RF-FSO Dual-Path UAV Network for High Fidelity Multi-Viewpoint Scalable 360° Video Streaming

Mahmudur Khan, Jacob Chakareski, and Sabyasachi Gupta

***Abstract*—We explore a novel RF-FSO dual-path UAV network for remote scene aerial scalable 360° video capture and streaming, to enable future virtual human teleportation. One UAV captures the 360° video viewpoint and constructs a scalable tiling representation of the data comprising a base layer and an enhancement layer. The base layer is sent by the UAV to a ground-based remote server using a direct RF link. The enhancement layer is relayed by the UAV to the server over a multi-hop path comprising directed UAV to UAV FSO links. The viewport-specific content from the two layers is then integrated at the server to construct high fidelity content to stream to a remote VR user. The dual-path connectivity ensures both reliability and high fidelity remote immersion. We formulate an optimization problem to maximize the delivered immersion fidelity which depends on the content capture rate, FSO and RF link rates, effective routing path selection, and fast UAV deployment. The problem is mixed integer programming and we formulate an optimization framework that captures the optimal solution at lower complexity. Our experimental results demonstrate an up to 6 dB gain in delivered immersion fidelity over a state-of-the-art method and for the first time enable 12K-120fps 360° video streaming at high fidelity.**

I. Introduction

360° video streaming systems are gaining popularity by enabling new digital immersive experiences in diverse application areas such as gaming and entertainment, education, weather sciences, and healthcare. Networked unmanned-aerial-vehicle (UAV) systems are another technology that is expected to play an increasingly prominent role in advancing our society. Together with 360° video streaming, a UAV network can help in monitoring immersively forest fires and forest ecology, remote sensing, disaster relief, broadcasting sports events, and observing behind enemy lines. The current state of the world (online classes, work from home, telemedicine, etc.) due to the *COVID-19* pandemic aptly illustrates the importance of remote 360° video communication and immersion. Compared to traditional video streaming, 360° video streaming presents further challenges such as ultra high data rate and ultra low latency. Thus, only low-quality/low resolution 360° videos can be streamed over traditional radio frequency (RF)-based wireless networks as they cannot support these challenging requirements for streaming high quality 360° content. Hence, we explore the use of free-space-optical (FSO) communication since it can enable much higher data transmission rates.

In this paper, we present for the first time a novel RF-FSO dual-path UAV network system that synergistically integrates with our scalable 360° tiling design, which will help overcome the above challenges and enable high fidelity aerial 360° video streaming (Fig. 1) to a remote virtual reality (VR) user. The UAV network is spatially distributed over a remote scene of interest where one UAV captures the present 360° video viewpoint of the remote user. All the UAVs are linked to a ground-based remote server through a common omnidirectional RF channel and they are connected to each other via high speed directed FSO links. One UAV can also directly communicate with the server using an FSO link, due to its proximity to the server. The capture UAV constructs a scalable tiling representation of the captured data comprising a base layer and an enhancement layer. The base layer is directly sent over the RF channel to the server. The enhancement layer is delivered to the server over multiple hops comprising directed UAV to UAV high speed FSO links. The VR user only sees a small portion of the entire 360° panorama, denoted as viewport (see Fig. 2, far right). Thus, leveraging the tiling aspect of our scalable representation, the server integrates the viewport-specific content from the two layers to construct a viewport-driven immersive representation of the present 360° viewpoint of the user. The omnidirectional RF link ensures reliability by enabling consistent streaming of the base layer. Streaming the enhancement layer over the high speed FSO links enables high fidelity immersion.

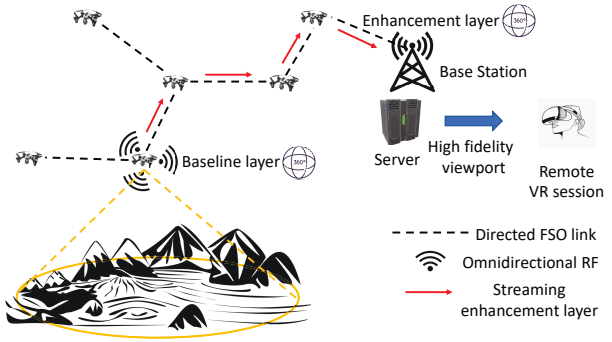

Fig. 1: RF-FSO dual-path UAV virtual human teleportation system. RF streams base layer and multi-hop FSO streams enhancement layer of 360° content.

