GENERAL INSTRUCTION

- **Authors: We cannot accept new source files as corrections for your article. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.**

QUERIES

Q1. Author: Please confirm or add details for any funding or financial support for the research of this article.
Q2. Author: Please check and confirm that author name Mohammad Ghanbari in byline and biography are correct as set.
Q3. Author: Please check and confirm that postal codes of affiliations of authors Mohammad Kazemi and Mohammad Ghanbari are correct as set.
Q4. Author: Please check and confirm that page range in Refs. [3], [9], and [19], are correct as set.
Q5. Author: Please update with further details in Ref. [11], if available.
Q6. Author: Please provide page range in Refs. [22], [29], and [36], if available.
Q7. Author: Please check and confirm that Ref. [46] is correct as set. Also, provide location in the said reference, if available.
Q8. Author: Please provide complete educational details of author Mohammad Ghanbari in biography.

Intra Coding Strategy for Video Error Resiliency: Behavioral Analysis

Mohammad Kazemi, Mohammad Ghanbari, *Life Fellow, IEEE*, and Shervin Shirmohammadi, *Fellow, IEEE*

***Abstract*—One challenge in video transmission is to deal with packet loss. Since the compressed video streams are sensitive to data loss, the error resiliency of the encoded video becomes important. When video data is lost and retransmission is not possible, the missed data should be concealed. But loss concealment causes distortion in the lossy frame which also propagates into the next frames even if their data are received correctly. One promising solution to mitigate this error propagation is intra coding. There are three approaches for intra coding: intra coding of a number of blocks selected randomly or regularly, intra coding of some specific blocks selected by an appropriate cost function, or intra coding of a whole frame. But Intra coding reduces the compression ratio; therefore, there exists a trade-off between bitrate and error resiliency achieved by intra coding. In this paper, we study and show the best strategy for getting the best rate-distortion performance. Considering the error propagation, an objective function is formulated, and with some approximations, this objective function is simplified and solved. The solution demonstrates that periodical I-frame coding is preferred over coding only a number of blocks as intra mode in P-frames. Through examination of various test sequences, it is shown that the best intra frame period depends on the coding bitrate as well as the packet loss rate. We then propose a scheme to estimate this period from curve fitting of the experimental results, and show that our proposed scheme outperforms other methods of intra coding especially for higher loss rates and coding bitrates.**

***Index Terms*—Error resilient video coding, video error concealment, intra coding.**

I. Introduction

NOWADAYS, real-time digital video transmission over networks is very popular. Due to the tremendous volume of the raw video data, video compression is inevitable. But delivering compressed data over wired/wireless channels is challenging since the underlying networks are not always reliable and some data losses are usually experienced during transmission.

The erroneous and unreceived data corrupts the decompression process and the video fidelity. In this condition, using error concealment techniques alleviates the problem to some extent [1]. At high compression rates, data loss is more destructive; the higher the compression, the more the sensitivity to data loss. For this reason, High Efficiency Video Coding (HEVC), the latest standard video codec, is less error resilient than H.264/AVC [2]. Therefore, this makes HEVC video communication over error prone channels a challenging problem for researchers and practitioners in this field.

In error concealment techniques, the non-received data are estimated from the received ones. This is done by exploiting the spatio-temporal correlations among the available data at the area of missing information. However, the replaced data will not be exactly the same as the actual data; therefore, there exists a mismatch/distortion between them. If the recovered frame was used as the prediction reference at the encoder, its reconstructed erroneous part would propagate into the next frames at the decoder. In video coding, a large portion of compression comes from inter frame coding, but inter frame coding increases inter dependency and causes error propagation. In contrast, although intra coding is less efficient for compression, it mitigates the error propagation problem and could be used as a strong error resiliency tool, since it does not use prediction from the other frames [3]–[6]. Therefore, by intra coding, there is a trade-off between error resiliency and compression ratio. That is, intra mode for a block must be selected with sufficient care. For this reason, loss resiliency through intra mode is discussed in several works, as described in Section II.

In this paper we show both analytically and experimentally that the best strategy for intra coding is to code some selected frames deliberately in intra mode. For doing so, considering the transmission distortion, the decoder side distortion is formulated and an objective function is developed. Through optimization of this objective function, the optimal solution indicates that coding a whole frame as an I-frame is preferred over coding a certain number of blocks in the frames as intra mode. Our investigations show that the best approach to exploit the error resiliency of intra coding is to reduce the intra period instead of distributing the intra coded blocks among the frames. We had solved a similar problem for Multiple Description Coding (MDC) in [32], but the treatment of a single stream is different from the multi streams of MDC. MDC rarely deals with concealment distortion, since most of the time at least one description is available. Therefore,

Manuscript received January 28, 2019; revised June 17, 2019, August 15, 2019, October 5, 2019, and November 24, 2019; accepted November 25, 2019. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sanjeev Mehrotra. *(Corresponding author: Mohammad Kazemi.)*
Q1

Q2 M. Kazemi is with the Department of Electrical Engineering, University of Isfahan, Isfahan 81746-73441, Iran (e-mail: m.kazemi@eng.ui.ac.ir).

Q3 M. Ghanbari is with the School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran 1417466191, Iran, and also with the School of Computer Science and Electronic Engineering, University of Essex, CO4 3SQ Colchester, U.K. (e-mail: ghan@ut.ac.ir).

S. Shirmohammadi is with the Distributed and Collaborative Virtual Environments Research Laboratory, School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: shervin@eecs.uottawa.ca).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2019.2957991

the objective functions and approaches taken for solving the problem here are different from [32].

The rest of the paper is structured as follows. Section II presents the related work, while in Section III the objective function is formulated. Through experimental results, this function is first simplified and then solved in Sections IV and V. The performance comparison with the other approaches for error resilient intra coding is presented in Section VI and the paper is concluded in Section VII.

II. Related Work

In the works presented in [7], [8], a recursive algorithm called ROPE has been developed. In this algorithm, by pixel-wise operations, the encoder estimates the receiver-side expected distortion which is then used for intra/inter mode decision. However, this algorithm is too complex due to its pixel-by-pixel computations. Its extension for bursty loss channels is presented in [9], and the extension of error resilient mode decision to motion estimation and also considering intra-frame prediction for intra modes is presented in [10]. ROPE was also used in [11] to optimally decide between Motion Vector (MV) replication or intra coding mode. Using ROPE for motion estimation and reference frame generation is presented in [12], while [13] discusses the intra/inter mode selection in video transcoding, where the lossy frame and its propagated error within the ROPE algorithm is exploited. Finally, the extension of ROPE for including constrained intra prediction as candidate modes, in addition to inter and intra modes, is presented in [14].

In another proposal described in [15], frame-level channel distortion is analyzed, where through linear models/approximations and end-to-end distortion optimization, a scheme for intra mode selection and rate control is developed. Some models for transmission distortion are presented in [16] where parameters such as intra prediction, deblocking filtering, sub-pixel motion estimation and the effect of decoder side temporal error concealment are taken into account. With the same approach, an end-to-end distortion modeling and optimization method was presented in [17] which is then used to develop a faster algorithm for intra/skip mode decision [18], [19]. In [20], motion estimation and mode decision in HEVC are performed based on error propagation. Another algorithm is presented in [21] where due to high sensitivity to error propagation, the algorithm selects the intra mode for the Prediction Units (PUs) much more than is required, especially for lower content videos. Even though the authors try to solve this issue with updating some parameters, the intra rates are still high and this degrades its performance at low Packet Loss Rates (PLRs) and low bitrates.

A fast intra mode decision for loss resiliency is developed in [22] where, through a linear model, the distortion is estimated and an optimal value for Intra Refresh Rate (IRR) is obtained. IRR or simply the intra rate is the number of blocks coded in intra mode divided by the total number of blocks in the frame. A modified model for considering the role of IRR in bitrate and distortion is introduced in [23]. Using a linear model and considering motion activity and PLR, the optimal IRR and the intra coded MBs' pattern are discussed in [24]. It is shown in [25] that for low activity sequences, cyclic intra coding of MBs is more effective than periodic I-frames, and vice-versa for highly active videos. Combining cyclic intra-refreshing with unequal error protection is introduced in [26], [27], though intra-refresh is in conflict with multiple reference selection, as shown in [28]. Error propagation is formulated and the IRR is obtained in [29], then the selected MBs for intra coding are grouped into a common slice group where they are then protected with stronger channel codes.

The above mentioned intra coding research works can be categorized into two groups: those which discuss Selecting Intra Mode (SIM) and those which discuss Intra Refresh. In SIM methods, the cost function for inter/intra mode decision is modified to take into account the lossy channel and the transmission distortion; examples are [7]–[21]. In Intra Refresh, the intra rate is determined. Then, the intra coded blocks can be selected randomly, or selected with vertically or horizontally ordered columns/rows, provided that they do not overlap in the successive frames such that the blocks in all positions are intra refreshed after a while; examples are [22]–[29].

Our work is different form the above works since our formulation and optimization leads to a straightforward and specific solution: reduce the intra period (coding a whole frame in intra mode) to achieve the best error resiliency outcome of intra coding, instead of distributing the intra coded blocks within the frames of the GOP. Afterwards, the best intra period, which depends on the content and channel loss rate, is approximately but simply obtained from a function, without additional computational complexity. The experimental results confirm the efficient performance of the proposed scheme, for various loss rates and video contents.

Another tool which can help to prevent error propagating is Reference Picture Selection (RPS) which allows the encoder to select one or two frames from a list as inter-prediction references for each prediction block. Several reference frames are examined for the best rate-distortion coding. For error resiliency, this feature is usually in conjunction with decoder feedback which informs the encoder not to select the erroneously received frames as the prediction reference [42], [43]. However, this feedback information is not available in many applications; e.g., multicast and broadcast applications, or pre-recorded video on demand applications. Moreover, responding to various receivers concurrently is not practical, or the feedback messages might be received too late. RPS without a back channel and for error resiliency has been presented in [44]. In this work, the authors propose not to use a single frame as prediction reference of the PUs, but to select from a list of reference frames such that all frames in the list are selected uniformly. However, this method needs to consider a list of frames as candidate reference frames, so it has the complexity of the multi-reference prediction. For example, for five candidate reference frames, the computational complexity of Motion Estimation and Mode Decision becomes five times more. The required encoder/decoder buffer size becomes larger with the number of reference frames as well. In the error resiliency of the intra coding method proposed in our paper, the only required information is channel loss rate, without any additional complexity in the encoder/decoder. It is worth noting that full frame intra prediction provides random access to video stream, but it also generates large peaks of bitrates. However, such large spikes in the bitrate can be eased with either a few frames delay, which is acceptable in many applications, or compensated by statistical multiplexing with lower bitrates of P- and B-frames of other video streams. Therefore, there are many circumstances that intra coding is a feasible solution while RPS cannot be employed.

Our application scenario is distribution/broadcast of video without assuming any specific limits on delay or bandwidth, and assuming the PLR is known by the encoder. PLR can be estimated with or without back channel. The application scenario without back channel compromises the great majority of current video distribution/broadcast systems today. In such a scenario, we assume PLR is estimated by the service/network operator from the history of the channel for the specific weekday and time of the day, or is calculated offline, or is tested by small ping packets periodically, or by traffic modeling [30], or by one-way estimation methods that use message segment size, goodput, and delay [31] all estimated at the sender side, or by any other estimation means. Of course PLR can also be estimated with back channel, and this would be more accurate. In the latter scenario, our proposed method is applicable if this back channel cannot inform us of the lost packets immediately due to the delay in feedbacks or a long Round-Trip Time; therefore, retransmission of the lost data or Reference Picture Selection is not possible. This scenario is also assumed by other credible works [8], [13], [15], [17], [21]. Finally, since our method introduces a small delay of about 92 msec, as will be shown later in the paper, we assume that such small delay is acceptable for the target application.

III. The Objective Function for Error Resilient Intra Coding

For error resilient coding, the following two aspects of intra coding must be considered:

First aspect - intra coding prevents temporal error propagation, since it has no reference to the other frames. In intra coding of advanced video codecs, such as H.264/AVC and H.265/HEVC, pixels of the adjacent blocks are used as intra prediction references, and these references (in encoded form) together with the residual data are encapsulated and transmitted in a single packet. However, if the reference pixels had been encoded in inter mode, they themselves might be erroneous, even if the residual data is received correctly. In this case, temporally propagated errors can propagate spatially into the intra-coded blocks. To avoid this condition and exploit the error propagation prevention provided by intra coding, the option of "ConstrainedIntraPred" can be enabled, which restricts the intra mode to use only the pixels of adjacent intra coded blocks as prediction references. This way, the received intra coded PUs are correctly decodable.

Second aspect - in no loss conditions, inter mode is obviously used more often than intra mode, because inter-coded blocks have lower bitrates than intra coded ones. By enabling the "ConstrainedIntraPred" option, the compression efficiency of intra mode is reduced even more, but it is beneficial for error resiliency [16] when we do have losses.

Therefore, in deciding to code a block in intra mode, there is a trade-off between bitrate and error resiliency. In this section, an objective function is developed which, rather than optimizing the encoder side rate-distortion, the decoder side rate-distortion is optimized. In other words, taking into account the channel distortion, the receiver side distortion is estimated at the encoder which is then used as the objective function.

Intra/inter mode selection is conventionally carried out based on the following Lagrangian cost function [33], [34]:

$$cost = D_q + \lambda R \tag{1}$$

where D_q is the quantization distortion in Mean Squared Error (MSE), λ is the Lagrangian coefficient and R is the number of required bits. This cost function is computed for the candidate modes and the mode with the lowest cost is selected as the final mode. However, this cost function does not take the transmission distortion into account. To consider it, with the same approach as presented in [15]–[17], the rate overall-distortion in a frame is represented in (2). The assumption behind this equation is that PLR is known at the transmitter side.

$$D^{(1)} = (1 - PLR)\, D_q^{(1)} + PLR\ D_{conceal}^{(1)} \tag{2}$$

where $D^{(1)}$, $D_q^{(1)}$ and $D_{conceal}^{(1)}$ are the expected total distortion, the quantization distortion, and the error concealment distortion for frame 1, respectively. The *expected* distortion means the average distortion seen over a long enough duration, or equivalently over a variant enough packet loss pattern, the latter used in our simulation. Note that the concealment distortion, $D_{conceal}$ in (2), is the distortion when all packets of the frame are lost and the frame is error concealed. It is evident that the frame is transmitted by a single packet; however, as shown in the Appendix, this is also valid when the frame is encoded into n packets and the packets convey the same amount of information.

Frame 0 is the initial I-frame of the sequence which is assumed to be received correctly. For frame 1, depending on whether its packets are received or not, the distortion will be $D_q^{(1)}$ or $D_{conceal}^{(1)}$, respectively. For frame 2, it becomes:

$$D^{(2)} = (1 - PLR)\, D_q^{(2)} + PLR\, D_{conceal}^{(2)} + PLR\left[1 - \beta^{(2)}\right]\Delta^{(1)} \tag{3}$$

where $\beta^{(2)}$ is the intra rate of frame 2 and

$$\Delta^{(1)} = E\left[\left(F_q^{(1)} - F_{conceal}^{(1)}\right)^2\right] \tag{4}$$

is the mean squared difference between frame 1 decoded correctly ($F_q^{(1)}$) and loss concealed ($F_{conceal}^{(1)}$); i.e., $\Delta^{(1)}$ denotes the *Mismatched Distortion* for frame 1 caused by error concealment. We assume that only the previous frame is used as prediction reference, as happens most of the times in encoders. Enabling multi-frame prediction results in a slight improvement in quality but at the cost of significant computational cost.

Equations (3) and (4) show that the quality of frame 1 directly affects the quality of frame 2, and its effect is controlled by parameter $\beta^{(2)}$ in (3). At larger β, the impact of mismatched distortion is clearly reduced, since intra coded PUs do not refer to the previous frame. As shown in [15], the quantization and mismatched distortions are independent of each other and one can simply write:

$$D_{conceal}^{(2)} = D_q^{(2)} + \Delta^{(2)} \tag{5}$$

Substituting (5) into (3) gives:

$$\begin{aligned} D^{(2)} &= D_q^{(2)} + PLR\left[\Delta^{(2)} + \left(1-\beta^{(2)}\right)\Delta^{(1)}\right] \\ &= D_q^{(2)} + PLR\,\Delta_{accum}^{(2)} \end{aligned} \tag{6}$$

in which

$$\Delta_{accum}^{(2)} = \Delta^{(2)} + \left(1-\beta^{(2)}\right)\ \Delta^{(1)} \tag{7}$$

is the *Accumulated Mismatched Distortion* seen in frame 2. It is evident that for frame 1, $\Delta_{accum}^{(1)} = \Delta^{(1)}$ and then (7) can be rewritten as:

$$\Delta_{accum}^{(2)} = \Delta^{(2)} + \left(1-\beta^{(2)}\right)\ \Delta_{accum}^{(1)} \tag{8}$$

Following the above concept, the distortion for the nth frame is:

$$D^{(n)} = D_q^{(n)} + PLR\,\Delta_{accum}^{(n)} \tag{9}$$

where

$$\begin{aligned} \Delta_{accum}^{(n)} &= \Delta^{(n)} + \left(1-\beta^{(n)}\right)\Delta_{accum}^{(n-1)} \\ \Delta_{accum}^{(0)} &= 0 \end{aligned} \tag{10}$$

and

$$\Delta^{(n)} = E\left[\left(F_q^{(n)} - F_{conceal}^{(n)}\right)^2\right] \tag{11}$$

Therefore, the distortion over a GoP of N frames is as given in (12) (as already mentioned, the 0th frame of the GoP is excluded from the summation):

$$D_{GoP} = \sum_{i=1}^{N} D^{(i)} = \sum_{i=1}^{N}\left(D_q^{(i)}(\boldsymbol{\beta}) + PLR\,\Delta_{accum}^{(i)}(\boldsymbol{\beta})\right) \tag{12}$$

where $\boldsymbol{\beta} = [\beta^{(1)}, \beta^{(2)}, \beta^{(3)}, \ldots \beta^{(N)}]$ is the vector intra rates for the N frames of the GoP. Quantization Parameter (*QP*) is excluded from this formulation, since its variation is usually ± 1 units at the given bitrate, except for sudden changes; e.g., scene-cut or fast/non-translational motions which is difficult for compensation with inter prediction. Therefore, we can assume that *QP* does not have significant changes for the N frames under consideration.

With the aim of maximizing the received video quality, the objective function with a constraint on the overhead bitrate of intra coding is:

$$\begin{aligned} &\min_{\boldsymbol{\beta}}\left\{\sum_{i=1}^{N}\left(D_q^{(i)}(\boldsymbol{\beta}) + PLR\,\Delta_{accum}^{(i)}(\boldsymbol{\beta})\right)\right\} \\ &\text{s.t.} \sum_{i=1}^{N} R_{intra}^{(i)}(\boldsymbol{\beta}) \le R_{red} \end{aligned} \tag{13}$$

where $R_{intra}^{(i)}$ is the number of additional bits needed for intra encoding of the ith frame according to the intra rate of $\beta^{(i)}$; that is, if $\beta^{(i)} = 0$, then $R_{intra}^{(i)} = 0$ and no block is codded in intra mode for error resiliency. The term R_{red} in (13) is the total redundancy budget allowed for these N frames for intra coding, which in turn is related to the PLR and the required degree of error resiliency. Increasing the intra rates of frame i; i.e., $\beta^{(i)}$, reduces $\Delta_{accum}^{(i)}$ (see equation (10)) but in turn increases bitrate usage.

IV. Solving the Objective Function

In this section, a solution to the constrained problem of (13) is driven through approximation. The behaviors of terms in this equation are observed and approximated through matching them to the experimental results. The experiments settings, the simplifying approximations, and the solution to the problem are provided in subsections A, B and C, respectively.

A. The Error Concealment Strategy

An important part of distortion in (13) belongs to the distortions caused by error/loss concealment. Error concealment techniques can be categorized into spatial and temporal domain processing techniques. In the spatial domain, the lost area of the frame is concealed using the spatially neighboring pixels. These methods exploit the correlations that usually exist among the neighboring pixels. In the temporal processing techniques, the contents from the previous and/or the future frames are addressed by MVs and used for temporal replacement. The actual MVs are not available and must be estimated or recovered first by the temporal loss concealment methods.

If the lost area is large, spatial domain is not effective, as pixels are very far apart from each other to be useful. The reason for dealing with large lost areas in HEVC is the size of its Coding Tree Unit (CTU), which can be as large as 64×64 pixels. An integer number of CTUs are regarded as one slice and an integer number of slices are encapsulated into a single transmission packet. Therefore, packet losses in HEVC streams affect a significant portion of the picture area especially for smaller picture sizes. As a result, temporal error concealment in HEVC streams is more applicable than spatial concealment.

