

RESEARCH

Open Access

# Error-resilient video coding with end-to-end rate-distortion optimized at macroblock level

Jimin Xiao<sup>1,2</sup>, Tammam Tillo<sup>2\*</sup>, Chunyu Lin<sup>3</sup> and Yao Zhao<sup>3</sup>

## Abstract

Intra macroblock refreshment is an effective approach for error-resilient video coding. In this paper, in addition to intra coding, we propose to add two macroblock coding modes to enhance the transmission robustness of the coded bitstream, which are inter coding with redundant macroblock and intra coding with redundant macroblock. The selection of coding modes and the parameters for coding the redundant version of the macroblock are determined by the rate-distortion optimization. It is worth mentioning that the end-to-end distortion is employed in the optimization procedure, which considers the channel conditions. Extensive simulation results show that the proposed approach outperforms other error-resilient approaches significantly; for some video sequences, the average PSNR can be up to 4 dB higher than that of the Optimal Intra Refreshment approach.

**Keywords:** H.264/AVC, error resilience, end-to-end distortion, intra refreshment, redundant coding

## 1. Introduction

The H.264/AVC [1] video coding standard provides higher coding efficiency and stronger network adaptation capability in comparison with all the previously developed video coding standards. However, as previous video compression standards, it is based on a hybrid coding method, which uses transform coding with Motion-Compensated Prediction (MCP). Therefore, when the hybrid-coded video bit-stream is transmitted over packet loss networks, it suffers from error propagations and this leads to the well-known drifting phenomenon [2,3].

Due to the unreliable underlying networks, the development of error-resilient techniques is a crucial requirement for video communication over lossy networks. For applications that can tolerate long delay, channel-coding techniques, like Forward Error Correction (FEC), provide very significant reductions of transmission errors at a comparably moderate bitrate overhead. For the real-time applications, however, the effective use of FEC and re-transmission is limited. Here, the use of error resilience techniques in the source codec becomes important. Two categories of source coding approaches are

promising. One category is based on intra macroblock refreshment, and another one is redundant coding.

The intra macroblock refreshment approach is standard compatible, and it is a useful tool to combat network packet losses. It can be employed to weaken the inter picture dependency due to inter prediction, and eventually, cut-off the error propagations. The early intra macroblock refreshment algorithms are based on randomly inserting intra macroblocks [4] or periodically inserting intra contiguous macroblocks [5]. However, in both [4] and [5], the intra refresh frequency is determined in a heuristic way, and as the intra coding mode is costly, the trade-off between code efficiency and error resiliency need to be balanced. Zhang et al. [6] first treated this problem as optimal coding mode selection of macroblocks and proposed the well-known Recursive Optimal Per-pixel Estimate (ROPE) approach to determine where to insert intra macroblock. In [6], the expected end-to-end distortion for each pixel is calculated in recursive way, and then in the mode selection step, the expected end-to-end distortion is used in the rate-distortion optimization process. In [7], another flexible intra macroblock update algorithm was investigated to optimize the expected rate-distortion performance. In this approach, the end-to-end distortion is calculated by emulating the real channel behavior; therefore, the computation overhead is tremendous. The work in [6,7] is

\* Correspondence: tammam.tillo@xjtlu.edu.cn

<sup>2</sup>Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, 111 Ren Ai Road, Suzhou, People's Republic of China  
Full list of author information is available at the end of the article

loss-aware end-to-end rate-distortion optimized intra macroblock refreshment algorithm, which is currently the best known way for determining both the correct number and placement of intra macroblocks for error resilience.

Redundant coding is another effective tool for robust video communication over lossy network. In [8], an optimal algorithm is presented to determine whether one picture needs redundant version. In [9], redundant slice is optimally allocated based on the slice position in the GOP, and the primary and redundant slices are then interleaved to generate two equal importance descriptions using the MDC [10] diagram. Whereas in [11], the two descriptions are generated by splitting the video pictures into two threads, and then redundant pictures are periodically inserted into the two threads. In both [8] and [11], redundant coding are optimized in frame level, namely all the macroblocks in one frame is encoded with the same redundant coding parameters, whereas for [9], redundant information is allocated in slice level. In [12], redundant coding is optimized in macroblock level. However, in order to optimally tune the redundancy, this approach needs all the motion vector information in one GOP, which leads to a delay of one GOP; consequently, this work cannot be applied in real-time applications, such as video conference.

Intra macroblock refreshment can stop errors in the previous frames, while redundant coding is a way of preventing errors in the future frames. In order to take advantage of the two approaches, we propose to add two new encoding modes, namely inter coding with redundant macroblock and intra coding with redundant macroblock, in addition to the conventional intra and inter coding modes. This approach is called Hybrid Redundant Macroblock and Intra macroblock Refreshment (HRMIR). The redundant version macroblock is encoded with lower quality and rate, which is implemented by scaling the quantization parameter (QP). The selection of coding modes and the parameters for coding the redundant version of the macroblock are determined by the rate-distortion optimization procedure. It is worth noticing, the loss-aware end-to-end expected distortion is used for the RD optimization, and the end-to-end distortion is calculated with the ROPE [6] method. Since calculating the end-to-end distortion with the ROPE method causes no additional delay, the proposed approach is suitable for real-time applications.

The rest of the paper is organized as follows. In Section II, the method to calculate the loss-aware end-to-end distortion is presented. In Section III, the proposed HRMIR approach is introduced. In Section IV, extensive simulation results are given, which validate our approach. Finally, some conclusions are drawn in Section V.

## II. End-to-end distortion calculation

In an ideal error-free environment, the rate-distortion optimized intra/inter mode decision is an efficient tool to determine the macroblock mode based on the cost function defined in [13], and the cost function of any macroblocks is defined as

$$J_{MB} = D_{MB} + \lambda_{mode} \cdot R_{MB} \quad (1)$$

where  $\lambda_{mode}$  is the Lagrange multiplier,  $D_{MB}$  and  $R_{MB}$  are the encoding distortion and the bitrate in different encoding modes, respectively. This optimization mode is tailored for error-free environment, and no channel packet loss is considered here.

