Optimal Adaptive Streaming of A Scalable Multi-view Video via Rate Splitting and SIC

Wuyang Jiang[1], **Chencheng Ye**[2], **Lingzhi Zhao**[2], **Ying Cui**[2a)], **and Zhi Liu**[3]
[1] *Shanghai University of Engineering Science*
333 Longteng Rd., Shanghai 201620, China
[2] *Department of Electronic Engineering, Shanghai Jiao Tong University*
800 Dongchuan Rd., Shanghai 200240, China
[3] *School of Informatics and Engineering, The University of Electro-Communications*
Tokyo 182-8585, Japan

a) *cuiying@sjtu.edu.cn*

Abstract: This letter considers optimal adaptive wireless streaming of a scalable coding-based multi-view video (MVV) to multiple users. To improve quality of service (QoS), we propose a transceiver design based on rate splitting and successive interference cancellation (SIC) and optimize the design parameters to maximize the sum encoding rate of the enhanced versions of the requested views. We obtain a KKT point of the non-convex problem using complementary geometric programming (CGP) and successive convex approximation (SCA).
Keywords: Multi-view video, rate splitting, complementary geometric programming, successive interference cancellation.
Classification: Multimedia systems for communications

References

[1] J. Wu, Q. Zhao, N. Yang and J. Duan, "Augmented reality multi-view video scheduling under vehicle-pedestrian situations," in Proc. of ICCVE, Shenzhen, China, pp. 163-168, Oct. 2015. DOI:10.1109/ICCVE.2015.75.
[2] Q. Zhao, Y. Mao, S. Leng and Y. Jiang, "QoS-aware energy-efficient multicast for multi-view video in indoor small cell networks," in Proc. of GLOBECOM, Austin, TX, USA, pp. 4478–4483, Dec. 2014. DOI: 10.1109/GLOCOM.2014.7037513.
[3] W. Xu, Y. Cui and Z. Liu, "Optimal multi-view video transmission in multiuser wireless networks by exploiting natural and view synthesisenabled multicast opportunities," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1494-1507, Mar. 2020. DOI:10.1109/TCOMM.2019.2954523.
[4] Z. Li, C. Ye, Y. Cui, S. Yang and S. Shamai, "Rate splitting for multiantenna downlink: Precoder design and practical implementation," *IEEE J. Select. Areas Commun.*, vol. 38, no. 8, pp. 1910–1924, Aug. 2020. DOI: 10.1109/JSAC.2020.3000824.
[5] M. Chiang, C. W. Tan, D. P. Palomar, D. O'neill and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640-2651, Jul. 2007. DOI:10.1109/TWC.2007.05960.

1 Introduction

Wireless streaming of an MVV to multiple users arises in several applications such as entertainment, education, military, etc. Encoding rate adaptation and transmission design are crucial for an MVV with a much larger size than a traditional video. However, the existing works [1, 2, 3] do not consider encoding rate adaptation. Besides, the transmission designs in [1, 2, 3] rely on orthogonal multiple access schemes which are less spectrum efficient. In this letter, we would like to address the above issues. We consider optimal adaptive streaming of an MVV to multiple users in a wireless network. We adopt an efficient scalable coding structure to enable instantaneous viewing angle switch and handle heterogeneous channel conditions. To maximally improve QoS for all users, we consider nonorthogonal multiple access and propose a transceiver design based on rate splitting and SIC. We optimize the design parameters to maximize the sum encoding rate of the enhanced versions of all requested views. By transforming the challenging non-convex problem into a CGP and using SCA, we obtain a KKT point.

2 System Model

As illustrated in Fig. **1**, we consider adaptive wireless streaming of an MVV with N views from a single-antenna server to $K > 1$ single-antenna users requesting K different views among the N views, where $K \leq N$.[1] Let $\mathcal{K} \triangleq \{1, \ldots, K\}$ denote the set of user indices. There are two time units, group of pictures (GOP) duration (usually 0.5-3s) and (transmission) slot duration (usually 1-5ms). Each GOP lasts hundreds of (around 100-3000) slots. We consider one GOP and focus on optimizing the QoS for each user via encoding rate adaptation and transmission rate adaptation. The encoding rate adaptation is operated at the beginning of each GOP according to the channel statistics of the K users, whereas the transmission rate adaptation is conducted at the beginning of each slot based on the instantaneous channel conditions of the K users.

