A NOVEL SOURCE-CHANNEL CONSTANT DISTORTION MODEL AND ITS APPLICATION IN ERROR RESILIENT FRAME-LEVEL BIT ALLOCATION

Hua Yang

ECE Department
Univ. of California Santa Barbara
Santa Barbara, CA, 93106, USA

Ligang Lu

IBM T.J. Watson Research Center
1101 Kitchawan Road
Yorktown Heights, NY, 10598, USA

ABSTRACT

Compared with either heuristic or RD optimization based approaches, frame-level bit allocation strategies based on constant quality models are inadequately studied, which, however, are inherently more capable on achieving consistent video quality over the whole sequence for real-time single-pass rate control tasks. Moreover, no any existent models is error-resilient. In this paper, we propose a novel error resilient source-channel constant distortion (SCCD) model, whose uniquely selected frame complexity measure effectively accomodates the joint impact from both the non-stationary source video signal and the lossy channel. Compared with its existent counterparts, our model has a higher modelling accuracy, and can be generally applied in frame-level bit allocation to enhance error robustness of rate control. As an example, an error resilient frame-level bit allocation solution is proposed based on the model. Simulation results demonstrate that: for lossy video transmission, the proposed method can achieve more consistent and higher overall visual qualities than that of the other non-error-resilient schemes, which proves the remarkable potentials of the new model.

1. INTRODUCTION

Rate control is widely recognized as an important strategy to counter the inherent non-stationarity of video signal for uniform output quality [1]. It generally consists of bit allocation and bit achievement [2] [3]. In terms of frame-level bit allocation, the problem can be stated as: for a given target bit rate, how to efficiently allocate the available bits to frames of differing complexities such that a consistent video quality can be achieved over the whole video sequence.

Conventional frame-level bit allocation schemes assume error-free channel conditions, and thus, are not error-resilient. Besides those simple heuristic solutions as in H.263 TMN8 or MPEG-4 Annex L etc., most of the other conventional schemes are based on RD optimization. However, to achieve universally constant quality, the globally optimal solution is required, which entails a multi-pass optimization process [4]. This is infeasible for real-time video transmission, where single-pass rate control is necessary. On the other hand, a different way for frame-level bit allocation is based on constant quality models, where a proper amount of bits are allocated to each frame once and for all, based only on its own complexity value [1] [3]. In spite of its non-optimality, for real-time single-pass rate control tasks, a constant quality model based bit allocation scheme is inherently more capable in obtaining universally consistent quality. Nevertheless, to the best of our knowledge, not so much effort has been focused on this type of method so far.

Nowadays, more and more research efforts have been devoted to video-over-network applications, where inevitable packet loss is usually involved. The concerned quality now becomes the *decoder reconstruction* quality, or the so called *end-to-end* quality, but no longer the *encoder reconstruction* quality as in conventional error-free rate control problems. However, regarding all the existent works on error resilient rate control, they are either heuristic approaches [2] [5] or RD optimization based approaches [6] [7]. Again, constant quality model based solutions are ignored.

Therefore, our effort in this work is rightly focused on this inadequately studied problem. In particular, the objective is to find an effective error resilient constant quality model, such that more consistent end-to-end video quality can be achieved over the whole sequence via error resilient frame-level bit allocation. Specifically, we propose a novel source-channel constant distortion (SCCD) model. Its uniquely designed frame complexity measure effectively takes into account the joint impact of both the non-stationary source signal and the lossy channel. As such, unlike all the existent counterparts, our SCCD model is error resilient, and can be generally applied to enhance error robustness of rate control. As an example, we present a frame-level bit allocation solution by applying our model in an existent sequence-based rate control scheme [3]. Simulation results demonstrate the overall quality improvement of our pro-

This work was performed during a summer internship at IBM T.J. Watson Research Center.

posed scheme over the other non-error-resilient approaches, which evidently reveals the remarkable potentials of the proposed model.

The rest of the paper proceeds as follows. In Section 2, details of our SCCD model and its complexity measure are discussed. The proposed error resilient bit allocation scheme and corresponding results are provided in Section 3 and Section 4, respectively. Finally, we state our conclusions in Section 5.

2. SOURCE-CHANNEL CONSTANT DISTORTION MODEL

A constant quality model is a specified relationship between the number of bits of a frame and the magnitude of its complexity, when the entire sequence is coded of consistent quality. If the model has a good match with actual results, the right amount of bits will be allocated to each frame according to its complexity value, and finally, the whole sequence will have small quality variations.