The immersion fidelity and the quality of the delivered 360° content depend on the data capture rate, FSO and RF transmission rates, effective routing path selection for streaming the data, and fast deployment of UAVs to appropriate positions along the routing path. We formulate an optimization problem to maximize the delivered immersion fidelity given FSO and RF link capacities, available energy of the UAVs, and system latency constraints. The problem is mixed integer programming and we design a rigorous optimization framework that captures the optimal solution at lower complexity. We demonstrate through experiments that the proposed system can provide an up to 6 dB gain in delivered immersion fidelity over a state-of-the-art method and for the first time enable high fidelity 360° video streaming at 12K resolution and 120 fps.

This work was supported in part by NSF Awards ECCS-1711335, CCF-1528030, ECCS-1711592, CNS-1836909, and CNS-1821875.

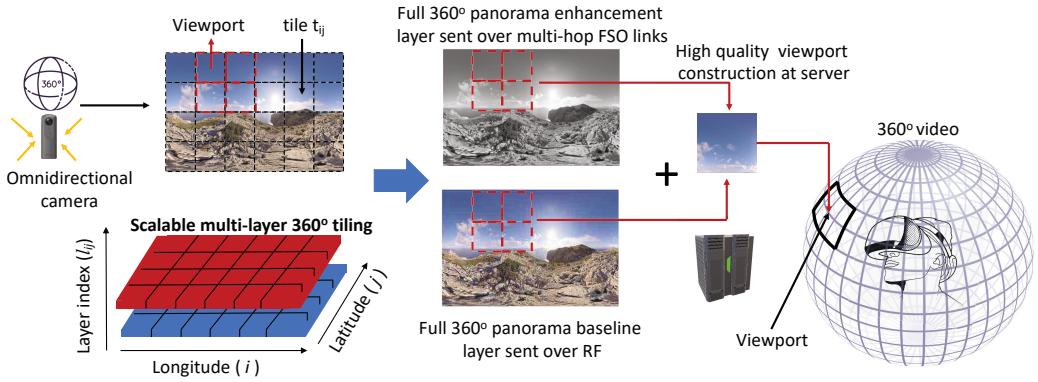

Fig. 2: A user's 360° viewpoint is represented as two embedded layers using scalable 360° tiling. The base layer of the entire 360° panorama is streamed over RF. Enhancement layer is sent over multi-hop FSO path. The viewport tiles from the two layers are then integrated at the server to enable high-fidelity immersion.

II. Related Work

Aerial 360° video capture and streaming for virtual human teleportation is an emerging topic. In [1], an aerial 360° video streaming system was proposed where multiple UAVs capture the remote scene content separately using multiple cameras. The captured content is sent to a ground-based server over an RF channel. There, the server synthesizes a remote scene VR immersion for a user. In [2], a system for capturing and live streaming regular video of a soccer game using a group of UAVs over an RF channel was proposed. Although the system in [1] enables aerial 360° video streaming, it cannot provide high immersion fidelity as it streams over an RF channel. On the other hand, the system in [2] enables only regular video streaming. In contrast, we present an RF-FSO dual-path aerial 360° video streaming system that provides high fidelity immersion. Thus, our system and application constraints are more challenging as they involve high speed and ultra-low latency requirements.

Providing area coverage via aerial sensors has been studied for different applications such as surveillance, exploration, or communication. In [3], placing UAVs with sensors at every location of interest was proposed to minimize the sensor travel time. In [4], two fast UAV network deployment methods were proposed to provide optimal wireless coverage. On the other hand, our goal is to maximize a user's quality of VR immersion by delivering high fidelity 360° content over a UAV network, which requires ultra-low latency and high speed data transfer. Hence, we focus on fast deployment of UAVs by minimizing the maximum delay to select the best FSO routing path and dynamically place the UAVs along the path, while considering constraints on the energy consumption of the individual UAVs.

III. System Framework

A. User application and remote location modeling

In our system framework, we consider that a VR user is physically located in an indoor VR arena, from where he/she can virtually teleport to a remote actual location. We model the virtual navigation space of the user as a grid which has a one-to-one correspondence with the remote scene grid (see Fig. 3). There are N_G points/locations on the grid and the user can teleport from one location to another using a hand-held device. Let P_S denote the probability of the user staying at his current location. Then, the probability of the user teleporting to a new location on the grid is $P_M = \frac{1-P_S}{N_G-1}$. We also define a time duration ΔT, during which the user navigates from his current location to a new location and must receive his present viewport on the headset. We model the remote scene locations to which the user can teleport as a grid (G_{remote}) with N_G nodes (see Fig. 3). The distance between two adjacent grid locations is d_{link}.