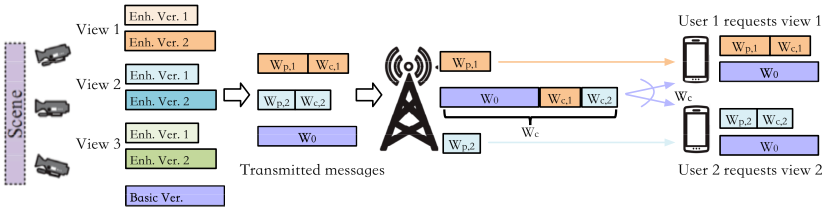


Fig. 1. System model. $N = 3$, $L = 2$, $K = 2$.

The N views are encoded into one basic version using multi-view video

[1]The proposed framework can be applied when more than one user request the same view by only considering the user with the worst channel.

coding and NL enhancement versions to enable instantaneous viewing angle switch and handle heterogeneous channel conditions. The basic version with encoding rate R_0 (in bit/s) provides the basic quality for the N views and serves as a prediction reference to reduce the encoding rates of the NL enhancement versions. Each view has L enhancement versions corresponding to L quality levels. Let $\mathcal{L} \triangleq \{1, \cdots, L\}$ denote the set of quality level indices. The ℓth ($\ell \in \mathcal{L}$) enhancement version has encoding rate R_ℓ and provides the lth highest quality, where $R_1 < \ldots < R_L$. In total, there are $L+1$ quality levels for each view. The basic version is transmitted to all K users to ensure that at any time, each user can watch the scene from any viewing angle with a basic quality guarantee. Besides, one enhanced version may be transmitted to provide enhanced quality of a certain level for each user, depending on the channel statistics.

We consider a discrete-time narrowband system and assume block fading, i.e., each channel does not change within one slot and changes in an independent and identically distributed (i.i.d.) manner over slots within each GOP. Assume that the server knows each user's channel distribution at the beginning of each GOP and each user's channel state at the beginning of each slot. We assume Additive White Gaussian Noise (AWGN) at each user. The transmission rate of an enhanced version of each requested view is adjusted according to the channel condition at each slot. The average transmission rate of an enhanced version of user k's requested view in the considered GOP, denoted by $\bar{r}_k$, which can be calculated according to the channel distribution in the considered GOP, determines the encoding rate of user k's requested view. Specifically, if there exists $\ell_k \in \mathcal{L}$ such that $R_{\ell_k} \leq \bar{r}_k < R_{\ell_k+1}$,[2] then the ℓ_k-th enhancement version of user k's requested view is transmitted to user k; otherwise, no enhancement version is transmitted to user k. Therefore, we focus on the optimal transceiver design for each slot in the following.

3 Transceiver Design

Note that orthogonal multiple access causes inefficient spectral utilization especially when users have heterogeneous channel conditions. To increase the transmission rate region and provide scalable visual quality for all users, we propose a transceiver design based on rate splitting and SIC [4]. Consider a particular slot. The message for user k's enhancement version, denoted as W_k, is split into a common part $W_{c,k}$ and a private part $W_{p,k}$. The message for the basic version, denoted as W_0, and the common messages of all users $W_{c,k}, k \in \mathcal{K}$ are combined into one supper common message, denoted by W_c. The super common message W_c and K private messages $W_{p,k}, k \in \mathcal{K}$ are independently encoded. W_c is delivered to all K users and $W_{p,k}$ is delivered to user k. Let $r_i \geq 0$ denote the achievable rate of message W_i, where $r_0 = R_0$ and

$$r_i \geq 0, \quad i \in \{(p,1), \ldots, (p,K), (c,1), \ldots, (c,K)\}. \tag{1}$$

[2]For notational convenience, we set $R_{L+1} = \infty$.