2.1. Source-Channel Frame Complexity Measure

A critical point in deriving a good constant quality model is to select a good complexity measure, which is generally a quantity to indicate the difficulty of rendering a certain frame with the same (or similar) quality as the others. Existent frame complexity measures include: the quantization scale in constant-bit-rate (CBR) encoding [1], and the DCT coefficient variance [8] or the mean absolute difference (MAD) [3] of the motion compensated residue of a frame. An apparent shortcoming of all the existent measures is that: what they intend to capture is only source coding complexity, but nothing about channel coding complexity, incurred by possible packet loss and consequent error propagation. Therefore, they are inefficient in the concerned video-over-lossy-network scenarios.

Herein, we propose a source-channel complexity measure based on expected mean squared difference (EMSD) of the motion compensated frame at the *decoder* given by (2). For comparison, (1) describes the existent MAD-based measure of [3]. (Note that the work of [3] on the MAD-based model is always referred as our competitive counterpart, as, to the best of our knowledge, its R-MAD model is so far the only explicitly proposed objective constant quality model.)

- Existent MAD-based complexity measure:

$$X_n = \frac{1}{A}\sum_i |f_n^i - \hat{f}_{n-1}^{i+mv_i}| \tag{1}$$

- Proposed EMSD-based complexity measure:

$$X_n = \sqrt{\frac{1}{A}\sum_i E(f_n^i - \tilde{f}_{n-1}^{i+mv_i})^2} \tag{2}$$

Here, X_n denotes the complexity of frame n. f_n^i is the original value of pixel i in frame n, and A is the total number of pixels in a frame. mv_i is the motion vector related to pixel i. $\hat{f}_{n-1}^{i+mv_i}$ and $\tilde{f}_{n-1}^{i+mv_i}$ denote the encoder and decoder reconstructed values of pixel $i + mv_i$ in frame $n-1$, respectively. We can see that: instead of using the prediction residue at the encoder as in [3], the proposed complexity measure depends on expectation of the prediction residue at the *decoder*, which additionally accomodates the impact from the lossy channel, and thus, renders our later derived SCCD model error resilient.

2.2. SCCD Model Derivation

Let R_n denote the number of allocated bits for frame n. We then need to derive a certain relationship between R_n and X_n, i.e., the SCCD model, according to actual coding results $\{X_n, R_n\}_{n=1}^N$, when all the N frames of a sequence are coded with the same designated distortion level D^*. The experiment can be briefly formulated as follows.

$$R_n^* = \min_{D_n=D^*,\{QP,IntraMB,FEC\}} R_n, \tag{3}$$

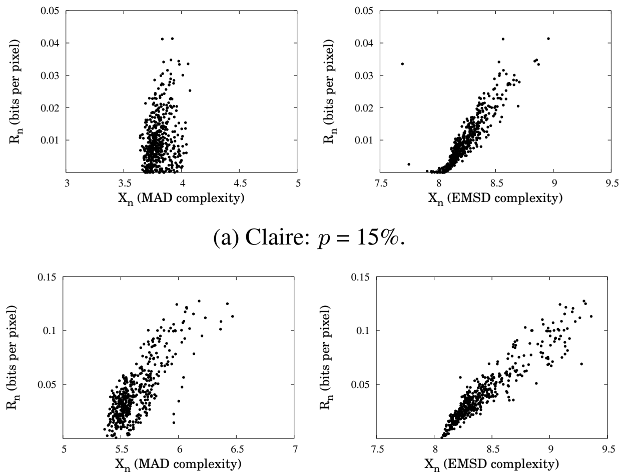
where

$$D_n = \frac{1}{A}\sum_i E(f_n^i - \tilde{f}_n^i)^2. \tag{4}$$

Note that to measure video quality, we assume the most commonly used objective quality measure, i.e., mean squared error (MSE) distortion. Moreover, to fit with the concerned scenario, the distortion herein is *end-to-end* MSE distortion given by (4). As shown in (3), R_n is reasonably defined as the minimum bit cost to maintain frame n at the desired distortion level D^*, and the assumed adjustable source and channel coding parameters are the quantization scale, Intra coded macroblocks (MB), and the protection level of forward error correction (FEC) channel coding. In experiment, for any possible $\{QP, FEC\}$ combination, a heuristic method is employed to select the MBs for Intra coding.

Besides, to compute the involved expectations in (2) and (4), we employ the recursive optimal per-pixel estimate (ROPE) method [9]. Thanks to its good performance, all the related end-to-end quantities in our problem can be accurately and efficiently estimated, which greatly contributes to the good performance of the proposed model.

