

where $C_1 = 4.2005$ and $C_2 = 13.7122$.

4. Experiment Results

4.1. Experimental Conditions

HM-16.7-SCM7.0 is an open-source encoder maintained by the HEVC standard development organization JCT-VC [27], which integrates the new coding tools in the SCC standard and greatly improves the coding efficiency of screen content sequences. The video sequences used are all selected from the test sequences recommended in the SCC official general test conditions. The frame structure of the encoder uses a low-latency profile to simulate real-time transmission. The resolution, frame rate, and target bit rate of the experimental sequence are shown in Table 1 [27]. Note that some sequences in Table 1 have the same information and bit rate, but their contents are entirely different. As a result, their names are in different forms. In addition, packet loss ratio refers to the ratio of the number of packets lost to the data groups sent during the test.

Table 1. Experimental video sequence information and corresponding target bit rate.

Sequence	Resolution	Total Frames	Color Mode	Frame Rate (Hz)	Target Bitrate (Kbps)	
					Low Bit Rate	High Bit Rate
Map	1280 × 720	600	4:4:4	60	1000	2000
Programming	1280 × 720	600	4:4:4	60	1000	2000
Robot	1280 × 720	300	4:4:4	30	512	1000
SlideShow	1280 × 720	500	4:4:4	20	128	256
Web_browsing	1280 × 720	300	4:4:4	30	128	256
Wordedit	1280 × 720	600	4:4:4	60	1000	2000
Console	1280 × 1080	600	4:4:4	60	1000	2000
Desktop	1280 × 1080	600	4:4:4	60	1000	2000

Note that the packet loss ratio refers to the ratio of the number of packets lost to the data groups sent during the test. The packet loss in the channel can be simulated by the packet loss simulator. The packet loss simulation software can be obtained from [28]. Packet loss simulation software can set

packet loss rate (packet loss probability) to simulate the packet loss when the video is transmitting in the channel.

The PSNR [29] can be evaluated based on:

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{23}$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2 \tag{24}$$

where MAX_I is the maximum possible pixel value of the image. Because each pixel of an image is represented by 8 bits, the value of MAX_I is 255. $I(i,j)$ represents the pixel value of the original image at the point (i,j). $K(i,j)$ represents the pixel value at point of the reconstructed image decoded and reconstructed by the decoder after the code stream is processed by the packet loss simulation software.

PSNR is the peak signal-to-noise ratio, usually after image compression, the output image will be somewhat different from the original image. A smaller PSNR value indicates that the distortion caused by the embedded information is small. Taking PSNR as the main objective quality evaluation standard, the higher the PSNR value is, the better the video quality gets. That is, PSNR is positively correlated with the algorithm performance. As a result, PSNR is considered as the evaluation index of the proposed scheme.

4.2. Experimental Results

Figure 4a,b, respectively exhibits the PSNR value distribution with the change of the web browsing and sideshow screen content sequences at the 256kbps bit rate. According to the experiment results, the proposed scheme is able to achieve a higher overall PSNR score, which illustrates the better video encoding effect. By judging the mutation frame, the abrupt frame is regarded as a more important frame and more redundancy is assigned to it for adaptive frame-level video error-resilient coding. Additional code stream or redundant information is considered to improve error-resilient performance with redundant coding and acceptable video quality is obtained in the case of frame transmission error as a consequence.
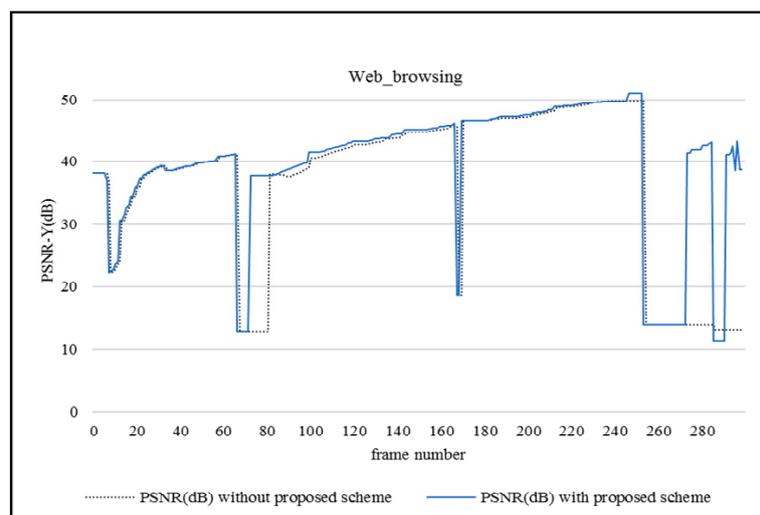
(**a**) Sequence of web browsing.