Actually, exploiting the temporal frames' MVs will provide higher quality error concealment. One simple yet efficient technique is the Motion Copy algorithm where the MV of the collocated block is simply used for motion compensated temporal replacement. If the collocated block is coded in intra mode, Zero MV is used. However, in the case of having a high percentage of intra coded blocks, this approach is not efficient due to lack of MVs for intra blocks. For intra coded blocks, one solution is to recover the MVs by Boundary Matching Algorithm (BMA). A suggestion is to combine Motion Copy for inter coded and reliable collocated blocks, and BMA for intra coded or unreliable collocated blocks, as presented in [35]. The blocks with high residual signals are labeled as unreliable blocks and their MVs are not used for MV recovery. In [35], loss concealment is performed in two stages: firstly, the lost area is replaced using the Motion Copy algorithm. Then, for the unreliable MVs,

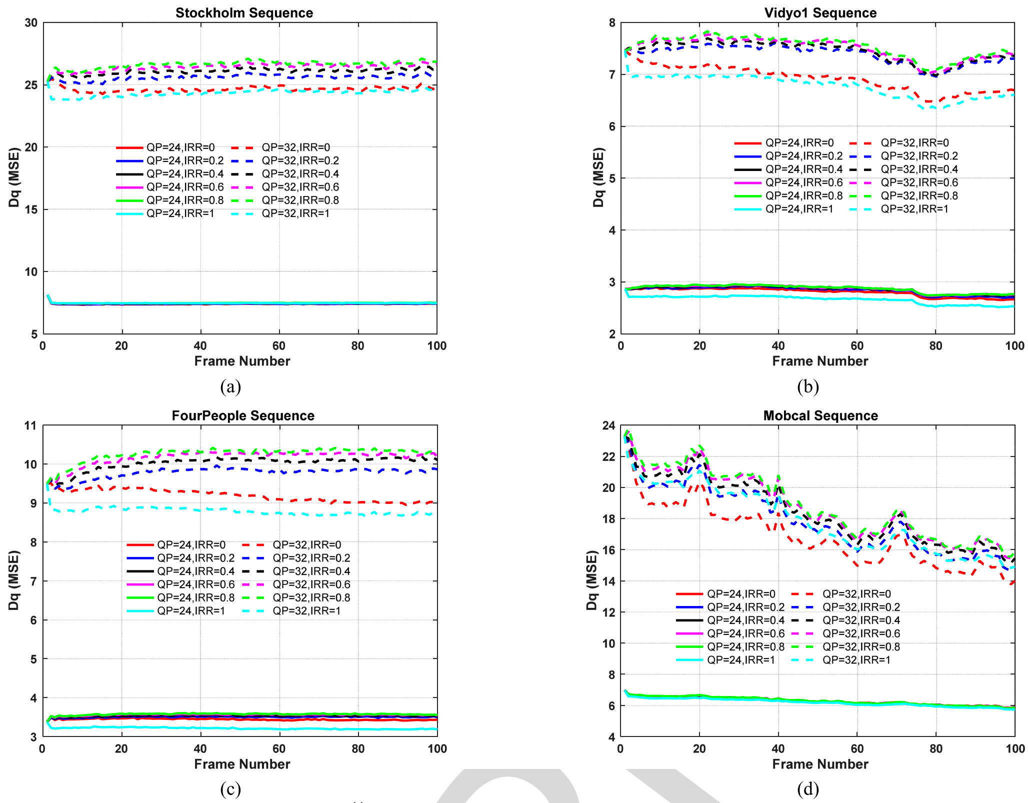


Fig. 1. The variation of $D_q^{(i)}(\beta^{(i)})$ with respect to IRR for test sequences and two values of *QP*.

MVs are obtained using BMA. Some other methods such as [36]–[37] are proposed for HEVC error concealment, but they work again based on the spatially close boundary pixels, which are not always available in the actual scenarios. The algorithm presented in [38] works based on MV extrapolation but with applying higher weights to the MVs belonging to the larger partitions. However, the problem is that this algorithm is based on the MVs of the blocks of the previous frame and therefore it is efficient when there are a few intra coded blocks. The method of [39] is proposed for error concealment of a sequence of successive MBs in H.264/AVC. This error concealment method does not need to know the MVs' neighboring spatially or temporally missed blocks, and instead estimates them by BMA. This method is also useful for large area losses of HEVC. The challenge in using BMA is the fact that the error concealment of one block will affect the error concealment of the following blocks as well. One solution, proposed in [39], is rank ordering the MBs for error concealment based on the texture of the available MBs in the surrounding of the lost area. A missed MB with a higher texture around it will be error concealed with higher priority. The criterion for the higher texture is the standard deviation of the luminance pixel values. Another solution for considering the interaction of loss concealed blocks is presented in [40], but it imposes significant computational complexity without considerable improvement.

In this paper, three techniques are used for error concealment: Motion Copy, the method presented in [35], and the method presented in [39]. The first two methods are appropriate when the blocks in the earlier frame are encoded mostly in inter mode, and the third method is suitable when the blocks in the earlier frame are encoded mostly or completely in intra mode. Then, the highest quality output is selected and used for the measurements.

It is worth noting that throughout the paper error concealment and loss concealment are used interchangeably, but in fact loss concealment is carried out. The reason is that in highly error prone networks, such as wireless networks, severely erroneous packets cannot be corrected and they are regarded as lost packets by the decoder. However, if the used entropy coder is symmetric, such as that of the H.263 codec, then parts of the data can be retrieved and the lossy area can be less than that of whole packet loss [45]. Since H.265/HEVC does not use symmetric entropy coder, then there would not be any retrieval of erroneous parts and the whole packet can be regarded as lost. Hence, loss concealment is a proper choice.

B. Simplifying the Objective Function

First, the quantization distortion does not significantly change with parameter $\beta^{(i)}$. That is, $D_q^{(i)}(\beta^{(i)})$ is approximately constant when $\beta^{(i)}$ varies from minimum $(\beta = 0)$ to maximum $(\beta = 1)$. This can be verified from the simulation results shown in Fig. 1. In this figure, four HD test video sequences are coded with HM16.0, the reference software of HEVC, at two values of *QP* and six values of IRR. The tested video sequences are *Stockholm*, *Vidyo1*, *FourPeople* and *Mobcal*. For the given intra rates, a sufficient number of PUs with sizes of 16×16 pixels

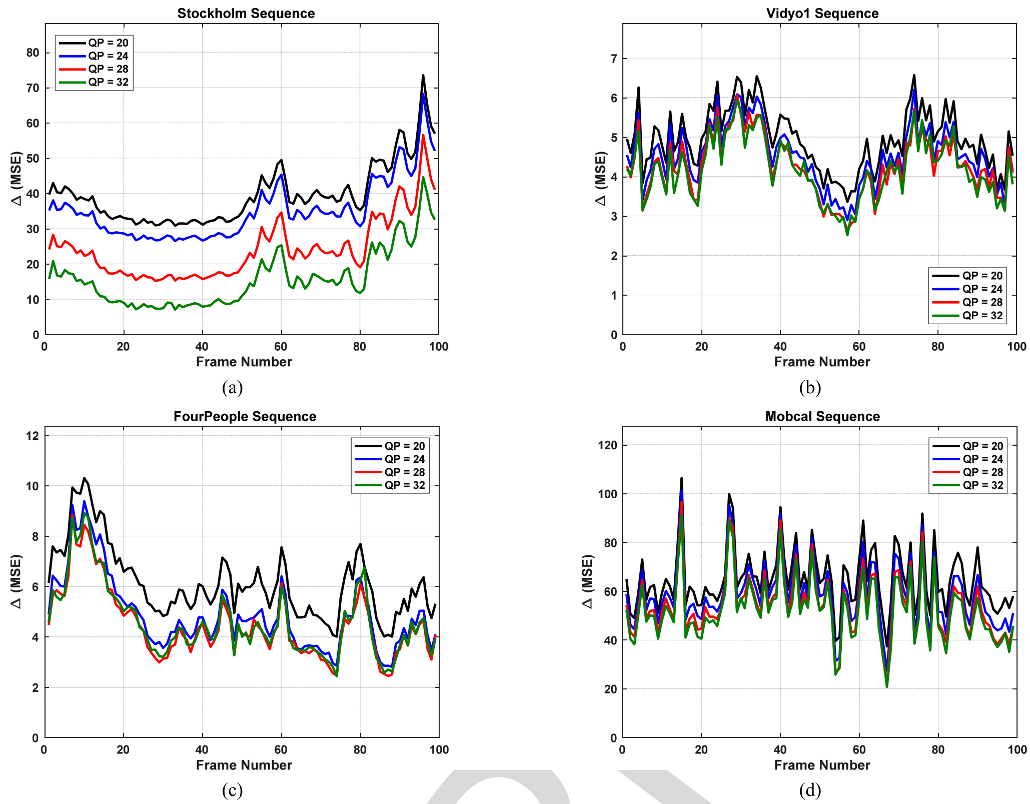


Fig. 2. The variation of mismatched distortion ($\Delta^{(i)}$) for the frames of test sequences for four values of QP.

are selected randomly and are forced to intra coding mode. As already mentioned, the randomly selected PUs do not overlap in the frames.

Fig. 1 shows the ignorable changes of $D_q^{(i)}$ with $\beta^{(i)}$. It can also be seen that the variation in $D_q^{(i)}$ for $QP = 24$ is less than that of $QP = 32$. The mathematical reason is that for smaller QPs, or smaller quantization step sizes (Q_{SS}), the high bitrate approximation is more accurate and the quantization distortion is nearly equal to $\frac{Q_{SS}^2}{12}$ [17]. This is fixed for various signals, independent of inter or intra coding. However, the difference in $D_q^{(i)}$s for various βs is still ignorable, even for $QP = 32$. Therefore, $D_q^{(i)}$ is fixed with the optimization arguments and equation (13) can be simplified as:

$$\min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{N} \left(PLR\ \Delta_{accum}^{(i)} (\boldsymbol{\beta}) \right) \right\}$$
$$\text{s.t.} \sum_{i=1}^{N} R_{intra}^{(i)} (\boldsymbol{\beta}) \leq R_{red} \tag{14}$$

An important term in (14) is $\Delta_{accum}^{(i)}$ which is the multiply-accumulated of mismatched distortions $\Delta^{(i)}$s, with the multiplication coefficient of $(1 - \beta^{(i)})$, as given in (10). Therefore, the variation of $\Delta^{(i)}$ per frames is important in the behavior of the objective function of (13). To measure $\Delta^{(i)}s$, the frames are first error concealed with the strategy given in IV.A, and then $\Delta^{(i)}$ is calculated by (11). The results are shown in Fig. 2 where it can be seen that, most of the times and with a good approximation, the frames of a sequence have close mismatched distortions, that is:

$$\Delta^{(1)} \cong \Delta^{(2)} \cong \cdots \cong \Delta \tag{15}$$

Even though it might not be valid for all frames, the variations are smooth in the windows of N frames, as large as the usually used GoP sizes (30–60 frames). This assumption may not be much accurate; however, this assumption, by nature, is similar to the assumption made in Rate-Control (RC) algorithms. In RC algorithms, the goal is to control the total bitrate to be less than the given bound with minimum fluctuation in the quality. Therefore, for a real-time RC, the encoder must assume that the future frames have almost the same behavior in the view of compression properties. Even though this assumption is not always valid, it is very efficient and helpful in practice. Similar to RC algorithms, we can assume that the frames behave similarly in the view of mismatched distortion. Therefore, with the assumption of (15), $\Delta^{(i)}$ is fixed for the frames, and since the employed loss concealment strategy is not much sensitive to the intra/inter coding, it is also fixed with $\boldsymbol{\beta}$. Therefore, by substituting the recursive formula given in (8), the objective function of (14) can be expanded as follows:

$$\min_{\boldsymbol{\beta}} \left\{ PLR.\ \Delta \left[N + \sum_{i=1}^{N} \left(\sum_{j=1}^{N-(i-1)} \left(\prod_{k=0}^{i-1} \left(1 - \beta^{(j+k)} \right) \right) \right) \right] \right\}$$
$$\text{s.t.} \sum_{i=1}^{N} R_{intra}^{(i)} (\boldsymbol{\beta}) \leq R_{red} \tag{16}$$

where PLR and Δ are assumed constant during optimization.

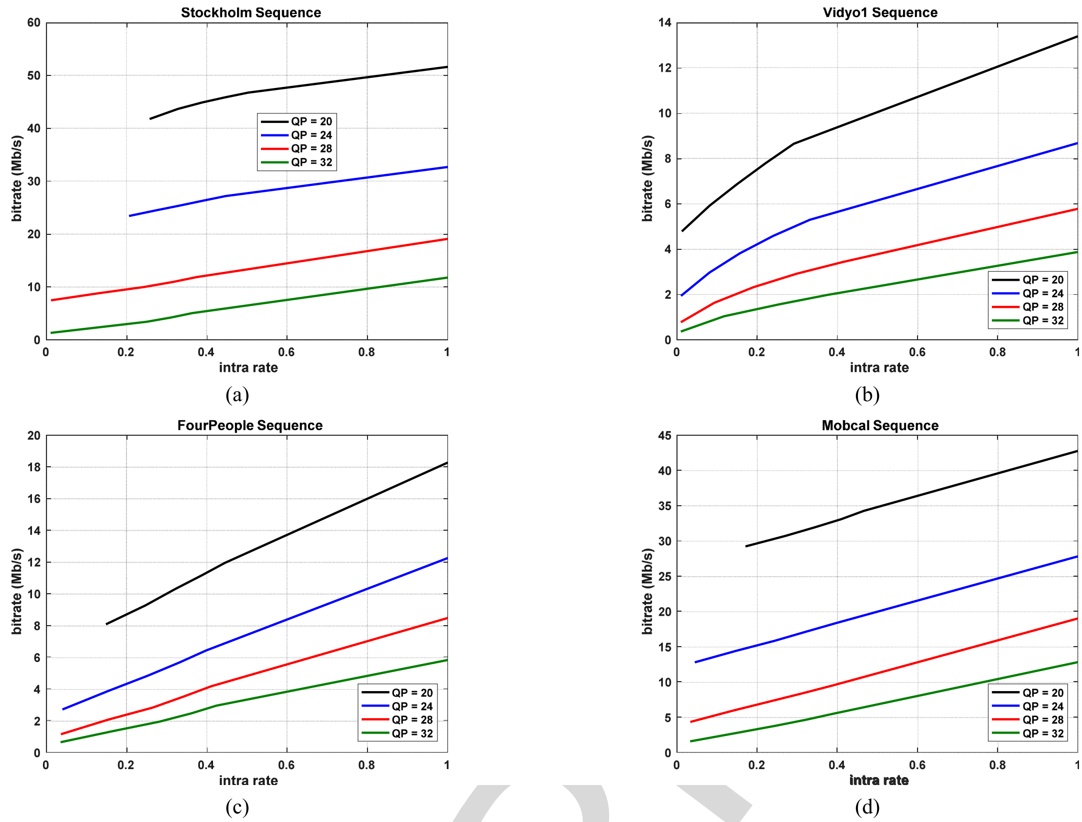


Fig. 3. The variation of bitrate with intra rate for various *QP*s and test sequences.

Now, for the constraint of (16), we do another simulation: the variation of frame bits when β varies from 0 to 1, as depicted in Fig. 3. In this figure, the average bitrates needed for sending the encoded video frames are measured and shown against the intra rate.

From Fig. 3, we can generally assume that the bitrates are increased almost linearly with β. That is:

$$R^{(i)} = R_0^{(i)} + R_{intra}^{(i)} = R_0^{(i)} + \alpha^{(i)}\,\beta^{(i)}$$
$$\Rightarrow\ R_{intra}^{(i)} = \alpha^{(i)}\,\beta^{(i)} \tag{17}$$

where $R_0^{(i)}$ is the bitrate of the i^{th} frame for $\beta = 0$. The figure shows that the curves almost have the same slope; that is, they have the same α defined in (17). For this reason, the constraint term in (16) can be stated as:

$$\sum_{i=1}^{N}\left(\alpha\,\beta^{(i)}\right) \le R_{red} \tag{18}$$

or equivalently

$$\sum_{i=1}^{N}\beta^{(i)} \le \beta_{red} \tag{19}$$

where β_{red} is the intra rate budget; i.e., the sum of total intra rates allowed to be assigned to these N frames. Therefore, the objective function of (16) is simplified as

$$\min_{\boldsymbol{\beta}}\left\{PLR.\,\Delta\left[N + \sum_{i=1}^{N}\left(\sum_{j=1}^{N-(i-1)}\left(\prod_{k=0}^{i-1}\left(1-\beta^{(j+k)}\right)\right)\right)\right]\right\}$$
$$\text{s.t.}\ \sum_{i=1}^{N}\beta^{(i)} \le \beta_{red} \tag{20}$$

Since we can assume that PLR and Δ are fixed during optimization, the problem in (20) can be rewritten as:

$$ErrorPro_{min} = \min_{\boldsymbol{\beta}}\left\{\sum_{i=1}^{N}\left(\sum_{j=1}^{N-(i-1)}\left(\prod_{k=0}^{i-1}\left(1-\beta^{(j+k)}\right)\right)\right)\right\}$$
$$\text{s.t.}\ \sum_{i=1}^{N}\beta^{(i)} \le \beta_{red} \tag{21}$$

It is worth mentioning that for simplicity of deriving the objective function, without loss of generality, there are no B-frames. This is because, in general, some B-frames maybe used as prediction reference like P-frames causing error propagation, while others are not used as the reference and hence do not propagate the errors but they become erroneous. Modifying our formulations to highlight this matter makes the equations more complicated without giving the required information to the reader.

TABLE I
THE OUTPUT OF THE OBJECTIVE FUNCTION OF (21) FOR SOME VALUES OF β_{red}

Range of β_{red}	Intra rate of the frames															Constraint
	$\beta^{(1)}$	$\beta^{(2)}$	$\beta^{(3)}$	$\beta^{(4)}$	$\beta^{(5)}$	$\beta^{(6)}$	$\beta^{(7)}$	$\beta^{(8)}$	$\beta^{(9)}$	$\beta^{(10)}$	$\beta^{(11)}$	$\beta^{(12)}$	$\beta^{(13)}$	$\beta^{(14)}$	$\beta^{(15)}$	
$\beta_{red} \leq 1$	0	0	0	0	0	0	0	a_1	0	0	0	0	0	0	0	$a_1 = \beta_{red}$
$1 < \beta_{red} \leq 1.45$	0	0	0	a_1	0	0	0	1	0	0	0	a_2	0	0	0	$a_1 + a_2 = \beta_{red} - 1$
$1.45 < \beta_{red} \leq 2$	0	0	0	0	a_1	0	0	0	0	1	0	0	0	0	0	$a_1 = \beta_{red} - 1$
	0	0	0	0	0	1	0	0	0	0	a_1	0	0	0	0	
$2 < \beta_{red} \leq 2.71$	0	0	0	0	1	0	0	0	0	1	0	0	a_1	0	0	$a_1 = \beta_{red} - 2$
	0	0	0	0	1	0	0	a_1	0	0	1	0	0	0	0	
	0	0	a_1	0	0	1	0	0	0	0	1	0	0	0	0	
$2.71 < \beta_{red} \leq 3$	0	0	0	a_1	0	0	0	1	0	0	0	1	0	0	0	$a_1 = \beta_{red} - 2$
	0	0	0	1	0	0	0	a_1	0	0	0	1	0	0	0	
	0	0	0	1	0	0	0	1	0	0	0	a_1	0	0	0	
$3 < \beta_{red} \leq 3.6$	0	a_1	0	1	0	a_2	0	1	0	a_3	0	1	0	a_4	0	$a_1 + a_2 + a_3 + a_4 = \beta_{red} - 3$

C. Solving the Simplified Objective Function

In this subsection, the solution for the objective function given in (21) is discussed. The problem is actually minimizing the error propagation at a given intra rate budget. As an example, this problem is solved for $N = 15$ and various values of β_{red}; the results are tabulated in Table I.

As can be seen from this table, when $\beta_{red} \leq 1$, the best frame for intra coding is the middle frame. If there was more intra budget; i.e., $1 < \beta_{red} \leq 1.45$, $\beta^{(4)}$ and $\beta^{(12)}$ begin to grow irrespective of whether the intra rate is allocated to the 4th frame or the 12th frame. However, when intra rate budget exceeds 1.45, the optimization function given in (21) recommends other frames for intra coding to be chosen; e.g., frames 5 and 10, where frame 10 is coded wholly as I-frame, and frame 5 has partially intra coded blocks. Equivalently, another package is frames 6 and 11, where frame 6 is now selected for I-frame coding. One can see that these two packages produce the same obstacle against the error propagation.