However, when the compressed video is transmitted over error-prone network, in addition to the distortion caused by source coding, there is channel distortion, which is caused by packet loss of the underlying network. Loss-aware end-to-end distortion, which encompasses both of the two categories distortion, is used in the proposed HRMIR approach to make better RD optimization. There are many methods to calculate the end-to-end distortion, in ROPE [6], end-to-end distortion for each pixel is calculated in recursive way. Recent advances in ROPE further expand its capability to accommodate sub-pixel prediction [14] and burst packet loss [15]. In [16], a block-based approach generates and recursively updates a block-level distortion map for each frame; therefore, the end-to-end distortion is calculated in block-level. Besides calculating end-to-end in the pixel domain, compressed-domain methods are introduced in [17]. It is important to note that, for the sake of complexity reduction, we apply ROPE [6] with full-pixel level accuracy in our HRMIR approach. For the sub-pixel version ROPE method [14], the computation of the second moment needs a large amount of storage capacity and computational power, which renders the whole process utterly formidable. Furthermore, constrained intra prediction is applied, so there is no error propagation in the intra prediction.

Let  $f_n^i$  denote the original value of pixel  $i$  in frame  $n$ , and let  $\hat{f}_n^i$  and  $\tilde{f}_n^i$  denote its encoder and decoder reconstruction, respectively. Because of possible packet loss in the channel,  $\tilde{f}_n^i$  can be modeled at the encoder side as a random variable. In the ROPE approach, the  $D_{MB}$  is redefined as the overall expected decoder distortion in one macroblock.

$$D_{MB} = \sum_{i \in MB} d_n^i \quad (2)$$

$$\begin{aligned} d_n^i &= E \left\{ (f_n^i - \tilde{f}_n^i)^2 \right\} \\ &= (f_n^i)^2 - 2 \cdot f_n^i \cdot E \left\{ \tilde{f}_n^i \right\} + E \left\{ (\tilde{f}_n^i)^2 \right\} \end{aligned} \quad (3)$$

The overall expected mean-squared-error (MSE) distortion of a pixel is  $\hat{d}_n^i$ ; obviously, it is determined by the first and second moments of the decoder reconstruction. ROPE provides an optimal recursive algorithm to accurately calculate the two moments for each pixel in a frame.

Let us assume that packet loss events are independent for simplicity, and the packet loss rate (PLR)  $p$  is available at the encoder, usually the encoder can get the statistics of packet loss through RTCP [18]. To make it more general, we will not impose any limitations on the slice shape and size, so the motion vectors from neighboring macroblocks are not always available in the error concealment stage. Therefore, the decoder may not be able to use motion vector from neighboring macroblocks for concealment. Accordingly, we assume the decoder copies reconstructed pixels from the previous frame for concealment. The prediction at the encoder only employs the previous reconstructed frame. The recursive formulate of ROPE is as follows.

- Pixel in the intra macroblock

$$E \left\{ \tilde{f}_n^i \right\} = (1 - p) \hat{f}_n^i + pE \left\{ \tilde{f}_{n-1}^i \right\} \quad (4)$$

$$E \left\{ \left( \tilde{f}_n^i \right)^2 \right\} = (1 - p) \left( \hat{f}_n^i \right)^2 + pE \left\{ \left( \tilde{f}_{n-1}^i \right)^2 \right\} \quad (5)$$

- Pixel in the inter macroblock

$$E \left\{ \tilde{f}_n^i \right\} = (1 - p) \left( \hat{e}_n^i + E \left\{ \tilde{f}_{n-1}^{i+mv} \right\} \right) + pE \left\{ \tilde{f}_{n-1}^i \right\} \quad (6)$$

$$E \left\{ \left( \tilde{f}_n^i \right)^2 \right\} = (1 - p) \left( \left( \hat{e}_n^i \right)^2 + 2\hat{e}_n^i E \left\{ \tilde{f}_{n-1}^{i+mv} \right\} + E \left\{ \left( \tilde{f}_{n-1}^{i+mv} \right)^2 \right\} \right) + pE \left\{ \left( \tilde{f}_{n-1}^i \right)^2 \right\} \quad (7)$$

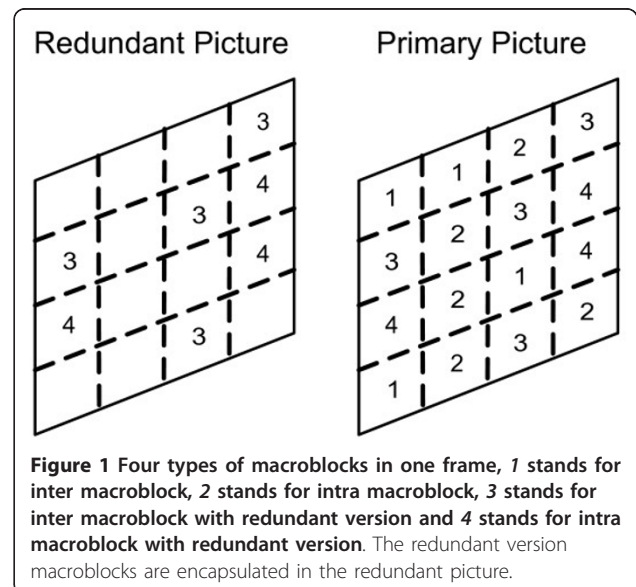
where inter coded pixel  $i$  is predicted from pixel  $i + mv$  in the previous frame. The prediction residual  $\hat{e}_n^i$  is quantized to  $\hat{e}_n^i$ .

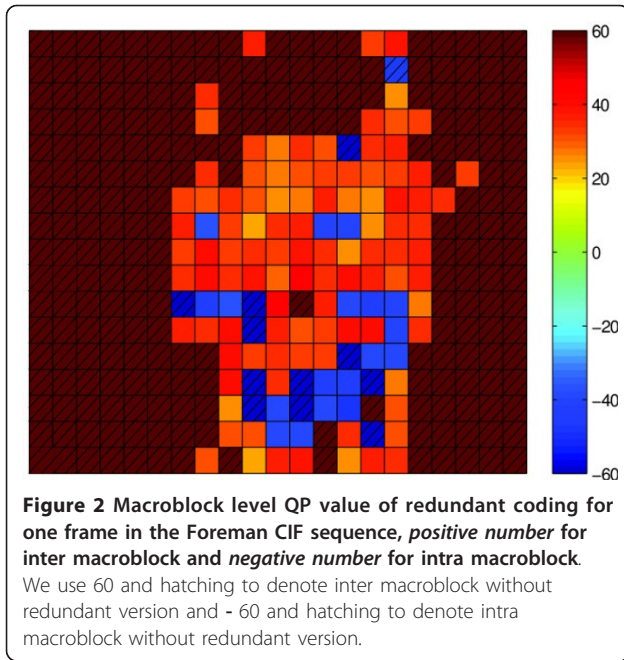
### III. The proposed HRMIR approach

As redundant coding and intra macroblock refreshment are both powerful tools for error resiliency video communication, in the proposed approach, they are hybridly applied to further protect the video stream. With the

Hybrid Redundant Macroblock and Intra macroblock Refreshment (HRMIR) approach, all the macroblocks of one frame are divided into four types, namely intra macroblock, inter macroblock, inter macroblock with redundant version and intra macroblock with redundant version. The redundant version macroblocks are encapsulated in the redundant picture. It is important to note that the concept of redundant slice is part of the H.264/AVC standard. In order to make the proposed approach fully compatible with the H.264/AVC standard, for those macroblocks without redundant version, SKIP mode could be used. Let us take macroblocks in Figure 1 as an example, suppose that the last macroblock in the first row is an inter macroblock with redundant version; accordingly, the redundant macroblock is stored in the redundant picture. Therefore, for macroblock with redundant version, if the macroblock in the primary picture is lost due to packet loss, the redundant version can be used to replace the macroblock. On the contrary, for intra macroblock and inter macroblock without redundant version, there will be no redundant information to be sent in the redundant picture.

