

Joint Optimization of Bitrate Selection and Beamforming for Holographic Video Cooperative Streaming in VLC Systems

Wanli Wen^{ID}, *Member, IEEE*, Xiaoling Peng^{ID}, Chen Chen^{ID}, *Member, IEEE*, Yunjian Jia^{ID}, *Member, IEEE*, Jianqing Li^{ID}, *Senior Member, IEEE*, and Harald Haas^{ID}, *Fellow, IEEE*

Abstract—Holographic video streaming requires ultrahigh channel capacity, which might not be achieved by the existing radio frequency-based wireless networks. To address this challenge, we propose a holographic video cooperative streaming framework by integrating coordinated multipoint transmission and beamforming technologies in visible light communication (VLC) systems. This framework enables simultaneous video streaming with an ultrahigh data rate for multiple users in the VLC system, resulting in a more efficient and effective streaming process. By mathematically modeling the streaming framework, we formulate a joint bitrate selection and beamforming problem, aiming to maximize the average video quality experienced by all users. The problem is a non-convex mixed-integer problem and is NP-hard in general. We propose an algorithm with polynomial time complexity for the problem using an alternative optimization technique along with an appropriate rounding operation. Numerical results demonstrate the superiority of the proposed joint bitrate selection and beamforming solution over baselines.

Index Terms—Holographic video streaming, visible light communication, coordinated multipoint transmission, bitrate selection, beamforming.

I. INTRODUCTION

WITH the advancement of wireless communications and augmented/virtual reality technologies, network and content providers are now offering users access to video content with three degrees of freedom (3-DoF), commonly referred to as panoramic video. By using smart devices like head-mounted displays (HMDs) and smartphones, users can immerse themselves in a virtual environment and interact with the content by changing their viewing direction (e.g.,

by looking up/down or left/right, or tilting their head). However, panoramic video has a limitation in that users are unable to move freely within the virtual environment, which reduces the overall level of immersion. In comparison, holographic video provides users with a more immersive 6-DoF viewing experience by allowing them to physically change their position and perspective, making it an increasingly popular subject of interest in academia and industry. Unfortunately, the high degree of freedom in holographic video creates a significant amount of data transmission, posing a major challenge for holographic video streaming. The current radio frequency (RF)-based wireless networks may not be able to provide sufficient channel capacity to support the ultrahigh data rate requirement (e.g., 1 Gbps [1]), which negatively impacts users' quality of experience (QoE). To tackle this issue, one approach is to enhance the channel capacity and the other one is to reduce the data rate requirement.

To improve channel capacity, advanced communication technologies such as Visible Light Communication (VLC), Coordinated Multipoint (CoMP), and beamforming can be employed in a synergistic manner. VLC has emerged as a complementary technology to RF-based networks, owing to its abundant, license-free light spectrum resources, making it well-suited for high-data-rate applications like video streaming. For example, authors in [2] and [3] proposed a WiFi-VLC dual connectivity system to support traditional video streaming [2] and holographic video streaming [3], respectively. CoMP and beamforming can enhance channel capacity through distinct techniques – CoMP coordinates data transmission from multiple access points (APs), while beamforming directs signals towards users and efficiently utilizes communication resources. Several studies, such as [4], [5], [6], and [7], have highlighted the potential of integrating CoMP and beamforming in VLC systems to significantly boost channel capacity. However, the current literature including [2], [3], [4], [5], [6], and [7] has not yet investigated whether and how this integration can benefit holographic video streaming.

To reduce the data rate requirement for video streaming, some works such as [8], [9], [10], [11], and [12] have proposed dividing a video into several tiles, transmitting only the tiles of interest to the users, and optimally allocating bitrate to the tiles and/or efficiently utilizing communication resources. The authors in [8] and [9] proposed partitioning panoramic and holographic videos into tiles, respectively. In particular, in [8], each tile is partitioned uniformly and a maximization problem of video quality is formulated to enhance QoE, while in [9], each tile is encoded with a different bitrate selection based on its distance from the user. The results gained in [8] and [9]

Manuscript received 4 August 2023; revised 23 August 2023; accepted 28 August 2023. Date of publication 1 September 2023; date of current version 11 October 2023. This work is sponsored by the National Natural Science Foundation of China under Grants 62201101, 61971077, and 62271091, the Project funded by China Postdoctoral Science Foundation under Grant 2022M720020, the Natural Science Foundation of Chongqing, China under Grant cstc2021jcyj-msxmX0458, the Chongqing Technology Innovation and Application Development Special Key Project under Grant CSTB2022TIAD-KPX0059, and the open research fund of National Mobile Communications Research Laboratory, Southeast University under Grant 2022D06. The associate editor coordinating the review of this letter and approving it for publication was A. Kaushik. (Wanli Wen and Xiaoling Peng are co-first authors.) (Corresponding authors: Chen Chen; Yunjian Jia.)