For some other regions of β_{red}, the selected frames are given in Table I. One important point is changing the intra coded frame candidates imposed by the objective function of (21). The reason is that, if β_{red} is between two integers K_1 and K_2; that is $K_1 < \beta_{red} < K_2$, the optimizers may decide to add another frame for intra coding in addition to K_1 frames (e.g., one frame between them), or decide to select K_2 frames for intra coding and reduce the intra rate of one of them to comply with the bound of β_{red}. Clearly, if β_{red} is close to K_1, the former case happens, and the latter case happens when β_{red} is close enough to K_2. However, as shown in Fig. 4, $ErrorPro_{min}$ behaves continuously at these border points of β_{red}. In each interval shown by broken lines, the intra coding frame candidates are the same where one or more appropriate frames of these candidates consume the allocated intra coding budget. As already mentioned, the slope of decay in $ErrorPro_{min}$ in each interval is constant. If β_{red} becomes larger than 7, the frames are alternately coded as I-frame; that is the GOP structure is IPIPIP, and now all P-frames have the same priority for intra rate for all β_{red} amounts; therefore; there are no broken lines in Fig. 4 for $\beta_{red} > 7$.

This solution proves that to achieve the best error resiliency for intra coding, the best strategy is to concentrate on intra

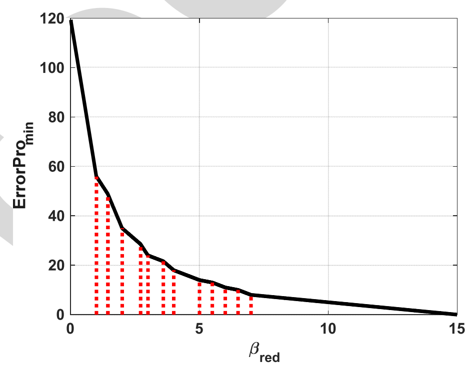


Fig. 4. Behavior of $ErrorPro_{min}$ with β_{red}.

coding the blocks in the middle frame of the GoP, such that the entire frame is coded as an I-frame. If the intra rate budget (or accordingly bitrate budget) allows, more frames can still be coded in intra mode. In other words, the output of the objective function is to reduce the intra period; this strategy leads to smaller error propagation and hence higher video quality for lossy channels, compared to the case that intra coded blocks are distributed among the frames.

V. THE OPTIMAL VALUE FOR THE INTRA PERIOD

As shown in Table I, at a larger β_{red}, the number of I-frames in the GoP can increase. This is in favor of mitigation of error propagation; however, the required bitrate for sending the video is increased since the compression ratio is decreased.

Having more I-frames is justified in channels with higher loss rates and vice versa. Therefore, PLR and the coding bitrate affect the best value for β^*_{red}. As shown in Table I, β^*_{red} is directly related to Intra Period (*IP*); therefore, the problem of finding $\beta_{red-opt}$ is equivalent to finding an optimal value for *IP*, denoted as IP^*. However, to find IP^* analytically, one must know the rate-distortion behavior of the frames of the GoP under consideration; that is, the behavior of future frames must be known *a priori*, which is not possible unless it is estimated based on the frames' history similar to the work presented in [41]. This

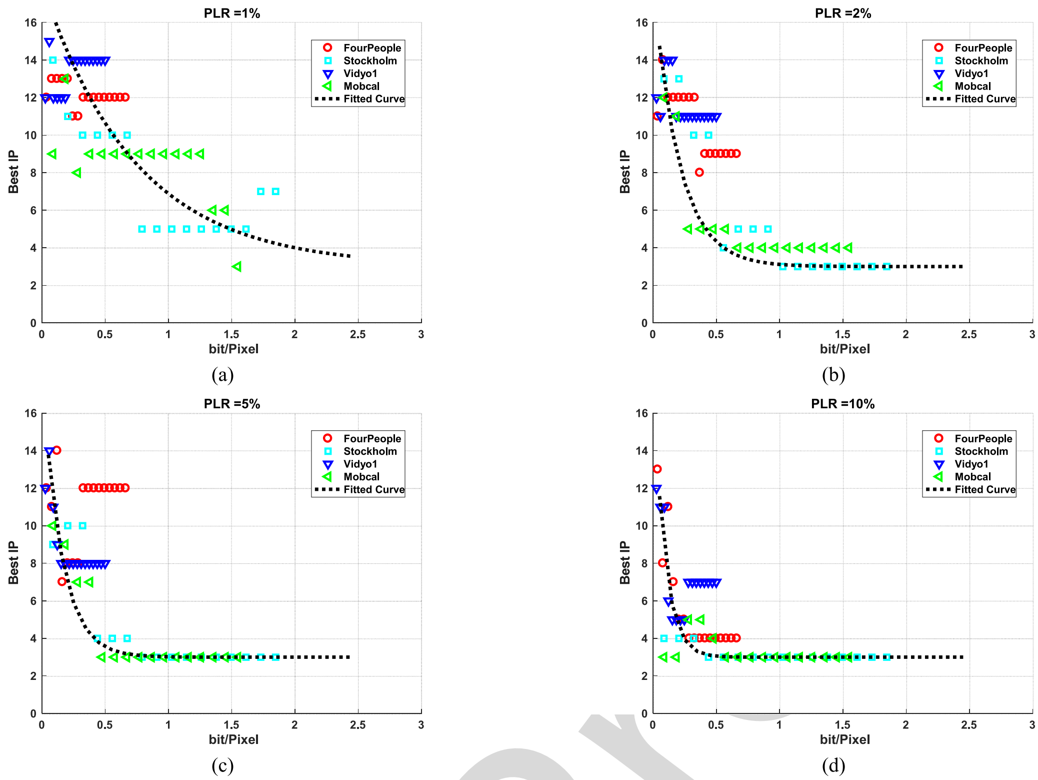


Fig. 5. The best *IP* versus bit/pixel for various PLRs and test sequences.

way, the problem can be solved using classical constraint optimization approaches. However, the complexity of the problem and non-trivial solutions have encouraged us to use an empirical approach. Therefore, IP^* has been found through experimental measurements, as follows.

The videos are encoded in Slice Mode, and each slice contains an integer number of CTUs in raster scan order. Each coded slice can be as large as 1500 bytes, meeting the Maximum Transmission Unit (MTU) of the network, and transmitted as a single packet. The channels experience a burst form of loss generated by Elliot-Gilbert model [46]. At each PLR, 40 packet loss patterns with an average burst length of three packets are generated and applied on the bit streams.

The video sequences are encoded at various values of IP; $IP = M$ means every *M*th frame of the sequence is coded as an I-frame. For example, for $IP = 3$, there are two P-frames after each I-frame, and this pattern is repeated throughout the sequence. In a GoP of 30 frames, the videos are encoded with $IP = 1, 2, \ldots, 15$ (for GoP of N frames, $IP > \frac{N}{2}$ is not reasonable). The compressed bit stream is subjected to a specific PLR, and the decoded video is loss concealed (as given in IV.A) and the resulting quality is measured. Video quality is measured in terms of Video Quality Model (VQM) [47] and its average index taken over the loss patterns is calculated. VQM is a video quality assessment method which considers both spatial and temporal distortions, so it is quite suitable to our case. For each test video and at the given PLR, the best IP which provides the best quality (i.e., the lowest VQM index, since higher quality is equivalent to lower VQM index) at the corresponding bitrate is selected. Fig. 5 shows the best *IP*s as a function of bit/pixel for four sequences.

It can be seen that the best *IP* becomes smaller at higher bitrates and higher PLRs. Even though some points are not close to the others, they can be fitted on decaying exponential functions, as shown in Fig. 5. The fitted curves can be formulated with the following equations:

$$IP^* = 3 + 15 \exp\left(-\frac{R}{R_0}\right)$$
$$R_0 = 0.15 + 1.4575 \ \exp\left(-\frac{PLR}{0.01}\right) \tag{22}$$

where R is the bit per pixel. Clearly, IP^* obtained from (22) must be rounded to the nearest integer number. Even though the decimal values are also applicable, our empirical approach and curve fitting is not accurate enough for extracting decimal values for frames' intra rates.

One issue is the fact that coding a frame fully in Intra mode might cause sudden changes in the bitrate and hence more congestion in the lossy channels. However, for numerous applications, like video broadcast, streaming, multicasting etc., one needs to play the video at almost any time during transmission. This facility can only be provided by Intra coded frames. On the issue of increased I-frame bitrates, one should note that in these applications, normally several video flows are multiplexed, such that higher bitrates of I-frames coincide with lower bitrates of many P- and B-frames of the other flows and are easily smoothed out. Despite this, even for a single video flow, some traffic shaping, such as coarser quantization parameter for I-frames can be applied to reduce the bits; however, this solution may lead to quality flicker due to lower qualities of the I-frames if the QPs of I-frame and P-frames are much different. For high motion

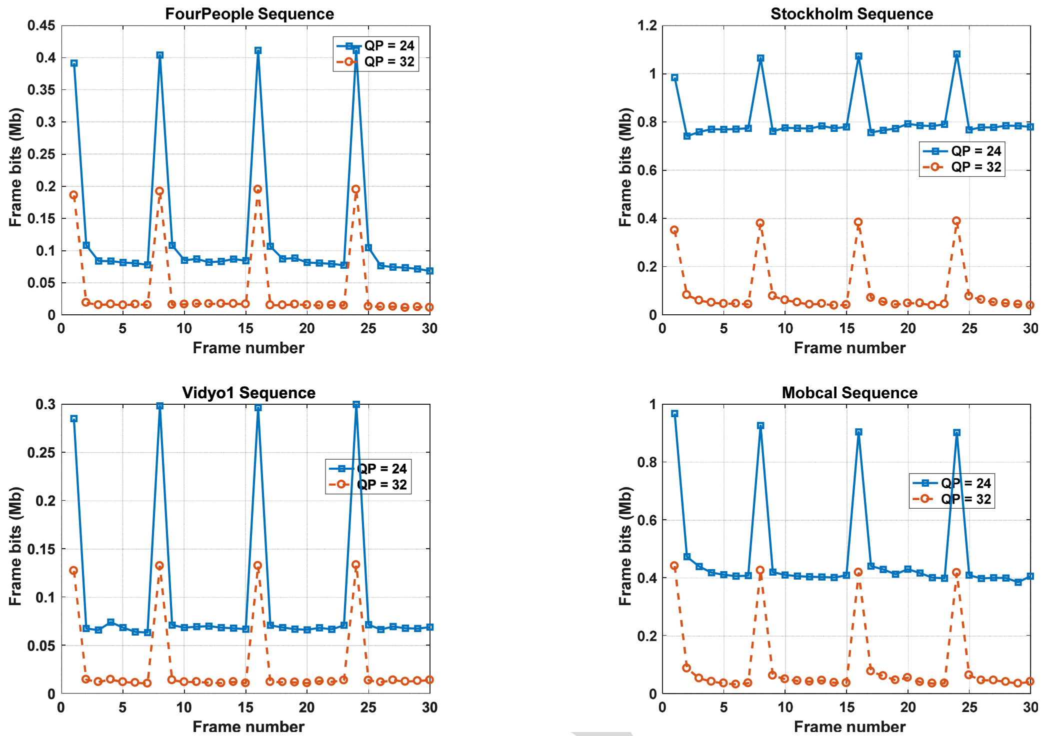


Fig. 6 The sudden changes in the frame bits when a frame is encoded as I-frame. Every 8 frames, one frame is Intra coded entirely.

and high texture videos, the difference between I and P frames' bits are not so large and this approach might be sufficient there. If this was not the case, another solution is to have the same QP for I and P frames but use an encoder smoothing buffer to regulate the bitrate (e.g., traffic shaping), of course at the cost of a few frames delay in video display. How this delay can solve the problem is explained below. Let us assume that intra period is M frames. If the frame rate of the video is FPS, the average bitrate required by the channel is:

$$R = FPS\frac{((M-1)\,R_P + R_I)}{M} \quad bits/sec \tag{23}$$

where R_P and R_I are number of bits needed for coding the P-frames and I-frames, respectively. Now if the I-frame has k times more bits than the P-frames, then:

$$\begin{aligned} R &= \frac{FPS\ (M-1+k)\,R_P}{M} \\ &= \frac{FPS\ (M-1+k)\,R_I}{kM} \quad bits/\sec \quad (k>1) \end{aligned} \tag{24}$$

At each 1/*FPS* second, the total sent bits are:

$$\begin{aligned} R &= \left(1 + \frac{k-1}{M}\right) R_P \\ &= \left(\frac{M+k-1}{kM}\right) R_I \quad bits/\sec \quad (k>1) \end{aligned} \tag{25}$$

Therefore, more than one P-frame or less than one I-frame is transmitted at each 1/*FPS*. That is, compared to the case where all frames have the same number of bits (i.e., $k = 1$), delivering P-frames is faster and delivering I-frames is slower. However, the issue that may arise here is the transmit and receive buffers' overflow and underflow in a live streaming application. It can be shown that, with display latency as large as $M(k-1)/(M-1+k)$ frames, there is no overflow or underflow in the buffers and continuous playing of the video is preserved (see Appendix B in [32] for the proof). This latency increases with M; therefore, a smaller M chosen for higher PLRs leads to lower latencies.

The value of k is content dependent; Fig. 6 shows the number of bites of I and P frames, with $M = 8$ for four sequences and two *QP*s. One can see that k is about 1.5 for *Stockholm* at $QP = 24$; that is, k is small and the delay is not significant. For example, for PLRs of 5%, if M is around of 4 as shown in Fig. 5, this gives a latency of about 0.5 frames. However, the ratio k becomes larger at $QP = 32$. And also, M is typically larger for lower bitrates; i.e., for $QP = 32$. Therefore, the incurred delay is more challenging here; for example, for *FourPeople* at $QP = 32$, k is about 10 as shown in Fig. 6, which is relatively very high. Now for $M = 15$ (as inferred from Fig. 5), the delay becomes about 5.5 frames. For *FPS* of 60, it leads to a delay less than 100 ms which is acceptable for many applications. For smaller delays, we can combine the above two approaches; that is, applying coarser quantizer and forcing a delay. The coarser quantizer to I-frames leads to a smaller k which in turn leads to a smaller delay.

VI. Performance Comparison

The analysis explained in the previous sections shows that using I-frames instead of applying IRR is more efficient as an error resiliency tool and gives higher quality in dealing with transmission of encoded videos over lossy channels. The

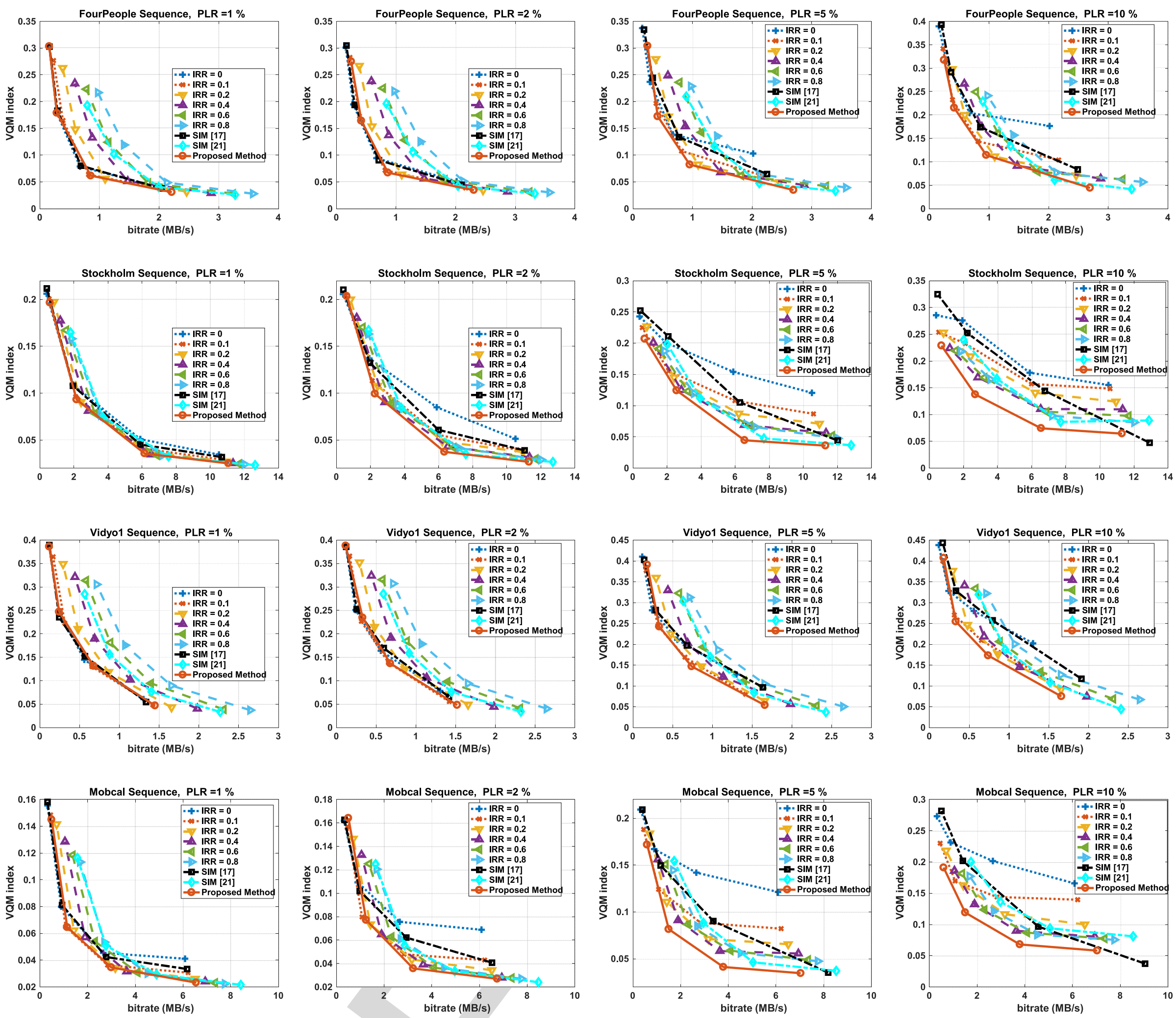


Fig. 7. Performance comparison of the proposed method for various PLRs and *FourPeople*, *Stockholm*, *Vidyo1* and *Mobcal* sequences for burst length of three packets.

suitable I-frame period is given by (22). As already mentioned in Section II, there are two other options for intra coding, SIM and Intra Refresh. For performance comparison, our proposed method is compared against two SIM methods, [17] and [21], where the PUs are selected based on an objective function for intra or inter coding. Note that as explained in Section II, there are also two options for selecting the blocks to be forced for intra coding in the Intra Refresh scheme. They can be selected randomly or in a regular manner, such as a column of intra blocks moving frame by frame from left to right. Our experiments showed that the latter option, called Periodic Intra Refresh (PIR) or cyclic intra-refresh generally gives superior performance in terms of rate-distortion. Therefore, we have included the results of PIR in Figs 7–9. Since there are no appropriate recent related works on the best value of IRR, we examine PIR with several possible values of IRR for all examined PLRs; these are $\{0, 0.1, 0.2, 0.4, 0.6, 0.8\}$. Note that $\text{IRR} = 0$ is equivalent to not paying any attention to channel loss at the encoder. With the experimental settings given in Section V, these results are shown in Figs 7–9, Fig. 9 is for the average burst length of six packets.

Despite of the simplifications and approximations made in our method through analysis and curve fitting, it can be seen from Figs 7–8 that our proposed method outperforms the others in many cases. For lower PLRs and smaller bitrates, the proposed method provides actually no gain. In these regions, since the video is less sensitive to packet loss, the curves are actually close to each other. The algorithm of [21] picks many PUs for intra coding; therefore, it applies intra rate much more than required but with a slight gain in quality in lower PLRs. For this reason, this algorithm does not work well for low PLRs. In the cases of higher PLR*s* and higher bitrates, one can see the VQM quality index of our proposed method is better than the others which is sometimes significant. A reminder that the smaller VQM index means higher quality. Light content video sequences, such as *FourPeople* and *KrisenAndSara*, as already mentioned are less sensitive to data loss; hence the VQM curves are again similar while ours are still marginally better.

