ENHANCED ERROR RESILIENCE OF VIDEO COMMUNICATIONS FOR BURST LOSSES USING AN EXTENDED ROPE ALGORITHM

Yiting Liao and Jerry D. Gibson

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA, 93106
Email: {yiting, gibson}@ece.ucsb.edu

ABSTRACT

Video communications over wireless networks suffers various patterns of losses, including burst losses that cause great degradation in video quality. In this paper, we propose an algorithm based on recursive optimal per-pixel estimate (ROPE) to accurately estimate the overall distortion accounting for the loss pattern. The estimated distortion is applied to the rate-distortion (RD)-based mode selection to provide the optimal tradeoff between intra and inter coding. Simulation results show that in lossy networks, the proposed extended ROPE algorithm achieves gains in average PSNR and $\mathrm{PSNR}_{r,f}$ by up to 0.8 dB and 3.6 dB respectively. This shows that the proposed algorithm can enhance error resilience for video communications over networks with burst losses.

Index Terms— video communications, error resilience, packet loss, rate-distortion

1. INTRODUCTION

One challenging problem for video communications over wireless networks is to provide error resilience for reliable communications. A number of techniques have been proposed to enhance the error robustness of video communications over such lossy networks, such as intra/inter mode selection [1], reference picture selection [2] [3], and multiple description coding [4]. Some of these approaches achieve a significant improvement using the rate-distortion (RD) based mode selection methods. For these techniques, distortion caused by channel losses is considered as well as the distortion due to compression.

In reference [1], an algorithm called "Recursive Optimal Perpixel Estimate" (ROPE) is proposed to estimate the overall distortion due to quantization, error propagation, and error concealment and uses rate-distortion optimization to choose the best intra/inter mode for each macroblock (MB). In the estimation model, they only consider a simple packet loss pattern, that is, every packet can be lost with a packet loss probability p and the loss or reception of each packet is independent. However, references [5] and [6] have shown that not only average packet loss rate but also the specific pattern of the loss affects the expected distortion; specifically, [6] shows that burst length has a great impact on the distortion. Because of the likelihood of burst losses in video communications over wireless networks, we propose an extended ROPE algorithm that accurately estimates the overall distortion under different loss patterns at the encoder. It can be applied for RD-based optimal mode selection. We demonstrate that this extended ROPE method achieves better objective and subjective video quality than ROPE under different loss patterns, which enhances error resilience for video communications over wireless networks.

The paper is organized as follows: Section 2 introduces the packet loss model and the RD-based optimal mode selection method. Our extended ROPE algorithm for burst losses is proposed in Section 3 and simulation results in Section 4 show the effectiveness of the method.

This research has been supported by the California Micro Program, Applied Signal Technology, Cisco, Sony-Ericsson and Qualcomm, Inc, and by NSF Grant Nos. CCF-0429884, CNS-0435527, and CCF-0728646.

2. BACKGROUND

In this section, we introduce the packet loss model for the wireless networks and outline RD-based optimal mode selection method.

2.1. Packet Loss Model

In order to investigate the extended ROPE for burst losses, we first illustrate the underlying packet loss model as shown in Fig. 1. Time is

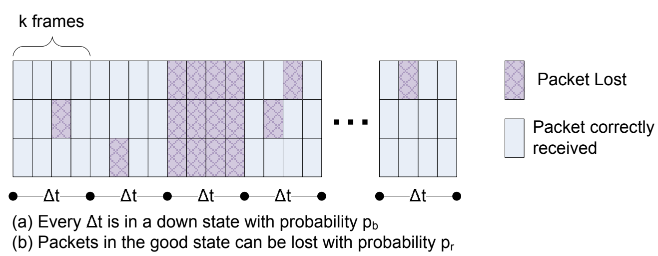

Fig. 1. Packet Loss Model

divided into Δt intervals and each interval corresponds to k frames. Each interval may be either in a good state with probability $(1-p_b)$ or in a down state with probability p_b, which is independent and identically distributed. The packets transmitted in a down state are all lost while the packets transmitted in the good state may suffer a random packet loss. Therefore, the packet loss model can be determined by three parameters: the average burst loss rate p_b, the burst length k (frames), and the random packet loss rate in a good state p_r. The total packet loss rate p in the network can be calculated by:

