JOINT SOURCE-CHANNEL RATE CONTROL FOR MULTIVIEW VIDEO STREAMING OVER ERASURE CHANNELS

Weiliang Xu, Junni Zou[*]

School of Communication Engineering
Shanghai University, China

Hongkai Xiong

Department of Electronic Engineering
Shanghai Jiao Tong University, China

ABSTRACT

To enable robust multiview video streaming over current wireless broadcast infrastructures, this paper proposes a joint source and channel coding optimization framework for video streaming based upon multiview video coding technique. It utilizes random linear coding (RLC) with ladder-shaped global coding coefficient matrix (GCM) for unequal erasure protection. An accurate expression of the decoding probability for RLC with ladder-shaped GCM is derived. A fast two-layer local search algorithm is developed to solve the source and channel coding rate allocation problem, where the source and channel rate of each layer can be adjusted independently for the overall video quality. Simulation results demonstrate significant video quality improvement to existing rate allocation schemes.

Index Terms— Interactive multiview video streaming, scheduling, fairness, asymmetric bargaining game

1. INTRODUCTION

Multiview video (MVV) consists of multiple video sequences that are captured simultaneously by multiple cameras from different viewpoints. The state of the art multiview video coding (MVC) standard [1] compresses MVV efficiently by exploiting both temporal and interview redundancies. It requires to transmit the entire set of multiview sequences to the receiver thus is very suitable for video broadcast applications such as cable TV broadcast. In this paper, we consider MVC content distribution in current wireless broadcast system, where MVC stream is broadcast from the server to multiple receivers with diverse channel conditions.

Wireless communication is prone to transmission errors due to attenuation, fading and interference [2]. A class of rateless codes, such as random linear coding (RLC) [3], have been proved to have the ability to provide erasure protection with simple implementation. One challenge for RLC is that a full-rank global coding coefficient matrix (GCM) is required to successfully decode all source packets. As RLC is operated at

intermediate nodes, it is very difficult to guarantee a full-rank GCM receivers. To reduce the impact of the GCM's rank deficiency on video transmission, RLC with ladder-shaped GCM [4] [5] is proposed. It improves the decodability of video data at receivers by enabling partial decoding of a block.

Since different packets may contribute differently to the decrease of video distortion, RLC with unequal error protection (UEP) schemes [5]-[8] have been proposed recently. UEP RLC provides better protection for most important packets, and thus guarantee their higher decoding probabilities. Joint source-channel coding (JSCC) is an effective method for robust video transmission over wireless lossy channels. The channel coding in the existing JSCC works on MVV streaming is designed on the basis of either forward error correction techniques [9] [10], or RLC with generalized GCM [8] [11]. In this paper, we consider a joint source-channel coding problem for MVV broadcast, in which views are classified into different layers in terms of MVC prediction structure, and RLC with ladder-shaped GCM is employed to provide unequal erasure protection. Our main contributions are summarized as follows:

- We propose a joint source and channel coding rate optimization framework that utilizes RLC with ladder-shaped GCM for unequal erasure protection. Constrained by the channel capacity, the source and channel rate allocation problem is modeled as an aggregate video quality maximization problem.

- We derive an accurate expression of the decoding probability for RLC with ladder-shaped GCM, in which the linear dependency of the encoding vectors for the redundant packets, ignored in [4], is completely considered.

- We develop a fast two-layer local search algorithm to solve the proposed JSCC rate allocation problem. Unlike the existing local search algorithms, it can adjust the source rate and channel rate of each layer independently, so as to be more close to the global optimum.

The remainder of this paper is organized as follows: Sec. II provides an accurate expression of the decoding probability for RLC with ladder-shaped GCM. Sec. III formulates

the JSCC optimization problem, and then presents a fast two-layer local search algorithm. Simulation results are discussed in Sec. IV. Finally, Sec. V concludes the paper.