Figure 4. *Cont.*
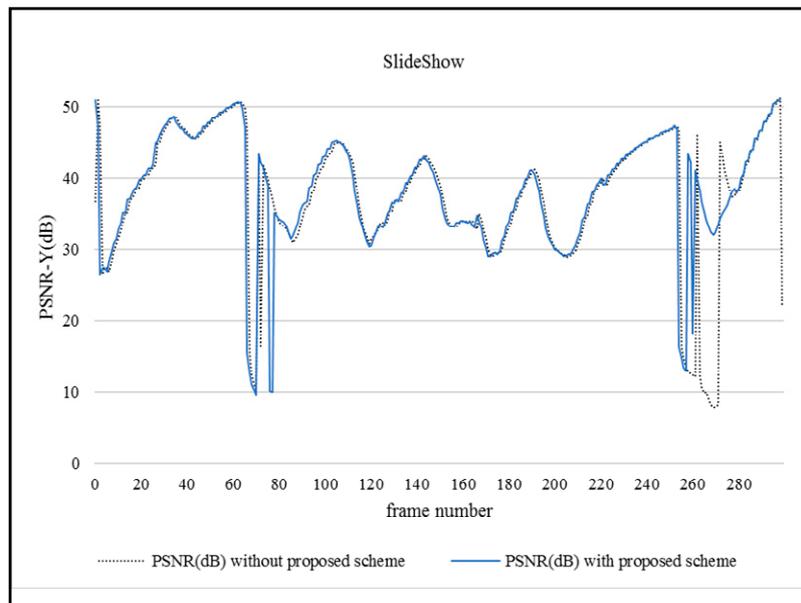
(**b**) Sequence of slideshow.

Figure 4. PSNR-Y comparison without and with the proposed scheme.

Figure 5 gives the average PSNR comparison without and with the proposed scheme. From Figure 5, the average PSNR score with the proposed scheme is higher than that without the proposed scheme [22]. These results accord with the analysis in Section 3, which validates the effectiveness of the proposed scheme in this paper.
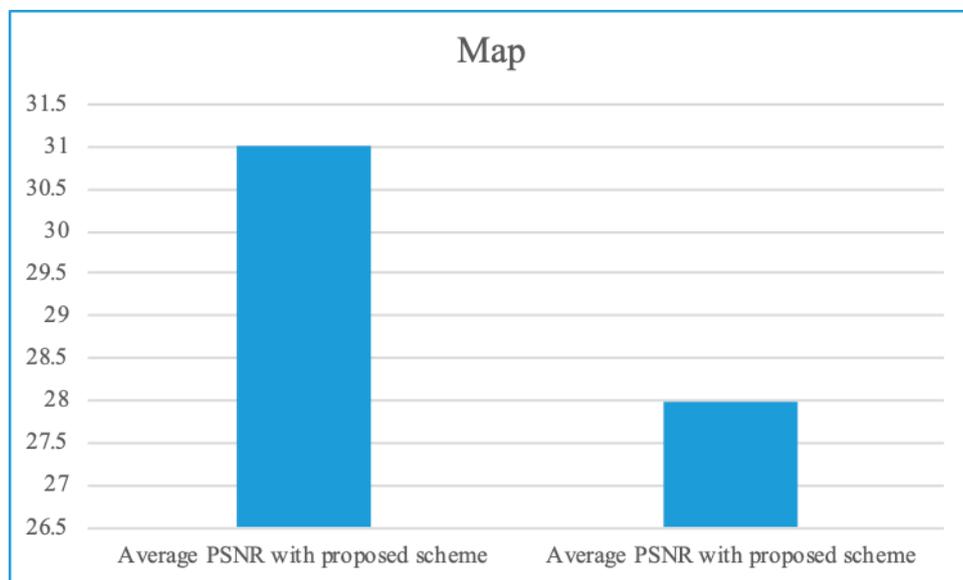
(**a**) Sequence of map.