It is worth noticing that, in general, the redundant version macroblock is encoded with lower bit rate than primary one, so the video quality is also lower than primary one. In our approach, this is implemented by setting a relative larger quantization parameter (QP) for redundant version macroblock. Like the selection of the coding type for each macroblock, the selection of the appropriate QP value for redundant macroblock is also optimized in the end-to-end RD optimization process. Figure 2 shows the QP value for redundant frame in the Foreman CIF sequence, where the QP of primary





macroblock is 22. In order to present all information in one figure, we use positive number for inter macroblock and negative number for intra macroblock. The valid QP range is (1-51) in H.264/AVC, so we use 60 to denote inter macroblock without redundant version and -60 to denote intra macroblock without redundant version. For example, if a macroblock in Figure 2 has a value -34, this means it is an intra macroblock with QP 34, whereas for a macroblock with value 34, it is an inter macroblock with QP 34. It can be seen that most of the background areas are encoded with inter coding without redundant version, because these areas are relatively static, and with the temporal replacement concealment algorithm, losing these areas will not lead to huge distortion. On the contrary, the parts of foreground, which is the Foreman face area in this frame, are strongly protected with intra coding and/or redundant coding. Note both the macroblock type and QP value are optimized in the RD optimization process, which are presented in the next section.

#### A. The HRMIR rate-distortion optimization

As in the other encoding approaches, in the HRMIR rate-distortion optimization process, the encoder selects the coding option  $O^*$  for current macroblock, so that the Lagrangian cost function is minimized.

$$O^* = \arg \min_{o \in \Gamma_{\text{HRMIR}}} (D_{\text{MB}}(o) + \lambda_{\text{mode}} R_{\text{MB}}(o)) \quad (8)$$

where  $D_{\text{MB}}(o)$  is the expected end-to-end distortion for mode  $o$ ,  $R_{\text{MB}}(o)$  is the rate for this mode and  $\lambda_{\text{mode}}$

is the Lagrangian multiplier.  $\Gamma_{\text{HRMIR}}$  is a set of encoding options, which includes all encoding modes. For the original ROPE approach, the available encoding modes includes intra mode  $I$  and inter mode  $P$ , so  $\Gamma_{\text{ROPE}} = \{I, P\}$ . However, in our HRMIR approach, there are two new modes. They are intra mode with redundant version macroblock and inter mode with redundant version macroblock. For simplicity, let us use  $I_r^u$  and  $P_r^v$  to denote the two new modes, respectively, with  $r$  standing for redundant coding,  $u$  representing the candidate QP value in the intra redundant coding and  $v$  representing the candidate QP value in the inter redundant coding. Therefore, for the HRMIR approach, the set of encoding options become  $\Gamma_{\text{HRMIR}} = \{I, P, I_r^u, P_r^v\}$ . In general, the QP value of redundant coding is larger than that of primary coding. Let us use  $QP_I$  and  $QP_P$  to denote the primary QP value of intra and inter coding, respectively. In the redundant coding, candidate QP value is  $u \in \{u | QP_I \leq u \leq 51\}$  and  $v \in \{v | QP_P \leq v \leq 51\}$ , where 51 is the maximum QP value in H.264/AVC [1].

#### B. The HRMIR end-to-end distortion and rate

When calculating the expected end-to-end distortion, we can still use the Equations 4, 5 for intra macroblock without redundant coding, and Equations 6, 7 for inter macroblock without redundant coding. Whereas for intra macroblock with redundant coding, first and second moments of the decoder reconstruction are as follows.

$$E \left\{ \tilde{f}_n^i \right\} = (1 - p) \hat{f}_n^i + p(1 - p) \hat{f}_n^{i,u} + p^2 E \left\{ \tilde{f}_{n-1}^i \right\} \quad (9)$$

$$E \left\{ \left( \tilde{f}_n^i \right)^2 \right\} = (1 - p) \left( \hat{f}_n^i \right)^2 + p(1 - p) \left( \hat{f}_n^{i,u} \right)^2 + p^2 E \left\{ \left( \tilde{f}_{n-1}^i \right)^2 \right\} \quad (10)$$

where in the primary coding  $f_n^i$  is quantized to  $\hat{f}_n^i$ , and in the redundant coding, it is quantized to  $\hat{f}_n^{i,u}$ , here  $u$  is the redundant QP value.

Similarly, for inter macroblock with redundant coding, first and second moments of the decoder reconstruction are as follows.

$$E \left\{ \tilde{f}_n^i \right\} = (1 - p) \left( \hat{e}_n^i + E \left\{ \tilde{f}_{n-1}^{i+mv} \right\} \right) + p(1 - p) \left( \hat{e}_n^{i,v} + E \left\{ \tilde{f}_{n-1}^{i+mv(v)} \right\} \right) + p^2 E \left\{ \tilde{f}_{n-1}^i \right\} \quad (11)$$

$$\begin{aligned}
 E \left\{ \left( \tilde{f}_n^i \right)^2 \right\} &= (1-p) \left( \left( \hat{e}_n^i \right)^2 + 2\hat{e}_n^i E \left\{ \tilde{f}_{n-1}^{i+mv} \right\} \right. \\
 &\quad \left. + E \left\{ \left( \tilde{f}_{n-1}^{i+mv} \right)^2 \right\} \right) \\
 &+ p(1-p) \left( \left( \hat{e}_n^{i,v} \right)^2 + 2\hat{e}_n^{i,v} E \left\{ \tilde{f}_{n-1}^{i+mv(v)} \right\} \right. \\
 &\quad \left. + E \left\{ \left( \tilde{f}_{n-1}^{i+mv(v)} \right)^2 \right\} \right) \\
 &+ p^2 E \left\{ \left( \tilde{f}_{n-1}^i \right)^2 \right\}
 \end{aligned} \quad (12)$$

where in the primary coding, pixel  $i$  is predicted from pixel  $i + mv$  in the previous frame, the prediction residual  $e_n^i$  is quantized to  $\hat{e}_n^i$ . In the redundant coding, the redundant QP value is  $v$ , pixel  $i$  is predicted from pixel  $i + mv(v)$  in the previous frame, the prediction residual  $e_n^{i,v}$  is quantized to  $\hat{e}_n^{i,v}$ .