Wanli Wen is with the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China, and also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: wanli_wen@cqu.edu.cn).

Xiaoling Peng and Jianqing Li are with the School of Computer Science and Engineering, Macau University of Science and Technology, Macau, China (e-mail: 2009853CII30001@student.must.edu.mo; jqli@must.edu.mo).

Chen Chen and Yunjian Jia are with the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China (e-mail: c.chen@cqu.edu.cn; yunjian@cqu.edu.cn).

Harald Haas is with the Technology Innovation Centre and the Department of Electronic and Electrical Engineering, University of Strathclyde, G1 1RD Glasgow, U.K. (e-mail: harald.haas@strath.ac.uk).

Digital Object Identifier 10.1109/LCOMM.2023.3310974

1558-2558 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

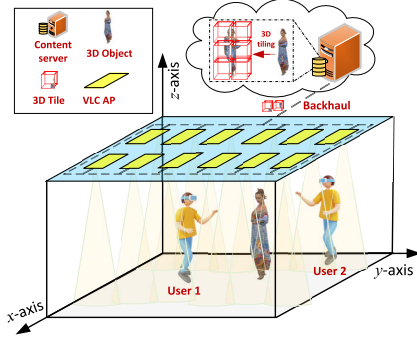


Fig. 1. System model.

are independent of any video characteristics such as saliency. The saliency of the video corresponds to the areas that are most likely to be viewed by users. Any distortions occurring in the higher-saliency areas are more noticeable and can be perceived as more bothersome, leading to a deterioration in QoE [10]. To address this issue, [11] investigated the influence of saliency in panoramic video on QoE and proposed a saliency-based bitrate selection and communication resource allocation scheme, with the aim of maximizing the average video quality of all users. Building on these findings, [12] extended the research to the context of holographic video and introduced a saliency-based bitrate selection scheme. Note that, the outcomes of the existing works are mostly restricted to RF-based wireless networks with limited resources, hence there is substantial scope for enhancing QoE.

Based on the above discussions, our primary motivation behind this study is to fill the gap in the literature and comprehensively investigate the potential benefits of integrating CoMP and beamforming for an enhanced holographic video streaming experience. We believe that our work may have a substantial impact and provide valuable insights, contributing to advancements in the field of holographic video communications. Our contributions are threefold: 1) We propose a holographic video cooperative streaming framework, which can enable simultaneous video streaming with an ultrahigh data rate for multiple users, resulting in a more efficient and effective streaming process. 2) We rigorously formulate a joint bitrate selection and beamforming problem with the aim of maximizing the average video quality experienced by all users. This is a non-convex mixed-integer problem and is NP-hard in general. 3) We develop an algorithm with polynomial time complexity for the problem using an alternative optimization technique along with an appropriate rounding operation.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, we present the system model, describing the key elements and interactions in the network. Subsequently, we formulate a joint bitrate selection and beamforming design problem aimed at maximizing the average video quality experienced by all users.

A. System Model

As shown in Fig. 1, we consider an indoor VLC system, which consists of a content center, N VLC APs, and K users. For ease of illustration, we denote $\mathcal{N} \triangleq \{1, 2, \dots, N\}$ and $\mathcal{K} \triangleq \{1, 2, \dots, K\}$ as the sets of the APs and the users, respectively. The content server stores a holographic video

that depicts a 3D object in the room, and connects to AP $n \in \mathcal{N}$ via a high-speed and low-latency backhaul link, such as optical fiber. By wearing a wireless HMD device that integrates a photo-detector (PD), user $k \in \mathcal{K}$ is streaming the holographic video from the content server via the VLC APs in $\mathcal{N}$. We examine a time-slotted system and focus on the streaming operation in an arbitrary slot (called the *typical slot*).

3D tiling is adopted to improve the transmission efficiency of the holographic video [12]. In particular, the holographic video is evenly divided into T cuboids, each referred to as a 3D tile. Let $\mathcal{T} \triangleq \{1, 2, \dots, T\}$ denote the tile set. To reduce the data rate requirement, we transmit only the tiles within the field of view (FoV) of each user from the APs. Let $\mathcal{T}_k \subseteq \mathcal{T}$ denote the set of tiles within user k 's FoV. We presume that $\mathcal{T}_k$ does not change during the typical slot. Given the user heterogeneity (e.g., spatial locations of users and channel conditions), each tile is pre-encoded into L quality levels. Define $\mathcal{L} \triangleq \{1, 2, \dots, L\}$ to be the quality level set. A higher quality level corresponds to a larger bitrate. Let μ_l (in bps) represent the bitrate of a tile with quality level $l \in \mathcal{L}$ and $\mathcal{B} \triangleq \{\mu_l\}_{l \in \mathcal{L}}$ the bitrate set, where $\mu_1 \leq \mu_2 \leq \dots \leq \mu_L$. We denote by $b_{t,k} \in \mathcal{B}$ the bitrate selection variable for user k concerning tile t , which satisfies

$$b_{t,k} \in \mathcal{B}, \quad t \in \mathcal{T}_k, \quad k \in \mathcal{K}, \quad (1)$$

$$b_{t,k} = 0, \quad t \in \mathcal{T} \setminus \mathcal{T}_k, \quad k \in \mathcal{K}. \quad (2)$$