Figure 5. *Cont.*
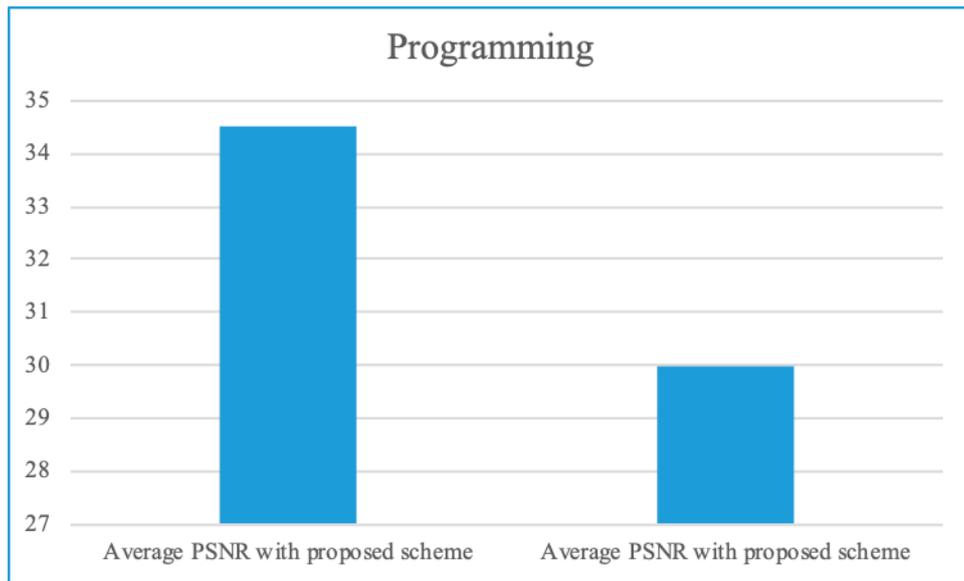
(**b**) Sequence of programming.

Figure 5. Average PSNR comparison without and with the proposed scheme.

To further verify the performance of the proposed scheme, the average PSNR is compared with three classical video error-resilience coding schemes [11,13,15], as it is shown in Figure 6. Obviously, the average PSNR score of the proposed scheme is the highest, which indicates the proposed adaptive frame-level video error prevention coding algorithm can improve the transmission quality of video on an error-prone network channel by judging the abrupt frames and allocating more redundancy to them.
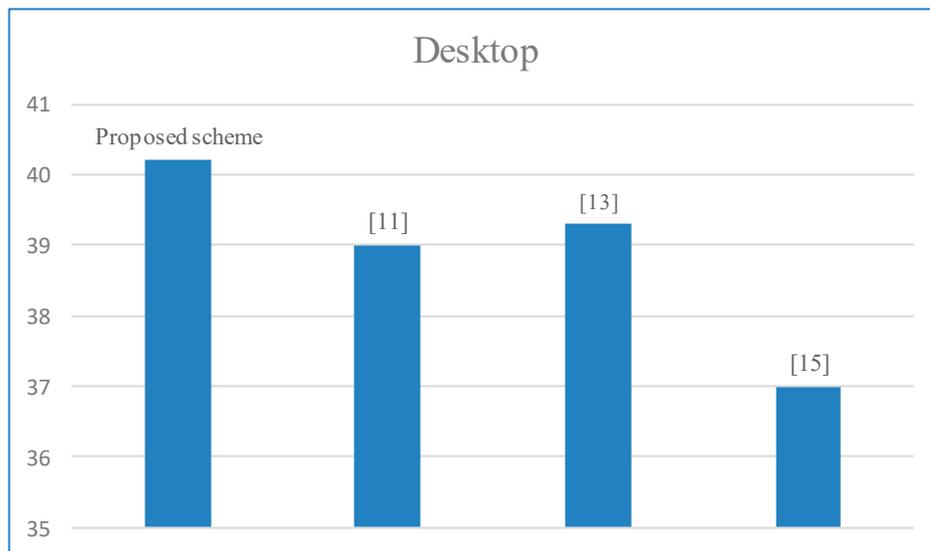
(**a**) Sequence of desktop.