For those intra and inter macroblocks with redundant coding, the probability of receiving the primary macroblock is  $1 - p$ . The probability of receiving the redundant macroblock while losing the primary information is  $p(1 - p)$ , and the probability of losing both the primary and redundant macroblocks is  $p^2$ . With all those probabilities, we can easily get Equations 9, 10, 11, 12 for macroblock with redundant version. It is important to note that when the macroblock is encoded with redundant version, namely  $0 \in \{I_r^u, P_r^v\}$ , the total bit rate  $R_{MB}(o)$  is calculated by summing up the bit rate used for both primary and redundant coding.

### C. Lagrange multiplier selection

The Lagrange multiplier  $\lambda_{mode}$  in (8) controls the rate-distortion trade-off. For the error-prone environment, extensive experimental evidence suggests that there is no significant performance difference between using the Lagrange multiplier tailored to the error-free or the error-prone environment. This argument has also been confirmed in [7]. So  $\lambda_{mode}$  is set as the one tailored to error-free environment.

$$\lambda_{mode} = 0.85 \times 2^{(QP-12)/3} \quad (13)$$

where  $QP$  is the quantization parameter.

### D. Computation complexity reduction

In the HRMIR rate-distortion optimization procedure, in order to find the optimal QP value for redundant coding, we need to calculate the rate-distortion cost for all possible redundant QP value; therefore, the computation complexity is tremendous. For example, let us assume the primary QP value is 22, in the RDO procedure described in Section III-A, the encoding options are  $\Gamma_{HRMIR} = \{I, P, I_r^u, P_r^v\}$ , then both  $I_r^u$  and  $P_r^v$  have  $(51 - 22 + 1)$  possible redundant QP values, here 51 is the

maximum QP value in H.264/AVC. Therefore,  $\Gamma_{HRMIR}$  includes 62 encoding options (both  $I_r^u$  and  $P_r^v$  have 30 QP values plus intra/inter coding without redundant version).

By lowering the number of encoding options, the computation complexity will be reduced. Let us set the redundant QP increase step as  $QP_{step}$ , then the candidate QP value would be  $u \in \{u | u = QP_I + K \times QP_{step}, u \leq 51, K = 0, 1, 2, \dots\}$  and  $v \in \{v | v = QP_P + K \times QP_{step}, v \leq 51, K = 0, 1, 2, \dots\}$ .

In Figure 3, the trade-off between PSNR and computation complexity is reported. It is observed that when the value of  $QP_{step}$  is set as 5 and 10, the PSNR is lower than that when the  $QP_{step}$  is 1. However, the PSNR decrease is very limited. The computation overhead for the  $QP_{step} = 5$  case is nearly 1/5 of that for the  $QP_{step} = 1$  case, but the resulting decrease of PSNR is less than 0.3 dB. Even when the  $QP_{step}$  value is set to 10, the PSNR penalty is less than 0.5 dB. The indication of this property of HRMIR is significant, which means it is possible to deploy this approach in hand-device, where the computation resource is limited, by setting relatively large  $QP_{step}$  value.

## IV. Simulation result

Our simulation setting builds on the JM9.4 H.264 codec [19]. We use constrained intra prediction and CABAC for entropy coding, and fixed QP value is used for all of our simulations. One row of macroblocks per slice is used to create slices. For each sequence, only the first frame is coded as I-frame, and the rest are coded as P-frames; the reference frame number is 1. In order to have fair comparison with the Optimal Intra approach

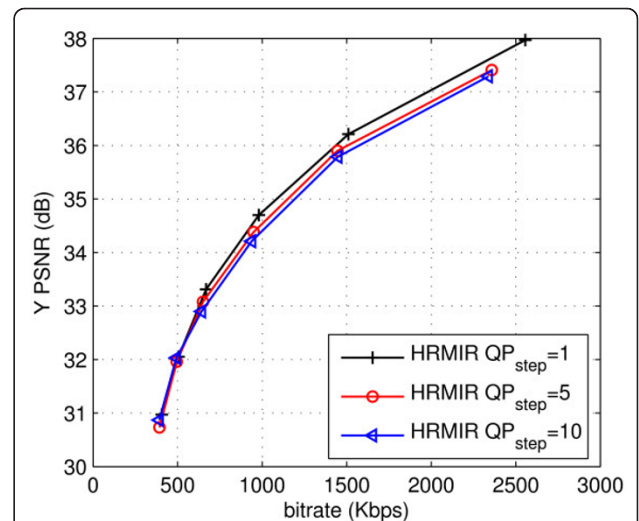


Figure 3 PSNR versus bit rate for the Foreman sequence,  $QP_{step}$  of HRMIR is set to 1, 5, 10. PLR is set to 10%, and GOP is 30.

[6], it is assumed that the I-frame is transmitted over secure channel. We use the average luminance PSNR to assess the objective video quality; the mean squared error (mse) is averaged over 200 trials, then the value of PSNR is calculated based on the averaged mse. A random packet loss generator is used to drop the packets according to the required packet loss rate. For the lost slices, temporal replacement concealment is used, which means the pixel value of lost slice is copied from the same position in the previous frame. To evaluate the proposed HRMIR approach, extensive experiments have been conducted, and as benchmark, we use conventional Optimal Intra Refreshment [6] and RS-MDC [9] for comparison.

