A Power-Optimized Joint Source Channel Coding for Scalable Video Streaming over Wireless Channel

Qian Zhang, Wenwu Zhu, Zu Ji, and Ya-Qin Zhang*
Microsoft Research, China
5F, Bejing Sigma Center, No. 49, Zhichun Road
Haidian District, Beijing 100080, P. R. China

ABSTRACT

Scalable video streaming over wireless link is a very challenging task due to the time-varying characteristics of wireless channel and limited battery resource in the handheld devices. This paper first describes a general approach for joint source channel coding (JSCC) with optimal power consumption. Then, an end-to-end power-optimized architecture for video streaming is presented. Therein the channel-adaptive hybrid UEP (Unequal Error Protection) and ARQ (Automatic Repeat reQuest) scheme is proposed to achieve the minimal power consumption. Simulation results demonstrate the effectiveness of our proposed scheme.

I. INTRODUCTION

With tremendous increase in wireless access, video communication over wireless is of great interest recently. However, this is a very challenging task because of the inherent unreliability of the wireless channel, which suffers from fading or shadowing effects. Meanwhile, there is limited battery life in the mobile host. To cope with errors in wireless channel, in addition to developing error-resilient source coding, error control is essential for robust video transmission. Due to the complexity of the source and channel coding algorithms, the processing power can become a significant component of the battery drain.

This paper aims to take the power constraint into account while designing an efficient error control scheme for video streaming. Note that the power consumption discussed in this work is for streaming case and mainly consists of source decoding and channel decoding in the mobile host.

JSCC (Joint Source Channel Coding) is efficient for video delivery over wireless channel. UEP (Unequal Error Protection) scheme improves video quality by partitioning the bit stream into different classes of priority [1]. However, fixed-error-correction rate for different priority class adopted in this scheme is inefficient since the characteristics of time-varying channel has not been considered. Hybrid ARQ scheme [2] achieves both delay bound and rate effectiveness by limiting retransmission number. However, it assumes that the maximum number of retransmission is fixed and known a priori. This may not reflect the time-varying nature of delay. Cheung et al. [3] proposed an optimal bit allocation scheme for joint source channel coding of scalable video. RCPC (Rate Compatible Punctured Convolutional) and UEP are combined in the scheme, but no ARQ is addressed. In our previous work, we proposed an end-to-end architecture to deliver the MEPG-4 scalable video over wireless with channel-adaptive delay-constrained hybrid ARQ and FEC, where minimal distortion is achieved by optimal bit allocation [4]. However, in all the above schemes no power constraint had been considered.

There are several existing work related to the power consumption in literature. Havinga et al. [5] proposed an energy-efficient error control scheme while source side had not been considered. Lan and Tewfik [6] considered the problem of minimizing the total consumed energy of a wireless system subject to a quality-of-service constraint. Appadwedula et al. [7] proposed an efficient wireless image transmission scheme under a total power constraint. However in [6, 7], just image is considered on the source side. No specific error control scheme had been discussed in [7]. Meanwhile, time-varying channel conditions had not been taken into account in [6, 7].

In this paper, we propose a general approach to design an efficient JSCC scheme for video streaming with the optimal power consumption. Varying channel conditions, video characteristics, and specific error control scheme are jointly considered to achieve the minimal power consumption in the mobile host.

II. GENERAL APPROACH FOR POWER-OPTIMIZED JSCC

The design of low-power JSCC for the mobile host requires exploiting the tradeoff between data, redundancy, processing power based on the characteristic of time-varying channel. In this section we describe the general approach for JSCC, which can achieve the minimal power consumption.

Note that each media has its own source rate and distortion relation given by $D_s = f(R_s)$ (see curve ① in Figure 1). The optimization problem highly depends on this R-D function. According to the rate-distortion theory, the lower the source coding-rate, the larger the distortion. While media transmitting over channel, channel distortion would result from channel transmission error, which can be presented as $D_c = f'(f_b(i), P_b(i), R_c, R_s)$, where $f_b(i)$ is the distortion due to the error of block i, $P_b(i)$ is the probability that the i^{th} block is corrupted, R_c is the channel coding rate, and R_s is the source coding rate. Since $P_b(i)$ is related to the channel bit error rate (BER) and bursty length, different channel conditions may have different distortion impacts. More specifically, as shown in Figure 1, $R_2 \neq R_1 \Rightarrow \Delta D_2 \neq \Delta D_1$.