Figure 6. *Cont.*
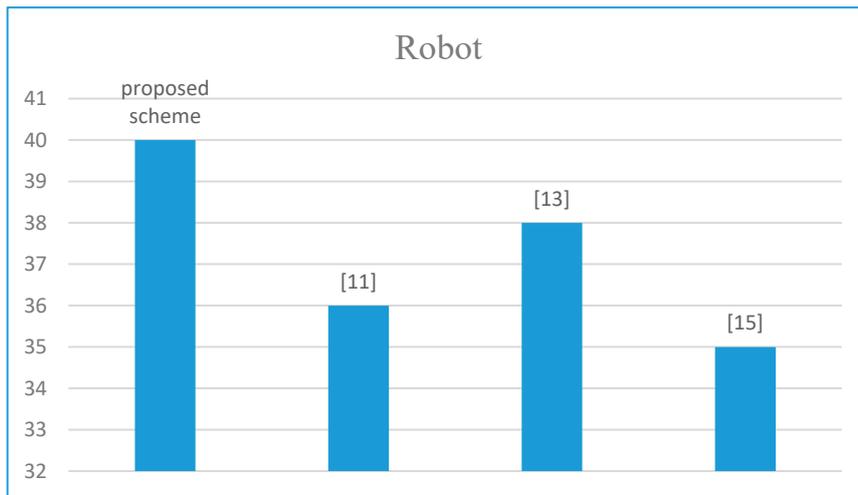
(**b**) Sequence of robot.
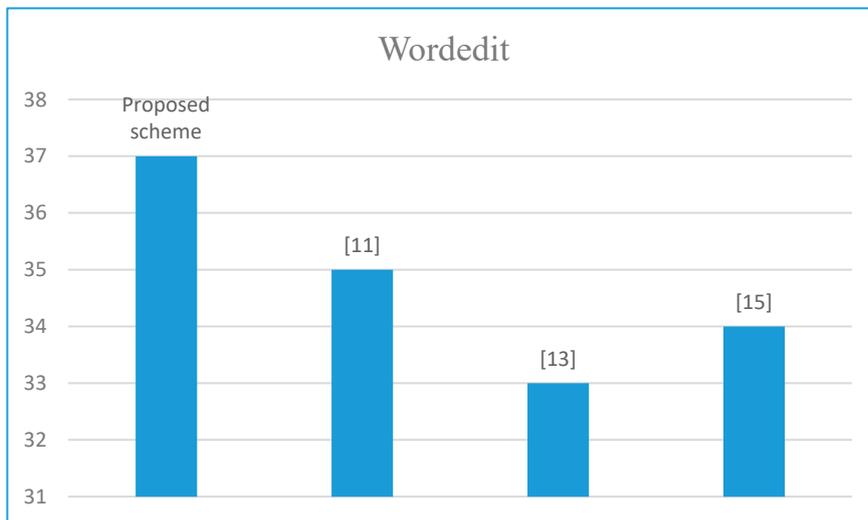
(**c**) Sequence of wordedit.
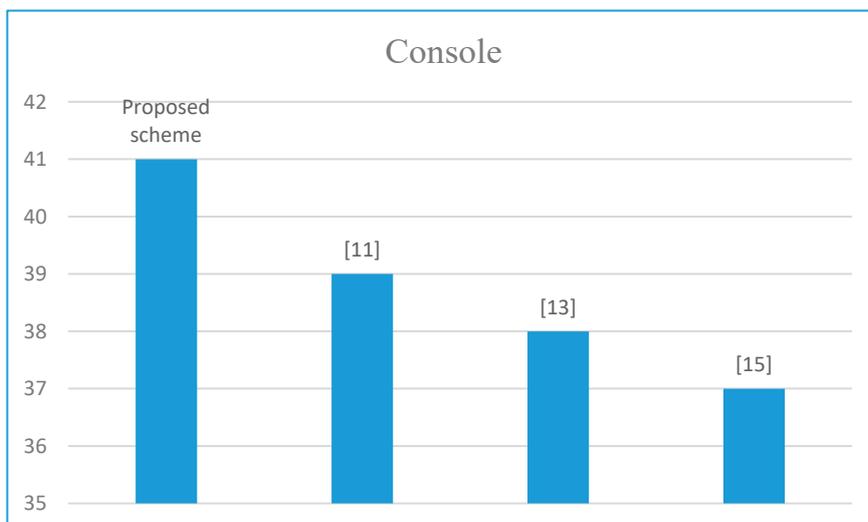
(**d**) Sequence of console.