For the PLR*s* of 5% and 10%, the results of applying average burst length of six packets are shown in Fig. 9. It can be seen that the performance of our proposed method is still better than the others. Actually, the loss pattern does not significantly affect our

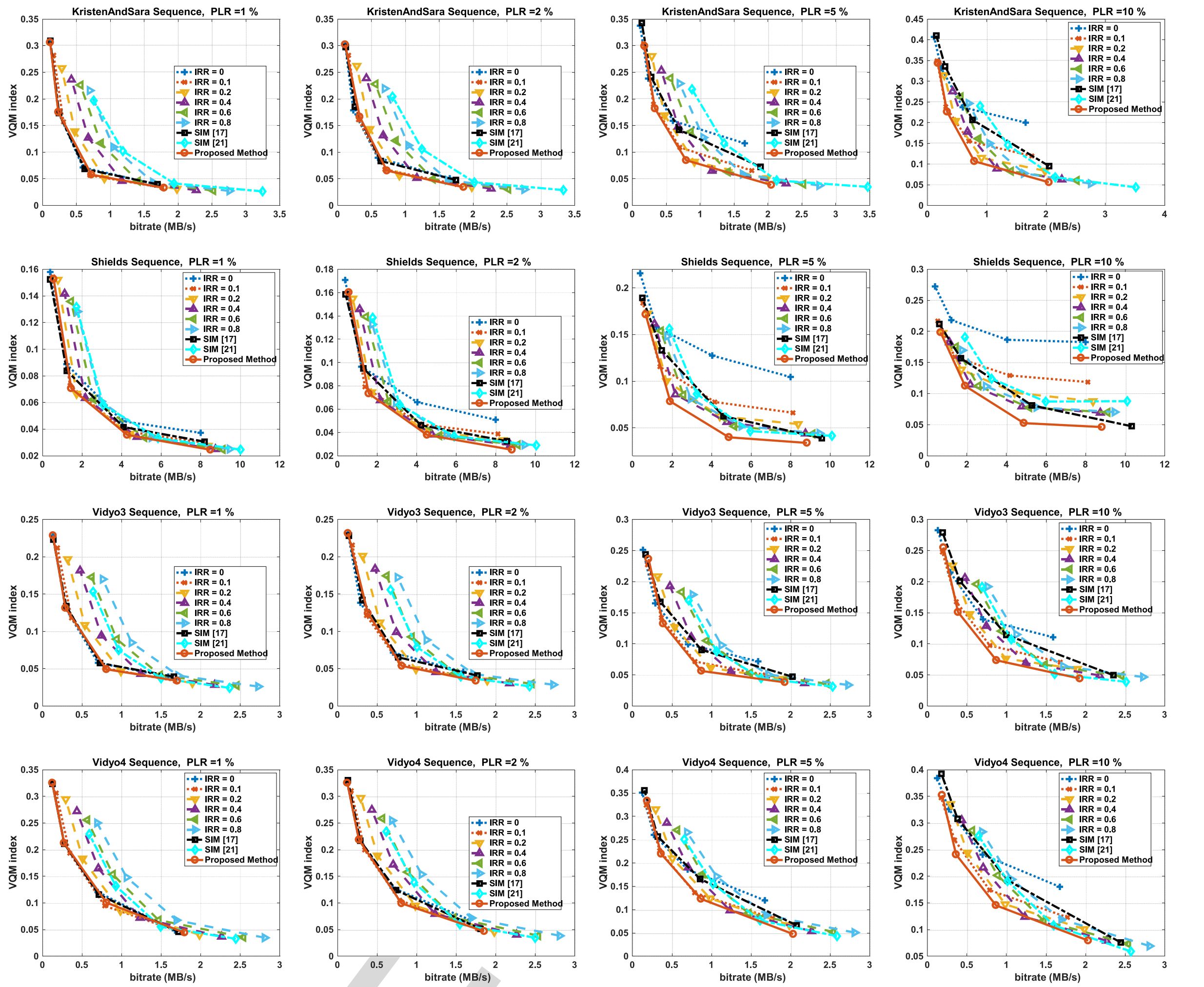


Fig. 8 Performance comparison of the proposed method for various PLRs and *Kristen and Sara*, *Shields*, *Vidyo3* and *Vidyo4* sequences for burst length of three packets.

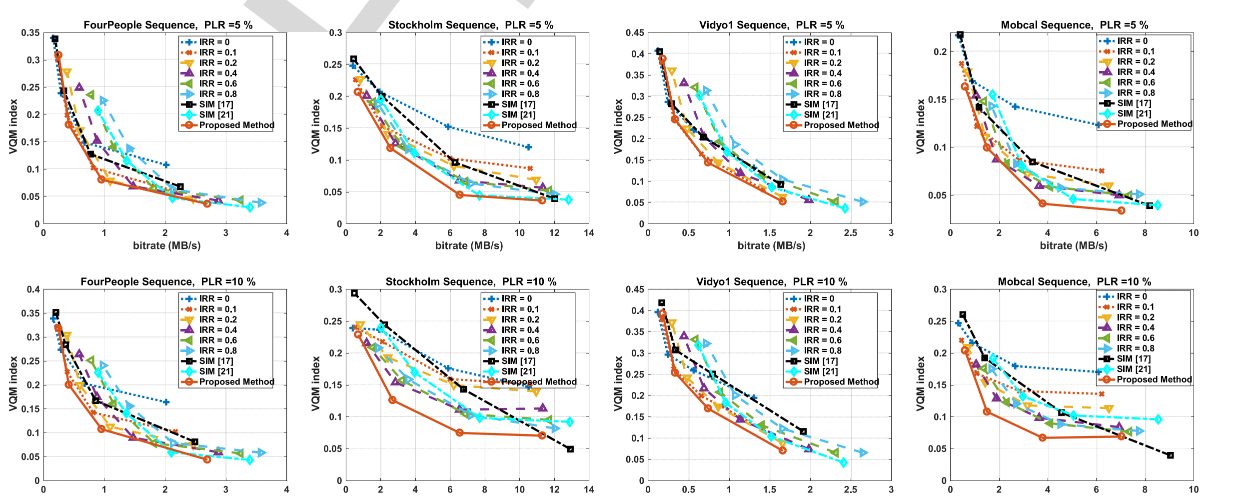


Fig. 9. Performance comparison of the proposed method for various PLRs and *FourPeople*, *Stockholm*, *Vidyo1* and *Mobcal* sequences for burst length of six packets.

results due to our loss concealment procedure applied to lossy bitstreams generated by any of competing methods.

The burst loss leads to larger lossy areas in the pictures, which is usually handled by the Motion Copy algorithm. Note that burst loss will lead to the loss of consecutive frames in low bitrate and low resolution videos, while it is not so destructive for HD and beyond.

VII. Conclusion

In this paper, the best strategy for intra coding as an error resiliency tool is presented. It was proposed to encode some frames entirely in intra mode, rather than the conventional approach where some blocks or PUs are selected in PIR manner (with a specific intra rate) or by a cost function to be coded in intra mode. Considering the error propagation, the receiver side distortion is formulated and it is simplified with some observations. The simplified objective function has a straightforward solution: $\beta^{(i^*)} = 1$, where i^* is the index of the frames symmetrically positioned in the GoP, and the number of I-frames depends on β_{red} or equivalently the available bitrate budget for intra coding. The output of the objective function is to reduce the *IP* as much as possible and as long as the bitrate overhead of intra coding is justified at the given channel loss rate.

The optimal IP varies with the coding bitrate as well as the PLR as shown in Fig. 5. We have fitted a curve to the experimental points obtained from examining various test sequences, as given in (22). With the IP^* selected by (22), experimental results show that the proposed method achieves lower VQM index compared to the conventional SIM and PIR methods.

Appendix

Assume that frame 1 is transmitted through n packets. If m packets are lost, the average distortion after error concealment is:

$$D_m^{(1)} = \frac{(n-m)}{n} D_q^{(1)} + \frac{m}{n}\ D_{conceal}^{(1)} \qquad (26)$$

If each packet is lost with a probability of PLR, the probability of losing m packets is

$$\begin{aligned} PLR_m &= C\left(n,m\right)\ PLR^m \left(1-PLR\right)^{n-m} \\ &= \binom{n}{m}\ PLR^m \left(1-PLR\right)^{n-m} \end{aligned} \qquad (27)$$

where $C(n, m)$ is the number of m-combinations from n packets. Therefore, the expected distortion of frame 1 is as given by (28):

$$\begin{aligned} D^{(1)} &= \sum_{m=0}^{n} \left(PLR_m\ D_m^{(1)}\right) \\ &= \frac{D_q^{(1)}}{n} \sum_{m=0}^{n} \left[(n-m)\left(\binom{n}{m} PLR^m (1-PLR)^{n-m}\right)\right] \\ &\quad + \frac{D_{conceal}^{(1)}}{n} \sum_{m=0}^{n} \left[m\ \left(\binom{n}{m} PLR^m (1-PLR)^{n-m}\right)\right] \end{aligned} \qquad (28)$$

Both summations in (28) are the expected values of a *Binomial* distribution with probabilities of $(1 - PLR)$ and PLR, respectively. That is $D^{(1)}$ becomes

$$\begin{aligned} D^{(1)} &= \frac{D_q^{(1)}}{n}\left[n\left(1-PLR\right)\right] + \frac{D_{conceal}^{(1)}}{n}\left[n\ PLR\right] \\ &= \left(1-PLR\right) D_q^{(1)} + PLR\ \ D_{conceal}^{(1)} \end{aligned} \qquad (29)$$

which is the same as equation (2).

References

[1] Y. Wang and Q. F. Zhu, "Error control and concealment for video communication: A review," *Proc. IEEE*, vol. 86, no. 5, pp. 974–997, May 1998.

[2] M. T. Pourazad Oztas, P. Nasiopoulos, and V. C. M. Leung, "A study on the HEVC performance over lossy networks," in *Proc. 19th IEEE Int. Conf. Electron. Circuits Syst.*, 2012, pp. 785–788.

[3] M. Fleury, S. Moiron, and M. Ghanbari, "Innovations in video error resilience and concealment," *Recent Patents Signal Process.* vol. 1, no. 2, pp. 124–134, 2011. Q4

[4] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, "Review of error resilient coding techniques for real-time video communications," *IEEE Signal Process. Mag.*, vol. 17, no. 4, pp. 61–82, Jul. 2000.

[5] M. Fleury, M. Altaf, S. Moiron, N. Qadri, and M. Ghanbari, "Source coding methods for robust wireless video streaming," *Multimedia Netw. Coding*, pp. 175–207, 2012.

[6] M. Fleury, I. A Ali, and M. Ghanbari, "Video intra coding for compression and error resilience: A review," *Recent Patents Signal Process.*, vol. 4, pp. 32–43, 2014.

[7] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE JSAC*, vol. 18, no. 6, pp. 966–976, Jul. 2000.

[8] H. Yang and K. Rose, "Advances in recursive per-pixel end-to-end distortion estimation for robust video coding in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 7, pp. 845–856, Jul. 2007.

[9] Y. Liao and J. D. Gibson, "Rate-Distortion based mode selection for video coding over wireless networks with burst losses," in *Proc. Int. Packet Video Workshop*, Seattle, WA, USA, 2009, pp. 1–10.

[10] S. Wan and E. Izquierdo, "Rate-distortion optimized motion-compensated prediction for packet loss resilient video coding," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 1327–1338, May 2007.

[11] J. Xiao, T. Tillo, C. Lin, and Y. Zhao, "Joint redundant motion vector and intra macroblock refreshment for video transmission," *EURASIP J. Image Video Process.*, 2011. Q5

[12] H. Yang and K. Rose, "Optimizing motion compensated prediction for error resilient video coding," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 108–118, Jan. 2010.

[13] H. Shu and L. P. Chau, "Intra/Inter macroblock mode decision for error-resilient transcoding," *IEEE Trans. Multimedia*, vol. 10, no. 1, pp. 97–104, Jan. 2008.

[14] B. Li, T. Nanjundaswamy, and K. Rose, "An error-resilient video coding framework with soft reset and end-to-end distortion optimization," in *Proc. IEEE Int. Conf. Image Process.*, Beijing, China, 2017, pp. 1910–1914.

[15] Z. He, J. Cai, and C. W. Chang, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 12, no. 6, pp. 511–523, Jun. 2002.

[16] Y. Wang, Z. Wu, and J. M. Boyce, "Modeling of transmission-loss-induced distortion in decoded video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 6, pp. 716–732, Jun. 2006.

[17] Y. Zhang, W. Gao, Y. Lu, Q. Huang, and D. Zhao, "Joint source-channel rate-distortion optimization for H.264 video coding over error-prone networks," *IEEE Trans. Multimedia.* vol. 9, no. 3, pp. 445–454, Apr. 2007.

[18] Y. Zhang, and P. C. Cosman, "Fast mode decision for H.264 video coding in packet loss environment," in *Proc. Int. Conf. Image Process.*, Brussels, Belgium, 2011, pp. 3217–3220.

[19] Y. Wei, Y. Zhang, and J. Yan, "Fast mode decision for error resilient video coding," in *Proc. 16th Int. Workshop Multimedia Signal Process.*, Jakarta, Indonesia, 2014, pp. 1–6.

[20] G. Kulupana, D. S. Talagala, H. K. Arachchi, and A. Fernando, "Error resilience aware motion estimation and mode selection for HEVC video transmission," in *Proc. IEEE Int. Conf. Consum. Electron.*, Jan. 2016, pp. 85–86.

[21] G. Kulupana, D. S. Talagala, H. K. Arachchi, and A. Fernando, "End user video quality prediction and coding parameters selection at the encoder for robust HEVC video transmission," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 11, pp. 3367–3381, Nov. 2019, doi: 10.1109/TCSVT.2018.2879956.

[22] R. D. C. Cruz, "Low complexity adaptive intra-refresh rate for real-time wireless video transmission," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Conf.*, 2014. Q6

[23] R. Dela, C. Angelo, and D. C. Rhandley, "Error resilient joint source-channel adaptive intra-refresh rate for wireless video transmission," in *Proc. Int. Conf. Digital Signal Process.*, 2014, pp. 509–514.

[24] H. Chen, C. Zhao, M. T. Sun, and A. Drake, "Adaptive intra-refresh for low-delay error-resilient video coding," *J. Vis. Commun. Image Representation*, vol. 31, pp. 294–304, 2015.

[25] I. Ali, M. Fleury, and M. Ghanbari, "Content-Aware intra-refresh for video streaming over lossy links," in *Proc. Int. Conf. Consum. Electron.*, Las Vegas, NV, USA, 2012, pp. 118–119.

[26] S. Moiron and M. Ghanbari, "Enhanced error resiliency for video with cyclic intra-refresh lines," in *Proc. IEEE Int. Conf. Image Process.*, Brussels, Belgium, 2011, pp. 3229–3232.

[27] I. A. Ali, S. Moiron, M. Fleury, and M. Ghanbari, "Packet prioritization for H.264/AVC video with cyclic intra-refresh line," *J. Vis. Commun. Image Representation*, vol. 24, no. 4, pp. 486–498, 2013.

[28] S. Moiron, I. Ali, M. Ghanbari, and M. Fleury, "Limitations of multiple reference frames with cyclic intra-refresh line for H.264/AVC," *Electron. Lett.*, vol. 47, no. 2, pp. 103–104, Jan. 2011.

[29] J. Xu and Z. Wu, "Joint adaptive intra refreshment and unequally error protection algorithms for robust transmission of H.264/AVC video," in *Proc. IEEE Int. Conf. Multimedia Expo.*, 2006.

[30] F. Wan, "Traffic modeling and performance analysis for IPTV systems," Master's thesis, Southeast Univ., Jiangsu, China, 2005.

[31] S. Basso *et al.*, "Estimating packet loss rate in the access through application-level measurements," in *Proc. Assoc. Comput. Machinery's Special Interest Group Data Commun. Workshop Meas. Stack,* Helsinki, Finland, Aug. 17, 2012, pp. 7–12.

[32] M. Kazemi, R. Iqbal, and S. Shirmohammadi, "Joint intra and multiple description coding for packet loss resilient video transmission," *IEEE Trans. Multimedia.* vol. 20, no. 4, pp. 781–795, Apr. 2018.

[33] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 23–50, Nov. 1998.

[34] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74–90, Nov. 1998.

[35] K. H. Choi and D. H. Kim, "A downhill simplex approach for HEVC error concealment in wireless IP networks," in *Proc. IEEE Int. Conf. Consum. Electron.*, Las Vegas, NV, USA, 2016, pp. 143–146.

[36] Y. L. Chang, Y. A. Reznick, Z. Chen, and P. C. Cosman, "Motion compensated error concealment for HEVC based on block-merging and residual energy," in *Proc. Int. Packet Video Workshop*, San Jose, CA, USA, 2013, pp. 1–6.

[37] Y.-T. Peng and P. C. Cosman, "Weighted boundary matching error concealment for HEVC using block partition decisions," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 921–925.

[38] T. Lin, N. Yang, and R. Syu, "Error concealment algorithm for HEVC coded video using block partition decisions," in *Proc. IEEE Int. Conf. Signal Process. Commun. Comput.*, Aug. 2013.

[39] X. Qian, G. Liu, and H. Wang, "Recovering connected error region based on adaptive error concealment order determination," *IEEE Trans. Multimedia*, vol. 11, no. 4, pp. 683–695, Jun. 2009.

[40] W. Lie, C. Lee, C. Yeh, and Z. Gao, "Motion vector recovery for video error concealment by using iterative dynamic programming optimization," *IEEE Trans. Multimedia*, vol. 16, no. 1, pp. 216–227, Jan. 2014.

[41] M. Kazemi, R. Iqbal, and S. Shirmohammadi, "Redundancy allocation based on weighted mismatch-rate slope for multiple description video coding," *IEEE Trans. Multimedia*, vol. 19, no. 1, pp. 54–66, Jan. 2017.

[42] H. Hadizadeh and I. Bajic, "Rate-distortion optimized pixel-based motion vector concatenation for reference picture selection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 8, pp. 1139–1151, Aug. 2011.

[43] W. Tu and E. Steinbach, "Proxy-Based reference picture selection for error resilient conversational video in mobile networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 2, pp. 151–164, Feb. 2009.

[44] J. Carreira, P. A. Assuncao, S. M. M. de Faria, E. Ekmekcioglu, and A. Kondoz, "A two-stage approach for robust HEVC coding and streaming," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 8, pp. 1960–1973, Aug. 2018.

[45] E. Khan, S. Lehmann, H. Gunji, and M. Ghanbari, "Iterative error detection and correction of H.263 coded video for wireless networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 12, pp. 1294–1307, Dec. 2004.

[46] M. Ghanbari, "Standard codecs: Image compression to advanced video coding," *Appendix E*, 3rd ed. IET publ., 2011. Q7

[47] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans. Broadcast.*, vol. 50, no. 3, pp. 312–322, Sep. 2004.

Mohammad Kazemi received the B.Sc. degree from the Isfahan University of Technology, Isfahan, Iran, in 2003, and the M.Sc. and Ph.D. degrees from the Sharif University of Technology (SUT), Tehran, Iran, both in electrical engineering, in 2005 and 2012, respectively. He currently works with the Electrical Engineering Department, University of Isfahan, Iran, as an Assistant Professor. His current research interests include the areas of image/video processing, coding, and video transmission over lossy channels. He is also interested in digital systems design, and applications of artificial intelligence for image/video analysis.

Mohammad Ghanbari (IEEE M'78–SM'97–F'01–LF'14) is a Professor with the School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran, as well as an Emeritus Professor with the School of Computer Science and Electronic Engineering, University of Essex, Colchester, U.K. He has Q8 authored or coauthored 8 books and has registered for 13 international patents and authored over 750 technical papers on various aspects of video networking, many of which have had fundamental influences in this field. These include video/image compression, layered/scalable video coding, video over networks, video transcoding, motion estimation, and video quality metrics. Mr. Ghanbari's book Video Coding: An Introduction to Standard Codecs (IET Press, 1999) received the Rayleigh Prize as the Best Book of the Year 2000 by IET. He has also received several prizes, such as Reeves Prize for Best Paper Award in 1995 and 14th Khwarizmi International award for work on video networking in 2001.

Shervin Shirmohammadi (M'04–SM'04–F'17) received the Ph.D. degree in electrical engineering from the University of Ottawa, ON, Canada, where he is currently a Professor with the School of Electrical Engineering and Computer Science. He is Director of the Distributed and Collaborative Virtual Environment Research Laboratory, doing research in multimedia systems and networks, specifically video systems, gaming systems, and multimedia-assisted healthcare systems. The results of Dr. Shirmohammadi's research, funded by more than $14 million from public and private sectors, have led to over 350 publications, 3 Best Paper awards, over 70 researchers trained at the Postdoctoral, Ph.D., and Master's levels, over 20 patents and technology transfers to the private sector, and a number of awards. He is the Editor-in-Chief for the IEEE Transactions on Instrumentation and Measurement, an Associate Editor for the IEEE Transactions on Circuits and Systems for Video Technology, and an Associate Editor for *ACM Transactions on Multimedia Computing Communications and Applications*, having been numerously recognized as the Associate Editor for the Year by all three journals. He is a Key Member of the IEEE MMTC Multimedia Cloud Computing Interest Group, member of the IEEE CS Technical Committee on Multimedia Computing (TCMC), and member of the IEEE ComSoc Technical Committee on Multimedia Communications (MMTC). He is a winner of the 2019 George S. Glinski Award for Excellence in Research, a Lifetime Senior Member of the ACM, a University of Ottawa Gold Medalist, and a licensed Professional Engineer in Ontario.