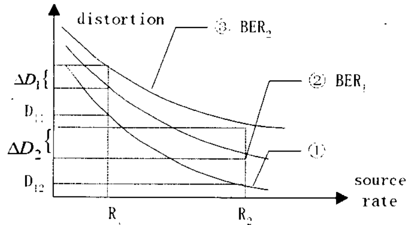


Figure 1. Rate-distortion relation with different channel conditions.

In wireless link, the end-to-end distortion, D_T, is composed of the source distortion and the channel distortion. The source distortion is caused by media rate control D_s; while the channel

---

[*] This work is performed while he is a visiting student at Microsoft Research China. He is from Department of E.E., Tsinghua University, Beijing, P. R. China.

0-7803-6685-9/01/$10.00©2001 IEEE

distortion results from channel random transmission error and burst fading error D_c . Then, the end-to-end distortion is represented as $D_T = D_s + D_c$.

Based on the above rate-distortion relation, it is essential to adopt some error protection schemes such that reducing the distortion caused by channel transmission. Different error control schemes may have different impacts on the rate-distortion relation. In general, a JSCC scheme may affect the rate-distortion relation as illustrated in Figure 2. In a given total available rate, with increase of source rate, the end-to-end distortion is decreased. However, after an optimal source rate, with increase of source rate, the end-to-end distortion is increased accordingly.

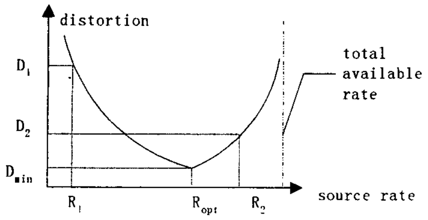


Figure 2. Rate-distortion relation with JSCC scheme.

Up to now, we have discussed the rate-distortion relation. However, another important factor, power consumption, has not been addressed yet. In general, Figure 3 depicts the relation between distortion and complexity as well as the relation between complexity and power consumption. As shown in Figure 3, with higher complexity, smaller distortion can be achieved; while with higher complexity, more power would be consumed. That is, $C_2 > C_1 \Rightarrow D_2 < D_1$ and $C_2 > C_1 \Rightarrow P_2 > P_1$.

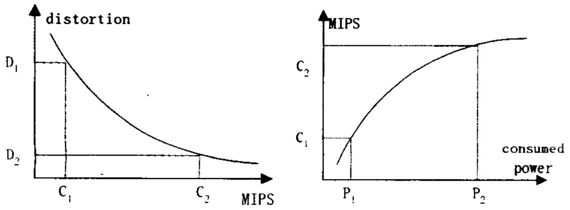


Figure 3. Distortion-complexity and complexity-power relations.

Similar to the rate part, the processing power consumed in the source decoding and channel decoding side is different. The processing power consumed in the source decoding can be described as

$$P_s = h(R_s), \tag{1}$$

where R_s is the source coding rate; while the processing power consumed in the channel decoding can be represented as

$$P_c = h'(R_s, R_c), \tag{2}$$

where R_c is the channel coding rate.

Taking the power consumption into consideration, a desirable JSCC scheme is to keep the quality of service to remain fixed while the total power consumed is minimized to save battery life. This optimization problem is formulated as

$$\min_{R_s, R_c} (P_s + P_c) \text{ subject to } D_T \le D_0 \text{ and } (R_s + R_c) \le R_0. \tag{3}$$

III. POWER-OPTIMIZED HYBRID UEP AND DELAY-CONSTRAINED ARQ

In this section, we use PFGS (Progressive Fine Granularity Scalable) as the source codec, channel-adaptive hybrid UEP and delay-constrained ARQ as the error control scheme to derive our proposed power-optimized JSCC scheme.

3.1 An end-to-end architecture

Figure 4 depicts the block diagram of our powered-optimized hybrid UEP and delay-constrained ARQ for PFGS over wireless. PFGS is a layered scalable video codec. It encodes raw video into two layers: one is the base layer (BL) and the other is the enhance layer (EL). With the video characteristic in mind, these layers are packetized and protected against channel errors according to their importance and channel conditions with different error control schemes. Specifically, for the BL stream, the sender will decide the retransmission time and the protection degree of each re-transmission according to the delay constraint and channel condition. While for the EL stream, only FEC is added according to their importance and channel condition.

The channel decoder reconstructs packets through channel decoding process. For EL, the output of channel decoder is directed for source decoding; while for BL, if residual error still exists, receiver sends a request to the sender for retransmission if the delay constraint is met.

In this end-to-end architecture, the power-optimized bit allocation is performed based on the dynamically estimated channel condition. In this paper we aim at modeling wireless channel characterized by slow- and highly-correlated fading. In this type of channel, the input signal is first multiplied by a Rayleigh fading process and then added by a AWGN process. As in [4], we use the two-state Markov model to approximately represent the fading channel.