Figure 6. Average PSNR comparisons among multiple schemes.

According to the simulations in Figures 4–6, the proposed algorithm improves bit rate error and encoding efficiency, which is very helpful for video encoding transmission in multimedia applications, especially in the case of poor network conditions like the situation of packet loss. At the same time, since the proposed method does not perform redundant coding for each frame, the bit rate is small. During redundant embedding, more redundant slices are embedded in the abrupt frame and the region of interest is considered to achieve a better error-resilience and ensure the performance of video R-D performance.

5. Conclusions

In this paper, an error-resilience video coding scheme is proposed based on an adaptive redundancy allocation strategy for quality improvement of video transmission on error-prone network channels. By adjusting the number of coded redundant frames, the overall end-to-end rate-distortion cost is optimized. In addition, to obtain further optimization of the bit rate and redundancy allocation, the abrupt frame can be distinguished according to the characteristics of the screen content encoding while the region of interest is considered for redundant allocation. On the premise of error-resilient transmission, the proposed scheme achieves better adaptability to different network situations and scenarios and guarantees rate-distortion performance. In the future work, more optimization factors will be considered for a higher video transmission quality.

References

1. Sjoberg, R.; Chen, Y.; Fujibayashi, A.; Hannuksela, M.M.; Samuelsson, J.; Tan, T.K.; Wang, Y.K.; Wenger, S. Overview of HEVC High-Level Syntax and Reference Picture Management. *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *22*, 1858–1870. [CrossRef]
2. Ni, Z.; Ma, L.; Zeng, H.; Chen, J.; Cai, C.; Ma, K.K. ESIM: Edge Similarity for Screen Content Image Quality Assessment. *IEEE Trans. Image Process.* **2017**, *26*, 4818–4831. [CrossRef] [PubMed]
3. Zhu, W.; Ding, W.; Xu, J.; Shi, Y.; Yin, B. Screen Content Coding Based on HEVC Framework. *IEEE Trans. Multimed.* **2014**, *16*, 1316–1326. [CrossRef]
4. Pu, W.; Karczewicz, M.; Joshi, R.; Seregin, V.; Zou, F.; Sole, J.; Sun, Y.C.; Chuang, T.D.; Lai, P.; Liu, S. Palette Mode Coding in HEVC Screen Content Coding Extension. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2016**, *6*, 420–432. [CrossRef]
5. Pourazad, M.T. HEVC: The New Gold Standard for Video Compression: How Does HEVC Compare with H.264/AVC? *IEEE Consum. Electron. Mag.* **2012**, *1*, 36–46. [CrossRef]
6. Jiang, L. "Subjective-Driven Complexity Control Approach for HEVC" *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *26*, 91–106.
7. Carreira, J.F.; Assunção, P.A.; de Faria, S.M.; Ekmekcioglu, E.; Kondoz, A. A Two-stage Approach for Robust HEVC Coding and Streaming. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *28*, 1960–1973. [CrossRef]
8. Zhou, Y.; Zhang, Q.; Ning, S. A Dynamic-Mode Redundant Coding for Error-Resilient Video Transcoding. In Proceedings of the 2013 International Conference on Information Technology and Applications, Chengdu, China, 16–17 November 2013; pp. 12–125.
9. Ferré, P.; Agrafiotis, D.; Bull, D. A video error resilience redundant slices algorithm and its performance relative to other fixed redundancy schemes. *Signal Process. Image Commun.* **2010**, *25*, 163–178. [CrossRef]
10. Xu, J.; Wu, Z. A perceptual Sensitivity Based Redundant Slices Coding Scheme for Error-Resilient Transmission H.264/AVC Video. In Proceedings of the International Conference on Communications, Guilin, China, 25–28 June 2006.

11. Carreira, J.; Assuncao, P.; Faria, S.; Ekmekcioglu, E.; Kondoz, A. A robust video encoding scheme to enhance error concealment of intra frames. In Proceedings of the 2017 IEEE International Symposium on Circuits and Systems (ISCAS), Baltimore, MD, USA, 28–31 May 2017; pp. 1–4.