$$p = p_b + (1-p_b)p_r = p_b + p_r - p_b p_r \qquad (1)$$

2.2. RD-based Optimal Mode Selection

Video standards such as H.264 provide different Intra and Inter modes to encode a MB. To decide the best mode for each MB, a Lagrangian optimization technique is used to minimize the distortion subject to a rate constraint [7]. The optimal mode that minimizes the Lagrangian cost in the following equation is chosen to code the MB,

$$\min_{mode}(J_{\mathrm{MB}}) = \min_{mode}(D_{MB} + \lambda R_{\mathrm{MB}}) \tag{2}$$

where λ is the Lagrangian multiplier, R_{MB} denotes the bits needed for coding the MB in the specific mode, and D_{MB} represents the distortion of the MB. In the next section, we propose the extended ROPE algorithm to estimate the decoder-reconstructed pixel value with both random and burst losses, and use it to calculate the overall distortion D_{MB} for mode selection.

3. EXTENDED ROPE WITH BURST LOSSES

Reference [1] proposed the ROPE algorithm to estimate the overall distortion of each pixel caused by quantization, error propagation, and error concealment. Based on the notation in Table 1, the distortion of each MB can be calculated by

Table 1. Notations

	Definitions
d_n^i	Distortion of pixel i in frame n
f_n^i	Original value of pixel i in frame n
$\hat{f}_n^i$	Encoder-reconstructed value of pixel i in frame n
$\tilde{f}_n^i$	Decoder-reconstructed value of pixel i in frame n (after error concealment)
$\hat{e}_n^i$	Quantized residue of pixel i in frame n (Inter)

$$\begin{aligned} D_{\mathrm{MB}} &= \sum_{i \in \mathrm{MB}} d_n^i = \sum_{i \in \mathrm{MB}} E[(f_n^i - \tilde{f}_n^i)^2] \\ &= \sum_{i \in \mathrm{MB}} \left({f_n^i}^2 - 2 f_n^i E[\tilde{f}_n^i] + E[{\tilde{f}_n^i}{}^2] \right) \end{aligned} \tag{3}$$

The equations used to estimate the first and second moments of $\tilde{f}_n^i$ can be found in [1]. We notice that only independent packet loss rate is considered when estimating the error propagation due to packet loss in this formulation. However, [5] and [6] show that loss pattern, especially burst loss, has a significant impact on the distortion, which leads us to derive the extended ROPE algorithm with burst losses. The extended ROPE algorithm can more accurately estimate the reconstructed pixel value by considering the random and burst loss separately.

As mentioned in Section 2.1, the burst loss rate p_b, burst length k (frames), and random packet loss rate p_r are considered in the extended ROPE algorithm. Temporal-copy error concealment is used to recover the lost video segment. Using the notations in Table 1, we calculate the first and second moments of $\tilde{f}_n^i$ in Intra or Inter mode with burst losses as follows.

For an intra-coded MB,

$$\begin{aligned} E[\tilde{f}_n^i] &= (1-p_r)(1-p_b)(\hat{f}_n^i) + (1-p_b)p_r E[\tilde{f}_{n-1}^i] \\ &\quad + p_b E[\tilde{f}_{n-(n \bmod k)}^i] \end{aligned} \tag{4}$$

$$\begin{aligned} E[(\tilde{f}_n^i)^2] &= (1-p_r)(1-p_b)(\hat{f}_n^i)^2 + (1-p_b)p_r E[(\tilde{f}_{n-1}^i)^2] \\ &\quad + p_b E[(\tilde{f}_{n-(n \bmod k)}^i)^2] \end{aligned} \tag{5}$$

For an inter-coded MB,

$$\begin{aligned} E[\tilde{f}_n^i] &= (1-p_r)(1-p_b)(\hat{e}_n^i + E[(\tilde{f}_{n-1}^j)]) \\ &\quad + (1-p_b)p_r E[\tilde{f}_{n-1}^i] + p_b E[\tilde{f}_{n-(n \bmod k)}^i] \end{aligned} \tag{6}$$

$$\begin{aligned} E[(\tilde{f}_n^i)^2] &= (1-p_r)(1-p_b)E[(\hat{e}_n^i + \tilde{f}_{n-1}^j)^2] + (1-p_b)p_r \\ &\quad E[(\tilde{f}_{n-1}^i)^2] + p_b E[(\tilde{f}_{n-(n \bmod k)}^i)^2] \end{aligned} \tag{7}$$

Equations (4)-(7) all contain three terms: The first term calculates the reconstructed value when the packet is correctly received. The second term estimates the reconstructed pixel value for a random packet loss. The last term estimates the reconstructed pixel value under burst losses.