3.2 Rate-distortion and rate-power relations for PFGS

As mentioned above, PFGS encodes raw video into BL and EL. While BL carries the most important information and EL carries less important information. The EL bit stream can be truncated anywhere. The higher layer information relies on the corresponding one in the lower layers. Therefore, if any residual error occurs in the lower layers, the corresponding information bits will be discarded no matter if they are correct or not.

In this paper, we use the CPU computational time to indicate the consumed processing power. The rate-power relation for source decoding is illustrated in Figure 5 (a).

3.3 The impact of hybrid UEP and delay-constrained ARQ on the rate-distortion and power-rate relations for PFGS

Here we use the hybrid UEP and delay-constrained ARQ we proposed in [4]. Let $R(t)$ represent the channel rate available for transmission at time t. Let $R_S(t)$, $R_{ARQ}(t)$, and $R_{FEC}(t)$ denote the PFGS source rate, ARQ rate for BL, and FEC rate for EL at time t, respectively. Then those rates should satisfy the following constraint: $R_S(t) + R_{ARQ}(t) + R_{FEC}(t) \le R(t)$.

We adopt hybrid ARQ and FEC for BL to reduce the residual error. The rate and distortion expressions for the BL are calculated as follows:

$$R_{ARQ} = R_{s\_base} + R(1, t_1) + \sum_{l=2}^{N_{\max}} \Big(\prod_{j=1}^{l-1} P_{block}(j)\Big) R(l, t_l), \tag{4}$$

$$D(R_{ARQ}) = D_s(R_{s\_base}) + \Big(\prod_{l=1}^{N_{\max}-1} P_{block}(l)\Big) D_{error}(N_{\max}), \tag{5}$$

where $N_{\max}$ is the number of maximal transmission times of the current packet, R_{s_base} represents the rate of BL, $R(l, t_l)$ represents the rate spent in the l^{th} transmission and t_l is the error protection capability of the l^{th} transmission, $P_{block}(j)$ represents the failure probability of the l^{th} delivery, $D_s(R_{s-base})$ denotes source perceptual distortion-rate function, and $D_{error}(N_{\max})$ represents the distortion resulting from channel error in the ${N_{\max}}^{th}$ transmission.

After bit resource needed in BL is determined, the rest of bits are allocated between the EL source and EL channel protection. We only adopt channel-adaptive UEP for the EL for the sake of simplicity. The rate and distortion expression for the EL is calculated as follows:

$$D(R_{FEC}) = D_s(R_{FEC}) + \sum_{m=1}^{M}(w_m \times \sum_{l=1}^{L_m}(D_{error}(l,m) \times P(l,m))), \quad (6)$$

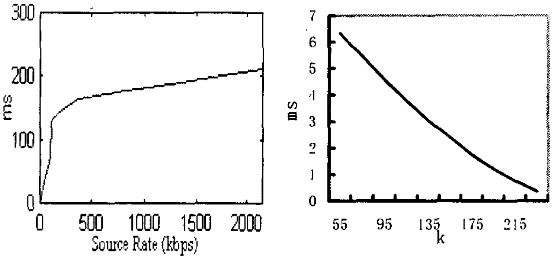
where $D_s(R_{FEC})$ represents source perceptual distortion-rate function, $D_{error}(l,m)$ represents the channel distortion that the l^{th} symbol in the m^{th} layer is corrupted, M denotes the number of layers, w_m is the weight of the m^{th} layer, and L_m denotes the number of symbols in the m^{th} layer. To avoid overlapped channel distortion, $P(l,m)$ is the probability that the l^{th} symbol of the m^{th} layer is corrupted while the corresponding symbols in the previous layers are correct.

We used Reed-Solomon (RS) codes for error control. An RS code is represented as *RS (n, k)*, where k is the length of source symbols and n-k is the length of protection symbols. It is known that an RS code usually can correct up to $t = \left\lfloor \frac{n-k}{2} \right\rfloor$ symbol errors. The packet failure probability of an RS code is defined as $P_{block} = 1 - \sum_{j=0}^{t} P(n,j)$, where $P(n,j)$ represents the probability of the j^{th} errors in the n^{th} transmission.

The majority of the power consumption of the RS codec is due to the RS decoder. It was found that the energy consumption of the RS decoder per codeword is

$$\varepsilon_{dec/codeword} = (4tn + 10t^2)\varepsilon_{mult} + (4tn + 6t^2)\varepsilon_{add} + 3\varepsilon_{inv}, \quad (7)$$

where $\varepsilon_{mult}, \varepsilon_{add}, \varepsilon_{inv}$ represent the energy consumed in the $m \times m-bit$ multiplier, *m-bit* addition, and *m-bit* inversion, respectively [7].