Let $\mathbf{b} \triangleq (\mathbf{b}_k)_{k \in \mathcal{K}}$ with $\mathbf{b}_k \triangleq (b_{t,k})_{t \in \mathcal{T}}$ denote the bitrate selection design.

As mentioned earlier, due to the huge data volume, holographic video streaming requires an ultrahigh channel capacity. To meet this challenge, we propose integrating CoMP and beamforming into the VLC system, which allows all the VLC APs in $\mathcal{N}$ to cooperatively transmit the holographic video to the users in $\mathcal{K}$ using the same frequency channel of bandwidth B (in Hz) during the typical slot. Such integration has great potential to enhance the channel capacity and provide efficient transmission of the large data volume. Let $w_{n,k} \in \mathbb{R}$ denote the transmission beamformer on AP n for user k and $\mathbf{w} \triangleq (w_{n,k})_{n \in \mathcal{N}, k \in \mathcal{K}}$ the beamforming design, which satisfies

$$\sum_{k \in \mathcal{K}} w_{n,k}^2 \leq P_n, \quad n \in \mathcal{N}. \quad (3)$$

Here, P_n represents the maximal transmission power of AP n . Note that the equation (3) enforces a more stringent VLC signal amplitude constraint in comparison to those introduced in [4], [5], and [6]. Then, the signal-to-interference-plus-noise ratio (SINR) at user k can be expressed as

$$\text{SINR}_k(\mathbf{w}) \triangleq \frac{|\sum_{n \in \mathcal{N}} h_{n,k} w_{n,k}|^2}{\sum_{k' \in \mathcal{K} \setminus \{k\}} |\sum_{n \in \mathcal{N}} h_{n,k} w_{n,k'}|^2 + N_0 B}, \quad (4)$$

where $h_{n,k}$ denotes the VLC channel direct current gain from the n -th AP to the k -th user, and N_0 (in dBm/Hz) is the noise power spectral density. Based on (4), we can calculate the channel capacity $C_k(\mathbf{w})$ (in bps) of user k as $C_k(\mathbf{w}) \triangleq B \log_2(1 + \text{SINR}_k(\mathbf{w}))$. To ensure that user k can successfully receive the tiles within its FoV, the following constraint should be satisfied

$$C_k(\mathbf{w}) \geq \sum_{t \in \mathcal{T}_k} b_{t,k}, \quad k \in \mathcal{K}. \quad (5)$$

B. Problem Formulation

In this letter, we would like to maximize the average video quality experienced by all users in the system by jointly optimizing the bitrate selection $\mathbf{b}$ and beamforming design $\mathbf{w}$. Note that the quality of holographic video is positively correlated with the users' QoE [13] and is closely tied to three crucial factors: video bitrate, the distance between the user and the content, and video saliency [11], [12], [14]. First, as pointed out in [12] and [14], the video quality is logarithmically (rather than linearly) proportional to video bitrate, since it saturates at higher bitrates. Second, the holographic video quality is inversely proportional to the spatial distance between the user and the holographic video [12], i.e., the farther the 3D object is from the user, the worse the viewing experience of the user. Third, video saliency highlights the regions that the user is more likely to watch, and the distortions on the higher-saliency regions are more easily to be perceived as more obnoxious, thus worsening the user's viewing experience [11]. Therefore, based on the above discussions, we define the average video quality experienced by all users as

$$Q(\mathbf{b}) \triangleq \frac{1}{K} \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \frac{\delta_{t,k}}{D_{t,k}} \alpha \ln \left(\frac{\beta}{\mu_L} b_{t,k} + \gamma \right). \quad (6)$$

Here, $\delta_{t,k} \in [0, 1]$ is the saliency score for tile t , satisfying $\sum_{t \in \mathcal{T}} \delta_{t,k} = 1$ for all $k \in \mathcal{K}$; $D_{t,k}$ denotes the distance between tile t and user k ; α , β , and γ are some positive constants¹. Then, considering the constraints in (1), (2), (3), (5), and the objective function in (6), we formulate the following optimization problem:

Problem 1 (Average Video Quality Maximization):

$$\begin{aligned} & \max_{\mathbf{w}, \mathbf{b}} \quad Q(\mathbf{b}) \\ & \text{s.t.} \quad (1), (2), (3), (5). \end{aligned}$$

Due to the integer variable $\mathbf{b}$ and the non-convex constraint in (5), Problem 1 is a non-convex MINLP problem, which is NP-hard in general. Note that, for non-convex problems, there is generally no guarantee of obtaining an optimal solution, and the primary goal in solving such optimization problems is to find a suboptimal solution, as detailed in the following section. In addition, for practical implementation purposes, Problem 1 should be solved at the content server side in a centralized manner and the optimization results will be used to guide the bitrate selection at the server and the beamforming at the APs.

III. PROBLEM FEASIBILITY AND SOLUTION

In this section, we first derive the sufficient and necessary condition for the feasibility of Problem 1, and then develop an algorithm with polynomial time complexity for Problem 1.

A. Problem Feasibility

We have the following lemma to address the feasibility condition of Problem 1.

¹The specific values of these constants depend on the types of devices used by the users, such as TVs, smartphones, or laptops [14].

Lemma 1 (Feasibility Condition of Problem 1): Problem 1 is feasible if and only if the following problem is feasible.

$$\begin{aligned} & \text{find} \quad \mathbf{w} \\ & \text{s.t.} \quad (3), \\ & \quad C_k(\mathbf{w}) \geq |\mathcal{T}_k| \mu_1, \quad k \in \mathcal{K}. \end{aligned} \quad (7)$$

Proof: On the one hand, if Problem 1 is feasible, then we have $C_k(\mathbf{w}) \geq \sum_{t \in \mathcal{T}_k} b_{t,k} \geq |\mathcal{T}_k| \mu_1$, for all $k \in \mathcal{K}$, so the problem in (7) is feasible. On the other hand, if the problem in (7) has a feasible solution, denoted by $\mathbf{w}^\circ$, then we can construct a solution, denoted by $(\mathbf{w}^\circ, \mathbf{b}^\circ)$, where for all $k \in \mathcal{K}$, $b_{t,k}^\circ = \mu_1$ if $t \in \mathcal{T}_k$ and $b_{t,k}^\circ = 0$, otherwise. It is easy to verify that $\mathbf{w}^\circ$ satisfies (3), $\mathbf{b}^\circ$ satisfies (1) and (2), and $(\mathbf{w}^\circ, \mathbf{b}^\circ)$ satisfies (5), indicating that $(\mathbf{w}^\circ, \mathbf{b}^\circ)$ is a feasible solution of Problem 1. Thus, we complete the proof. ■

Lemma 1 indicates that the feasibility of Problem 1 is independent of its bitrate constraints in (1) and (2). Since the problem in (7) is a convex feasibility problem, we can solve it by using an interior-point (IPT) method to check Problem 1's feasibility. For the rest of this work, we presume that Problem 1 is feasible unless otherwise stated.

B. Problem Solution

First, by relaxing $\mathbf{b}$ to be continuous, i.e.,

$$\mu_1 \leq b_{t,k} \leq \mu_L, \quad t \in \mathcal{T}_k, \quad k \in \mathcal{K}, \quad (8)$$

we obtain a continuous version of Problem 1 as follows.

Problem 2 (Continuous Version of Problem 1):

$$\begin{aligned} & \max_{\mathbf{w}, \mathbf{b}} \quad Q(\mathbf{b}) \\ & \text{s.t.} \quad (2), (3), (5), (8). \end{aligned}$$

Problem 2 is non-convex but it is convex with respect to $\mathbf{w}$ (resp. $\mathbf{b}$) if $\mathbf{b}$ (resp. $\mathbf{w}$) is fixed. Thus, we can solve Problem 2 iteratively, where $\mathbf{w}$ and $\mathbf{b}$ will be optimized alternatively in each iteration. Specifically, let $i = 0, 1, 2, \dots$ denote the iteration index. Then, at iteration i , we optimize $\mathbf{w}$ and $\mathbf{b}$ by solving the following two subproblems, i.e.,

$$\begin{aligned} & \mathbf{w}^{(i+1)} \triangleq \arg \max_{\mathbf{w}} \quad Q(\mathbf{b}^{(i)}) \\ & \text{s.t.} \quad (3), \end{aligned}$$

$$C_k(\mathbf{w}) \geq \sum_{t \in \mathcal{T}_k} b_{t,k}^{(i)}, \quad k \in \mathcal{K}. \quad (9)$$

$$\begin{aligned} & \mathbf{b}^{(i+1)} \triangleq \arg \max_{\mathbf{b}} \quad Q(\mathbf{b}) \\ & \text{s.t.} \quad (2), (8), \end{aligned}$$

$$C_k(\mathbf{w}^{(i)}) \geq \sum_{t \in \mathcal{T}_k} b_{t,k}, \quad k \in \mathcal{K}. \quad (10)$$

Here, $(\mathbf{w}^{(i+1)}, \mathbf{b}^{(i+1)})$ denotes the solution at iteration i .