4. SIMULATION RESULTS

4.1. Simulation Settings

In this section, we examine the performance of our method under burst losses. Based on the packet loss model in Section 2.1, we set $p_r = 0$ and run simulations under different burst loss rate p_b and burst length k (frames). In [8] [9], the common test conditions for video transmission over wireless networks are discussed and for a RTP size of 100 bytes, the common condition of packet loss rate is in the range of 0%-20%. Reference [8] also mentions that packet loss rates of 10% show annoying artifacts even when high error resilience strength is used. Therefore, for our simulation, a burst loss rate in the range of 1% -10% is chosen. We know that in wireless networks the losses can be very bursty and can cause a loss of multiple frames [5]. Here a burst length between 2-10 frames is chosen, since for a burst length longer than 10 frames, the distortion can be very large and we may reoptimize the network to avoid this kind of long burst losses.

The original ROPE and extended ROPE methods are implemented by modifying JM 13.2, which is the reference software for H.264. The video sequence is encoded and packetized to RTP format. The video sequences used for simulations contain 300 frames with a frame rate 30 fps and are in QCIF (176 $\times$ 144) format. For original ROPE, only the total packet loss rate calculated by Eq.(1) is considered at the encoder for the optimized mode selection, while the extended ROPE considers all three parameters in the packet loss model when encoding. A burst loss generator is used to generate the burst loss pattern based on the model in Section 2.1 and drops packets accordingly. For a specified burst loss rate and burst length, 500 realizations are simulated to show the performance of the two methods.

In order to analyze the performance of the decoded video sequences, the average PSNR of all frames over all realizations is presented. Also a video quality indicator $PSNR_{r,f}$ proposed in [10] is used. $PSNR_{r,f}$ can capture the performance loss due to damaged frames in a single video sequence (f%) and also the specific quality that a user would experience in multiple uses of the channel(r%). A set of typical values for r and f is $f = 90\%, r = 85\%$.

4.2. Performance Evaluation

The Foreman sequence is encoded using ROPE and extended ROPE at a fixed bitrate of 300 kbps under different burst loss rates and burst lengths. The burst loss generator generates 500 realizations for each set of burst loss rate and burst length. Table 2 and Table 3 show the average PSNR of the Foreman sequence under different burst loss rates and different burst lengths, respectively. The results in the tables show that extended ROPE achieves better PSNR than ROPE under different burst loss rates and patterns.

Table 2. Average PSNR (dB) of the Foreman sequence under different burst loss rates at a burst length 5

Burst Loss Rate	1%	3%	5%	8%	10%
ROPE	34.53	32.81	31.82	30.90	30.29
Extended ROPE	34.62	33.40	32.44	31.58	31.09

Table 3. Average PSNR (dB) of the Foreman sequence under different burst lengths at a burst loss rate 5%

Burst Length	2	4	6	8	10
ROPE	32.31	31.94	31.89	32.09	32.17
Extended ROPE	32.52	32.57	32.41	32.30	32.22

We introduce $PSNR_{r,f}$ to evaluate the video quality because average PSNR treats all frames equally, which is not as effective as $PSNR_{r,f}$ to represent the perceptual video quality [10]. $PSNR_{r,f}$ is defined as the PSNR achieved by $f\%$ of the frames for $r\%$ of the realizations, which shows the video quality guaranteed for $r\%$ of realizations of $f\%$ frames. We are motivated to use $PSNR_{r,f}$ because of two findings [10] [11]: (1) The bad-quality frames dominate users' experience with the video; (2) For PSNRs higher than a certain threshold, increasing PSNR does not help to enhance the perceptual quality.

Figure 2 and Figure 3 plot the $PSNR_{r=85\%,f=90\%}$ of ROPE and extended ROPE for the Foreman sequence under different burst loss rates and different burst lengths respectively. The figures show that the PSNRs achieved by 90% of the frames in 85% of the realizations for extended ROPE under different burst loss rates and patterns are 1.0-3.6 dB higher than original ROPE. The results show that even though ROPE and extended ROPE both achieve high average PSNR across all frames and realizations, extended ROPE implies much better perceptual quality because a lower percentage of the frames experiences extremely bad quality which can dominate human experience with the video.