In the first set of experiment, frame-by-frame average PSNR is reported for Foreman and Bus CIF video sequences. We compare HRMIR results with Optimal Intra [6] and RS-MDC [9]. In this experiment, constant QP value is used. For the HRMIR approach, QP is set to 22 and 28 for Foreman and Bus, respectively, while for the other two approaches, the encoded bitrate is close to but no less than that of HRMIR approach. In Figure 4, full-pixel accuracy motion estimation (ME) is used, whereas in Figure 5, motion estimation with 1/4 pixel accuracy is adopted. In both full-pixel and sub-pixel motion estimation environments, the video quality of HRMIR and RS-MDC is similar at the beginning of several video frames for both the Foreman and Bus sequences. However, the video quality of RS-MDC decreases much faster than that of HRMIR; therefore, HRMIR outperforms RS-MDC significantly with frame number increasing. This result indicates that for those P-frames relatively far away from the intra frame, only providing redundant coding is not enough to protect the video quality effectively. Meanwhile, when comparing HRMIR with Optimal Intra, for most of the frames, PSNR of HRMIR is higher than that of Optimal Intra. Another advantage of the HRMIR approach is that the video quality for each frame is more stable than the other two approaches, which is an essential characteristic of subjective high-quality video. When the encoder adopts sub-pixel ME, the accuracy of the end-to-end distortion calculated with the ROPE [6] method is compromised, and eventually, the optimal procedure in Section III-A becomes sub-optimal. However, comparing results in Figure 4 with that in Figure 5, it is found that in both full-pixel ME and sub-pixel ME environments, HRMIR outperforms Optimal Intra and RS-MDC, and the superiority of HRMIR over the other two approaches remains almost unchanged in the sub-pixel ME environment. Therefore, in the following experiments, we adopt the sub-pixel ME with the purpose of good performance in the sense of rate-distortion.

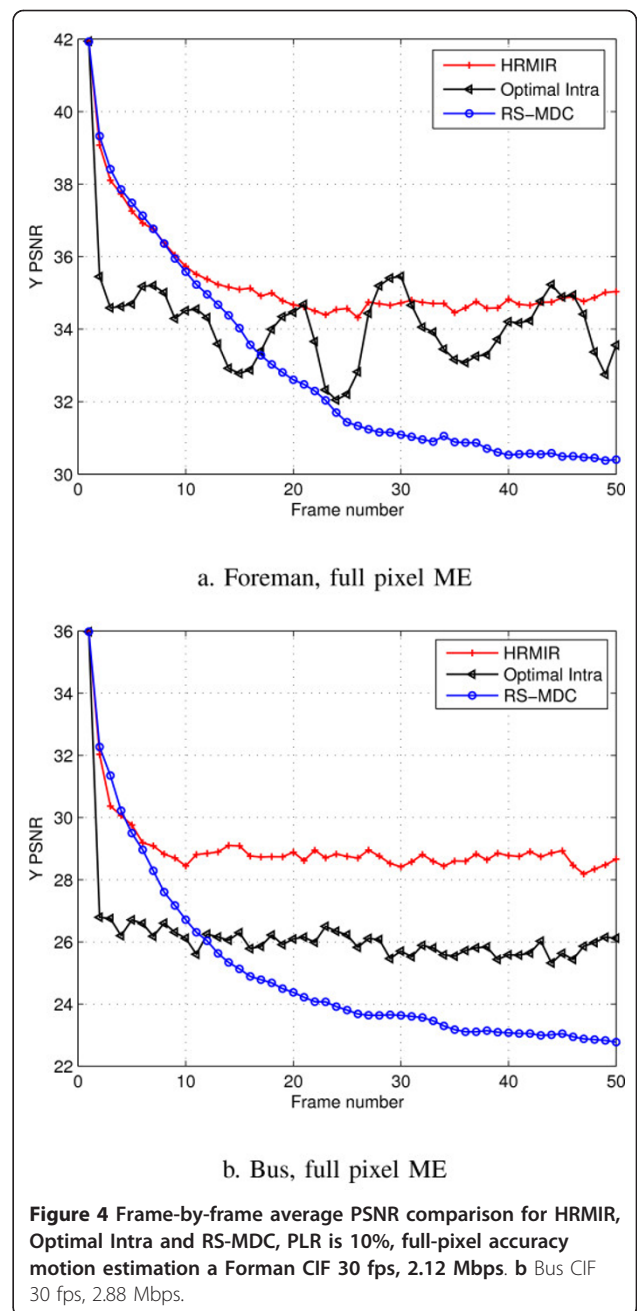
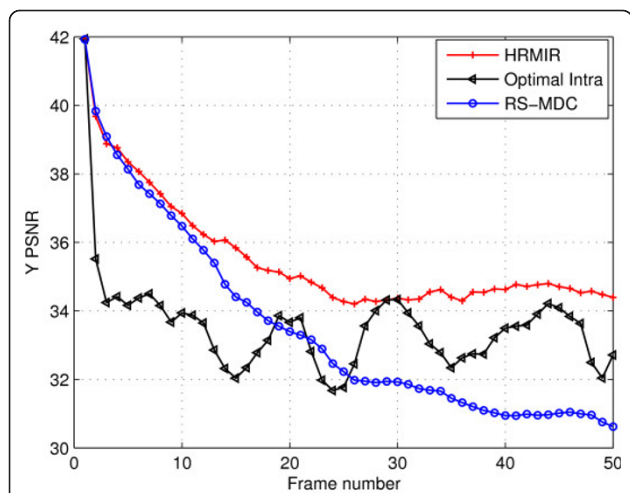
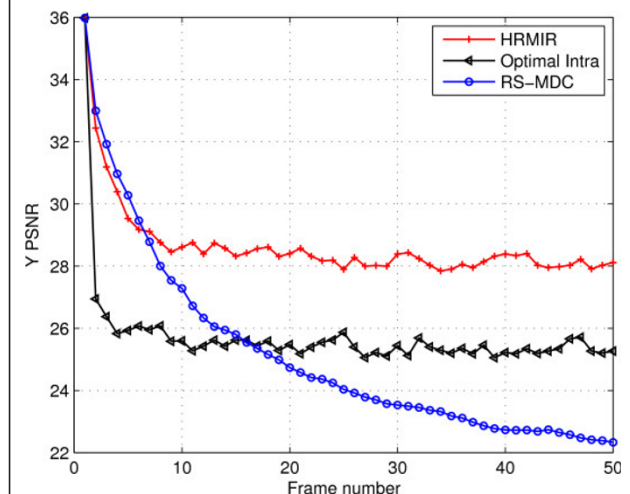


Figure 6 shows the video quality versus bit rates for CIF video sequences Foreman and Bus. Different QP values are selected in order to span a considerable range of coding rates. In Figure 6, we fix the PLR as 10% and GOP length is set to 15 and 30. It is observed that when GOP is 15, HRMIR has slight advantage over RS-MDC, whereas when the GOP is 30, HRMIR outperforms RS-MDC significantly. In Figure 7, we fix the GOP length as 30 and PLR is set to 5 and 10%. It is interesting to see that when the PLR is 10%, the superiority of HRMIR over RS-MDC is larger than the case that when