Next, we focus on solving the problems in (9) and (10). Since the objective function is independent of $\mathbf{w}$ and all constraints are convex, the problem in (9) is a convex feasibility problem. The convexity of the problem in (10) follows from the fact that the constraints are linear and the objective function $Q(\mathbf{b})$ is concave. Thus, we can obtain optimal solutions of these two problems by using an IPT method. The details of the iterative algorithm to solve Problem 2 are summarized in Algorithm 1. Note that due to the optimality of $\mathbf{w}^{(i+1)}$ and $\mathbf{b}^{(i+1)}$, the sequence $\{\mathbf{w}^{(i+1)}, \mathbf{b}^{(i+1)}\}_{i=0}^{\infty}$ will converge

Algorithm 1 Algorithm to Solve Problem 2

- 1: Set $i \leftarrow 0$ and $\mathbf{b}^{(0)} \leftarrow \mu_1 \mathbf{I}$ for all $k \in \mathcal{K}$ with $\mathbf{I}$ denoting a vector with all components being 1.
- 2: **repeat**
- 3: Obtain $\mathbf{w}^{(i+1)}$ by solving the problem in (9).
- 4: Obtain $\mathbf{b}^{(i+1)}$ by solving the problem in (10).
- 5: Set $i \leftarrow i + 1$.
- 6: **until** some convergence condition is met.
- 7: Set $(\mathbf{w}^*, \mathbf{b}^*) \leftarrow (\mathbf{w}^{(i)}, \mathbf{b}^{(i)})$.

Algorithm 2 Algorithm to Solve Problem 1

- 1: Obtain the stationary point $(\mathbf{w}^*, \mathbf{b}^*)$ via Algorithm 1.
- 2: For each $k \in \mathcal{K}$, set $b_{t,k}^\dagger \leftarrow \mu_1$ if $t \in \mathcal{T}_k$ and $b_{t,k}^\dagger \leftarrow 0$ if $t \in \mathcal{T} \setminus \mathcal{T}_k$.
- 3: For each $k \in \mathcal{K}$, sort the sequence $\{b_{t,k}^*\}_{t \in \mathcal{T}_k}$ in ascending order to form an ordered sequence $\{b_{(t),k}^*\}_{t \in \mathcal{T}_k}$.
- 4: For each $k \in \mathcal{K}$, $(t) \in \mathcal{T}_k$, and $l \in \{l^*, l^* - 1, \dots, 1\}$, set $b_{(t),k}^\dagger \leftarrow \mu_l$, if $C_k(\mathbf{w}) \geq \sum_{t \in \mathcal{T}_k} b_{(t),k}^*$, where $l^* \triangleq \arg \min_{l \in \mathcal{L}} |b_{(t),k}^* - \mu_l|$.

to a stationary point, denoted by $(\mathbf{w}^*, \mathbf{b}^*)$, of Problem 2 [15]. In addition, the computational complexity of Algorithm 1 is dominated by solving (9) and (10). If these problems are solved via an IPT method, the total complexity of Algorithm 1 is on the order of $\mathcal{O}(\max\{(NK)^3, (NK)^2(N+K), \Xi_1\} + \max\{K^3, 2K^3, \Xi_2\})$ [16], where Ξ_1 (resp. Ξ_2) denotes the cost of evaluating the first and second derivatives of the objective and constraint functions of the problem in (9) (resp. (10)).

Finally, the algorithm to solve Problem 1 is detailed in Algorithm 2, which consists of four steps. Specifically, Step 1 runs Algorithm 1 to obtain $(\mathbf{w}^*, \mathbf{b}^*)$ of Problem 2. Note that $(\mathbf{w}^*, \mathbf{b}^*)$ may not be a feasible solution of Problem 1 since $\mathbf{b}^*$ may be continuous. Thus, based on $(\mathbf{w}^*, \mathbf{b}^*)$, it is required to construct a feasible solution, denoted by $(\mathbf{w}^\dagger, \mathbf{b}^\dagger)$, of Problem 1, as detailed in Steps 2 to 4. Particularly, Step 2 initializes $b_{t,k}^\dagger$ to μ_1 (the lowest bitrate) if tile t is inside user k 's FoV and zero, otherwise. Note that this step can ensure that the constraint in (2) is satisfied. Step 3 sorts the bitrate sequence $\{b_{t,k}^*\}_{t \in \mathcal{T}_k}$ in ascending order to form an ordered sequence, denoted by $\{b_{(t),k}^*\}_{t \in \mathcal{T}_k}$, where $b_{(1),k}^* \leq b_{(2),k}^* \leq \dots \leq b_{(|\mathcal{T}_k|),k}^*$. Step 4 updates $b_{t,k}^\dagger$ based on the distances between $b_{(t),k}^*$ and the candidate bitrates in $\mathcal{B}$, which ensures that the constraints in (1) and (5) are satisfied. Since Steps 2 to 4 involve only some simple algebraic calculations, the computational complexity of Algorithm 2 is approximately equal to that of Algorithm 1.