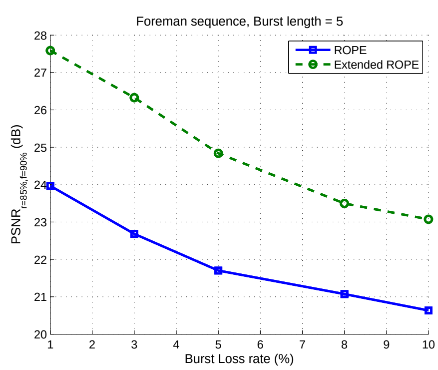

Fig. 2. $PSNR_{r=85\%,f=90\%}$ versus burst loss rate at a fixed burst length 5 for the Foreman sequence

We also examine the performance of four video sequences Carphone, Foreman, Mother-daughter, and Salesman under burst loss rate 5% and burst length 5. The average PSNR and $PSNR_{r=85\%,f=90\%}$ are presented in Table 4 on the next page.

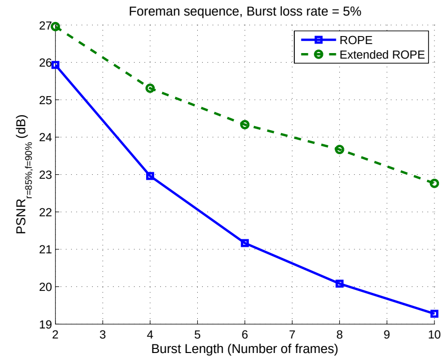

Fig. 3. $PSNR_{r=85\%,f=90\%}$ versus burst length at a fixed burst loss rate 5% for the Foreman sequence

The results support the claim that extended ROPE outperforms ROPE under burst losses. It is important to notice how different $PSNR_{r,f}$ is than PSNR.

It is possible that the burst loss rate and burst length assumed at the encoder are different from the actual burst loss rate and burst length in the network. In order to investigate the robustness of extended ROPE, we assume that the burst loss rate and burst length known at the encoder are 5% and 5 respectively, while the actual burst loss rate in the network varies between 1% to 10% and the actual burst length is in the range of 2-10.

Table 5 and Table 6 show the average PSNR of ROPE and extended ROPE under different burst loss rates and patterns when there is a mismatch between the design network conditions and actual network conditions. We see that extended ROPE has higher PSNR than ROPE in most cases. Only when actual burst loss rate is 1%, extended ROPE achieves lower PSNR. The reason may be that extended ROPE wastes more bits to enhance the error robustness while the actual burst loss rate is lower than the expected burst loss rate.

Table 5. Average PSNR (dB) of Foreman sequence under different burst loss rates at a fixed burst length 5 for the Foreman sequence, burst loss rate and burst length known at the encoder is 5% and 5

Burst Loss Rate	1%	3%	5%	8%	10%
ROPE	34.38	32.98	31.82	30.25	29.29
Extended ROPE	33.71	33.02	32.44	31.57	30.97

Table 6. Average PSNR (dB) of the Foreman sequence under different burst lengths at a fixed burst loss rate 5% for foreman sequence, burst loss rate and burst length known at the encoder is 5% and 5

Burst Length	2	4	6	8	10
ROPE	32.28	32.01	31.91	32.09	32.36
Extended ROPE	32.56	32.45	32.35	32.38	32.37

Figure 4 shows $PSNR_{r=85\%,f=90\%}$ versus actual burst loss rate in the network when the burst loss rate known at the encoder is fixed at 5% and Fig. 5 shows $PSNR_{r=85\%,f=90\%}$ versus actual burst length in the network when the burst length known at the encoder is fixed at 5. In both figures, we see that extended ROPE has higher $PSNR_{r=85\%,f=90\%}$ than ROPE in all cases, which shows that extended ROPE provides better video quality than ROPE when mismatch between expected and actual the network conditions occurs.

Table 4. Average PSNR and $PSNR_{r,f}$ for different video sequences at a burst loss rate 5% and burst length 5

$PSNR_{r,f}$ (dB)	Carphone		Foreman		Mother-daughter		Salesman	
	ROPE	E-ROPE	ROPE	E-ROPE	ROPE	E-ROPE	ROPE	E-ROPE
Average PSNR	34.11	34.32	31.82	32.44	39.89	40.28	40.29	40.33
$r = 85\%, f = 90\%$	23.84	25.48	21.70	24.84	30.84	34.37	32.89	35.04

*E-ROPE denotes the extended ROPE method.