GENERAL INSTRUCTION

- **Authors: We cannot accept new source files as corrections for your article. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.**

QUERIES

Q1. Author: Please confirm or add details for any funding or financial support for the research of this article.
Q2. Author: Please check and confirm that author name Mohammad Ghanbari in byline and biography are correct as set.
Q3. Author: Please check and confirm that postal codes of affiliations of authors Mohammad Kazemi and Mohammad Ghanbari are correct as set.
Q4. Author: Please check and confirm that page range in Refs. [3], [9], and [19], are correct as set.
Q5. Author: Please update with further details in Ref. [11], if available.
Q6. Author: Please provide page range in Refs. [22], [29], and [36], if available.
Q7. Author: Please check and confirm that Ref. [46] is correct as set. Also, provide location in the said reference, if available.
Q8. Author: Please provide complete educational details of author Mohammad Ghanbari in biography.

Intra Coding Strategy for Video Error Resiliency: Behavioral Analysis

Mohammad Kazemi, Mohammad Ghanbari, *Life Fellow, IEEE*, and Shervin Shirmohammadi, *Fellow, IEEE*

***Abstract*—One challenge in video transmission is to deal with packet loss. Since the compressed video streams are sensitive to data loss, the error resiliency of the encoded video becomes important. When video data is lost and retransmission is not possible, the missed data should be concealed. But loss concealment causes distortion in the lossy frame which also propagates into the next frames even if their data are received correctly. One promising solution to mitigate this error propagation is intra coding. There are three approaches for intra coding: intra coding of a number of blocks selected randomly or regularly, intra coding of some specific blocks selected by an appropriate cost function, or intra coding of a whole frame. But Intra coding reduces the compression ratio; therefore, there exists a trade-off between bitrate and error resiliency achieved by intra coding. In this paper, we study and show the best strategy for getting the best rate-distortion performance. Considering the error propagation, an objective function is formulated, and with some approximations, this objective function is simplified and solved. The solution demonstrates that periodical I-frame coding is preferred over coding only a number of blocks as intra mode in P-frames. Through examination of various test sequences, it is shown that the best intra frame period depends on the coding bitrate as well as the packet loss rate. We then propose a scheme to estimate this period from curve fitting of the experimental results, and show that our proposed scheme outperforms other methods of intra coding especially for higher loss rates and coding bitrates.**

***Index Terms*—Error resilient video coding, video error concealment, intra coding.**

I. Introduction

Nowadays, real-time digital video transmission over networks is very popular. Due to the tremendous volume of the raw video data, video compression is inevitable. But delivering compressed data over wired/wireless channels is challenging since the underlying networks are not always reliable and some data losses are usually experienced during transmission.

The erroneous and unreceived data corrupts the decompression process and the video fidelity. In this condition, using error concealment techniques alleviates the problem to some extent [1]. At high compression rates, data loss is more destructive; the higher the compression, the more the sensitivity to data loss. For this reason, High Efficiency Video Coding (HEVC), the latest standard video codec, is less error resilient than H.264/AVC [2]. Therefore, this makes HEVC video communication over error prone channels a challenging problem for researchers and practitioners in this field.

In error concealment techniques, the non-received data are estimated from the received ones. This is done by exploiting the spatio-temporal correlations among the available data at the area of missing information. However, the replaced data will not be exactly the same as the actual data; therefore, there exists a mismatch/distortion between them. If the recovered frame was used as the prediction reference at the encoder, its reconstructed erroneous part would propagate into the next frames at the decoder. In video coding, a large portion of compression comes from inter frame coding, but inter frame coding increases inter dependency and causes error propagation. In contrast, although intra coding is less efficient for compression, it mitigates the error propagation problem and could be used as a strong error resiliency tool, since it does not use prediction from the other frames [3]–[6]. Therefore, by intra coding, there is a trade-off between error resiliency and compression ratio. That is, intra mode for a block must be selected with sufficient care. For this reason, loss resiliency through intra mode is discussed in several works, as described in Section II.

In this paper we show both analytically and experimentally that the best strategy for intra coding is to code some selected frames deliberately in intra mode. For doing so, considering the transmission distortion, the decoder side distortion is formulated and an objective function is developed. Through optimization of this objective function, the optimal solution indicates that coding a whole frame as an I-frame is preferred over coding a certain number of blocks in the frames as intra mode. Our investigations show that the best approach to exploit the error resiliency of intra coding is to reduce the intra period instead of distributing the intra coded blocks among the frames. We had solved a similar problem for Multiple Description Coding (MDC) in [32], but the treatment of a single stream is different from the multi streams of MDC. MDC rarely deals with concealment distortion, since most of the time at least one description is available. Therefore,

Manuscript received January 28, 2019; revised June 17, 2019, August 15, 2019, October 5, 2019, and November 24, 2019; accepted November 25, 2019.
Q1 The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sanjeev Mehrotra. *(Corresponding author: Mohammad Kazemi.)*

Q2 M. Kazemi is with the Department of Electrical Engineering, University of Isfahan, Isfahan 81746-73441, Iran (e-mail: m.kazemi@eng.ui.ac.ir).
Q3 M. Ghanbari is with the School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran 1417466191, Iran, and also with the School of Computer Science and Electronic Engineering, University of Essex, CO4 3SQ Colchester, U.K. (e-mail: ghan@ut.ac.ir).
S. Shirmohammadi is with the Distributed and Collaborative Virtual Environments Research Laboratory, School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: shervin@eecs.uottawa.ca).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.
Digital Object Identifier 10.1109/TMM.2019.2957991

the objective functions and approaches taken for solving the problem here are different from [32].

The rest of the paper is structured as follows. Section II presents the related work, while in Section III the objective function is formulated. Through experimental results, this function is first simplified and then solved in Sections IV and V. The performance comparison with the other approaches for error resilient intra coding is presented in Section VI and the paper is concluded in Section VII.

II. Related Work

In the works presented in [7], [8], a recursive algorithm called ROPE has been developed. In this algorithm, by pixel-wise operations, the encoder estimates the receiver-side expected distortion which is then used for intra/inter mode decision. However, this algorithm is too complex due to its pixel-by-pixel computations. Its extension for bursty loss channels is presented in [9], and the extension of error resilient mode decision to motion estimation and also considering intra-frame prediction for intra modes is presented in [10]. ROPE was also used in [11] to optimally decide between Motion Vector (MV) replication or intra coding mode. Using ROPE for motion estimation and reference frame generation is presented in [12], while [13] discusses the intra/inter mode selection in video transcoding, where the lossy frame and its propagated error within the ROPE algorithm is exploited. Finally, the extension of ROPE for including constrained intra prediction as candidate modes, in addition to inter and intra modes, is presented in [14].

In another proposal described in [15], frame-level channel distortion is analyzed, where through linear models/approximations and end-to-end distortion optimization, a scheme for intra mode selection and rate control is developed. Some models for transmission distortion are presented in [16] where parameters such as intra prediction, deblocking filtering, sub-pixel motion estimation and the effect of decoder side temporal error concealment are taken into account. With the same approach, an end-to-end distortion modeling and optimization method was presented in [17] which is then used to develop a faster algorithm for intra/skip mode decision [18], [19]. In [20], motion estimation and mode decision in HEVC are performed based on error propagation. Another algorithm is presented in [21] where due to high sensitivity to error propagation, the algorithm selects the intra mode for the Prediction Units (PUs) much more than is required, especially for lower content videos. Even though the authors try to solve this issue with updating some parameters, the intra rates are still high and this degrades its performance at low Packet Loss Rates (PLRs) and low bitrates.

A fast intra mode decision for loss resiliency is developed in [22] where, through a linear model, the distortion is estimated and an optimal value for Intra Refresh Rate (IRR) is obtained. IRR or simply the intra rate is the number of blocks coded in intra mode divided by the total number of blocks in the frame. A modified model for considering the role of IRR in bitrate and distortion is introduced in [23]. Using a linear model and considering motion activity and PLR, the optimal IRR and the intra coded MBs' pattern are discussed in [24]. It is shown in [25] that for low activity sequences, cyclic intra coding of MBs is more effective than periodic I-frames, and vice-versa for highly active videos. Combining cyclic intra-refreshing with unequal error protection is introduced in [26], [27], though intra-refresh is in conflict with multiple reference selection, as shown in [28]. Error propagation is formulated and the IRR is obtained in [29], then the selected MBs for intra coding are grouped into a common slice group where they are then protected with stronger channel codes.

The above mentioned intra coding research works can be categorized into two groups: those which discuss Selecting Intra Mode (SIM) and those which discuss Intra Refresh. In SIM methods, the cost function for inter/intra mode decision is modified to take into account the lossy channel and the transmission distortion; examples are [7]–[21]. In Intra Refresh, the intra rate is determined. Then, the intra coded blocks can be selected randomly, or selected with vertically or horizontally ordered columns/rows, provided that they do not overlap in the successive frames such that the blocks in all positions are intra refreshed after a while; examples are [22]–[29].

Our work is different form the above works since our formulation and optimization leads to a straightforward and specific solution: reduce the intra period (coding a whole frame in intra mode) to achieve the best error resiliency outcome of intra coding, instead of distributing the intra coded blocks within the frames of the GOP. Afterwards, the best intra period, which depends on the content and channel loss rate, is approximately but simply obtained from a function, without additional computational complexity. The experimental results confirm the efficient performance of the proposed scheme, for various loss rates and video contents.

Another tool which can help to prevent error propagating is Reference Picture Selection (RPS) which allows the encoder to select one or two frames from a list as inter-prediction references for each prediction block. Several reference frames are examined for the best rate-distortion coding. For error resiliency, this feature is usually in conjunction with decoder feedback which informs the encoder not to select the erroneously received frames as the prediction reference [42], [43]. However, this feedback information is not available in many applications; e.g., multicast and broadcast applications, or pre-recorded video on demand applications. Moreover, responding to various receivers concurrently is not practical, or the feedback messages might be received too late. RPS without a back channel and for error resiliency has been presented in [44]. In this work, the authors propose not to use a single frame as prediction reference of the PUs, but to select from a list of reference frames such that all frames in the list are selected uniformly. However, this method needs to consider a list of frames as candidate reference frames, so it has the complexity of the multi-reference prediction. For example, for five candidate reference frames, the computational complexity of Motion Estimation and Mode Decision becomes five times more. The required encoder/decoder buffer size becomes larger with the number of reference frames as well. In the error resiliency of the intra coding method proposed in our paper, the only required information is channel loss rate, without any additional complexity in the encoder/decoder. It is worth noting that full frame intra prediction provides random access to video stream, but it also generates large peaks of bitrates. However, such large spikes in the bitrate can be eased with either a few frames delay, which is acceptable in many applications, or compensated by statistical multiplexing with lower bitrates of P- and B-frames of other video streams. Therefore, there are many circumstances that intra coding is a feasible solution while RPS cannot be employed.

Our application scenario is distribution/broadcast of video without assuming any specific limits on delay or bandwidth, and assuming the PLR is known by the encoder. PLR can be estimated with or without back channel. The application scenario without back channel compromises the great majority of current video distribution/broadcast systems today. In such a scenario, we assume PLR is estimated by the service/network operator from the history of the channel for the specific weekday and time of the day, or is calculated offline, or is tested by small ping packets periodically, or by traffic modeling [30], or by one-way estimation methods that use message segment size, goodput, and delay [31] all estimated at the sender side, or by any other estimation means. Of course PLR can also be estimated with back channel, and this would be more accurate. In the latter scenario, our proposed method is applicable if this back channel cannot inform us of the lost packets immediately due to the delay in feedbacks or a long Round-Trip Time; therefore, retransmission of the lost data or Reference Picture Selection is not possible. This scenario is also assumed by other credible works [8], [13], [15], [17], [21]. Finally, since our method introduces a small delay of about 92 msec, as will be shown later in the paper, we assume that such small delay is acceptable for the target application.

III. The Objective Function for Error Resilient Intra Coding

For error resilient coding, the following two aspects of intra coding must be considered:

First aspect - intra coding prevents temporal error propagation, since it has no reference to the other frames. In intra coding of advanced video codecs, such as H.264/AVC and H.265/HEVC, pixels of the adjacent blocks are used as intra prediction references, and these references (in encoded form) together with the residual data are encapsulated and transmitted in a single packet. However, if the reference pixels had been encoded in inter mode, they themselves might be erroneous, even if the residual data is received correctly. In this case, temporally propagated errors can propagate spatially into the intra-coded blocks. To avoid this condition and exploit the error propagation prevention provided by intra coding, the option of "ConstrainedIntraPred" can be enabled, which restricts the intra mode to use only the pixels of adjacent intra coded blocks as prediction references. This way, the received intra coded PUs are correctly decodable.

Second aspect - in no loss conditions, inter mode is obviously used more often than intra mode, because inter-coded blocks have lower bitrates than intra coded ones. By enabling the "ConstrainedIntraPred" option, the compression efficiency of intra mode is reduced even more, but it is beneficial for error resiliency [16] when we do have losses.

Therefore, in deciding to code a block in intra mode, there is a trade-off between bitrate and error resiliency. In this section, an objective function is developed which, rather than optimizing the encoder side rate-distortion, the decoder side rate-distortion is optimized. In other words, taking into account the channel distortion, the receiver side distortion is estimated at the encoder which is then used as the objective function.

Intra/inter mode selection is conventionally carried out based on the following Lagrangian cost function [33], [34]:

$$cost = D_q + \lambda R \tag{1}$$

where D_q is the quantization distortion in Mean Squared Error (MSE), λ is the Lagrangian coefficient and R is the number of required bits. This cost function is computed for the candidate modes and the mode with the lowest cost is selected as the final mode. However, this cost function does not take the transmission distortion into account. To consider it, with the same approach as presented in [15]–[17], the rate overall-distortion in a frame is represented in (2). The assumption behind this equation is that PLR is known at the transmitter side.

$$D^{(1)} = (1 - PLR)\, D_q^{(1)} + PLR\ D_{conceal}^{(1)} \tag{2}$$

where $D^{(1)}$, $D_q^{(1)}$ and $D_{conceal}^{(1)}$ are the expected total distortion, the quantization distortion, and the error concealment distortion for frame 1, respectively. The *expected* distortion means the average distortion seen over a long enough duration, or equivalently over a variant enough packet loss pattern, the latter used in our simulation. Note that the concealment distortion, $D_{conceal}$ in (2), is the distortion when all packets of the frame are lost and the frame is error concealed. It is evident that the frame is transmitted by a single packet; however, as shown in the Appendix, this is also valid when the frame is encoded into n packets and the packets convey the same amount of information.

Frame 0 is the initial I-frame of the sequence which is assumed to be received correctly. For frame 1, depending on whether its packets are received or not, the distortion will be $D_q^{(1)}$ or $D_{conceal}^{(1)}$, respectively. For frame 2, it becomes:

$$\begin{aligned} D^{(2)} = {} & (1 - PLR)\, D_q^{(2)} \\ & + PLR\, D_{conceal}^{(2)} + PLR\left[1 - \beta^{(2)}\right]\Delta^{(1)} \end{aligned} \tag{3}$$

where $\beta^{(2)}$ is the intra rate of frame 2 and

$$\Delta^{(1)} = E\left[\left(F_q^{(1)} - F_{conceal}^{(1)}\right)^2\right] \tag{4}$$

is the mean squared difference between frame 1 decoded correctly ($F_q^{(1)}$) and loss concealed ($F_{conceal}^{(1)}$); i.e., $\Delta^{(1)}$ denotes the *Mismatched Distortion* for frame 1 caused by error concealment. We assume that only the previous frame is used as prediction reference, as happens most of the times in encoders. Enabling multi-frame prediction results in a slight improvement in quality but at the cost of significant computational cost.

Equations (3) and (4) show that the quality of frame 1 directly affects the quality of frame 2, and its effect is controlled by parameter $\beta^{(2)}$ in (3). At larger β, the impact of mismatched distortion is clearly reduced, since intra coded PUs do not refer to the previous frame. As shown in [15], the quantization and mismatched distortions are independent of each other and one can simply write:

$$D_{conceal}^{(2)} = D_q^{(2)} + \Delta^{(2)} \tag{5}$$

Substituting (5) into (3) gives:

$$\begin{aligned} D^{(2)} &= D_q^{(2)} + PLR\left[\Delta^{(2)} + \left(1-\beta^{(2)}\right)\Delta^{(1)}\right] \\ &= D_q^{(2)} + PLR\,\Delta_{accum}^{(2)} \end{aligned} \tag{6}$$

in which

$$\Delta_{accum}^{(2)} = \Delta^{(2)} + \left(1-\beta^{(2)}\right)\ \Delta^{(1)} \tag{7}$$

is the *Accumulated Mismatched Distortion* seen in frame 2. It is evident that for frame 1, $\Delta_{accum}^{(1)} = \Delta^{(1)}$ and then (7) can be rewritten as:

$$\Delta_{accum}^{(2)} = \Delta^{(2)} + \left(1-\beta^{(2)}\right)\ \Delta_{accum}^{(1)} \tag{8}$$

Following the above concept, the distortion for the nth frame is:

$$D^{(n)} = D_q^{(n)} + PLR\,\Delta_{accum}^{(n)} \tag{9}$$

where

$$\begin{aligned} \Delta_{accum}^{(n)} &= \Delta^{(n)} + \left(1-\beta^{(n)}\right)\Delta_{accum}^{(n-1)} \\ \Delta_{accum}^{(0)} &= 0 \end{aligned} \tag{10}$$

and

$$\Delta^{(n)} = E\left[\left(F_q^{(n)} - F_{conceal}^{(n)}\right)^2\right] \tag{11}$$

Therefore, the distortion over a GoP of N frames is as given in (12) (as already mentioned, the 0th frame of the GoP is excluded from the summation):

$$D_{GoP} = \sum_{i=1}^{N} D^{(i)} = \sum_{i=1}^{N}\left(D_q^{(i)}(\boldsymbol{\beta}) + PLR\,\Delta_{accum}^{(i)}(\boldsymbol{\beta})\right) \tag{12}$$

where $\boldsymbol{\beta} = [\beta^{(1)}, \beta^{(2)}, \beta^{(3)}, \ldots \beta^{(N)}]$ is the vector intra rates for the N frames of the GoP. Quantization Parameter (*QP*) is excluded from this formulation, since its variation is usually ± 1 units at the given bitrate, except for sudden changes; e.g., scene-cut or fast/non-translational motions which is difficult for compensation with inter prediction. Therefore, we can assume that *QP* does not have significant changes for the N frames under consideration.

With the aim of maximizing the received video quality, the objective function with a constraint on the overhead bitrate of intra coding is:

$$\begin{aligned} &\min_{\boldsymbol{\beta}}\left\{\sum_{i=1}^{N}\left(D_q^{(i)}(\boldsymbol{\beta}) + PLR\,\Delta_{accum}^{(i)}(\boldsymbol{\beta})\right)\right\} \\ &\text{s.t.} \sum_{i=1}^{N} R_{intra}^{(i)}(\boldsymbol{\beta}) \le R_{red} \end{aligned} \tag{13}$$

where $R_{intra}^{(i)}$ is the number of additional bits needed for intra encoding of the ith frame according to the intra rate of $\beta^{(i)}$; that is, if $\beta^{(i)} = 0$, then $R_{intra}^{(i)} = 0$ and no block is codded in intra mode for error resiliency. The term R_{red} in (13) is the total redundancy budget allowed for these N frames for intra coding, which in turn is related to the PLR and the required degree of error resiliency. Increasing the intra rates of frame i; i.e., $\beta^{(i)}$, reduces $\Delta_{accum}^{(i)}$ (see equation (10)) but in turn increases bitrate usage.

IV. Solving the Objective Function

In this section, a solution to the constrained problem of (13) is driven through approximation. The behaviors of terms in this equation are observed and approximated through matching them to the experimental results. The experiments settings, the simplifying approximations, and the solution to the problem are provided in subsections A, B and C, respectively.

A. The Error Concealment Strategy

An important part of distortion in (13) belongs to the distortions caused by error/loss concealment. Error concealment techniques can be categorized into spatial and temporal domain processing techniques. In the spatial domain, the lost area of the frame is concealed using the spatially neighboring pixels. These methods exploit the correlations that usually exist among the neighboring pixels. In the temporal processing techniques, the contents from the previous and/or the future frames are addressed by MVs and used for temporal replacement. The actual MVs are not available and must be estimated or recovered first by the temporal loss concealment methods.

If the lost area is large, spatial domain is not effective, as pixels are very far apart from each other to be useful. The reason for dealing with large lost areas in HEVC is the size of its Coding Tree Unit (CTU), which can be as large as 64×64 pixels. An integer number of CTUs are regarded as one slice and an integer number of slices are encapsulated into a single transmission packet. Therefore, packet losses in HEVC streams affect a significant portion of the picture area especially for smaller picture sizes. As a result, temporal error concealment in HEVC streams is more applicable than spatial concealment.