2. DECODING PROBABILITY FORMULATION

We adopt ladder-shaped GCM as our channel coding scheme to cope with the rank deficiency problem in RLC with generalized GCM. Assume that the source message is classified into L layers according to their contributions to video quality. Let $\mathbf{S} = (s_1, \ldots s_L)$ represents the number of source packets in each layer. The structure of the ladder-shaped GCM is shown in Fig. 1, where each submatrix M_i contains global encoding vectors corresponding to the i-th layer, the white area is filled by zeros. The first part of M_i have s_i rows are all linearly independent, and the second part have k_i rows for generating the redundant packets. Let $\mathbf{K} = (k_1, \ldots k_L)$ denote the number of redundant packets, then the total number of the encoded packets for all the layers can be represented as $\mathbf{N} = (n_1, \ldots n_L)$, where $n_i = s_i + k_i$. The UEP coding strategy includes the first l layers' source packets into the encoded packets of all subsequent layers. The decoding probability of the l-th layer is non-zero only if: (i) for the first l layers, the number of the received encoded packets are at least equal to the number of the source packets, or (ii) any higher layers can be decoded with a probability greater than zero.
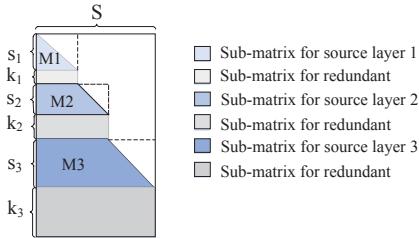

Fig. 1. The structure of the ladder-shaped GCM

Further, let $\mathbf{A} = (\alpha_1, \ldots \alpha_L)$ represent the number of successfully received encoded packets associated with each layer. The source message is transmitted to the receiver at a packet erasure rate p. Let $P_{1:l}$ be the decoding probability of the first l layers regardless of the impact of their higher layers. For $L = 1$ and $l = 1$, the least number of received packets is s_1, then $P_{1:1}$ is given by

$$P_{1:1}(n_1, s_1) = \sum_{\alpha_1=s_1}^{n_1} P_R(n_1, \alpha_1) \cdot P_F(s_1, \alpha_1) \quad (1)$$

where

$$P_R(n_1, \alpha_1) = \binom{n_1}{\alpha_1} (1-p)^{\alpha_1} p^{n_1 - \alpha_1} \quad (2)$$

represents the probability of receiving α_1 encoded packets while n_1 encoded packets are generated at the source node.

Here, $P_F(s_1, \alpha_1)$ is the probability that the GCM can reach the rank of s_1 when receiving α_1 encoded packets.

In [4], $P_F(s_1, \alpha_1)$ is regarded as 1 if $\alpha_1 \geq s_1$. However, even if the first s_i encoding vectors of layer i are linearly independent, the encoding vectors of the redundant packets may be linearly dependent. It means that the value of $P_F(s_1, \alpha_1)$ may be less than 1 even if we have $\alpha_1 \geq s_1$.

Let $P_D(d, k)$ represent the probability that d out of k redundant packets are linearly independent for Galois field $GF(2^q)$ [12], then we have

$$P_D(d, k) = \prod_{i=0}^{d-1} \left[1 - \frac{1}{q^{k-i}} \right] \quad (3)$$

For the worst case, if $\alpha_1 \leq k_1$ all the received packets are redundant packets, otherwise k_1 redundant packets are included in the received packets. The minimum probability of $P_F(s_1, \alpha_1)$ is

$$P_F^{\min}(s_1, \alpha_1) = \begin{cases} P_D(s_1, \alpha_1), & \alpha_1 \leq k_1 \\ P_D(n_1 - \alpha_1, k_1), & k_1 < \alpha_1 < n_1 \\ 1, & \alpha_1 = n_1 \end{cases} \quad (4)$$