The experiment results are provided in Fig. 1. (p denotes the packet loss rate.) It is easy to observe that: under lossy channel conditions, data points of R-EMSD demonstrate a more apparent regularity from their spatial distributions than that of the R-MAD result. This favorable result is



(a) Claire: p = 15%.

(b) Salesman: p = 10%.

Fig. 1. R_n vs. X_n. Frame rate: 30f/s, PSNR level: 30dB.

mainly attributed to the unique source-channel complexity measure. The derived SCCD model is then given by:

$$R_n = \begin{cases} K(X_n - d) + h_n & , \ if X_n \geq d \\ h_n & , \ otherwise. \end{cases} \quad (5)$$

where

$$d = \sqrt{D^*}. \quad (6)$$

h_n is a quantity to separately identify the bit cost for coding the motion vectors, header and other syntax of a frame, as they are mostly not dependent on frame complexity. d denotes the lower bound of X_n, meaning that: when $X_n < d$, it is enough to simply spent h_n bits in order to keep the frame at the desired distortion level D^*. Here, a nice property of our model is that: the lower bound d is exactly the square root of D^*, and hence, sequence-independent (as also verified in Fig. 1). This well justifies its usage in (5), which definitely renders an important contribution on the good modelling accuracy. In contrast, as the property is primarily attributed to the formal similarity between the EMSD in (2) and MSE distortion in (4), it is not available in the existent R-MAD model, where the lower bound of X_n is sequence-dependent, and thus, cannot be effectively utilized.

3. ERROR RESILIENT FRAME-LEVEL BIT ALLOCATION

The proposed SCCD model can be generally applied to replace its non-error-resilient counterparts in any existent constant quality based rate control scheme, and enhance error resilience. Herein, simply shown as an example to evaluate the performance of our model, we propose an error resilient frame-level bit allocation solution via modifying the sequence-based scheme of [3].

Assume a CBR rate control task with a target bit rate R_T. Let C denote the number of transmitted bits during a frame interval, and thus, $C = R_T/framerate$. Then, for frame n, while the allocated number of bits R_n is computed by (5), the involved K and d are updated by:

$$d^* = \sqrt{\bar{D}_{n-1}} \quad (7)$$

$$K^* = \frac{C - \bar{h}_{n-1}}{\bar{X}_{n-1} - d^*} \quad (8)$$

Here, $\bar{D}_{n-1}$, $\bar{X}_{n-1}$ and $\bar{h}_{n-1}$ are the averaged values of the corresponding coding results from all the previous $(n-1)$ frames, i.e., $\{D_i, X_i, h_i\}_{i=1}^{n-1}$. The usage of C, instead of $\bar{R}_{n-1}$, is to "help correct the bit allocation error and prevent error propagation resulted from some bad bit allocation" [3]. Finally, the resultant R_n is adjusted in the regular way to meet the CBR buffer constrains so as to avoid buffer overflow and underflow [3].

In simulation, we also need to specify a bit achievement algorithm to accurately and efficiently utilize the allocated number of bits on source and channel coding. As the purpose herein is simply to evaluate the performance of various bit allocation schemes, algorithm complexity is not a factor of primary concern. We devise a scheme as follows.

$$D_n^* = \min_{R_n = R^*, \{QP, IntraMB, FEC\}} D_n \quad (9)$$

Here, R^* is the allocated number of bits to frame n, i.e., the target number of bits for bit achievement. Note that the only difference between the scheme in (9) and the one in (3) is that: in bit achievement, we now intend to minimize D_n by exactly consuming R^* bits.

4. SIMULATION RESULTS

Our simulation is carried out with UBC H.263+ codec. Since MBs in a frame are adaptively Intra updated according to the lossy channel conditions, in the coded sequence, only the first frame is an I frame, and all the others are P frames. We assume a $1.5s$ buffer size, and initial buffer level is one half of the buffer size. Our SCCD model based rate control solution, and the existent R-MAD model based scheme are labelled as "SCCD RC" and "RMAD RC", respectively. We also implement the bit allocation scheme in MPEG-4 Annex L denoted as "MPG4 RC", and a naive bit allocation scheme "CNST RC", which allocates a constant number of bits for each frame. Simulation results are shown as follows.

It is easy to see from Fig. 2 that: while all the other three non-error-resilient schemes have similar output qualities, the proposed SCCD model based scheme achieves a more consistent PSNR quality across the whole sequence.