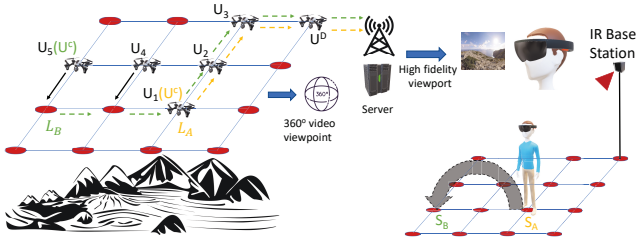

Fig. 3: Routing path selection for streaming the 360° enhancement layer and dynamic placement of UAVs along the path.

B. UAV modeling

We deploy a network of N_D UAVs $\mathcal{U} = \{U_1, U_2, ..., U_{N_D}\}$ over the remote scene of interest, where they hover at different locations on the grid at the same height. We equip each UAV with an omnidirectional camera, a multi-element FSO transceiver (adopted from our earlier work in [5]), and an RF antenna. They can establish high speed FSO communication links with neighbor UAVs hovering at one hop distance. All the UAVs can directly communicate with a ground-based server (see Section III-C) using RF links. The capacity of each FSO link is C_{FSO} and that of an RF link is C_{RF}. We designate each UAV in multiple ways during a given time period ΔT based on the task assigned to it by the server. One UAV is selected to capture the 360° content from a grid location corresponding to the present user location and is designated as U^C. It constructs a scalable tiling representation of the captured data comprising a base layer and an enhancement layer (more on this in Section

III-D). It streams the base layer to the server over the RF channel directly and the enhancement layer to a neighbor UAV over a high speed FSO link. A set of UAVs $\mathbf{U^R}$ is assigned the task of relaying the content from one UAV to another over FSO links. Finally, one UAV designated as U^D hovers at a grid location closest to the server and relays the content directly to the server over a high speed FSO link. Some UAVs may not be assigned any task. To ensure that a UAV can stay in service for the whole session, it is allowed to consume E_{max} amount of energy every time it moves to a new location.

C. Server operation

The server tracks the navigation actions of the user and uses this tracking information to compute in concert an optimal routing path along G_{remote} and new UAV position coordinates along the path, that meet the overall system latency constraints in capturing and streaming the selected 360° video viewpoint. This control information is relayed to the UAVs using the uplink RF channel. At the new coordinates, the UAVs establish FSO links with their neighbors to form the routing path.

The user can only view a small portion of the 360° content, denoted as viewport (Fig. 2). Thus, for efficient resource utilization, the server integrates the viewport-specific content from the received base layer and enhancement layer to construct a viewport-driven immersive representation of the remote scene, and streams it to the VR user in the arena.

D. Scalable multi-layer 360° tiling

We explore scalable 360° video tiling that synergistically integrates with the RF-FSO dual-path streaming. We partition each panoramic 360° video frame into a set of tiles. We denote a block of consecutive 360° video frames compressed together with no reference to other frames, as a group of pictures (GoP). We denote the set of tiles at the same spatial location (i, j) in a GoP as a GoP-tile. We construct two embedded layers of increasing immersion fidelity for each GoP-tile (Fig. 2) using the scalable extension of the latest video compression standard (SHVC [6]). The first layer is denoted as the base layer, and the second layer is denoted as the enhancement layer.

E. 360° viewpoint capture

Let V_c denote the 360° aerial view captured by UAV U^C from a given grid location. R^{FSO} denotes the assigned transmission rate to stream the enhancement layer over an FSO link and R^{RF} denotes the transmission rate to stream the base layer over the RF channel. R_c represents the temporal data rate that U^C uses to capture the 360° view V_c. The base layer associated with V_c is encoded at R^{RF} and the enhancement layer is encoded at R^{FSO}, and $R_c = R^{RF} + R^{FSO}$. We accurately characterize the reconstruction error of the delivered content at the server as $D_c(R^{RF}+R^{FSO}) = ae^{b(R^{RF}+R^{FSO})}$, which is monotonically decreasing with $(R^{RF} + R^{FSO})$ [1].