12. Schmidt, J.C.; Rose, K. Jointly optimized mode decisions in redundant video streaming. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 797–800.

13. Xiao, J.; Tillo, T.; Lin, C.; Zhao, Y. Error resilient video coding with end-to-end ratedistortion optimized at macroblock level. *Eurasip J. Adv. Signal Process.* **2011**, *2011*, 80. [CrossRef]

14. Wang, Y.; Wenger, S. Error resilient video coding techniques. *IEEE Signal Process. Mag.* **2000**, *17*, 61–82. [CrossRef]

15. Al-Jobouri, L.; Fleury, M.; Ghanbari, M. Error Resilient IPTV for an IEEE 802.16e Channel. *Wirel. Eng. Technol.* **2011**, *2*, 70–79. [CrossRef]

16. Milani, S.; Calvagno, G. "Multiple Description Distributed Video Coding Using Redundant Slices and Lossy Syndromes". *IEEE Signal Process Lett.* **2010**, *17*, 51–54.

17. Peng, Q.; Zhang, L.; Chen, J.X. Overview of error concealment for video transmission. *J. Southwest Jiaotong Univ.* **2009**, *44*, 473–483.

18. Zhu, C.; Wang, Y.K.; Li, H. Adaptive Redundant Picture for Error Resilient Video Coding. In Proceedings of the IEEE International Conference on Image Processing, San Antonio, TX, USA, 16 September–19 October 2007.

19. Wu, Z.; Boyce, J.M. Adaptive Error Resilient Video Coding Based on Redundant Slices of H.264/AVC. In Proceedings of the IEEE International Conference on Multimedia & Expo, Beijing, China, 2–5 July 2007.

20. Majid, M.; Owais, M.; Anwar, S.M. Visual saliency based redundancy allocation in HEVC compatible multiple description video coding. *Multimed. Tools Appl.* **2018**, *77*, 20955–20977. [CrossRef]

21. Piñol, Pablo, Martinez-Rach M, Garrido P, et al. Error Resilient Coding Techniques for Video Delivery over Vehicular Networks. *Sensors* **2018**, *18*, 3495. [CrossRef] [PubMed]

22. Zheng, X.; Zarcone, R.; Paiton, D.; Sohn, J.; Wan, W.; Olshausen, B.; Wong, H.S.P. Error-Resilient Analog Image Storage and Compression with Analog-Valued RRAM Arrays: An Adaptive Joint Source-Channel Coding Approach. In Proceedings of the IEEE International Electron Devices Meeting, San Francisco, CA, USA, 1–5 December 2018.

23. Yang, G.; Jiao, S.; Liu, J.P.; Lei, T.; Yuan, X. Error diffusion method with optimized weighting coefficients for binary hologram generation. *Appl. Opt.* **2019**, *58*, 5547–5554. [CrossRef] [PubMed]

24. Tillo, T.; Grangetto, M.; Olmo, G. Redundant Slice Optimal Allocation for H.264 Multiple Description Coding. *IEEE Trans. Circuits Syst. Video Technol.* **2008**, *18*, 59–70. [CrossRef]

25. Li, B.; Li, H.; Li, L.; Zhang, J. Rate control by R-lambda model for HEVC. In Proceedings of the Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 11th Meeting, Shanghai, China, 10–19 October 2012; 2012; pp. 1–5.

26. Li, B.; Zhang, D.; Li, H.; Xu, J. QP Determination by Lambda Value. In Proceedings of the JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 9th Meeting, Geneva, Switzerland, 27 April–7 May 2012.

27. Sullivan, G.J.; Ohm, J.-R.; Han, W.-J.; Wiegand, T. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *22*, 1649–1668. [CrossRef]

28. Wenger, S. Nal Unit Loss Software.Document JCTVC-H0072, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC). 2012. Available online: http://phenix.int-evry.fr/jct/doc_end_user/current_document.php?id=4373 (accessed on 25 April 2020).

29. Liao, K.; Lin, C.; Zhao, Y.; Gabbouj, M. DR-GAN: Automatic Radial Distortion Rectification Using Conditional GAN in Real-Time. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *30*, 725–733. [CrossRef]