We now investigate the case that $L > 1$ and $l \leq L$. Let $\alpha_l^{\min}$ denote the least number of received encoded packets for $P_{1:l}$ to be a non-zero value, $S_l = \sum_{i=1}^{l} s_i$ and $A_l = \sum_{i=1}^{l} \alpha_i$ are the number of source packets and received encoded packets comprising the first l layers. Clearly, for $l = 1$, we have $\alpha_1^{\min} = S_1$. For the case with $2 \leq l \leq L$, $\alpha_l^{\min}$ can be obtained using the the following recursion:

$$\alpha_l^{\min} = s_l + \max(\alpha_{l-1}^{\min} - \alpha_{l-1}, 0) \quad (5)$$

Thus, the minimum decoding probability of receiving the first l layers but not higher can be written as

$$P_{1,l}^{min}(\mathbf{N}, \mathbf{S}) \simeq \sum_{\alpha_1=0}^{n_1} \cdots \sum_{\alpha_{l-1}=0}^{n_{l-1}} \sum_{\alpha_l=\alpha_l^{\min}}^{n_l}$$
$$\sum_{\alpha_{l+1}=0}^{\alpha_{l+1}^{\min}-1} \cdots \sum_{\alpha_L=0}^{\alpha_L^{\min}-1}$$
$$\left[\prod_{i=1}^{L} \binom{n_i}{\alpha_i} (1-p)^{\alpha_i} p^{n_i - \alpha_i} \right] \cdot P_F^{min}(\mathbf{S}, \mathbf{A}) \quad (6)$$

where the first line counts the events of receiving the possible number of encoded packets to decode the first l layers, while the second line lists all the possible cases that the subsequent layers $l + 1, \ldots, L$ can not be decoded. We seek $P_F^{\min}(\mathbf{S}, \mathbf{A})$ an extension of $P_F^{\min}(s_1, \alpha_1)$ and have

$$P_F^{\min}(\mathbf{S}, \mathbf{A}) \simeq P_F^{\min}(S_l, A_l)$$
$$= \begin{cases} P_D(S_l, A_l), & A_l \leq K_l \\ P_D(N_l - A_l, K_l), & K_l < A_l < N_l \\ 1, & A_l = N_l \end{cases} \quad (7)$$

Therefore, the average decoding probability can be calculated by

$$\bar{P}_{1,l}(\mathbf{N}, \mathbf{S}) = \frac{P_{1,l}^{\min}(\mathbf{N}, \mathbf{S}) + P_{1,l}^{\max}(\mathbf{N}, \mathbf{S})}{2} \quad (8)$$

where $P_{1,l}^{\max}(\mathbf{N}, \mathbf{S})$ is the corresponding maximum value that can be obtained by replacing $P_F^{min}(S_l, A_l)$ with 1 in (6).

3. JOINT SOURCE AND CHANNEL CODING

At the server side, the input MVV is encoded using MVC. The resulting bit stream is packetized in equal length and classified into multiple layers of different importance according to the prediction structure. RLC with ladder-shaped GCM is employed to generate redundant packets. Then, the encoded packets are broadcast to the receivers. At the receiver side, Gaussian elimination is used for RLC decoding, followed by a MVC decoder. In this work, we aim at the overall video quality by optimal source and channel rate allocation strategy.

3.1. Optimal Source and Channel Rate Allocation

We utilize RLC to encode the MVV at group of pictures (GOP) level. Denote $\mathbf{S} = \{s_1, \ldots, s_L\}$ the number of equal length source packets included in one GOP for each view, and $\mathbf{N} = \{n_1, \ldots, n_L\}$ the corresponding number of encoded packets generated at the server including $\mathbf{K} = \{k_1, \ldots, k_L\}$ redundant packets. Assume the size of a source packet is b bits, and a GOP is transmitted in t seconds, then the source rate of view i can be calculated as