IV. Problem formulation

Our objective is to maximize the delivered immersion fidelity given FSO and RF link capacities, available energy of the UAVs, and system latency constraints. This is equivalent to minimizing the reconstruction error of the delivered 360° content due to the one-to-one mapping between the former and the latter. We formulate our optimization problem as

$$\min_{\mathbf{X}_\phi, \mathbf{Y}_\phi, \boldsymbol{\tau}_\phi, \phi, R_\phi^{FSO}} \quad \sum_{i=1}^{N_D} Y_i D_c(R_\phi^{FSO} + R_c^{RF}), \tag{1}$$

$$\text{s.t.} \quad Y_i \in \{0,1\},\ X_{i,r} \in \{0,1\},\ \sum_{i=1}^{N_D} Y_i = 1, \tag{2}$$

$$\sum_{i=1}^{N_D} X_{i,r} = 1,\ r \in \{1,2,...,|\phi|\}, \tag{3}$$

$$\sum_{r=1}^{|\phi|} X_{i,r} = 1,\ i \in \{1,2,...,N_D\}, \tag{4}$$

$$\sum_{i=1}^{N_D} Y_i \kappa \left(\frac{||L_c - L_i^o||}{\tau_{i,c}} \right)^2 \leq E_{max}, \tag{5}$$

$$X_{i,r}\kappa \left(\frac{||L_{\varphi_r} - L_i^o||}{\tau_{i,\varphi_r}} \right)^2 \leq E_{max},$$
$$i \in \{1,2,...,N_D\},\ r \in \{1,2,...,|\phi|\}, \tag{6}$$

$$\frac{|\phi|\Delta T \sum_{i=1}^{N_D} Y_i R_\phi^{FSO}}{C_{FSO}} +$$
$$\max_{i \in \{1,2,...,N_D\}} \left(\sum_{r=1}^{|\phi|} X_{i,r}\tau_{i,\varphi_r} + Y_i \tau_{i,c} \right) \leq \Delta T, \tag{7}$$

$$\sum_{i=1}^{N_D} Y_i \left(\frac{\Delta T R_c^{RF}}{C_{RF}} + \tau_{i,c} \right) \leq \Delta T, \tag{8}$$

$$R_c^{RF} \leq C_{RF},\ R_\phi^{FSO} \leq C_{FSO}, \tag{9}$$

where $\phi \in \Phi$ denotes a routing path for streaming the 360° content enhancement layer from U^C to U^D, and Φ is the set of all possible routing paths. For a given routing path ϕ, the number of hops is $|\phi|$. The vertices in ϕ are $\{c, \varphi_1, \varphi_2, ..., \varphi_{|\phi|}\}$. $\mathbf{X}_\phi$ denotes a matrix of values $X_{i,r}$, where $i \in \{1,2,3,...,N_D\}$ and $r \in \{1,2,...,|\phi|\}$. $\mathbf{Y}_\phi$ denotes a vector of values Y_i. $\boldsymbol{\tau}_\phi$ denotes a matrix whose rows comprise the vectors $(\tau_{i,c}\tau_{i,\varphi_1}\tau_{i,\varphi_2}...\tau_{i,\varphi_{|\phi|}})$. $\tau_{i,c}$ denotes the time it takes UAV U_i to relocate from its current location L_i^o to capture location L_c. τ_{i,φ_r} denotes the time it takes UAV U_i to relocate from L_i^o to relay location L_{φ_r} along the path ϕ. R_ϕ^{FSO} is the transmission rate of an FSO link for path ϕ.

In the constraint in (2), $X_{i,r}$ denotes if a UAV is assigned to perform a relay operation ($X_{i,r} = 1$) or not ($X_{i,r} = 0$). Y_i denotes if a UAV is assigned the task of capturing the 360° viewpoint ($Y_i = 1$) or not ($Y_i = 0$). This constraint also indicates that only one UAV captures the 360° view by moving to the location corresponding to the user's location in the virtual navigation space. The constraint in (3) indicates that only one UAV can be associated with a given relay location along the routing path. The constraint in (4) indicates that a relay location cannot have more than one UAV associated with it. The constraints in (5) and (6) impose that the maximum energy consumed by a UAV while moving from its current

location to a new location along a new routing path must not exceed E_{max}. The constraint in (7) imposes that the total time required for U^C and $\mathbf{U^R}$ to move to their corresponding locations on a selected routing path, establish FSO links with neighbor UAVs, capture the 360° content, and stream the enhancement layer to U^D must not exceed ΔT. The constraint in (8) imposes that the time required for U^C to move to the capture location, capture the 360° viewpoint, and stream the base layer to the server over the RF downlink must not exceed ΔT. The constraint in (9) indicates that the encoding rate of the 360° content base layer is bounded by C_{RF} and that of the enhancement layer is bounded by C_{FSO}.