(a) (b)

Figure 5. The rate-power relation.
(a) Source decoding. (b) Channel decoding.

Similar to the source side, we use the computation time to indicate the consumed processing power. The rate-power relation for RS decoding is illustrated in Figure 5 (b).

3.4 Power-optimized source/channel bit allocation for PFGS

Having the rate-distortion and rate-power rations for source and RS decoding, now we will discuss, for a given total bit budget and the desired distortion range, how to add adaptive error control to each layer so as to minimize the consumed power while performing rate adaptation to channel condition. To achieve the minimal power consumption, it is necessary to optimally allocate bits between source and channel protection.

The processing power consists of the power consumed in the BL (P_{BL}) and EL (P_{EL}). That is

$$\begin{aligned} P_{BL} &= h(R_{s-base}) + h'(R_{s-base}, t_1) \\ &+ \sum_{l=2}^{N_{\max}}((1 - \prod_{j=1}^{l-1} P_{block}(j)) \times h'(R_{s-base}, t_l)), \end{aligned} \quad (8)$$

where $h(x)$ denotes the power consumed for the source decoding with the source rate of x, and $h'(x,y)$ denotes the power consumed for the channel decoding with the source rate of x and the error-protection capability of y. Mathematically,

$$P_{EL} = \sum_{m=1}^{M}(h(R_{s-m}) + h'(R_{s-m}, t_m)), \quad (9)$$

where R_{s-m} denotes the source rate in the m^{th} layer, and t_m denotes the error-protection capability of the m^{th} layer.

Now the objective is to minimize the end-to-end consumed power as follows:

$$\begin{aligned} &\min_{R_{s-base}, R_{s-m}, t_m} (P_{BL} + P_{EL}) \\ &\quad \text{subject to } D(R_{ARQ}) + D(R_{FEC}) \le D(t), \\ &\quad\quad \text{and } R_S(t) + R_{ARQ}(t) + R_{FEC}(t) \le R(t). \end{aligned} \quad (10)$$

Optimization methods such as Lagrange multiplier and penalty function methods can be used to solve the constrained non-linear optimization problem.

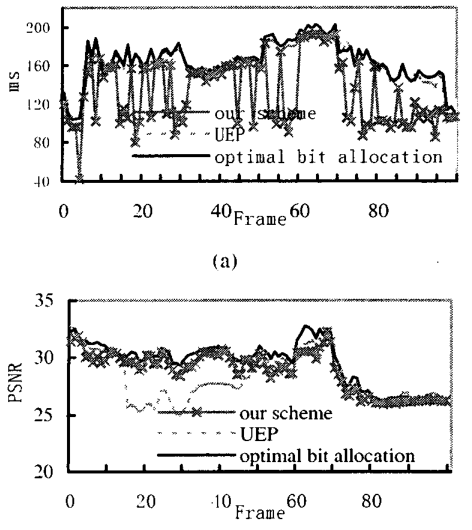
IV. SIMULATION RESULTS

This simulation is to demonstrate effectiveness of our proposed power-optimized JSCC scheme for PFGS. In this simulation we tested: (1) our channel-adaptive power-optimized bit allocation scheme for hybrid UEP and ARQ; (2) channel-adaptive bit allocation scheme for hybrid UEP and ARQ; (3) PFGS with UEP, which used fixed channel protection for each priority (25% protection for base layer, 10% protection for enhance layer). The testing video sequence *Foreman* is coded in CIF at a temporal resolution of 10 fps. The first frame was intra-coded and the remaining frames were inter-coded.

We conducted simulations under the channel bandwidth varying from 200kbps to 1Mbps. The bit error rate of tested channel varies from 1% to 4%.

Figure 3 shows comparison results for the test sequence *Foreman* at 256kbps available bandwidth. It can be seen that our proposed scheme requires less computational time than the other schemes almost in every frame. Meanwhile, the PSNR obtained in our scheme is little less than the one in optimal bit allocation scheme while higher than the one in fixed UEP scheme. Note that those performances vary with the different desired distortion tolerance range. In Figure 3, the desired distortion tolerance range is 60%.

Table 1 depicts comparison results of average computational time and PSNR for the whole sequence in these three schemes. Scheme 1 uses the optimal bit allocation scheme without considering power consumption. It needs the longest computational time while gains the highest PSNR. We use it as the comparison criteria. Scheme 2 uses the fixed UEP. As mentioned above, different desired distortion tolerance ranges may have different impacts on video quality and power consumption. Scheme 3, 4, 5, 6 used desired distortion range with 10%, 20%, 40%, 60%, respectively.