IV. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed solution (obtained by Algorithm 1 and Algorithm 2) and compare it with representative baselines.

A. Simulation Settings

We consider a $5 \text{ m} \times 10 \text{ m} \times 3 \text{ m}$ room model, in which there are $N = 32$ VLC APs and $K = 3$ users. We use a

TABLE I

THE POSITIONS OF THE APs (UNIT: m)							
$\mathbf{a}_1$	(1, 1, 3)	$\mathbf{a}_2$	(1, 2, 3)	$\mathbf{a}_3$	(1, 3, 3)	$\mathbf{a}_4$	(1, 4, 3)
$\mathbf{a}_5$	(1, 5, 3)	$\mathbf{a}_6$	(1, 6, 3)	$\mathbf{a}_7$	(1, 7, 3)	$\mathbf{a}_8$	(1, 8, 3)
$\mathbf{a}_9$	(2, 1, 3)	$\mathbf{a}_{10}$	(2, 2, 3)	$\mathbf{a}_{11}$	(2, 3, 3)	$\mathbf{a}_{12}$	(2, 4, 3)
$\mathbf{a}_{13}$	(2, 5, 3)	$\mathbf{a}_{14}$	(2, 6, 3)	$\mathbf{a}_{15}$	(2, 7, 3)	$\mathbf{a}_{16}$	(2, 8, 3)
$\mathbf{a}_{17}$	(3, 1, 3)	$\mathbf{a}_{18}$	(3, 2, 3)	$\mathbf{a}_{19}$	(3, 3, 3)	$\mathbf{a}_{20}$	(3, 4, 3)
$\mathbf{a}_{21}$	(3, 5, 3)	$\mathbf{a}_{22}$	(3, 6, 3)	$\mathbf{a}_{23}$	(3, 7, 3)	$\mathbf{a}_{24}$	(3, 8, 3)
$\mathbf{a}_{25}$	(4, 1, 3)	$\mathbf{a}_{26}$	(4, 2, 3)	$\mathbf{a}_{27}$	(4, 3, 3)	$\mathbf{a}_{28}$	(4, 4, 3)
$\mathbf{a}_{29}$	(4, 5, 3)	$\mathbf{a}_{30}$	(4, 6, 3)	$\mathbf{a}_{31}$	(4, 7, 3)	$\mathbf{a}_{32}$	(4, 8, 3)

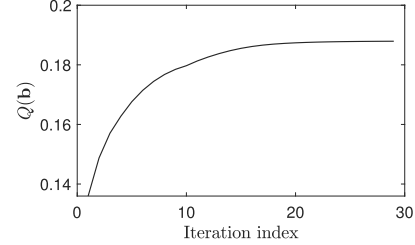


Fig. 2. The convergence of Algorithm 1.

Cartesian coordinate system, whose origin is shown in Fig. 1, to specify the positions of the APs and users. Specifically, the positions of the APs are detailed in Table I, where $\mathbf{a}_n \triangleq (x_{a,n}, y_{a,n}, z_{a,n})$ represents the 3D coordinates of AP n . The locations of the users are specified by $\mathbf{u}_1 = (1.5, 4.5, 0.85)$, $\mathbf{u}_2 = (2.5, 3.5, 0.85)$, and $\mathbf{u}_3 = (4, 7.5, 0.85)$, respectively, where $\mathbf{u}_k \triangleq (x_{u,k}, y_{u,k}, z_{u,k})$ represents the coordinate of user k with a unit of m. Assume that the VLC channel direct current gain $h_{n,k}$ follows the Lambertian radiation pattern, which is given by (11).

$$h_{n,k} = \begin{cases} \frac{m_n + 1}{2\pi d_{n,k}^2} \lambda A \cos^{m_n}(\phi_{n,k}) \kappa_1 \kappa_2 \cos(\theta_{n,k}), & \text{if } \theta_{n,k} \in [0, \Phi], \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