The problem in (1) is mixed integer programming, which is hard to solve optimally in practice [7]. The optimal solution can be achieved via an exhaustive search, which will require searching over all the paths in Φ, and for every path, searching all the possible UAV-to-location assignments and the respective achievable FSO and RF link rates. The complexity of this method is $\mathcal{O}(|\Phi|(N_D+L_D)!/N_D!)$, which is very high. Here, L_D is the number of locations on a path. Hence, we propose a lower complexity approach ($\mathcal{O}(KN_D^{2.5})$, see Section V-A) to solve (1), where we search over $K << |\Phi|$ shortest paths (in terms of number of hops in a path) instead of searching over all $|\Phi|$ paths. Our experimental results in Section VI show that the proposed approach provides the optimal solution to (1) with much lower complexity.

We first present an outline of our approach. We begin by determining the K-shortest routing paths. Then, for each such path ϕ_k, where $k \in \{1, 2, ..., K\}$ we determine the UAV-to-location assignment that minimizes the maximum deployment delay of the UAVs along the path. Using this UAV-to-location assignment we obtain the maximum FSO link rate $R^*_{\phi^k}$ for path ϕ^k which ensures that the 360° enhancement layer content will be delivered within ΔT. Finally, we choose the path ϕ^k for which $R^*_{\phi_k}$ is greatest as the optimum path, due to the analytical nature of D_c which is monotonically decreasing with the FSO link rate. This completes the solution of (1).

Now we describe the proposed approach in more detail. For the k-th shortest path ϕ^k ($k \in \{1, 2, ..., K\}$), the problem in (1) can be re-written as

$$\min_{\mathbf{X}_{\phi^k}, \mathbf{Y}_{\phi^k}, \boldsymbol{\tau}_{\phi^k}, R^{FSO}_{\phi^k}} \sum_{i=1}^{N_D} Y_i D_c(R^{FSO}_{\phi^k} + R^{RF}_c), \tag{10}$$

$$\text{s.t.} \quad (2)-(6), (8), (9),$$

$$\frac{|\phi^k|\Delta T \sum_{i=1}^{N_D} Y_i R^{FSO}_{\phi^k}}{C_{FSO}} + \max_{i\in\{1,2,...,N_D\}} \left(\sum_{r=1}^{|\phi^k|} X_{i,r}\tau_{i,\varphi^k_r} + Y_i\tau_{i,c} \right) \leq \Delta T, \tag{11}$$

We decompose problem (10) into two independent problems as shown below.

$$\mathbf{X}^*_{\phi^k}, \mathbf{Y}^*_{\phi^k}, \boldsymbol{\tau}^*_{\phi^k} = \arg\min_{\mathbf{X}_{\phi^k}, \mathbf{Y}_{\phi^k}, \boldsymbol{\tau}_{\phi^k}} \max \left(\sum_{r=1}^{|\phi^k|} X_{i,r}\tau_{i,\varphi^k_r} + Y_i\tau_{i,c} \right), \tag{12}$$

$$\text{s.t.} \quad (2)-(6),$$

and,

$$R^*_{\phi^k} = \min_{R^{FSO}_{\phi^k}} D_c(R^{FSO}_{\phi^k} + R^{RF}_c), \tag{13}$$

$$\text{s.t.} \quad \frac{|\phi^k|\Delta T R^{FSO}_{\phi^k}}{C_{FSO}} + \max(\boldsymbol{\tau}^*_{\phi_k}) \leq \Delta T,$$

where the solution to problem (12) will help determine the placement of UAVs along the k-th shortest path and the solution to problem (13) will determine the FSO link rate for that path. We obtain the solution to problem (12) in Section V-A. The method to achieve the solution to problem (13) is described in Section V-B. Finally, at the end of Section V we show how the optimum routing path is selected using the methods proposed in Sections V-A and V-B, thus completing the solution to problem (1).