Actually, exploiting the temporal frames' MVs will provide higher quality error concealment. One simple yet efficient technique is the Motion Copy algorithm where the MV of the collocated block is simply used for motion compensated temporal replacement. If the collocated block is coded in intra mode, Zero MV is used. However, in the case of having a high percentage of intra coded blocks, this approach is not efficient due to lack of MVs for intra blocks. For intra coded blocks, one solution is to recover the MVs by Boundary Matching Algorithm (BMA). A suggestion is to combine Motion Copy for inter coded and reliable collocated blocks, and BMA for intra coded or unreliable collocated blocks, as presented in [35]. The blocks with high residual signals are labeled as unreliable blocks and their MVs are not used for MV recovery. In [35], loss concealment is performed in two stages: firstly, the lost area is replaced using the Motion Copy algorithm. Then, for the unreliable MVs,

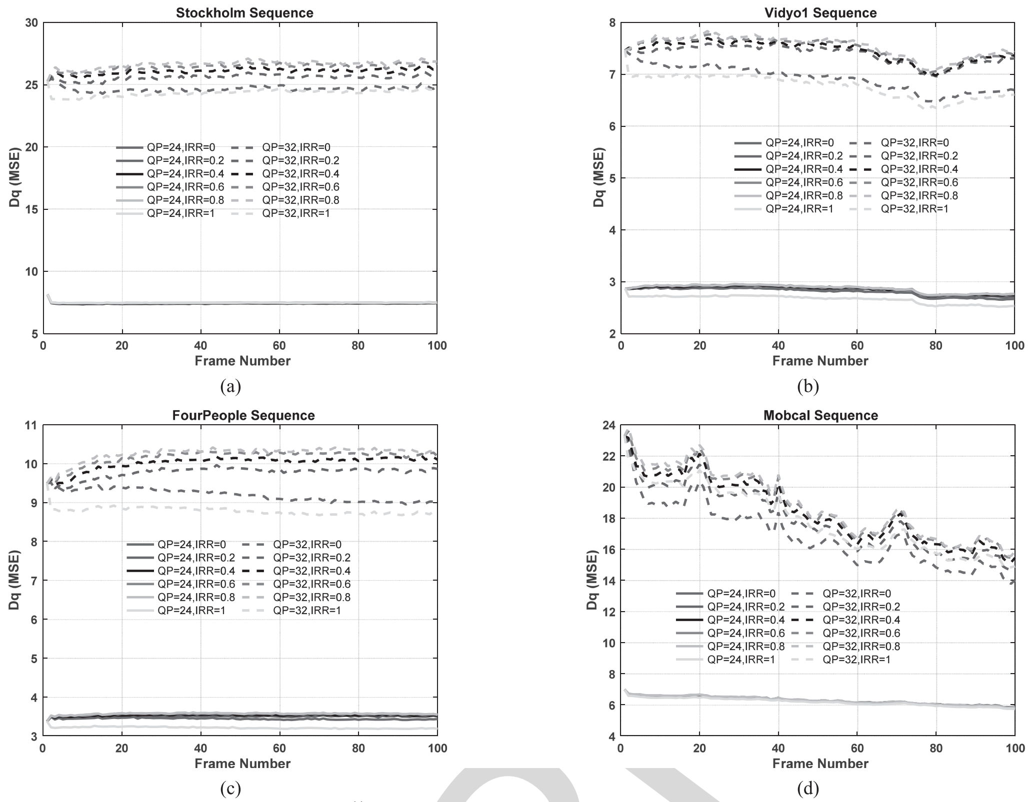


Fig. 1. The variation of $D_q^{(i)}(\beta^{(i)})$ with respect to IRR for test sequences and two values of *QP*.

MVs are obtained using BMA. Some other methods such as [36]–[37] are proposed for HEVC error concealment, but they work again based on the spatially close boundary pixels, which are not always available in the actual scenarios. The algorithm presented in [38] works based on MV extrapolation but with applying higher weights to the MVs belonging to the larger partitions. However, the problem is that this algorithm is based on the MVs of the blocks of the previous frame and therefore it is efficient when there are a few intra coded blocks. The method of [39] is proposed for error concealment of a sequence of successive MBs in H.264/AVC. This error concealment method does not need to know the MVs' neighboring spatially or temporally missed blocks, and instead estimates them by BMA. This method is also useful for large area losses of HEVC. The challenge in using BMA is the fact that the error concealment of one block will affect the error concealment of the following blocks as well. One solution, proposed in [39], is rank ordering the MBs for error concealment based on the texture of the available MBs in the surrounding of the lost area. A missed MB with a higher texture around it will be error concealed with higher priority. The criterion for the higher texture is the standard deviation of the luminance pixel values. Another solution for considering the interaction of loss concealed blocks is presented in [40], but it imposes significant computational complexity without considerable improvement.

In this paper, three techniques are used for error concealment: Motion Copy, the method presented in [35], and the method presented in [39]. The first two methods are appropriate when the blocks in the earlier frame are encoded mostly in inter mode, and the third method is suitable when the blocks in the earlier frame are encoded mostly or completely in intra mode. Then, the highest quality output is selected and used for the measurements.

It is worth noting that throughout the paper error concealment and loss concealment are used interchangeably, but in fact loss concealment is carried out. The reason is that in highly error prone networks, such as wireless networks, severely erroneous packets cannot be corrected and they are regarded as lost packets by the decoder. However, if the used entropy coder is symmetric, such as that of the H.263 codec, then parts of the data can be retrieved and the lossy area can be less than that of whole packet loss [45]. Since H.265/HEVC does not use symmetric entropy coder, then there would not be any retrieval of erroneous parts and the whole packet can be regarded as lost. Hence, loss concealment is a proper choice.

B. Simplifying the Objective Function

First, the quantization distortion does not significantly change with parameter $\beta^{(i)}$. That is, $D_q^{(i)}(\beta^{(i)})$ is approximately constant when $\beta^{(i)}$ varies from minimum $(\beta = 0)$ to maximum $(\beta = 1)$. This can be verified from the simulation results shown in Fig. 1. In this figure, four HD test video sequences are coded with HM16.0, the reference software of HEVC, at two values of *QP* and six values of IRR. The tested video sequences are *Stockholm*, *Vidyo1*, *FourPeople* and *Mobcal*. For the given intra rates, a sufficient number of PUs with sizes of 16×16 pixels

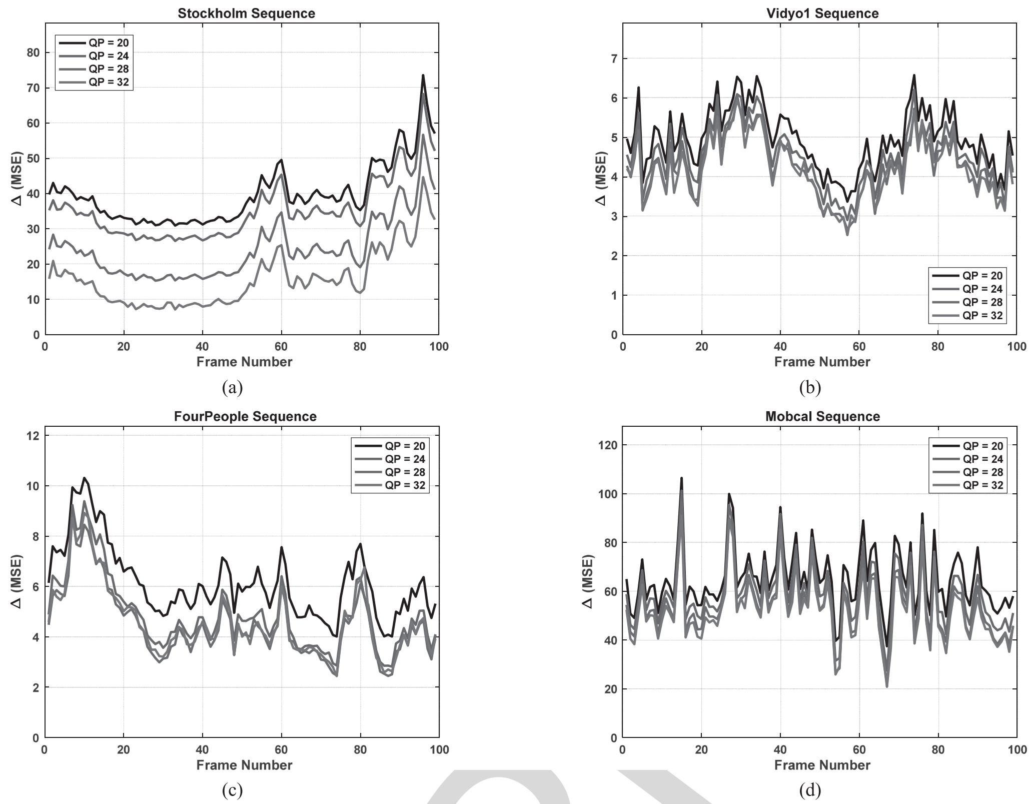


Fig. 2. The variation of mismatched distortion ($\Delta^{(i)}$) for the frames of test sequences for four values of QP.

are selected randomly and are forced to intra coding mode. As already mentioned, the randomly selected PUs do not overlap in the frames.

Fig. 1 shows the ignorable changes of $D_q^{(i)}$ with $\beta^{(i)}$. It can also be seen that the variation in $D_q^{(i)}$ for $QP = 24$ is less than that of $QP = 32$. The mathematical reason is that for smaller QPs, or smaller quantization step sizes (Q_{SS}), the high bitrate approximation is more accurate and the quantization distortion is nearly equal to $\frac{Q_{SS}^2}{12}$ [17]. This is fixed for various signals, independent of inter or intra coding. However, the difference in $D_q^{(i)}$s for various βs is still ignorable, even for $QP = 32$. Therefore, $D_q^{(i)}$ is fixed with the optimization arguments and equation (13) can be simplified as:

$$\min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{N} \left(PLR\, \Delta_{accum}^{(i)} (\boldsymbol{\beta}) \right) \right\}$$
$$\text{s.t.} \sum_{i=1}^{N} R_{intra}^{(i)} (\boldsymbol{\beta}) \leq R_{red} \tag{14}$$

An important term in (14) is $\Delta_{accum}^{(i)}$ which is the multiply-accumulated of mismatched distortions $\Delta^{(i)}$s, with the multiplication coefficient of $(1 - \beta^{(i)})$, as given in (10). Therefore, the variation of $\Delta^{(i)}$ per frames is important in the behavior of the objective function of (13). To measure $\Delta^{(i)}s$, the frames are first error concealed with the strategy given in IV.A, and then $\Delta^{(i)}$ is calculated by (11). The results are shown in Fig. 2 where it can be seen that, most of the times and with a good approximation, the frames of a sequence have close mismatched distortions, that is:

$$\Delta^{(1)} \cong \Delta^{(2)} \cong \cdots \cong \Delta \tag{15}$$

Even though it might not be valid for all frames, the variations are smooth in the windows of N frames, as large as the usually used GoP sizes (30–60 frames). This assumption may not be much accurate; however, this assumption, by nature, is similar to the assumption made in Rate-Control (RC) algorithms. In RC algorithms, the goal is to control the total bitrate to be less than the given bound with minimum fluctuation in the quality. Therefore, for a real-time RC, the encoder must assume that the future frames have almost the same behavior in the view of compression properties. Even though this assumption is not always valid, it is very efficient and helpful in practice. Similar to RC algorithms, we can assume that the frames behave similarly in the view of mismatched distortion. Therefore, with the assumption of (15), $\Delta^{(i)}$ is fixed for the frames, and since the employed loss concealment strategy is not much sensitive to the intra/inter coding, it is also fixed with $\boldsymbol{\beta}$. Therefore, by substituting the recursive formula given in (8), the objective function of (14) can be expanded as follows:

$$\min_{\boldsymbol{\beta}} \left\{ PLR.\, \Delta \left[N + \sum_{i=1}^{N} \left(\sum_{j=1}^{N-(i-1)} \left(\prod_{k=0}^{i-1} \left(1 - \beta^{(j+k)} \right) \right) \right) \right] \right\}$$
$$\text{s.t.} \sum_{i=1}^{N} R_{intra}^{(i)} (\boldsymbol{\beta}) \leq R_{red} \tag{16}$$

where PLR and Δ are assumed constant during optimization.

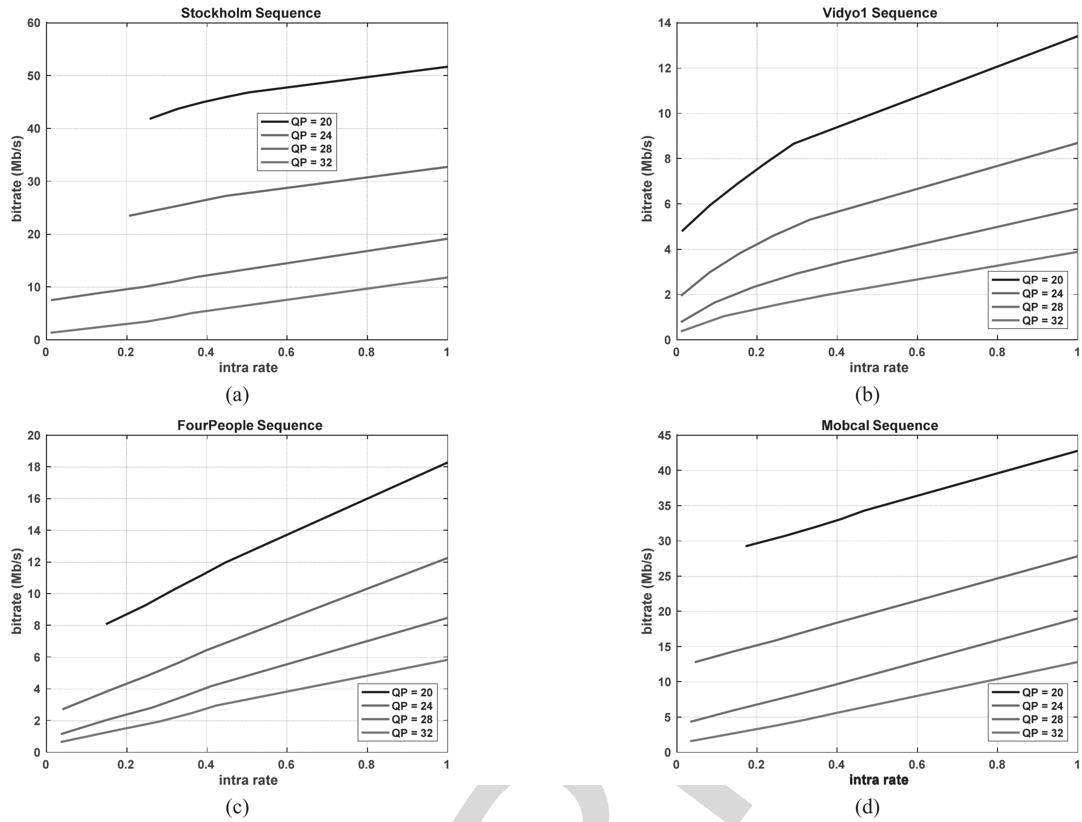


Fig. 3. The variation of bitrate with intra rate for various *QP*s and test sequences.

Now, for the constraint of (16), we do another simulation: the variation of frame bits when β varies from 0 to 1, as depicted in Fig. 3. In this figure, the average bitrates needed for sending the encoded video frames are measured and shown against the intra rate.

From Fig. 3, we can generally assume that the bitrates are increased almost linearly with β. That is:

$$R^{(i)} = R_0^{(i)} + R_{intra}^{(i)} = R_0^{(i)} + \alpha^{(i)}\ \beta^{(i)}$$
$$\Rightarrow\ R_{intra}^{(i)} = \alpha^{(i)}\ \beta^{(i)} \tag{17}$$

where $R_0^{(i)}$ is the bitrate of the i^{th} frame for $\beta = 0$. The figure shows that the curves almost have the same slope; that is, they have the same α defined in (17). For this reason, the constraint term in (16) can be stated as:

$$\sum_{i=1}^{N} \left(\alpha\ \beta^{(i)}\right) \leq R_{red} \tag{18}$$

or equivalently

$$\sum_{i=1}^{N} \beta^{(i)} \leq \beta_{red} \tag{19}$$

where β_{red} is the intra rate budget; i.e., the sum of total intra rates allowed to be assigned to these N frames. Therefore, the objective function of (16) is simplified as

$$\min_{\boldsymbol{\beta}} \left\{ PLR.\ \Delta \left[N + \sum_{i=1}^{N} \left(\sum_{j=1}^{N-(i-1)} \left(\prod_{k=0}^{i-1} \left(1 - \beta^{(j+k)}\right)\right)\right)\right]\right\}$$
$$\text{s.t.}\ \sum_{i=1}^{N} \beta^{(i)} \leq \beta_{red} \tag{20}$$

Since we can assume that PLR and Δ are fixed during optimization, the problem in (20) can be rewritten as:

$$ErrorPro_{min} = \min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{N} \left(\sum_{j=1}^{N-(i-1)} \left(\prod_{k=0}^{i-1} \left(1 - \beta^{(j+k)}\right)\right)\right)\right\}$$
$$\text{s.t.}\ \sum_{i=1}^{N} \beta^{(i)} \leq \beta_{red} \tag{21}$$

It is worth mentioning that for simplicity of deriving the objective function, without loss of generality, there are no B-frames. This is because, in general, some B-frames maybe used as prediction reference like P-frames causing error propagation, while others are not used as the reference and hence do not propagate the errors but they become erroneous. Modifying our formulations to highlight this matter makes the equations more complicated without giving the required information to the reader.

TABLE I
THE OUTPUT OF THE OBJECTIVE FUNCTION OF (21) FOR SOME VALUES OF β_{red}

Range of β_{red}	Intra rate of the frames															Constraint
	$\beta^{(1)}$	$\beta^{(2)}$	$\beta^{(3)}$	$\beta^{(4)}$	$\beta^{(5)}$	$\beta^{(6)}$	$\beta^{(7)}$	$\beta^{(8)}$	$\beta^{(9)}$	$\beta^{(10)}$	$\beta^{(11)}$	$\beta^{(12)}$	$\beta^{(13)}$	$\beta^{(14)}$	$\beta^{(15)}$	
$\beta_{red} \le 1$	0	0	0	0	0	0	0	a_1	0	0	0	0	0	0	0	$a_1 = \beta_{red}$
$1 < \beta_{red} \le 1.45$	0	0	0	a_1	0	0	0	1	0	0	0	a_2	0	0	0	$a_1 + a_2 = \beta_{red} - 1$
$1.45 < \beta_{red} \le 2$	0	0	0	0	a_1	0	0	0	0	1	0	0	0	0	0	$a_1 = \beta_{red} - 1$
	0	0	0	0	0	1	0	0	0	0	a_1	0	0	0	0	
$2 < \beta_{red} \le 2.71$	0	0	0	0	1	0	0	0	0	1	0	0	a_1	0	0	$a_1 = \beta_{red} - 2$
	0	0	0	0	1	0	0	a_1	0	0	1	0	0	0	0	
	0	0	a_1	0	0	1	0	0	0	0	1	0	0	0	0	
$2.71 < \beta_{red} \le 3$	0	0	0	a_1	0	0	0	1	0	0	0	1	0	0	0	$a_1 = \beta_{red} - 2$
	0	0	0	1	0	0	0	a_1	0	0	0	1	0	0	0	
	0	0	0	1	0	0	0	1	0	0	0	a_1	0	0	0	
$3 < \beta_{red} \le 3.6$	0	a_1	0	1	0	a_2	0	1	0	a_3	0	1	0	a_4	0	$a_1 + a_2 + a_3 + a_4 = \beta_{red} - 3$

C. Solving the Simplified Objective Function

In this subsection, the solution for the objective function given in (21) is discussed. The problem is actually minimizing the error propagation at a given intra rate budget. As an example, this problem is solved for $N = 15$ and various values of β_{red}; the results are tabulated in Table I.

As can be seen from this table, when $\beta_{red} \le 1$, the best frame for intra coding is the middle frame. If there was more intra budget; i.e., $1 < \beta_{red} \le 1.45$, $\beta^{(4)}$ and $\beta^{(12)}$ begin to grow irrespective of whether the intra rate is allocated to the 4th frame or the 12th frame. However, when intra rate budget exceeds 1.45, the optimization function given in (21) recommends other frames for intra coding to be chosen; e.g., frames 5 and 10, where frame 10 is coded wholly as I-frame, and frame 5 has partially intra coded blocks. Equivalently, another package is frames 6 and 11, where frame 6 is now selected for I-frame coding. One can see that these two packages produce the same obstacle against the error propagation.