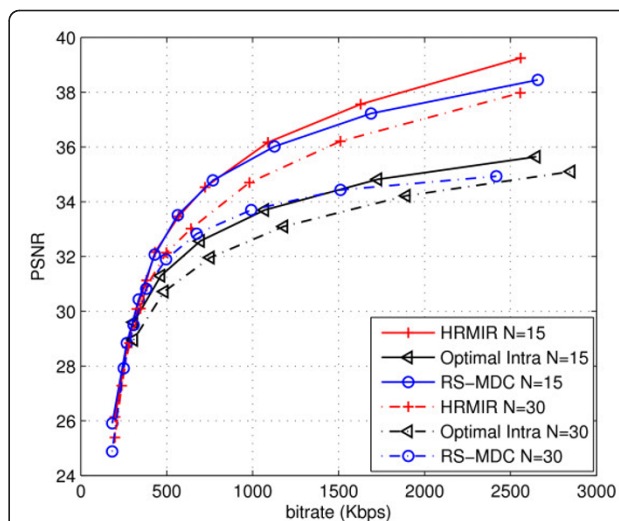


a. Foreman, 1/4-pixel ME

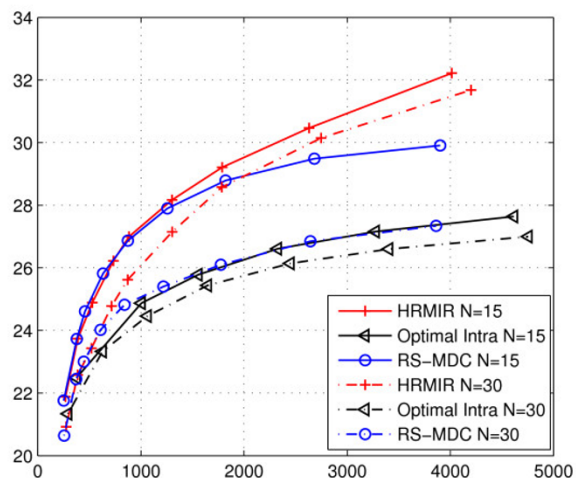


b. Bus, 1/4-pixel ME

**Figure 5** Frame-by-frame average PSNR comparison for HRMIR, Optimal Intra and RS-MDC, PLR is 10%, 1/4-pixel accuracy motion estimation. a Foreman CIF 30 fps, 1.48 Mbps. b Bus CIF 30 fps, 1.92 Mbps.



a. Foreman



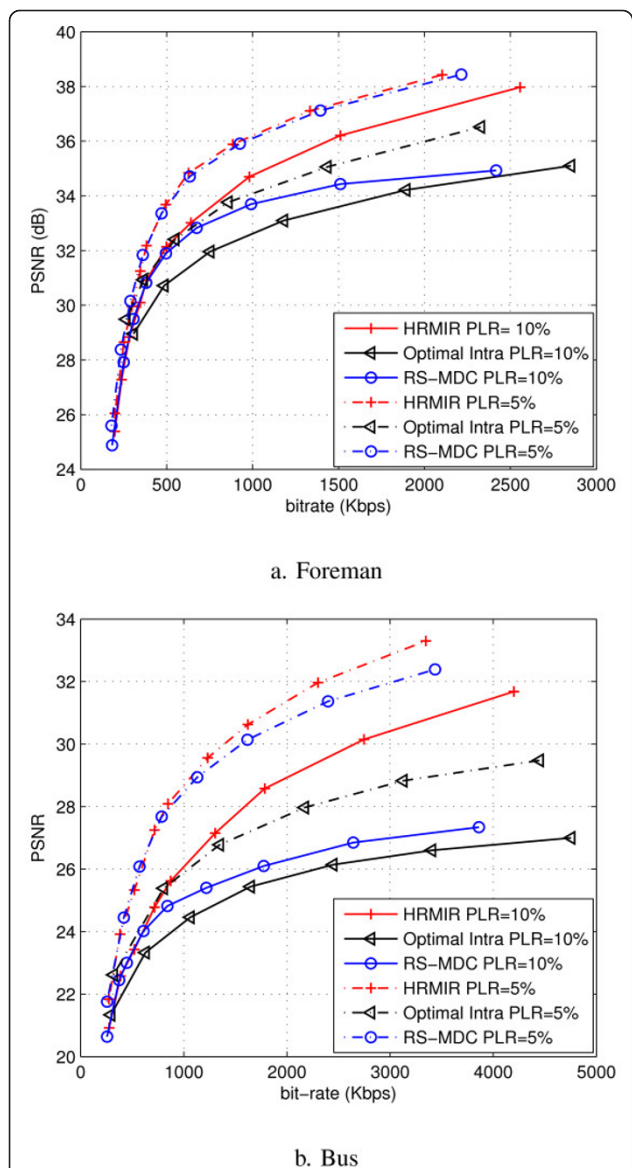
b. Bus

**Figure 6** PSNR versus bit rate for HRMIR, Optimal Intra and RS-MDC, PLR is 10%, GOP length  $N = 15$  and  $30$ , a CIF Foreman sequence, b CIF Bus sequence.

PLR is 5%. This phenomenon is because with long GOP and high packet loss rate, only providing redundant information cannot protect the video quality properly. Furthermore, for both the Foreman and Bus sequences, the HRMIR provides much higher PSNR than Optimal Intra in all the simulation environments. Let us take the Bus sequence for example, when PLR is 5% and GOP is 30, PSNR of HRMIR is about 4 dB higher than Optimal Intra with bitrate 2 Mbps. Note that in both Figures 6 and 7, when the bitrate is low, the PSNR of HRMIR and RS-MDC is nearly same; this is because in this case, very few Intra macroblocks are inserted, which makes

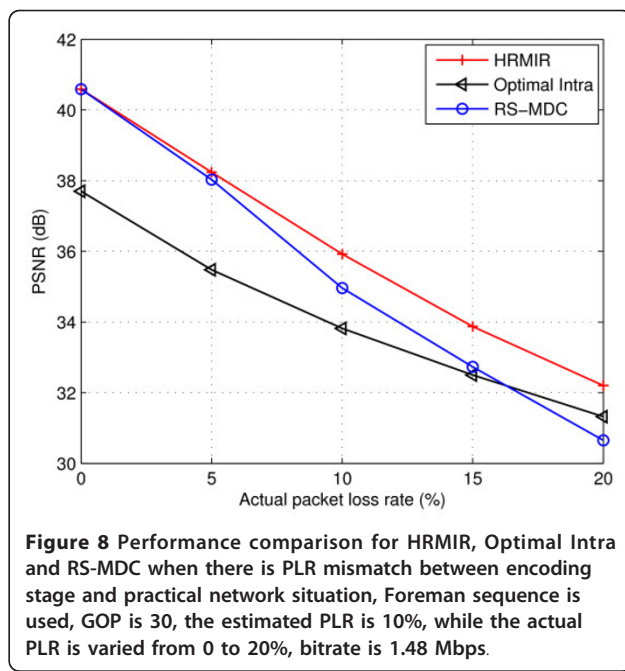
HRMIR approach similar as RS-MDC approach. Furthermore, as the QP values of different macroblocks in the proposed HRMIR approach are not identical, additional bits are needed to encode the residual QP value.