For all $k \in \mathcal{K}$, the achievable rate of message W_k for the considered slot is given by $r_{c,k}+r_{p,k}$. Denote $\mathcal{I} \triangleq \{c,(p,1),\cdots,(p,k)\}$. For all $i \in \mathcal{I}$, let $s_i \in \mathbb{C}$ denote the corresponding complex symbol of message W_i in an arbitrary channel use, with normalized power. Consider superposition coding (which allows a transmitter to send independent messages to multiple receivers simultaneously). The transmitted signal at the server is given by $x = P \sum\limits_{i\in\mathcal{I}} \alpha_i s_i$, where $\alpha_i \geq 0$ denotes the fraction of the total transmission power P allocated for transmitting symbol s_i and satisfies the power allocation constraint:

$$\sum_{i\in\mathcal{I}} \alpha_i \leq 1. \tag{2}$$

We consider SIC at each user. Each user k first decodes s_c by treating the interference from $s_{p,k}, k \in \mathcal{K}$ as noise and removes s_c from the received signal. Then, user k decodes $s_{p,k}$ by treating the interference from $s_{p,j}, j \in \mathcal{K}, j \neq k$ as noise. Let h_k denote the power of the channel between the server and user k. Let σ_k^2 denote the noise power at user k. The achievable rates satisfy:

$$r_0 + \sum_{k\in\mathcal{K}} r_{c,k} \leq \log_2\left(1+\frac{Ph_k\alpha_c}{Ph_k\sum\limits_{j\in\mathcal{K}}\alpha_{p,j}+\sigma_k^2}\right),\ k\in\mathcal{K}, \tag{3}$$

$$r_{p,k} \leq \log_2\left(1+\frac{Ph_k\alpha_{p,k}}{Ph_k\sum\limits_{j\in\mathcal{K},j\neq k}\alpha_{p,j}+\sigma_k^2}\right),\ k\in\mathcal{K}. \tag{4}$$

4 Problem Formulation and Solution

In this section, we would like to maximize the sum achievable rate of the enhanced versions of the K requested views by optimizing the power $\boldsymbol{\alpha} \triangleq (\alpha_i)_{i\in\mathcal{I}}$ allocation and achievable rate $\boldsymbol{r} \triangleq (r_i)_{i\in\mathcal{I}\setminus\{0\}}$ allocation.

Problem 1 (Rate Splitting and SIC)

$$\begin{aligned} \max_{\boldsymbol{\alpha},\boldsymbol{r}\succeq\mathbf{0}} \quad & \sum_{k\in\mathcal{K}} (r_{c,k}+r_{p,k}) \\ \text{s.t.} \quad & (1),(2),(3),(4). \end{aligned}$$

Problem 1 is non-convex, as the constraint functions in (3) and (4) are non-convex. The goal of solving a non-convex problem is usually to design an iterative algorithm to obtain a KKT point (which satisfies the necessary conditions for optimality if strong duality holds). In the following, we propose an iterative algorithm to obtain a KKT point of Problem 1 using CGP and SCA. Let $\tilde{r}_i = 2^{r_i}, i \in \mathcal{I}$ and define $\tilde{\mathbf{r}} \triangleq (\tilde{r}_i)_{i\in\mathcal{I}}$. We can equivalently[3] transform Problem 1 into:

[3]For tractability, in Problem 2, we consider $\boldsymbol{\alpha}, \tilde{\boldsymbol{r}} \succ \mathbf{0}$ instead of $\boldsymbol{\alpha}, \tilde{\boldsymbol{r}} \succeq \mathbf{0}$, which does not change the optimal value or affect the numerical solution.