$$r_i = \frac{b \cdot s_i}{t}, \quad i = 1, \ldots, L. \quad (9)$$

The stream of encoded packets then is broadcast to M receiver classes characterized with their packet erasure rates $\mathbf{p} = (p_1, \ldots p_M)$. Denote ξ_i the fraction of receiver population associated with class i and $\sum_{i=1}^{M} \xi_i = 1$. We can formulate the source and channel rate allocation as a constrained optimization problem:

$$\max_{\mathbf{N}, \mathbf{S}} \frac{1}{L} \cdot \sum_{i=1}^{M} \xi_i Q^i(\mathbf{N}, \mathbf{S}, p_i)$$
$$s.t. \quad \sum_{i=1}^{L} n_i \leq N_c \quad (10)$$

where the source and channel rate are jointly optimized for maximizing the average video quality with respect to different receiver classes. Q^i is the video quality of the i-th receiver class that can be computed by

$$Q^i(\mathbf{N}, \mathbf{S}, p_i) = \sum_{l=1}^{L} \bar{P}_{1,l}(\mathbf{N}, \mathbf{S}, p_i) \cdot Q_{1:l}(\mathbf{S}) \quad (11)$$

where $\bar{P}_{1,l}(\mathbf{N}, \mathbf{S}, p_i)$ represents the average decoding probability obtained by (8) for the packet erasure rate equal to p_i, and $Q_{1:l}(\mathbf{S})$ denotes the sum video quality of the first l views which can be computed via expressions derived in [13]. The constraint condition means that the overall data rate should not exceed a given available bandwidth, and N_c is the maximum number of packets can be transmitted on the channel.

3.2. Two-layer Local Search Algorithm

Due to high computational complexity of full search, local search becomes the alternative for the problem in (10). The existing local search algorithms often transform the joint source and channel rate allocation to a single variable optimization problem with several coding rate combinations [15] or keep the maximum rate for each layer fixed [16], resulting in the missing of too many possible combinations of source and channel rate. In this paper, we propose a two-layer local search algorithm, in which the source rate and channel rate can be adjusted independently.

The two-layer local search algorithm starts from an initial source distribution and S_L is equal to the maximum value N_c. For the inner-layer search, the optimization is essentially the same as channel coding with a given vector of source packets. Set an initial distribution of the encoded packets in terms of the fraction of source packets for each layer and the maximal number of packets to be transmitted. Every step the neighbors of multiple feasible distribution vectors are examined by (10). If any neighbor has a higher average video quality, it is included in the set of feasible distribution vectors. This procedure is repeated until the set of feasible distribution vectors is $\emptyset$, then we get a possible optimal pair of source and channel rate allocation. In the outer layer, the same process is conducted to search the neighbors of the source distribution.

After these two stages, we achieve a solution for the case that S_L is equal to N_c. Since lower source rate can increase the decoding probability while lead to a worse video quality, there exits a tradeoff between these two metrics. We next attempt to reduce the overall source rate and find the global optimal solution.

4. SIMULATION RESULTS

We encode the "Ballroom" sequence with 640x480 resolution by using H.264/AVC reference software. The sequence consists of three views. The GOP size is 8 and the frame rate is 30 fps. The source packets are classified into three layers organized along the decrease of their importance. The view encoding order is 0-2-1, where view 0 has the highest importance for it is intra-coded. The packet size including the network coding header is set to 1500 bytes and the field size is $GF(2^8)$ [17].

Fig. 2 shows the comparison of our RLC with ladder-shaped GCM and expanding-window RLC presented in [8].