Here, $m_n = -\frac{\ln(2)}{\ln(\cos(\Psi_n))}$ is the Lambertian emission order with $\Psi_n = 60^\circ$ being the semi-angle at half power of AP n ; $\lambda = 0.5$ and $A = 19.6 \text{ mm}^2$ are the responsivity and the active area of the PD of each user, respectively; $d_{n,k} \triangleq \|\mathbf{a}_n - \mathbf{u}_k\|_2$ is the distance between AP n and user k ; $\phi_{n,k} = \arccos\left(\frac{z_{a,n} - z_{u,k}}{d_{n,k}}\right)$ is the irradiance angle; $\theta_{n,k}$ denotes the incident angle, which is equal to $\phi_{n,k}$ (since we assume that the detecting surface of the PD of each user faces upwards vertically); $\kappa_1 = 0.9$ is the gain of optical filter; $\kappa_2 = \frac{\varphi}{\sin^2(\Phi)}$ depicts the gain of optical lens with $\varphi = 1.5$ and $\Phi = 60^\circ$ denoting the refractive index and the half-angle FoV of the optical lens, respectively. The above parameter settings are referenced from [5] and [7]. Additionally, we set $L = 10$, $\mu_l = 10l \text{ Mbps}$, for all $l \in \mathcal{L}$, $T = 64$, $P_n = 22 \text{ Watt}$, for all $n \in \mathcal{N}$, $N_0 = -174 \text{ dBm/Hz}$, $B = 40 \text{ MHz}$, $\alpha = 0.5$, $\beta = 5$, and $\gamma = 1$ [12], [14]. The set $\mathcal{T}_k$ is formed by randomly selecting $|\mathcal{T}_k| = 35$ tiles from $\mathcal{T}$. The distance $D_{t,k}$ between tile t and user k follows a uniform distribution in the range $[0.5, 2.0] \text{ m}$. The saliency $\delta_{t,k}$ of tile t for user k follows the uniform distribution with $\delta_{t,k} \in [0, 1]$ satisfying $\sum_{t \in \mathcal{T}} \delta_{t,k} = 1$ for all $k \in \mathcal{K}$.

Using the above simulation settings, we evaluate the performance of the proposed Algorithm 1 and Algorithm 2

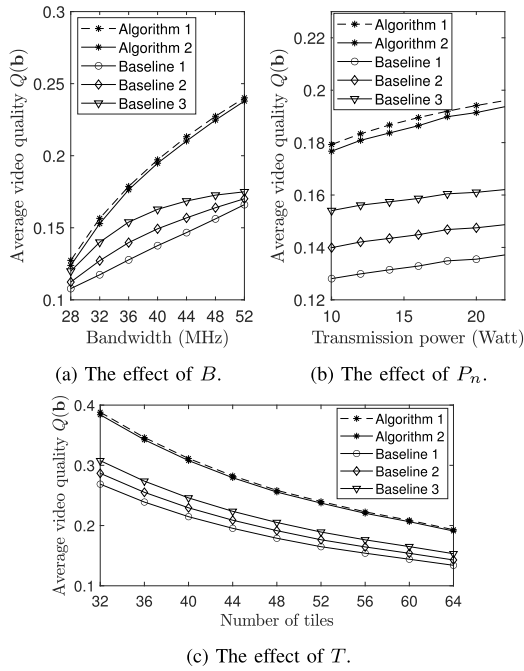


Fig. 3. Performance comparison.

and compare them with three baselines [8], [9], [11]. For ease of illustration, we denote the joint bitrate selection and beamforming scheme in relation to the baselines as $(\mathbf{w}^b, \mathbf{b}^b)$. The baselines run Algorithm 1 to obtain the beamforming design $\mathbf{w}^b$ and initialize the tiles in $\mathcal{T}_k$ to the lowest quality level, while the other tiles are set to zero bitrate (as in Steps 1 and 2 of Algorithm 2). Then, for user $k \in \mathcal{K}$, the quality level of each tile in $\mathcal{T}_k$ is then increased by one in a round-robin manner, with Baseline 1 starting with an arbitrary tile [8], Baseline 2 starting with the tile closest to user k [9], and Baseline 3 starting with the most salient tile [11], until the constraint $C_k(\mathbf{w}^b) \geq \sum_{t \in \mathcal{T}_k} b_{t,k}^b$ is no longer satisfied.

B. Performance Evaluation

Fig. 2 depicts the convergence behavior of the iterative algorithm, i.e., Algorithm 1. It is evident from the figure that Algorithm 1 demonstrates steady and rapid convergence, showcasing its effectiveness in optimizing the solution. The plot provides valuable insights into the performance of Algorithm 1 during the optimization process.