For some other regions of β_{red}, the selected frames are given in Table I. One important point is changing the intra coded frame candidates imposed by the objective function of (21). The reason is that, if β_{red} is between two integers K_1 and K_2; that is $K_1 < \beta_{red} < K_2$, the optimizers may decide to add another frame for intra coding in addition to K_1 frames (e.g., one frame between them), or decide to select K_2 frames for intra coding and reduce the intra rate of one of them to comply with the bound of β_{red}. Clearly, if β_{red} is close to K_1, the former case happens, and the latter case happens when β_{red} is close enough to K_2. However, as shown in Fig. 4, $ErrorPro_{min}$ behaves continuously at these border points of β_{red}. In each interval shown by broken lines, the intra coding frame candidates are the same where one or more appropriate frames of these candidates consume the allocated intra coding budget. As already mentioned, the slope of decay in $ErrorPro_{min}$ in each interval is constant. If β_{red} becomes larger than 7, the frames are alternately coded as I-frame; that is the GOP structure is IPIPIP, and now all P-frames have the same priority for intra rate for all β_{red} amounts; therefore; there are no broken lines in Fig. 4 for $\beta_{red} > 7$.

This solution proves that to achieve the best error resiliency for intra coding, the best strategy is to concentrate on intra coding the blocks in the middle frame of the GoP, such that the entire frame is coded as an I-frame. If the intra rate budget (or accordingly bitrate budget) allows, more frames can still be coded in intra mode. In other words, the output of the objective function is to reduce the intra period; this strategy leads to smaller error propagation and hence higher video quality for lossy channels, compared to the case that intra coded blocks are distributed among the frames.

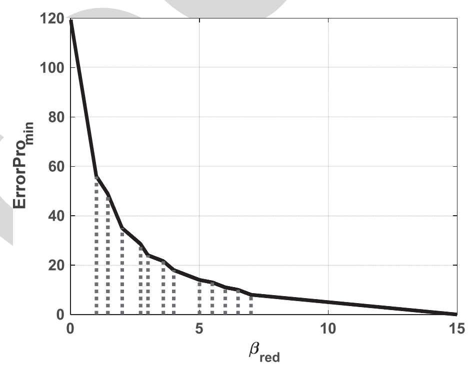


Fig. 4. Behavior of $ErrorPro_{min}$ with β_{red}.

V. THE OPTIMAL VALUE FOR THE INTRA PERIOD

As shown in Table I, at a larger β_{red}, the number of I-frames in the GoP can increase. This is in favor of mitigation of error propagation; however, the required bitrate for sending the video is increased since the compression ratio is decreased.

Having more I-frames is justified in channels with higher loss rates and vice versa. Therefore, PLR and the coding bitrate affect the best value for β^*_{red}. As shown in Table I, β^*_{red} is directly related to Intra Period (*IP*); therefore, the problem of finding $\beta_{red-opt}$ is equivalent to finding an optimal value for *IP*, denoted as IP^*. However, to find IP^* analytically, one must know the rate-distortion behavior of the frames of the GoP under consideration; that is, the behavior of future frames must be known *a priori*, which is not possible unless it is estimated based on the frames' history similar to the work presented in [41]. This

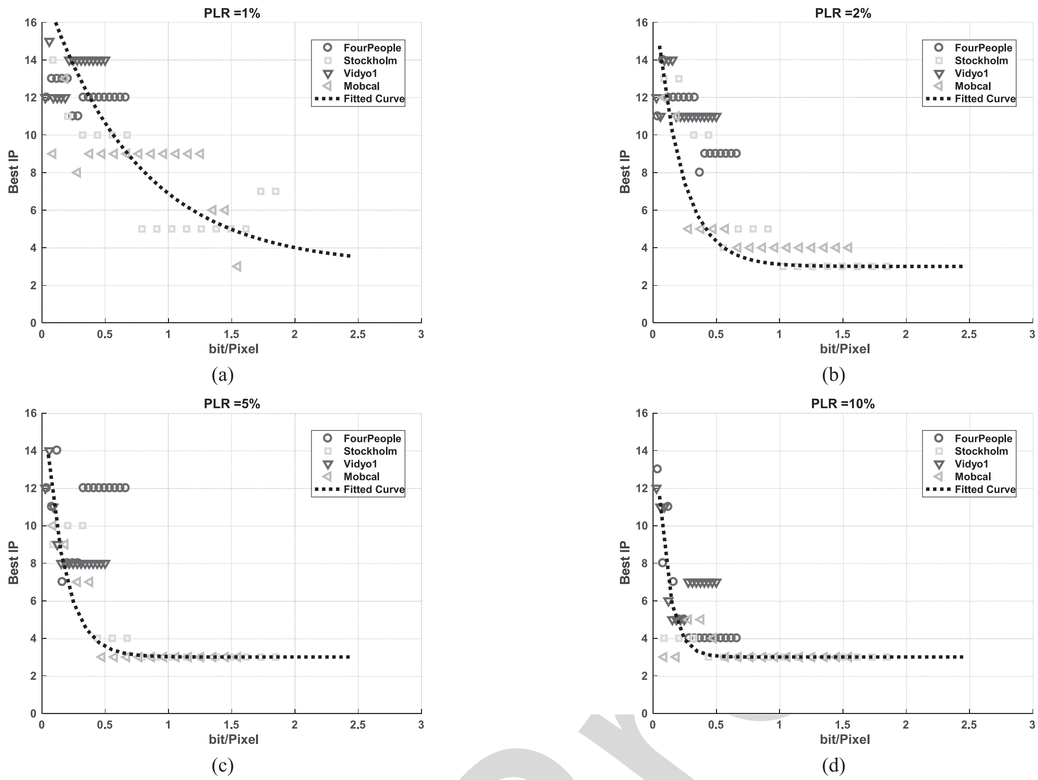


Fig. 5. The best *IP* versus bit/pixel for various PLRs and test sequences.

way, the problem can be solved using classical constraint optimization approaches. However, the complexity of the problem and non-trivial solutions have encouraged us to use an empirical approach. Therefore, IP^* has been found through experimental measurements, as follows.

The videos are encoded in Slice Mode, and each slice contains an integer number of CTUs in raster scan order. Each coded slice can be as large as 1500 bytes, meeting the Maximum Transmission Unit (MTU) of the network, and transmitted as a single packet. The channels experience a burst form of loss generated by Elliot-Gilbert model [46]. At each PLR, 40 packet loss patterns with an average burst length of three packets are generated and applied on the bit streams.

The video sequences are encoded at various values of IP; $IP = M$ means every *M*th frame of the sequence is coded as an I-frame. For example, for $IP = 3$, there are two P-frames after each I-frame, and this pattern is repeated throughout the sequence. In a GoP of 30 frames, the videos are encoded with $IP = 1, 2, \ldots, 15$ (for GoP of N frames, $IP > \frac{N}{2}$ is not reasonable). The compressed bit stream is subjected to a specific PLR, and the decoded video is loss concealed (as given in IV.A) and the resulting quality is measured. Video quality is measured in terms of Video Quality Model (VQM) [47] and its average index taken over the loss patterns is calculated. VQM is a video quality assessment method which considers both spatial and temporal distortions, so it is quite suitable to our case. For each test video and at the given PLR, the best IP which provides the best quality (i.e., the lowest VQM index, since higher quality is equivalent to lower VQM index) at the corresponding bitrate is selected. Fig. 5 shows the best *IP*s as a function of bit/pixel for four sequences.

It can be seen that the best *IP* becomes smaller at higher bitrates and higher PLRs. Even though some points are not close to the others, they can be fitted on decaying exponential functions, as shown in Fig. 5. The fitted curves can be formulated with the following equations:

$$IP^* = 3 + 15 \exp\left(-\frac{R}{R_0}\right)$$
$$R_0 = 0.15 + 1.4575\ \exp\left(-\frac{PLR}{0.01}\right) \tag{22}$$

where R is the bit per pixel. Clearly, IP^* obtained from (22) must be rounded to the nearest integer number. Even though the decimal values are also applicable, our empirical approach and curve fitting is not accurate enough for extracting decimal values for frames' intra rates.

One issue is the fact that coding a frame fully in Intra mode might cause sudden changes in the bitrate and hence more congestion in the lossy channels. However, for numerous applications, like video broadcast, streaming, multicasting etc., one needs to play the video at almost any time during transmission. This facility can only be provided by Intra coded frames. On the issue of increased I-frame bitrates, one should note that in these applications, normally several video flows are multiplexed, such that higher bitrates of I-frames coincide with lower bitrates of many P- and B-frames of the other flows and are easily smoothed out. Despite this, even for a single video flow, some traffic shaping, such as coarser quantization parameter for I-frames can be applied to reduce the bits; however, this solution may lead to quality flicker due to lower qualities of the I-frames if the QPs of I-frame and P-frames are much different. For high motion

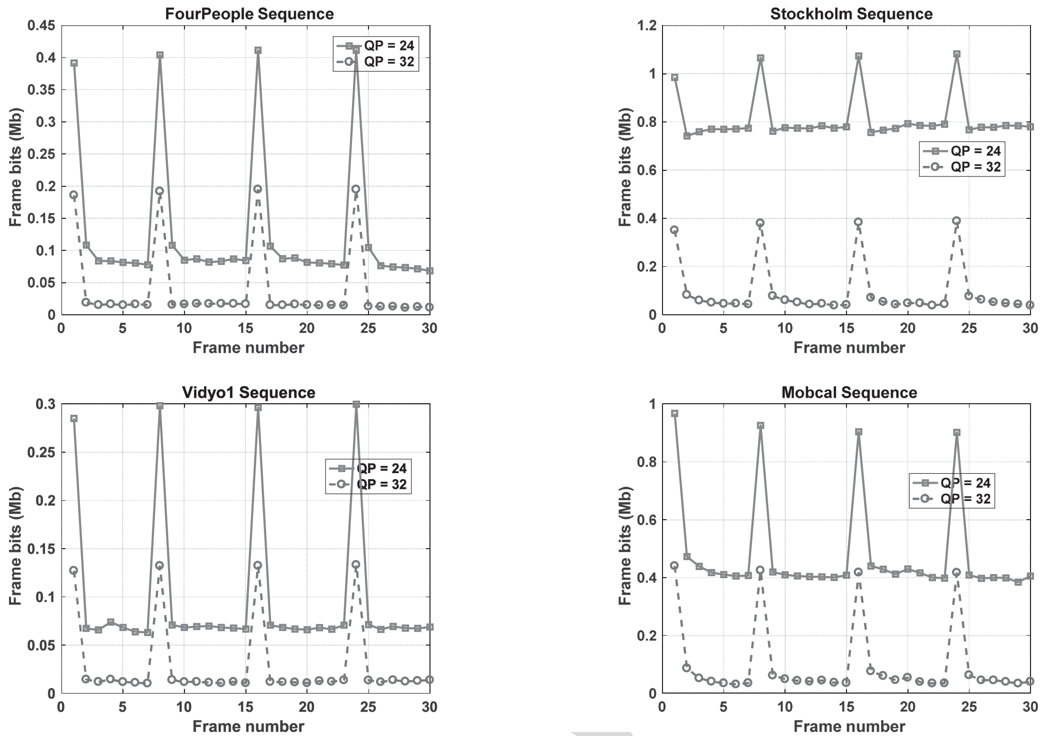


Fig. 6 The sudden changes in the frame bits when a frame is encoded as I-frame. Every 8 frames, one frame is Intra coded entirely.

and high texture videos, the difference between I and P frames' bits are not so large and this approach might be sufficient there. If this was not the case, another solution is to have the same QP for I and P frames but use an encoder smoothing buffer to regulate the bitrate (e.g., traffic shaping), of course at the cost of a few frames delay in video display. How this delay can solve the problem is explained below. Let us assume that intra period is M frames. If the frame rate of the video is FPS, the average bitrate required by the channel is:

$$R = FPS\frac{((M-1)\,R_P + R_I)}{M} \quad bits/sec \tag{23}$$

where R_P and R_I are number of bits needed for coding the P-frames and I-frames, respectively. Now if the I-frame has k times more bits than the P-frames, then:

$$\begin{aligned} R &= \frac{FPS\ (M-1+k)\,R_P}{M} \\ &= \frac{FPS\ (M-1+k)\,R_I}{kM} \quad bits/\sec \quad (k>1) \end{aligned} \tag{24}$$

At each 1/*FPS* second, the total sent bits are:

$$\begin{aligned} R &= \left(1+\frac{k-1}{M}\right)R_P \\ &= \left(\frac{M+k-1}{kM}\right)R_I \quad bits/\sec \quad (k>1) \end{aligned} \tag{25}$$

Therefore, more than one P-frame or less than one I-frame is transmitted at each 1/*FPS*. That is, compared to the case where all frames have the same number of bits (i.e., $k = 1$), delivering P-frames is faster and delivering I-frames is slower. However, the issue that may arise here is the transmit and receive buffers' overflow and underflow in a live streaming application. It can be shown that, with display latency as large as $M(k-1)/(M-1+k)$ frames, there is no overflow or underflow in the buffers and continuous playing of the video is preserved (see Appendix B in [32] for the proof). This latency increases with M; therefore, a smaller M chosen for higher PLRs leads to lower latencies.

The value of k is content dependent; Fig. 6 shows the number of bites of I and P frames, with $M = 8$ for four sequences and two *QP*s. One can see that *k* is about 1.5 for *Stockholm* at *QP = 24*; that is, *k* is small and the delay is not significant. For example, for PLRs of 5%, if M is around of 4 as shown in Fig. 5, this gives a latency of about 0.5 frames. However, the ratio k becomes larger at *QP = 32*. And also, M is typically larger for lower bitrates; i.e., for *QP* $= 32$. Therefore, the incurred delay is more challenging here; for example, for *FourPeople* at *QP* $= 32$, k is about 10 as shown in Fig. 6, which is relatively very high. Now for $M = 15$ (as inferred from Fig. 5), the delay becomes about 5.5 frames. For *FPS* of 60, it leads to a delay less than 100 ms which is acceptable for many applications. For smaller delays, we can combine the above two approaches; that is, applying coarser quantizer and forcing a delay. The coarser quantizer to I-frames leads to a smaller k which in turn leads to a smaller delay.

VI. Performance Comparison

The analysis explained in the previous sections shows that using I-frames instead of applying IRR is more efficient as an error resiliency tool and gives higher quality in dealing with transmission of encoded videos over lossy channels. The

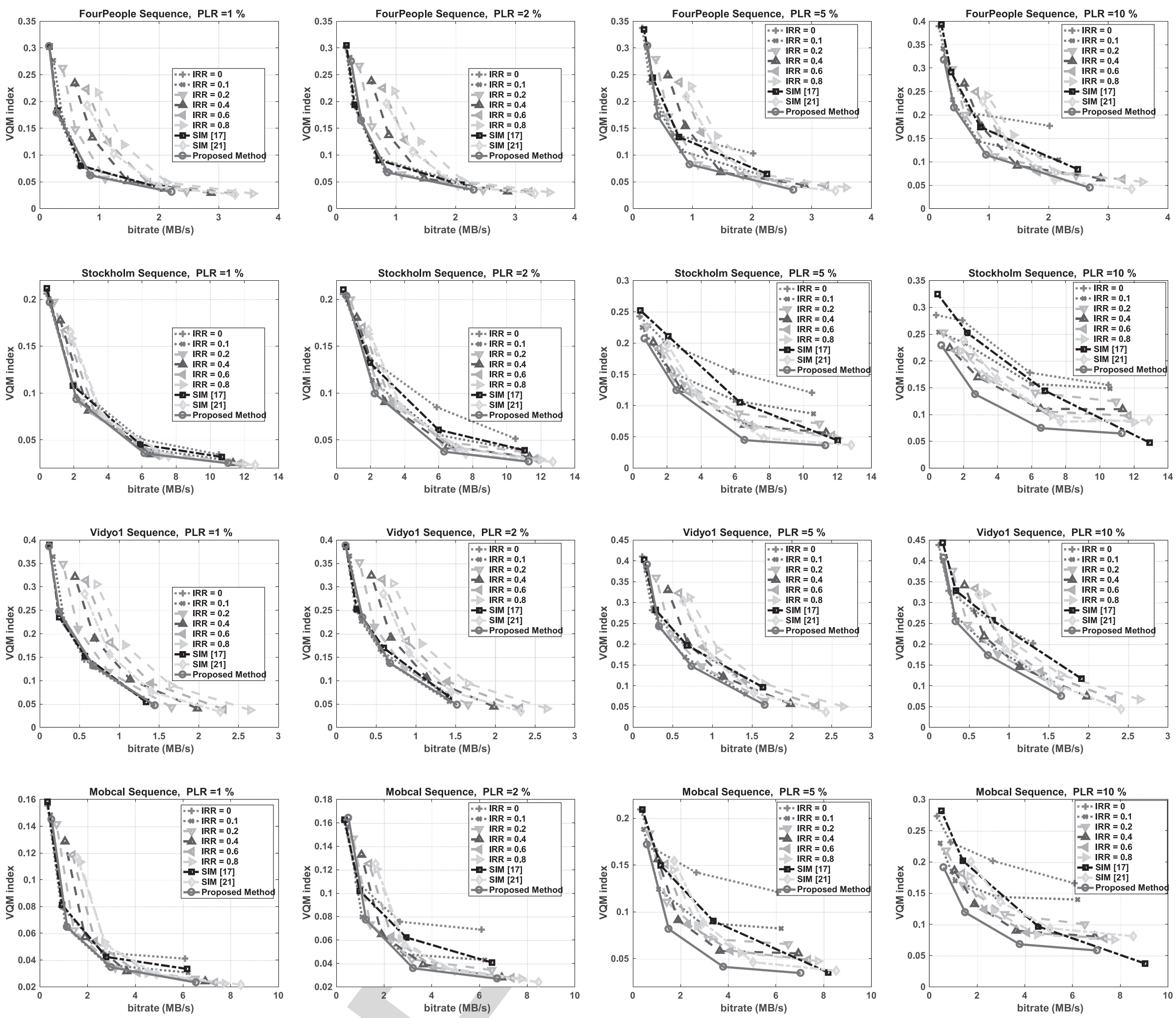


Fig. 7. Performance comparison of the proposed method for various PLRs and *FourPeople*, *Stockholm*, *Vidyo1* and *Mobcal* sequences for burst length of three packets.

suitable I-frame period is given by (22). As already mentioned in Section II, there are two other options for intra coding, SIM and Intra Refresh. For performance comparison, our proposed method is compared against two SIM methods, [17] and [21], where the PUs are selected based on an objective function for intra or inter coding. Note that as explained in Section II, there are also two options for selecting the blocks to be forced for intra coding in the Intra Refresh scheme. They can be selected randomly or in a regular manner, such as a column of intra blocks moving frame by frame from left to right. Our experiments showed that the latter option, called Periodic Intra Refresh (PIR) or cyclic intra-refresh generally gives superior performance in terms of rate-distortion. Therefore, we have included the results of PIR in Figs 7–9. Since there are no appropriate recent related works on the best value of IRR, we examine PIR with several possible values of IRR for all examined PLRs; these are {0, 0.1, 0.2, 0.4, 0.6, 0.8}. Note that IRR = 0 is equivalent to not paying any attention to channel loss at the encoder. With the experimental settings given in Section V, these results are shown in Figs 7–9, Fig. 9 is for the average burst length of six packets.

Despite of the simplifications and approximations made in our method through analysis and curve fitting, it can be seen from Figs 7–8 that our proposed method outperforms the others in many cases. For lower PLRs and smaller bitrates, the proposed method provides actually no gain. In these regions, since the video is less sensitive to packet loss, the curves are actually close to each other. The algorithm of [21] picks many PUs for intra coding; therefore, it applies intra rate much more than required but with a slight gain in quality in lower PLRs. For this reason, this algorithm does not work well for low PLRs. In the cases of higher PLR*s* and higher bitrates, one can see the VQM quality index of our proposed method is better than the others which is sometimes significant. A reminder that the smaller VQM index means higher quality. Light content video sequences, such as *FourPeople* and *KrisenAndSara*, as already mentioned are less sensitive to data loss; hence the VQM curves are again similar while ours are still marginally better.