It is observed that RLC with ladder-shaped GCM has a good robustness against packet loss. When the packet erasure rate increases from 0.02 to 0.14, the average video quality only experiences 1.41 dB performance degradation and still reaches 33.12 dB. A better performance is always achieved compared with expanding-window RLC due to a higher decoding probability of RLC with ladder-shaped GCM.
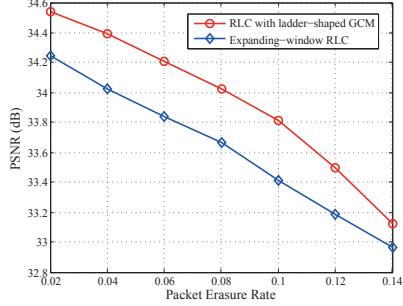

Fig. 2. Average video quality vs. packet erasure rate

Suppose there are two receiver classes with the packet erasure rate of 0.02 and 0.14. Consider two cases that the fraction of receiver population associated with class 1 is $\xi_1 = 0.1$ and $\xi_1 = 0.9$. Fig. 3 compares the source rates of these two cases with the available bandwidth. We can find that, given the same bandwidth, higher source rate is achieved when $\xi_1 = 0.9$ for the reason that the receivers with a better channel conditions, they can achieve a high decoding probability with less redundant packets.
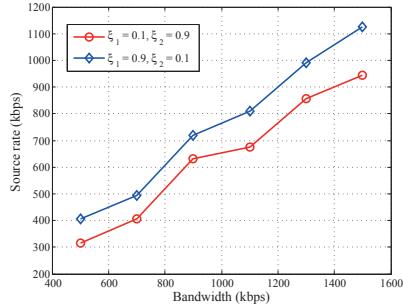

Fig. 3. Source rate vs. bandwidth for $\xi_1 = 0.1$ and $\xi_1 = 0.9$

Fig. 4 shows the average quality of each view, in which the proposed JSCC scheme is employed or just channel coding (CC) is adopted. The video quality achieved by decoding of view 0, 2 and 1 is 32.74 dB, 32.96 dB and 32.72 dB, respectively. The packet erasure rates is the same and $\xi_1 = 0.5$. We can see that JSCC significantly outperforms CC. When the bandwidth is 500 kbps, CC achieves a higher quality of view 0 than JSCC, but it fails to deliver the other two views as it allocates the whole rate to layer 1 for the purpose of decoding view 0 with a high probability. Similar result can also be found when the bandwidth is 700 kps.
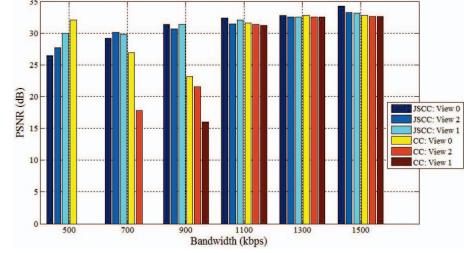

Fig. 4. Average video quality of each view vs. bandwidth for JSCC and CC

Finally, we compare the proposed two-layer search algorithm with the full search and the method that allocates an equal code rate for all the layers. The reference algorithm employs local search only once to search different distributions of source packets, while the number of the encoded packets is determined by the code rate that is iterated by a small step size. The packet erasure rates and the distribution of receiver population remain unchanged. It is observed in Fig. 5 that all these three algorithms achieve higher average quality along with the increasing bandwidth. Our two-layer local search algorithm gives a better performance than the scheme that allocates equal code rate to each layer and the performance degradation is negligible compared with full search.
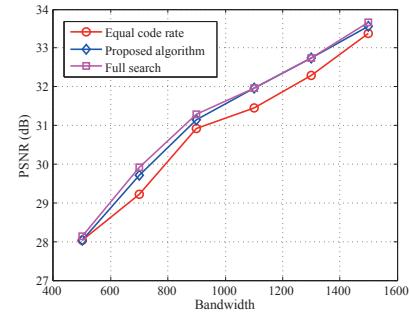

Fig. 5. PSNR comparison of three different algorithms

5. CONCLUSION

This paper proposed a framework of broadcasting robust MVV over wireless erasure channels. Our objective is to maximized the aggregate video quality of all the receivers by a joint source and channel coding rate optimization. We adopted RLC with ladder-shaped GCM for unequal erasure protection and presented an accurate expression of its decoding probability. Also, a two-layer local search algorithm was developed to find a solution in polynomial time. Simulation results showed that the proposed framework can maximize the overall video quality. Future work would be expanded to more complicated wireless multicast system and take the view popularity among the receivers into consideration.