Fig. 3 shows a comparison of the performance of the proposed algorithms against three baselines. The results are obtained by averaging 200 trials, each with a new realization of all random variables. The observations from Fig. 3 are as follows. First, Algorithm 1 serves as an upper bound for Algorithm 2, as it outputs a continuous bitrate selection. However, the performance loss is less than 2%, indicating that the designed rounding operation in Algorithm 2 is appropriate and effective. Second, our algorithms significantly outperform the baselines, demonstrating their superiority in providing a high viewing experience. This is due to the algorithms' ability to jointly select appropriate video bitrates and effectively utilize limited communication resources. Finally, the performance of all schemes increases with the bandwidth and transmission power, but decreases with the number of tiles. This can be

explained as follows. The increase of B and P_n leads to more available communication resources, thus enhancing the video quality experienced by all users. When the holographic video is divided into more tiles but the tiles within the user's FoV remain unchanged (we set $|\mathcal{T}|_k = 35$ in Fig. 3(c)), it becomes more difficult for the user to view the holographic video, resulting in a decrease in viewing experience.

V. CONCLUSION

In this letter, we proposed a holographic video cooperative streaming framework that integrates CoMP and beamforming in VLC systems. We formulated a joint bitrate selection and beamforming problem to maximize the video quality experienced by all users and proposed an algorithm with polynomial time complexity using the alternative optimization method along with appropriate rounding operation. Simulations demonstrated the superiority of our algorithm. For future work, we plan to develop more advanced algorithms to solve the joint bitrate selection and beamforming problem.

REFERENCES

- [1] Z. Liu et al., "Point cloud video streaming: Challenges and solutions," *IEEE Netw.*, vol. 35, no. 5, pp. 202–209, Sep. 2021.
- [2] A. M. Alenezi and K. A. Hamdi, "Reinforcement learning approach for content-aware resource allocation in hybrid WiFi-VLC networks," in *Proc. IEEE 93rd Veh. Technol. Conf.*, Apr. 2021, pp. 1–5.
- [3] J. Chakareski and M. Khan, "WiFi-VLC dual connectivity streaming system for 6DOF multi-user virtual reality," in *Proc. 31st ACM Workshop Netw. Operating Syst. Support Digit. Audio Video*, Jul. 2021, pp. 106–113.
- [4] H. Ma, L. Lampe, and S. Hranilovic, "Coordinated broadcasting for multiuser indoor visible light communication systems," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3313–3324, Sep. 2015.
- [5] T. V. Pham, H. Le-Minh, and A. T. Pham, "Multi-user visible light communication broadcast channels with zero-forcing precoding," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2509–2521, Jun. 2017.
- [6] T. V. Pham and A. T. Pham, "Coordination/cooperation strategies and optimal zero-forcing precoding design for multi-user multi-cell VLC networks," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4240–4251, Jun. 2019.
- [7] M. A. Arfaoui, A. Ghayeb, C. Assi, and M. Qaraqe, "CoMP-assisted NOMA and cooperative NOMA in indoor VLC cellular systems," *IEEE Trans. Commun.*, vol. 70, no. 9, pp. 6020–6034, Sep. 2022.
- [8] C. Guo, Y. Cui, and Z. Liu, "Optimal multicast of tiled 360 VR video," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 145–148, Feb. 2019.
- [9] J. Park, P. A. Chou, and J.-N. Hwang, "Rate-utility optimized streaming of volumetric media for augmented reality," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 9, no. 1, pp. 149–162, Mar. 2019.
- [10] H. Duan, W. Shen, X. Min, D. Tu, J. Li, and G. Zhai, "Saliency in augmented reality," 2022, *arXiv:2204.08308*.
- [11] W. Huang, L. Ding, H.-Y. Wei, J.-N. Hwang, Y. Xu, and W. Zhang, "QoE-oriented resource allocation for 360-degree video transmission over heterogeneous networks," 2018, *arXiv:1803.07789*.
- [12] J. Li, C. Zhang, Z. Liu, R. Hong, and H. Hu, "Optimal volumetric video streaming with hybrid saliency based tiling," *IEEE Trans. Multimedia*, vol. 25, pp. 2939–2953, 2023.
- [13] X. Chen, T. Tan, and G. Cao, "Macrotiling: Toward QoE-aware and energy-efficient 360-degree video streaming," *IEEE Trans. Mobile Comput.*, early access, Dec. 30, 2022, doi: [10.1109/TMC.2022.3233022](https://doi.org/10.1109/TMC.2022.3233022).
- [14] S. Wang, S. Bi, and Y. A. Zhang, "Adaptive wireless video streaming: Joint transcoding and transmission resource allocation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3208–3221, May 2022.
- [15] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss-Seidel method under convex constraints," *Oper. Res. Lett.*, vol. 26, no. 3, pp. 127–136, Apr. 2000.
- [16] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.