Table 1. Performance comparison of different bit allocation schemes. Akiyo: 30f/s, 100kb/s, p = 15%. Grandma: 10f/s, 64kb/s, p = 15%. News: 10f/s, 64kb/s, p = 10%. Claire: 10f/s, 64kb/s, p = 15%.

	$\overline{PSNR}$ (dB)				Std (dB)				$\overline{PSNR}_{min}$ (dB)			
	CNST	MPG4	RMAD	SCCD	CNST	MPG4	RMAD	SCCD	CNST	MPG4	RMAD	SCCD
Akiyo	35.24	35.33	35.28	35.52	0.71	0.70	0.73	0.38	33.73	33.78	33.79	34.66
Grandma	34.06	34.06	34.03	34.09	1.07	1.06	0.98	0.64	32.30	32.25	32.53	33.26
News	30.50	30.51	30.62	30.91	1.24	1.23	1.02	0.96	28.71	28.63	29.44	29.50
Claire	35.51	35.54	35.51	35.91	0.87	0.85	0.83	0.55	33.99	33.98	34.08	35.05

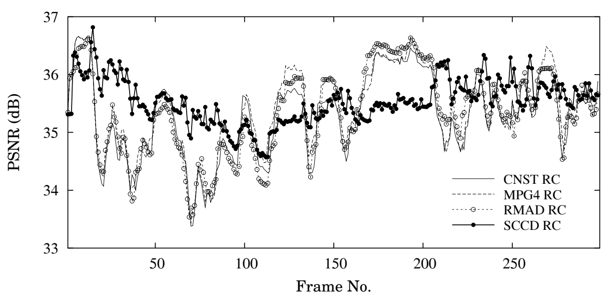


Fig. 2. Performance comparison: PSNR vs. Frame no. Akiyo: 30f/s, 100kb/s, p = 15%.

The performance improvement is further demonstrated in Table. 1, where the average PSNR ($\overline{PSNR}$), standard deviation of PSNR (Std), and average of the 10 minimum PSNR values ($\overline{PSNR}_{min}$) are provided. While "SCCD RC" always maintains a similar or higher $\overline{PSNR}$ than those of the others, noticeable performance improvement on Std and $\overline{PSNR}_{min}$ can also be observed. All these results demonstrate the distinguished potentials of the novel SCCD model for error resilient frame-level bit allocation.

5. CONCLUSIONS

In this paper, an important, but inadequately studied, problem of constant quality model based error resilient frame-level bit allocation is identified. A novel source-channel constant quality model is accordingly proposed. Due to its specially designed source-channel frame complexity measure, the proposed model is not only with good modelling accuracy but also uniquely error resilient. Simulation result of one proposed exemplar bit allocation solution based on this model confirms its remarkable potentials on improving error resilience of rate control.

6. REFERENCES

[1] P. H. Westerink, R. Rajagopalan, and C. A. Gonzales, "Two-pass MPEG-2 variable-bit-rate encoding," *IBM J. Res. Dev.*, vol. 43, no. 4, pp. 471-88, Jul.1999.

[2] H. Liu and M. E. Zarki, "Adaptive source rate control real-time wireless video transmission," *ACM/Baltzer Mobile Networks Applicat. J.*, vol. 3, no. 1, pp. 49-60, 1998.

[3] B. Xie and W. Zeng, "A sequence-based rate control framework for constant quality video," *submitted to IEEE Trans. Circuits Syst. Video Technol.*, 2003.

[4] J. Cai, Z. He, and C. W. Chen, "Optimal bit allocation for low bit rate video streaming applications," Proc. of IEEE ICIP 2002, Sept.2002.

[5] G. R. Rajugopal and R. H. M. Hafez, "Adaptive rate controlled, robust video communication over packet wireless networks," *ACM/Baltzer Mobile Networks and Applicat. J.*, vol. 3, no. 1, pp. 33-47, 1998.

[6] Z. He, J. Cai, and C. W. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," *IEEE Trans. on Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 511-23, June 2002.

[7] C. Y. Hsu, A. Ortega, and M. Khansari, "Rate control for robust video transmission over burst-error wireless channels," *IEEE J. on Select. Areas in Commun.*, vol. 17, no. 5, pp. 756-73, May 1999.

[8] J. Ribas-Corbera and S.-M. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 172-85, Feb. 1999.

[9] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal intra/inter mode switching for packet loss resilience," *IEEE J. Select. Areas in Commun.*, vol. 18, no. 6, p. 966-76, June 2000.