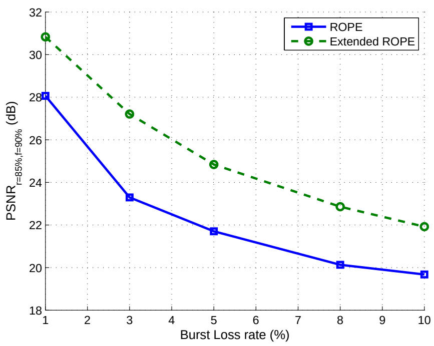

Fig. 4. $PSNR_{r=85\%,f=90\%}$ versus burst loss rate in the network at a fixed burst length 5 for the Foreman sequence, burst loss rate and burst length known at the encoder is 5% and 5

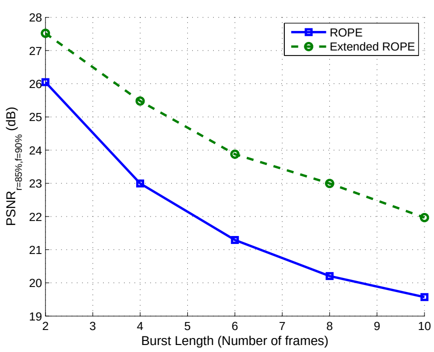

Fig. 5. $PSNR_{r=85\%,f=90\%}$ versus burst length in the network at a fixed burst loss rate 5% for the Foreman sequence, burst loss rate and burst length known at the encoder is 5% and 5

5. CONCLUSIONS

This paper proposes an extended ROPE algorithm that accurately estimates the distortion due to various loss patterns and applies it for optimal mode selection using rate-distortion optimization. We compare extended ROPE to ROPE under different random and burst loss patterns. The results show that extended ROPE achieves higher average PSNR and $PSNR_{r,f}$ than ROPE. We also examine the performance of extended ROPE and ROPE when there is a mismatch between the assumed and actual network conditions. The results verify the robustness of extended ROPE.

6. REFERENCES

[1] R. Zhang, S.L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *Selected Areas in Communications, IEEE Journal on*, vol. 18, no. 6, pp. 966–976, Jun 2000.

[2] T. Wiegand, N. Farber, K. Stuhlmuller, and B. Girod, "Error-resilient video transmission using long-term memory motion-compensated prediction," *Selected Areas in Communications, IEEE Journal on*, vol. 18, no. 6, pp. 1050–1062, Jun 2000.

[3] M. Budagavi and J.D. Gibson, "Multiframe video coding for improved performance over wireless channels," *Image Processing, IEEE Transactions on*, vol. 10, no. 2, pp. 252–265, Feb 2001.

[4] Y. Wang, A.R. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 57–70, Jan. 2005.

[5] Y.J. Liang, J.G. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: does burst-length matter?," *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, vol. 5, pp. V–684–7 vol.5, April 2003.

[6] Y.J. Liang, J.G. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: Effect of burst losses and correlation between error frames," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, no. 7, pp. 861–874, July 2008.

[7] T. Wiegand, M. Lightstone, D. Mukherjee, T.G. Campbell, and S.K. Mitra, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 6, no. 2, pp. 182–190, Apr 1996.

[8] G. Roth, R. Sjöberg, G. Liebl, T. Stockhammer, V. Varsa, and M. Karczewicz, "Common test conditions for RTP/IP over 3GPP/3GPP2," *ITU-T SG16 Doc. VCEG-M77, Austin, TX, USA, Apr*, 2001.

[9] S. Wenger, "On the equivalence of BERs and packet loss rates," *JVT-B025. doc, Jan*, 2002.

[10] J. Hu, S. Choudhury, and J. D. Gibson, "PSNR$_{r,f}$: Assessment of delivered AVC/H. 264 video quality over 802.11a WLANs with multipath fading," *Proceedings of the First Multimedia Communications Workshop*, 2006.

[11] Z. Wang, H.R. Sheikh, and A.C. Bovik, "Objective video quality assessment," *The Handbook of Video Databases: Design and Applications*, pp. 1041–1078, 2003.