For the PLR*s* of 5% and 10%, the results of applying average burst length of six packets are shown in Fig. 9. It can be seen that the performance of our proposed method is still better than the others. Actually, the loss pattern does not significantly affect our

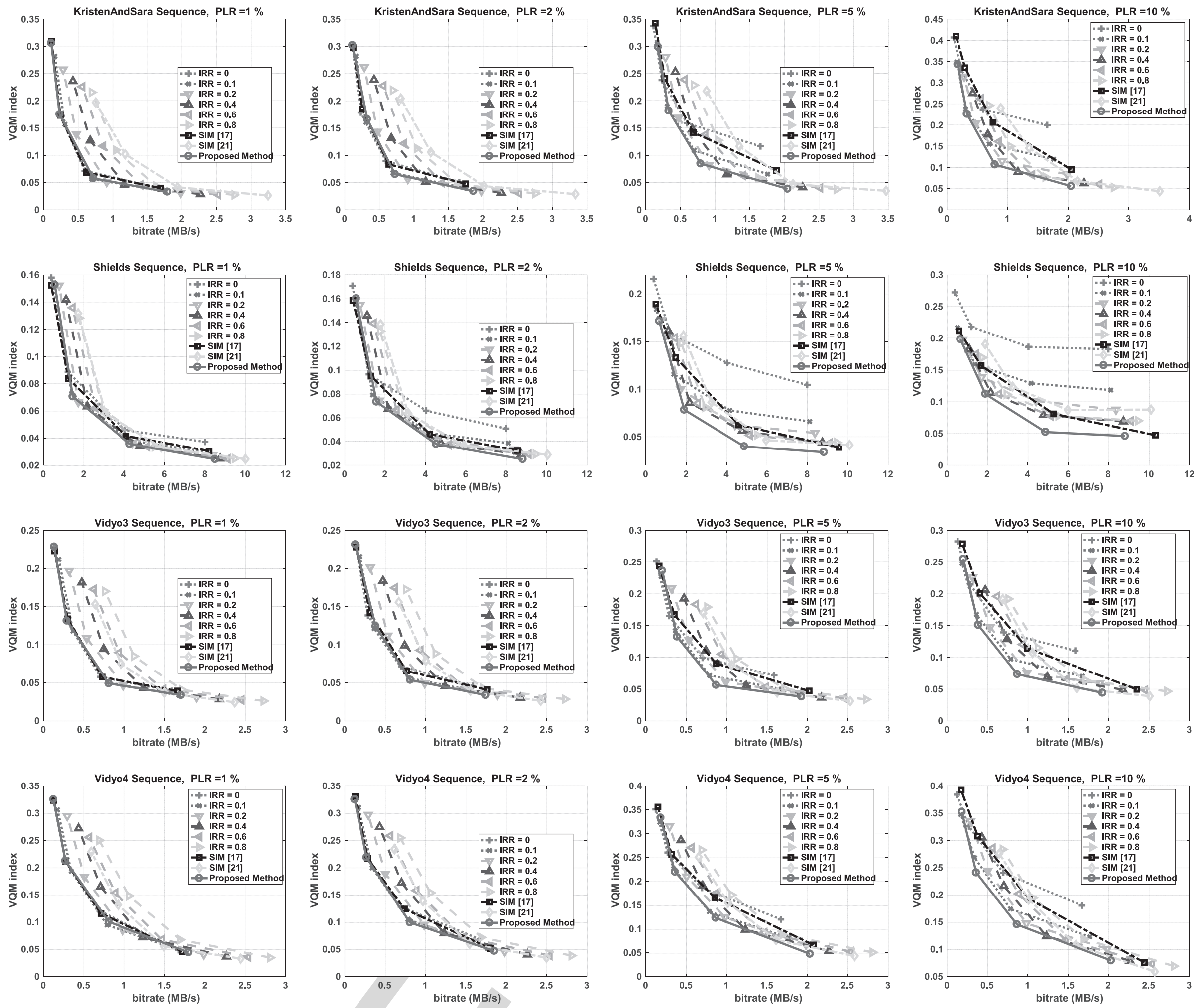


Fig. 8 Performance comparison of the proposed method for various PLRs and *Kristen and Sara*, *Shields*, *Vidyo3* and *Vidyo4* sequences for burst length of three packets.

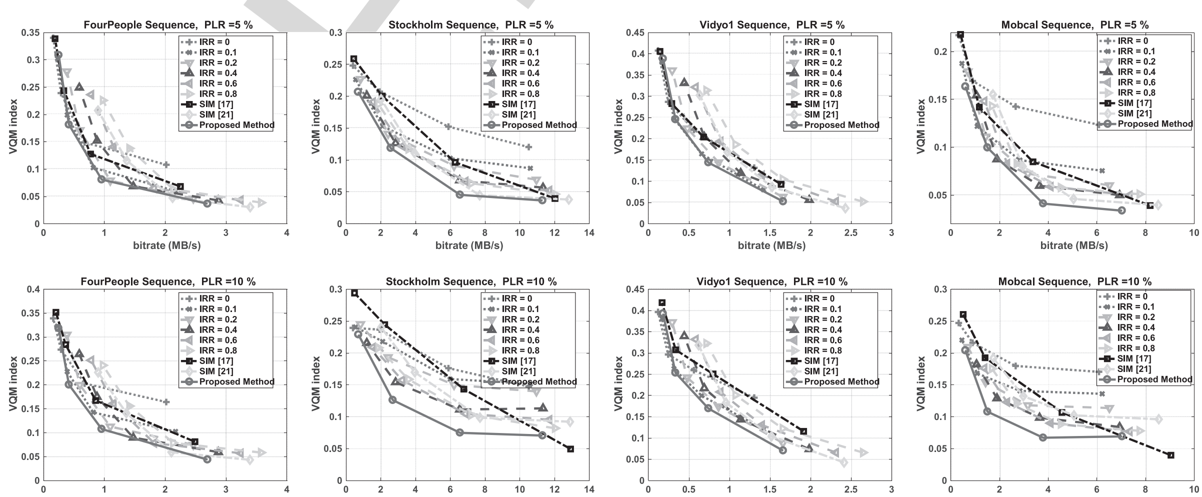


Fig. 9. Performance comparison of the proposed method for various PLRs and *FourPeople*, *Stockholm*, *Vidyo1* and *Mobcal* sequences for burst length of six packets.

results due to our loss concealment procedure applied to lossy bitstreams generated by any of competing methods.

The burst loss leads to larger lossy areas in the pictures, which is usually handled by the Motion Copy algorithm. Note that burst loss will lead to the loss of consecutive frames in low bitrate and low resolution videos, while it is not so destructive for HD and beyond.

VII. Conclusion

In this paper, the best strategy for intra coding as an error resiliency tool is presented. It was proposed to encode some frames entirely in intra mode, rather than the conventional approach where some blocks or PUs are selected in PIR manner (with a specific intra rate) or by a cost function to be coded in intra mode. Considering the error propagation, the receiver side distortion is formulated and it is simplified with some observations. The simplified objective function has a straightforward solution: $\beta^{(i^*)} = 1$, where i^* is the index of the frames symmetrically positioned in the GoP, and the number of I-frames depends on β_{red} or equivalently the available bitrate budget for intra coding. The output of the objective function is to reduce the *IP* as much as possible and as long as the bitrate overhead of intra coding is justified at the given channel loss rate.

The optimal IP varies with the coding bitrate as well as the PLR as shown in Fig. 5. We have fitted a curve to the experimental points obtained from examining various test sequences, as given in (22). With the IP^* selected by (22), experimental results show that the proposed method achieves lower VQM index compared to the conventional SIM and PIR methods.

Appendix

Assume that frame 1 is transmitted through n packets. If m packets are lost, the average distortion after error concealment is:

$$D_m^{(1)} = \frac{(n-m)}{n} D_q^{(1)} + \frac{m}{n}\ D_{conceal}^{(1)} \tag{26}$$

If each packet is lost with a probability of PLR, the probability of losing m packets is

$$\begin{aligned} PLR_m &= C\left(n, m\right)\ PLR^m \left(1 - PLR\right)^{n-m} \\ &= \binom{n}{m}\ PLR^m \left(1 - PLR\right)^{n-m} \end{aligned} \tag{27}$$

where $C(n, m)$ is the number of m-combinations from n packets. Therefore, the expected distortion of frame 1 is as given by (28):

$$\begin{aligned} D^{(1)} &= \sum_{m=0}^{n} \left(PLR_m\ D_m^{(1)}\right) \\ &= \frac{D_q^{(1)}}{n} \sum_{m=0}^{n} \left[(n-m)\left(\binom{n}{m} PLR^m (1-PLR)^{n-m}\right)\right] \\ &\quad + \frac{D_{conceal}^{(1)}}{n} \sum_{m=0}^{n} \left[m\ \left(\binom{n}{m} PLR^m (1-PLR)^{n-m}\right)\right] \end{aligned} \tag{28}$$

Both summations in (28) are the expected values of a *Binomial* distribution with probabilities of $(1 - PLR)$ and PLR, respectively. That is $D^{(1)}$ becomes

$$\begin{aligned} D^{(1)} &= \frac{D_q^{(1)}}{n} \left[n\left(1 - PLR\right)\right] + \frac{D_{conceal}^{(1)}}{n} \left[n\ PLR\right] \\ &= \left(1 - PLR\right) D_q^{(1)} + PLR\ \ D_{conceal}^{(1)} \end{aligned} \tag{29}$$

which is the same as equation (2).

References

[1] Y. Wang and Q. F. Zhu, “Error control and concealment for video communication: A review,” *Proc. IEEE*, vol. 86, no. 5, pp. 974–997, May 1998.

[2] M. T. Pourazad Oztas, P. Nasiopoulos, and V. C. M. Leung, “A study on the HEVC performance over lossy networks,” in *Proc. 19th IEEE Int. Conf. Electron. Circuits Syst.*, 2012, pp. 785–788.

[3] M. Fleury, S. Moiron, and M. Ghanbari, “Innovations in video error resilience and concealment,” *Recent Patents Signal Process.* vol. 1, no. 2, pp. 124–134, 2011. Q4

[4] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, “Review of error resilient coding techniques for real-time video communications,” *IEEE Signal Process. Mag.*, vol. 17, no. 4, pp. 61–82, Jul. 2000.

[5] M. Fleury, M. Altaf, S. Moiron, N. Qadri, and M. Ghanbari, “Source coding methods for robust wireless video streaming,” *Multimedia Netw. Coding*, pp. 175–207, 2012.

[6] M. Fleury, I. A Ali, and M. Ghanbari, “Video intra coding for compression and error resilience: A review,” *Recent Patents Signal Process.*, vol. 4, pp. 32–43, 2014.

[7] R. Zhang, S. L. Regunathan, and K. Rose, “Video coding with optimal inter/intra-mode switching for packet loss resilience,” *IEEE JSAC*, vol. 18, no. 6, pp. 966–976, Jul. 2000.

[8] H. Yang and K. Rose, “Advances in recursive per-pixel end-to-end distortion estimation for robust video coding in H.264/AVC,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 7, pp. 845–856, Jul. 2007.

[9] Y. Liao and J. D. Gibson, “Rate-Distortion based mode selection for video coding over wireless networks with burst losses,” in *Proc. Int. Packet Video Workshop*, Seattle, WA, USA, 2009, pp. 1–10.

[10] S. Wan and E. Izquierdo, “Rate-distortion optimized motion-compensated prediction for packet loss resilient video coding,” *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 1327–1338, May 2007.

[11] J. Xiao, T. Tillo, C. Lin, and Y. Zhao, “Joint redundant motion vector and intra macroblock refreshment for video transmission,” *EURASIP J. Image Video Process.*, 2011. Q5

[12] H. Yang and K. Rose, “Optimizing motion compensated prediction for error resilient video coding,” *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 108–118, Jan. 2010.

[13] H. Shu and L. P. Chau, “Intra/Inter macroblock mode decision for error-resilient transcoding,” *IEEE Trans. Multimedia*, vol. 10, no. 1, pp. 97–104, Jan. 2008.

[14] B. Li, T. Nanjundaswamy, and K. Rose, “An error-resilient video coding framework with soft reset and end-to-end distortion optimization,” in *Proc. IEEE Int. Conf. Image Process.*, Beijing, China, 2017, pp. 1910–1914.

[15] Z. He, J. Cai, and C. W. Chang, “Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding,” *IEEE Trans. Circuits. Syst. Video Technol.*, vol. 12, no. 6, pp. 511–523, Jun. 2002.

[16] Y. Wang, Z. Wu, and J. M. Boyce, “Modeling of transmission-loss-induced distortion in decoded video,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 6, pp. 716–732, Jun. 2006.

[17] Y. Zhang, W. Gao, Y. Lu, Q. Huang, and D. Zhao, “Joint source-channel rate-distortion optimization for H.264 video coding over error-prone networks,” *IEEE Trans. Multimedia.* vol. 9, no. 3, pp. 445–454, Apr. 2007.

[18] Y. Zhang, and P. C. Cosman, “Fast mode decision for H.264 video coding in packet loss environment,” in *Proc. Int. Conf. Image Process.*, Brussels, Belgium, 2011, pp. 3217–3220.

[19] Y. Wei, Y. Zhang, and J. Yan, “Fast mode decision for error resilient video coding,” in *Proc. 16th Int. Workshop Multimedia Signal Process.*, Jakarta, Indonesia, 2014, pp. 1–6.

[20] G. Kulupana, D. S. Talagala, H. K. Arachchi, and A. Fernando, “Error resilience aware motion estimation and mode selection for HEVC video transmission,” in *Proc. IEEE Int. Conf. Consum. Electron.*, Jan. 2016, pp. 85–86.

[21] G. Kulupana, D. S. Talagala, H. K. Arachchi, and A. Fernando, "End user video quality prediction and coding parameters selection at the encoder for robust HEVC video transmission," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 11, pp. 3367–3381, Nov. 2019, doi: 10.1109/TCSVT.2018.2879956.

[22] R. D. C. Cruz, "Low complexity adaptive intra-refresh rate for real-time wireless video transmission," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Conf.*, 2014.

[23] R. Dela, C. Angelo, and D. C. Rhandley, "Error resilient joint source-channel adaptive intra-refresh rate for wireless video transmission," in *Proc. Int. Conf. Digital Signal Process.*, 2014, pp. 509–514.

[24] H. Chen, C. Zhao, M. T. Sun, and A. Drake, "Adaptive intra-refresh for low-delay error-resilient video coding," *J. Vis. Commun. Image Representation*, vol. 31, pp. 294–304, 2015.

[25] I. Ali, M. Fleury, and M. Ghanbari, "Content-Aware intra-refresh for video streaming over lossy links," in *Proc. Int. Conf. Consum. Electron.*, Las Vegas, NV, USA, 2012, pp. 118–119.

[26] S. Moiron and M. Ghanbari, "Enhanced error resiliency for video with cyclic intra-refresh lines," in *Proc. IEEE Int. Conf. Image Process.*, Brussels, Belgium, 2011, pp. 3229–3232.

[27] I. A. Ali, S. Moiron, M. Fleury, and M. Ghanbari, "Packet prioritization for H.264/AVC video with cyclic intra-refresh line," *J. Vis. Commun. Image Representation*, vol. 24, no. 4, pp. 486–498, 2013.

[28] S. Moiron, I. Ali, M. Ghanbari, and M. Fleury, "Limitations of multiple reference frames with cyclic intra-refresh line for H.264/AVC," *Electron. Lett.*, vol. 47, no. 2, pp. 103–104, Jan. 2011.

[29] J. Xu and Z. Wu, "Joint adaptive intra refreshment and unequally error protection algorithms for robust transmission of H.264/AVC video," in *Proc. IEEE Int. Conf. Multimedia Expo.*, 2006.

[30] F. Wan, "Traffic modeling and performance analysis for IPTV systems," Master's thesis, Southeast Univ., Jiangsu, China, 2005.

[31] S. Basso *et al.*, "Estimating packet loss rate in the access through application-level measurements," in *Proc. Assoc. Comput. Machinery's Special Interest Group Data Commun. Workshop Meas. Stack,* Helsinki, Finland, Aug. 17, 2012, pp. 7–12.

[32] M. Kazemi, R. Iqbal, and S. Shirmohammadi, "Joint intra and multiple description coding for packet loss resilient video transmission," *IEEE Trans. Multimedia.* vol. 20, no. 4, pp. 781–795, Apr. 2018.

[33] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 23–50, Nov. 1998.

[34] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74–90, Nov. 1998.

[35] K. H. Choi and D. H. Kim, "A downhill simplex approach for HEVC error concealment in wireless IP networks," in *Proc. IEEE Int. Conf. Consum. Electron.*, Las Vegas, NV, USA, 2016, pp. 143–146.

[36] Y. L. Chang, Y. A. Reznick, Z. Chen, and P. C. Cosman, "Motion compensated error concealment for HEVC based on block-merging and residual energy," in *Proc. Int. Packet Video Workshop*, San Jose, CA, USA, 2013, pp. 1–6.

[37] Y.-T. Peng and P. C. Cosman, "Weighted boundary matching error concealment for HEVC using block partition decisions," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 921–925.

[38] T. Lin, N. Yang, and R. Syu, "Error concealment algorithm for HEVC coded video using block partition decisions," in *Proc. IEEE Int. Conf. Signal Process. Commun. Comput.*, Aug. 2013.

[39] X. Qian, G. Liu, and H. Wang, "Recovering connected error region based on adaptive error concealment order determination," *IEEE Trans. Multimedia*, vol. 11, no. 4, pp. 683–695, Jun. 2009.

[40] W. Lie, C. Lee, C. Yeh, and Z. Gao, "Motion vector recovery for video error concealment by using iterative dynamic programming optimization," *IEEE Trans. Multimedia*, vol. 16, no. 1, pp. 216–227, Jan. 2014.

[41] M. Kazemi, R. Iqbal, and S. Shirmohammadi, "Redundancy allocation based on weighted mismatch-rate slope for multiple description video coding," *IEEE Trans. Multimedia*, vol. 19, no. 1, pp. 54–66, Jan. 2017.

[42] H. Hadizadeh and I. Bajic, "Rate-distortion optimized pixel-based motion vector concatenation for reference picture selection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 8, pp. 1139–1151, Aug. 2011.

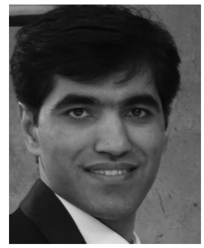
[43] W. Tu and E. Steinbach, "Proxy-Based reference picture selection for error resilient conversational video in mobile networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 2, pp. 151–164, Feb. 2009.

[44] J. Carreira, P. A. Assuncao, S. M. M. de Faria, E. Ekmekcioglu, and A. Kondoz, "A two-stage approach for robust HEVC coding and streaming," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 8, pp. 1960–1973, Aug. 2018.

[45] E. Khan, S. Lehmann, H. Gunji, and M. Ghanbari, "Iterative error detection and correction of H.263 coded video for wireless networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 12, pp. 1294–1307, Dec. 2004.

[46] M. Ghanbari, "Standard codecs: Image compression to advanced video coding," *Appendix E*, 3rd ed. IET publ., 2011.

[47] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans. Broadcast.*, vol. 50, no. 3, pp. 312–322, Sep. 2004.



Mohammad Kazemi received the B.Sc. degree from the Isfahan University of Technology, Isfahan, Iran, in 2003, and the M.Sc. and Ph.D. degrees from the Sharif University of Technology (SUT), Tehran, Iran, both in electrical engineering, in 2005 and 2012, respectively. He currently works with the Electrical Engineering Department, University of Isfahan, Iran, as an Assistant Professor. His current research interests include the areas of image/video processing, coding, and video transmission over lossy channels. He is also interested in digital systems design, and applications of artificial intelligence for image/video analysis.

Mohammad Ghanbari (IEEE M'78–SM'97–F'01–LF'14) is a Professor with the School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran, as well as an Emeritus Professor with the School of Computer Science and Electronic Engineering, University of Essex, Colchester, U.K. He has authored or coauthored 8 books and has registered for 13 international patents and authored over 750 technical papers on various aspects of video networking, many of which have had fundamental influences in this field. These include video/image compression, layered/scalable video coding, video over networks, video transcoding, motion estimation, and video quality metrics. Mr. Ghanbari's book Video Coding: An Introduction to Standard Codecs (IET Press, 1999) received the Rayleigh Prize as the Best Book of the Year 2000 by IET. He has also received several prizes, such as Reeves Prize for Best Paper Award in 1995 and 14th Khwarizmi International award for work on video networking in 2001.

Shervin Shirmohammadi (M'04–SM'04–F'17) received the Ph.D. degree in electrical engineering from the University of Ottawa, ON, Canada, where he is currently a Professor with the School of Electrical Engineering and Computer Science. He is Director of the Distributed and Collaborative Virtual Environment Research Laboratory, doing research in multimedia systems and networks, specifically video systems, gaming systems, and multimedia-assisted healthcare systems. The results of Dr. Shirmohammadi's research, funded by more than $14 million from public and private sectors, have led to over 350 publications, 3 Best Paper awards, over 70 researchers trained at the Postdoctoral, Ph.D., and Master's levels, over 20 patents and technology transfers to the private sector, and a number of awards. He is the Editor-in-Chief for the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, and an Associate Editor for *ACM Transactions on Multimedia Computing Communications and Applications*, having been numerously recognized as the Associate Editor for the Year by all three journals. He is a Key Member of the IEEE MMTC Multimedia Cloud Computing Interest Group, member of the IEEE CS Technical Committee on Multimedia Computing (TCMC), and member of the IEEE ComSoc Technical Committee on Multimedia Communications (MMTC). He is a winner of the 2019 George S. Glinski Award for Excellence in Research, a Lifetime Senior Member of the ACM, a University of Ottawa Gold Medalist, and a licensed Professional Engineer in Ontario.