Problem 2 (Equivalent Complementary GP of Problem 1)

$$\min_{\boldsymbol{\alpha},\tilde{\boldsymbol{r}}\succ\mathbf{0}} \quad \prod_{k\in\mathcal{K}} (\tilde{r}_{c,k}\tilde{r}_{p,k})^{-1}$$

$$\text{s.t.} \quad (2),$$

$$\tilde{r}_i^{-1} \leq 1, i\in\mathcal{I}, \tag{5}$$

$$\frac{\tilde{r}_0 \prod_{k\in\mathcal{K}} \tilde{r}_{c,k}\left(Ph_k\sum_{j\in\mathcal{K}}\alpha_{p,j}+\sigma_k^2\right)}{Ph_k\left(\sum_{j\in\mathcal{K}}\alpha_{p,j}+\alpha_c\right)+\sigma_k^2} \leq 1,\ k\in\mathcal{K}, \tag{6}$$

$$\frac{\tilde{r}_{p,k}\left(Ph_k\sum_{j\in\mathcal{K},j\neq k}\alpha_{p,j}+\sigma_k^2\right)}{Ph_k\sum_{j\in\mathcal{K}}\alpha_{p,j}+\sigma_k^2} \leq 1,\ k\in\mathcal{K}. \tag{7}$$

The constraints in (1), (3), and (4) are equivalent to the constraints in (5), (6), and (7), respectively. Problem 2 minimizes a posynomial subject to posynomial upper bound inequality constraints and a rational function of posynomials upper bound inequality constraint. Hence, Problem 2 is a CGP and can be solved using SCA [5]. Specifically, at iteration t, update $(\boldsymbol{\alpha}^{(t)},\tilde{\boldsymbol{r}}^{(t)})$ by solving the following approximate GP of Problem 2, which is parameterized by $(\boldsymbol{\alpha}^{(t-1)},\tilde{\boldsymbol{r}}^{(t-1)})$ obtained at iteration $t-1$.

Problem 3 (Approximate GP at Iteration t)

$$(\boldsymbol{\alpha}^{(t)},\tilde{\boldsymbol{r}}^{(t)}) \triangleq \arg\min_{\boldsymbol{\alpha},\tilde{\boldsymbol{r}}\succ\mathbf{0}} \quad \prod_{k\in\mathcal{K}} (\tilde{r}_{c,k}\tilde{r}_{p,k})^{-1}$$

$$\text{s.t.} \quad (2),(5),$$

$$\frac{\tilde{r}_0 \prod_{k\in\mathcal{K}} \tilde{r}_{c,k}\left(Ph_k\sum_{j\in\mathcal{K}}\alpha_{p,j}+\sigma_k^2\right)}{\prod_{j\in\mathcal{K}}\left(\frac{Ph_k\alpha_{p,j}}{\beta_{1,j,k}^{(t)}}\right)^{\beta_{1,j,k}^{(t)}}\left(\frac{Ph_k\alpha_c}{\beta_{c,k}^{(t)}}\right)^{\beta_{c,k}^{(t)}} C_{1,k}^{(t-1)\frac{\sigma_k^2}{C_{1,k}^{(t-1)}}}} \leq 1, \quad k\in\mathcal{K}, \tag{8}$$

$$\frac{\tilde{r}_{p,k}\left(Ph_k\sum_{j\in\mathcal{K},j\neq k}\alpha_{p,j}+\sigma_k^2\right)}{\prod_{j\in\mathcal{K}}\left(\frac{Ph_k\alpha_{p,j}}{\beta_{2,j,k}^{(t)}}\right)^{\beta_{2,j,k}^{(t)}} C_{2,k}^{(t-1)\frac{\sigma_k^2}{C_{2,k}^{(t-1)}}}} \leq 1, \quad k\in\mathcal{K}, \tag{9}$$