In all the previous experiments, the channel packet loss rate is assumed to be available at the encoder, and this can be implemented with the Real Time Control Protocol (RTCP) [18]. However, in practical situation, feedback packet loss rate information may be delayed from the decoder. Therefore, the packet loss rate used by the encoder in its RD optimization process may not



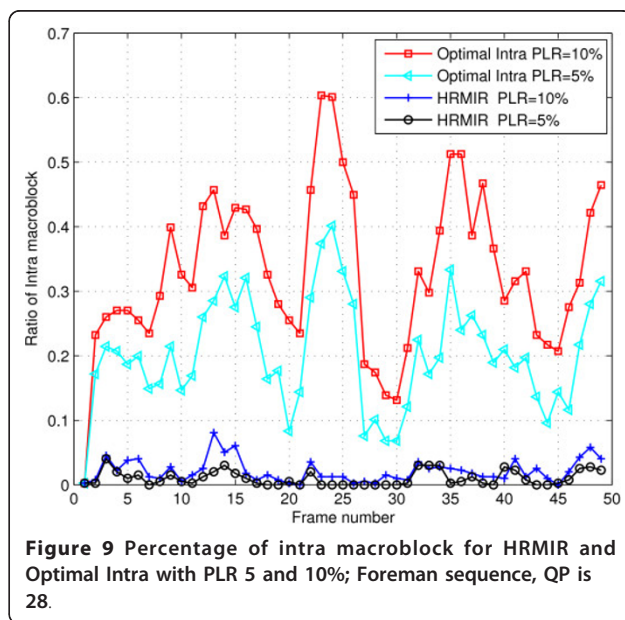
**Figure 7** PSNR versus bit rate for HRMIR, Optimal Intra and RS-MDC, PLR is 5 and 10%, GOP length  $N = 30$ , a CIF Foreman sequence, b CIF Bus sequence.

be exactly identical to the actual packet loss rate. To further evaluate the performances of the proposed HRMIR approach at the case when the estimated packet loss rate does not match the actual one, we use 10% as packet loss rate in the RD optimization process, whereas the actual packet loss rate is varied from 0 to 20%. In Figure 8, the HRMIR, Optimal Intra and RS-MDC approaches are all optimized for 10% packet loss rate. The encoded bitrate of HRMIR is 1.48 Mbps, whereas for the other two approaches, the encoded bitrate is close to but no less than the that of HRMIR approach. In the actual PLR range of [0-20]%, the PSNR of HRMIR is the highest among the three approaches,



which means when there is PLR mismatch, the HRMIR still can provide best video quality among the three approaches. Meanwhile, the gap between HRMIR and RS-MDC increases with actual PLR; therefore, when actual packet loss rate is high, RS-MDC fails to protect the video quality properly.

In Figure 9, we study how intra macroblocks are allocated in two different encoding approaches. CIF sequence Foreman is used, QP is set to 28, and the first 50 frames are used. Interestingly, the total percentage of intra macroblocks (both intra macroblocks with and





**Table 1 Percentage of intra macroblocks for HRMIR and Optimal Intra, QP is 28, first 50 frames are used, PLR is set to 3, 5, 10 and 20%**

Video	Approach	3%	5%	10%	20%
Foreman	HRMIR	0.71	1.02	2.14	5.87
	Optimal Intra	13.86	20.31	33.18	48.01
Bus	HRMIR	2.04	3.66	9.38	25.61
	Optimal Intra	53.41	64.91	78.07	89.49
Mobile	HRMIR	0.55	0.99	3.04	9.59
	Optimal Intra	26.53	41.27	66.72	84.69

without redundant coding) increases with the PLR in both the Optimal Intra and HRMIR approaches. This can be explained in the following manner, with high packet loss rate, the possibility of propagated mismatch error is high, then more intra macroblocks are required to cut-off the mismatch propagation. Meanwhile, with the same packet loss rate, the HRMIR approach allocates much less intra macroblocks than Optimal Intra. This is because there are two tools available for error-resilient coding with the HRMIR approach. Therefore, for some macroblocks, providing redundant coding leads to better usage of bitrate resource than intra coding. More statistics information about intra macroblock allocation can be found in Table 1.

Many papers [20-22] have addressed the actual network loss behavior, and most of them agree that Internet packet loss often exhibits finite temporal dependency, which means if current packet is lost, then the next packet is also likely to be lost. This leads to burst packets loss [20]; the average burst length for the Internet is two. Therefore, besides *i.i.d.* random packet loss model, we also use burst loss model for simulation,

and as indicated in [20], we set the average burst length as two. In Figure 10, the PSNR versus bitrate curves in burst loss environments are plotted. The results are similar with that in the *i.i.d.* case, and the proposed HRMIR approach can provide best video quality among the three approaches. The error-resilient performance of proposed HRMIR approach is robust on different error distribution models.

## V. Conclusions

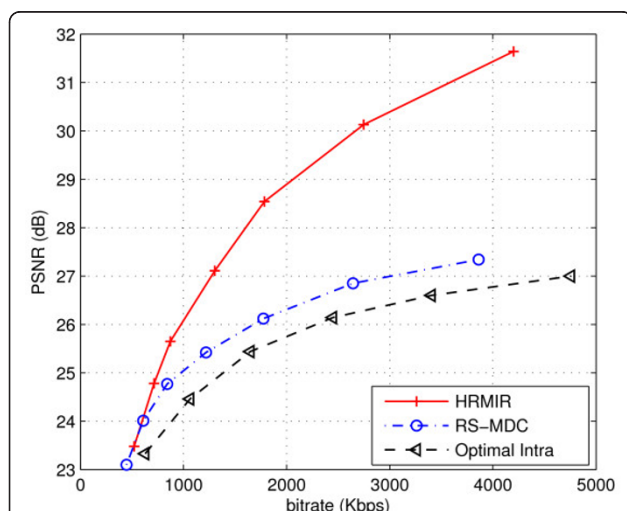
In this paper, a novel Hybrid Redundant Macroblock and Intra macroblock Refreshment approach has been proposed to combat packet loss. In the proposed approach, redundant coding and/or intra coding are optimally allocated in macroblock level. Whether to use redundant coding and/or intra coding and the quantization parameter of the redundant coding is all determined in the end-to-end rate-distortion optimization procedure. It is worth mentioning that, in the proposed approach, only information from the previously encoded frames is used to calculate the end-to-end distortion in the RDO process; therefore, no additional delay is caused, making the proposed approach suitable for real-time applications such as video conference. Extensive experimental results show that the proposed method provides better performance than other error-resilient source coding approaches. The performance gap between the proposed approach and the Optimal Intra Refreshment is huge, and in some simulation environments, the proposed approach can provide 4 dB higher PSNR than the conventional Optimal Intra Refreshment with the same bitrate. Our future work is to calculate the end-to-end distortion in sub-pixel accuracy; therefore, more accurate end-to-end distortion would be available, which would eventually lead to better resource allocation.