6. REFERENCES

[1] A. Vetro, T. Wiegand, and G. J. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard," *Proc. IEEE,* vol. 99, no. 4, pp. 626-642, Apr. 2011.

[2] H. Zhihai and X. Hongkai, "Transmission distortion analysis for real-time video encoding and streaming over wireless networks," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 16, pp. 1051-1062, 2006.

[3] D. Lun, M. Medard, R. Koetter, and M. Effros, "On coding for reliable communication over packet networks," *Physical Commun.*, vol. 1, no. 1, pp. 22-30, 2008.

[4] H. Wang, S. Xiao, C. Kuo, "Random linear network coding with ladder-shaped global coding matrix for robust video transmission," *Journal of Visual Communication and Image Representation*, vol. 22, no. 3, pp. 203-212, 2011.

[5] X. Zhang, C. Jing, F. Tang, et al., "Joint redundant and random network coding for robust video transmission over lossy networks," *Mobile Information Systems*, no. 8, pp. 213-230, 2012.

[6] D. Vukobratović and V. Stanković, "Unequal error protection random linear coding for erasure channels," in *IEEE Trans. Communications*, vol. 60, pp. 1243-1252, May 2012.

[7] N. Thomos, J. Chakareski, and P. Frossard, "Prioritized distributed video delivery with randomized network coding," in *IEEE Trans. Multimedia*, vol. 13, pp. 776-787, Aug. 2011.

[8] J. Chakareski, V. Velisavljevic, and V. Stankovic, "View-Popularity-Driven Joint Source and Channel Coding of View and Rate Scalable Multi-View Video," in *IEEE Journal of Selected Topics in Signal Processing*, 2015.

[9] J. Liu, Y. Liu, S. Ci, and R. Yao, "3d visual experience oriented crosslayer optimized scalable texture plus depth based 3d video streaming over wireless networks," in *Journal of Visual Communication and Image Representation*, vol. 25, no.5, July 2014, pp. 1209-1221.

[10] A. Vosoughi, P. Cosman, and L. B. Milstein, "Joint source-channel coding and unequal error protection for video plus depth," in *IEEE Signal Processing Letters*, vol. 22, August 2014, pp. 31-34.

[11] L. Toni, N. Thomos, and P. Frossard, "Interactive free viewpoint video streaming using prioritized network coding," in *International Workshop on Multimedia Signal Processing*, Pula, Italy, Oct. 2013.

[12] O. Trullols-Cruces, J. Barcelo-Ordinas, and M. Fiore, "Exact Decoding Probability Under Random Linear Network Coding," *IEEE Commun. Lett.,* vol. 15, no. 1, pp. 67-69, Jan. 2011.

[13] A. Fiandrotti, J. Chakareski, and P. Frossard, "Popularity-based rate allocation in multiview-video," in *Proc. Conf. Vis. Commun. Image Process.,* pp. 1-7, Jul. 2010.

[14] A. Tassi, I. Chatzigeorgiou and D.Vukobratović, "Resource-Allocation Frameworks for Network-Coded Layered Multimedia Multicast Services," *IEEE J. Sel. Areas Commun.,* vol. 33, no. 2, pp. 141-155, Feb. 2015.

[15] N. Ramzan, A. Amira, and C. Grecos, "Efficient transmission of multiview video over unreliable channels," in *IEEE International Conference on Image Processing (ICIP),* pp. 1885-1889, Sep. 2013.

[16] Stankovic V, Hamzaoui R and Xiong Z, "Real-time error protection of embedded codes for packet erasure and fading channels," in *IEEE Trans. Circuits Syst. Video Tech.,* vol. 14, no. 8, pp. 1064-1072, 2004.

[17] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proc. 2003 Allerton Conf. Commun., Control, Comput*.