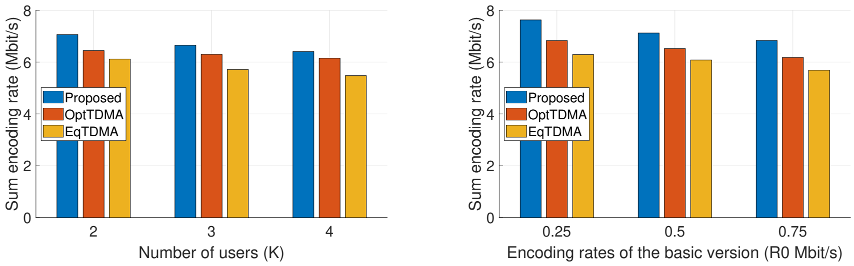
where $\beta_{i,j,k}^{(t)} \triangleq \frac{Ph_k\alpha_{p,j}^{(t-1)}}{C_{i,k}^{(t-1)}}$, $i=1,2,\ j,k\in\mathcal{K}$, $\beta_{c,k}^{(t)} \triangleq \frac{Ph_k\alpha_c^{(t-1)}}{C_{1,k}^{(t-1)}}$, $k\in\mathcal{K}$, $C_{1,k}^{(t-1)} \triangleq Ph_k\left(\sum_{j\in\mathcal{K}}\alpha_{p,j}^{(t-1)}+\alpha_c^{(t-1)}\right)+\sigma_k^2$, $C_{2,k}^{(t-1)} \triangleq Ph_k\sum_{j\in\mathcal{K}}\alpha_{p,j}^{(t-1)}+\sigma_k^2,\ k\in\mathcal{K}$.

Problem 3 is a standard GP, which can be readily transformed into a convex problem and solved efficiently. The details are summarized in Algorithm 1. Following the proof in [5], we can show its convergence.

Theorem 1 *For any feasible initial point $(\boldsymbol{\alpha}^{(0)},\tilde{\boldsymbol{r}}^{(0)})$, $(\boldsymbol{\alpha}^{(t)},\tilde{\boldsymbol{r}}^{(t)})$ obtained by Algorithm 1 converges to a KKT point of Problem 2, as $t\to\infty$.*

Algorithm 1 : Algorithm for Obtaining A KKT Point

1: **initialization**: Set $t = 0$ and choose any feasible initial point $(\boldsymbol{\alpha}^{(0)}, \tilde{\boldsymbol{r}}^{(0)})$ of Problem 2.
2: **repeat**
3: Set $t = t + 1$ and obtain $(\boldsymbol{\alpha}^{(t)}, \tilde{\boldsymbol{r}}^{(t)})$ by solving Problem 3 using CVX tools.
4: **until** Some convergence criteria is met.



(a) Sum encoding rate versus K at $R_0 = 0.5$Mbit/s. (b) Sum encoding rate versus R_0 at $K = 2$.

Fig. 2. Sum encoding rate versus K and R_0.

5 Numerical Results

In this section, we provide numerical results to illustrate the performance of the proposed scheme. We consider 100 slots in one GOP and set $N = 7$, $\sigma_k^2 = 1$ and $P = 30$dB. The bandwidth is set as 0.5MHz. We consider the spatially correlated Rayleigh-fading channel model in [4]. We evaluate the sum encoding rate over 100 random realizations of $h_k, k \in \mathcal{K}$ for the 100 slots. We consider two baseline schemes based on TDMA, namely OptTDMA and EqTDMA. In OptTDMA, the convex problem of optimal time allocation is solved using CVX tools. In EqTDMA, equal time allocation is considered. Then, the achievable rates of the messages for the enhanced versions of the K requested views in each slot are determined accordingly. For each scheme, the average achievable rates over the 100 slots are set as the encoding rates of the enhanced versions of the K requested views. Fig. **2** illustrates the sum encoding rate of the enhanced versions of the K requested views versus the number of users K and the encoding rate of the basic version R_0. The sum encoding rate of each scheme decreases with K and R_0, as more power is consumed for transmitting the basic version to all K users and less power is reserved for delivering the enhanced versions. Besides, the proposed solution outperforms the two baseline schemes, demonstrating the gain of the proposed nonorthogonal transmission scheme and optimal rate adaptation.

6 Conclusion

This letter investigated the adaptive wireless streaming of a scalable coding-based MVV to multiple users. We proposed a transceiver design based on rate splitting together with SIC and optimized the design parameters to maximize the overall QoS of all users. The results provide important design insights.