## VI. Competing interests

The authors declare that they have no competing interests.

## VII. Acknowledgements

This work was supported by National Natural Science Foundation of China (No. 60972085, No. 60903066), the Sino-Singapore JRP (No. 2010DFA11010)



**Figure 10 Performance comparison for HRMIR, Optimal Intra and RS-MDC when the packet loss is burst, PLR is 10%, burst length is two, Bus sequence is used, GOP is 30.**

and National Science Foundation of China for Distinguished Young Scholars (No. 61025013).

#### Author details

<sup>1</sup>Department of Electrical Engineering and Electronics, The University of Liverpool, Liverpool L69 3GJ, UK <sup>2</sup>Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, 111 Ren Ai Road, Suzhou, People's Republic of China <sup>3</sup>Institute of Information Science, Beijing Jiaotong University, Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing 100044, People's Republic of China

Received: 18 February 2011 Accepted: 30 September 2011

Published: 30 September 2011

#### References

1. T Wiegand, GJ Sullivan, G Bjøntegaard, A Luthra, Overview of the H.264/AVC video coding standard. *IEEE Trans Circuits Syst Video Technol.* **13**(7), 560–576 (2003)
2. S Wenger, H.264/AVC over IP. *IEEE Trans Circuits Syst Video Technol.* **13**(7), 645–656 (2003)
3. T Stockhammer, MM Hannuksela, T Wiegand, H.264/AVC in wireless environments. *IEEE Trans Circuits Syst Video Technol.* **13**(7), 657–673 (2003). doi:10.1109/TCSVT.2003.815167
4. G Cote, F Kossentini, Optimal intra coding of blocks for robust video communication over the internet. *Signal Process Commun.* **15**, 25–34 (1999). doi:10.1016/S0923-5965(99)00022-3
5. QF Zhu, L Kerofsky, Joint source coding, transport processing and error concealment for H.323-based packet video, in *Proceedings of the SPIE, VCIP 99*, vol. 3653. San Jose, CA, 52–62 (1999)
6. R Zhang, SL Regunathan, K Rose, Video coding with optimal inter/intra-mode switching for packet loss resilience. *IEEE J Sel Areas Commun.* **18**(6), 966–976 (2000). doi:10.1109/49.848250
7. T Stockhammer, D Kontopodis, T Wiegand, Rate-distortion optimization for JVT/H.26L coding in packet loss environment, in *Proceedings of Packet Video Workshop 2002*, Pittsburgh, PA (2002)
8. CB Zhu, YK Wang, MM Hannuksela, HQ Li, Error resilient video coding using redundant pictures. *IEEE Trans Circuits Syst Video Technol.* **19**(1), 3–14 (2009)
9. T Tillo, M Grangetto, M Olmo, Redundant slice optimal allocation for H.264 multiple description coding. *IEEE Trans Circuits Syst Video Technol.* **18**(1), 59–70 (2008)
10. Y Wang, SA Lin, Error-resilient video coding using multiple description motion compensation. *IEEE Trans Circuits Syst Video Technol.* **12**(6), 438–452 (2002). doi:10.1109/TCSVT.2002.800320
11. I Radulovic, P Frossard, YK Wang, M Hannuksela, A Hallapuro, Multiple description video coding with H.264/AVC redundant pictures. *IEEE Trans Circuits Syst Video Technol.* **20**(1), 144–148 (2010)
12. CY Lin, T Tillo, Y Zhao, B Jeon, Multiple description coding for H.264/AVC with redundancy allocation at macro block level. *IEEE Trans Circuits Syst Video Technol.* **21**(5), 589–600 (2011)
13. GJ Sullivan, T Wiegand, Rate-distortion optimization for video compression. *IEEE Signal Process Mag.* **15**(6), 74–90 (1998). doi:10.1109/79.733497
14. H Yang, 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.* **17**(7), 845–856 (2007)
15. Y Liao, JD Gibson, Enhanced error resilience of video communications for burst losses Using an extended ROPE algorithm, in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, 1853–1856 (2009)
16. Y Zhang, W Gao, Y Lu, Q Huang, D Zhao, Joint source-channel rate-distortion optimization for H.264 video coding over error-prone Networks. *IEEE Trans Multimedia* **9**(3), 445–454 (2007)
17. F Li, G Liu, Compressed-domain-based transmission distortion modeling for precoded H.264/AVC video. *IEEE Trans Circuits Syst Video Technol.* **19**(20), 1908–1914 (2009)
18. H Schulzrinne, S Casner, R Frederick, V Jacobson, *RTP: a transport protocol for real-time applications*. Internet Engineering Task Force–RFC 1889 (1996)
19. H.264/AVC JM Reference Software <http://iphome.hhi.de/suehring/tml/download>
20. D Loguinov, H Radha, End-to-end internet video traffic dynamics: statistical study and analysis, in *Proceedings of IEEE INFOCOM '02*, 723–732 (2002)
21. YJ Liang, JG Apostolopoulos, B Girod, Analysis of packet loss for compressed video: effect of burst losses and correlation between error frames. *IEEE Trans Circuits Syst Video Technol.* **18**(7), 861–874 (2008)
22. ZC Li, J Chakareski, XD Niu, YJ Zhang, WY Gu, Modeling and analysis of distortion caused by Markov-Model burst packet losses in video transmission. *IEEE Trans Circuits Syst Video Technol.* **19**(7), 917–931 (2009)

doi:10.1186/1687-6180-2011-80

Cite this article as: Xiao et al.: Error-resilient video coding with end-to-end rate-distortion optimized at macroblock level. *EURASIP Journal on Advances in Signal Processing* 2011 **2011**:80.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)