(a)

(b)

Figure 2. Comparison results for *Foreman* using three schemes at 256kbps.
(a) Computational time. (b) PSNR.

Table 1. Comparison results for sequence *Foreman*. QRR is the Quality Reduction Ratio and PSR is the Power Saving Ratio.

Schemes	256kbps				320kbps			
	PSNR (dB)	QRR (%)	Time (ms)	PSR (%)	PSNR (dB)	QRR (%)	Time (ms)	PSR (%)
1	29.59	0	164.11	0	30.26	0	172.88	0
2	28.13	4.92	160.07	2.47	28.44	6	165.90	4.04
3	29.47	0.43	157.41	4.08	30.06	0.67	167.66	3.02
4	29.33	0.91	155.35	5.33	29.91	1.17	165.41	4.32
5	29.09	1.72	145.51	11.33	29.56	2.33	161.49	6.59
6	28.88	2.42	132.32	19.37	29.32	3.12	152.28	11.91

It can be seen that we can achieve significant power saving ratio within a tolerable distortion range.

V. CONCLUSIONS

In this paper, we proposed a general approach for joint source channel coding (JSCC) with optimal power consumption. An end-to-end power-optimized architecture for video streaming is introduced in this paper. The channel-adaptive hybrid UEP and delay constrained ARQ scheme is described to achieve the minimal power consumption. Simulation results demonstrate that our proposed scheme can achieve significant power saving ratio within a tolerable distortion range.

Video streaming scenario is discussed in this paper where only source and channel decoding processing power is considered. While for video communication case, transmission power becomes a major issue. Power optimization for video communication is our further research.

ACKNOWLEDGEMENT

We thank Prof. Jianhua Lu from Tsinghua University (Beijing) for the support of this work. We also thank Prof. R. L. Legendijk from Delft Technology Univeristy for the initial discussion.

REFERENCES

[1] Uwe Horn, Bernd Girod, and Ben Belzer, "Scalable video coding for multimedia applications and robust transmission over wireless channels," in 7th *Int. Workshop on Packet Video*, 1996.

[2] Q. Zhang and S. A. Kassam, "Hybrid ARQ with selective combining for fading channels," *IEEE J. on Selected Areas in Comm.*, vol. 17, no. 5, pp. 867-880, 1999.

[3] G. Cheung and A. Zakhor, "Bit allocation for joint source/channel coding of scalable video," *IEEE Transactions on Image Processing*, March, 2000.

[4] G. J. Wang, Q. Zhang, W.W. Zhu, and Y.-Q. Zhang, "Channel-Adaptive Error Control for Scalable Video over Wireless Channel", *IEEE MoMuc 2000*, Oct. 2000.

[5] P. J. M. Havinga, "Energy efficiency of error correction on wireless systems", *IEEE Wireless Communications and Networking Conference (WCNC)*, Sept. 1999.

[6] T. H. Lan, A. H. Tewfik, "Adaptive low power multimedia wireless communications", *IEEE workshop on MMSP*, 1997.

[7] S. Appadwedula, M. Goel, D.L. Jones, K. Ramchandran, N.R. Shanbhag, "Efficient wireless image transmission under a total power constraint", *IEEE workshop on MMSP*, 1998.

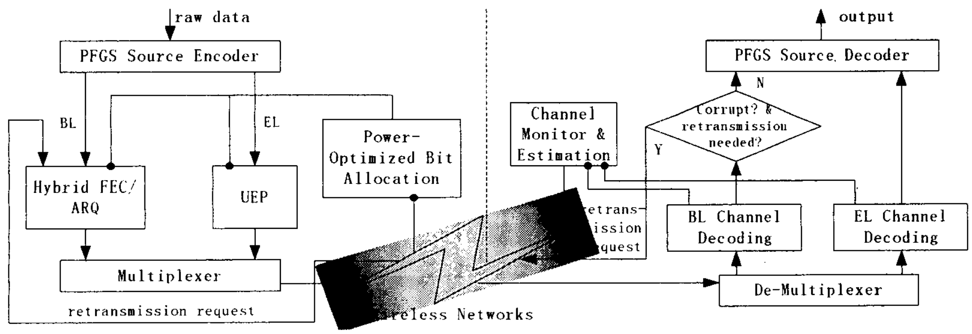


Figure 4. Architecture of power-optimized hybrid FEC/ARQ scheme for scalable video over wireless.