V. Optimization Framework

A. Dynamic placement of UAVs

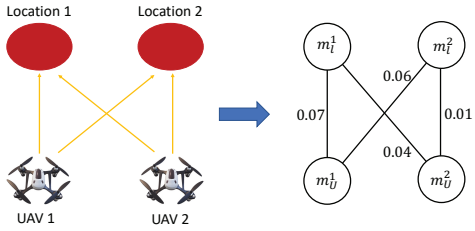

Fig. 4: Bipartite graph example for 2 locations and 2 UAVs.

We determine the K-shortest routing paths from U^c to U^D following an approach inspired by Yen's algorithm [8] that cannot be included here due to space limits. Once they are obtained, for each such path ϕ^k, where $k \in \{1, 2, 3, ..., K\}$ we study how to efficiently place the UAVs at its different grid locations. The 360° content enhancement layer cannot be streamed to the server until all links on the corresponding path are established, i.e., one UAV must move to each location on the path and establish FSO links with its neighbor UAVs. If the time required by UAV $u \in \{1, 2, ..., N_D\}$ to move from its current location to the new location $l \in \{1, 2, ..., L_D\}$ on a routing path is given by $\tau_{u,l}$, the streaming can start only after the UAV with $\max(\tau_{u,l})$ has reached its new location. Hence, we aim to find the UAV-to-location assignment that will minimize $\max(\tau_{u,l})$.

The optimal solution of the UAV-to-location assignment can be obtained by searching over all possible assignments. However, this requires searching over $(N_D+L_D)!/N_D!$ assignments and the time complexity is very high. Thus, we aim to solve the UAV-to-location assignment problem optimally at lower complexity by means of a graph-theoretic matching problem.

We formulate a weighted bipartite graph in which each UAV $u \in \{1, 2, ..., N_D\}$ and each location $l \in \{1, 2, ..., L_D\}$ are represented by vertices $m^1_u \in \mathcal{M}^1$ and $m^2_l \in \mathcal{M}^2$, respectively, and the weight of the edges (m^1_u, m^2_l) is expressed as $\omega_{(m^1_u, m^2_l)} = \tau_{u,l}$. We show a graph construction for two UAVs and two grid locations in Fig. 4. The two possible maximum matchings for this bipartite graph are $\{(m^1_l, m^1_u), (m^2_l, m^2_u)\}$ and $\{(m^1_l, m^2_u), (m^2_l, m^1_u)\}$.

The UAV-to-location assignment problem can be expressed as a bottleneck matching (BM) problem of the graph defined by the maximum matching whose largest edge weight is as small as possible, i.e.,

$$\min_{\pi \in \Pi} \max_{(m_u^1, m_l^2) \in \pi} \omega_{(m_u^1, m_l^2)}, \tag{14}$$

where Π comprises all the possible maximum matchings. For the graph in Fig. 4, the bottleneck matching is $\{(m_u^1, m_l^2), (m_u^2, m_l^1)\}$ and the corresponding assignment is: Location 1 is assigned to UAV 2 and Location 2 is assigned to UAV 1. We solve the problem in (14) using the BM algorithm proposed in [9] with complexity $\mathcal{O}(N_D^{2.5})$. Thus, the overall complexity of our proposed approach to solve (1) is $\mathcal{O}(KN_D^{2.5})$, where $K << |\Phi|$.

B. Determining the transmission rate of an FSO link

The reconstruction error of the enhancement layer decreases monotonically with the FSO link rate as observed from the analytical dependency introduced in Section III-E. Hence, the maximum value of $R_{\phi^k}^{FSO}$ for which the constraint in problem (13) holds, will minimize the reconstruction error of the 360° viewpoint given path ϕ^k and the corresponding UAV placements on it. This can be achieved by setting this constraint as an equality and rewriting it as:

$$R_{\phi^k}^* = (1 - \max \boldsymbol{\tau}_{\phi^k}^* / \Delta T) C_{FSO} / |\phi^k|. \tag{15}$$

Finally, we select the optimum routing path ϕ_{opt} from the K-shortest paths, and the respective optimum FSO link rate and the optimum UAV-to-location assignment for ϕ_{opt} as

$$R_{opt}^{FSO} = \max R_{\phi^k}^*, \tag{16}$$

$$\mathbf{X}_{opt}, \mathbf{Y}_{opt}, \boldsymbol{\tau}_{opt}, \phi_{opt} = \underset{\mathbf{X}_{\phi^k}^*, \mathbf{Y}_{\phi^k}^*, \boldsymbol{\tau}_{\phi^k}^*, \phi^k}{\arg\max} R_{\phi^k}^*.$$

The time required for the capture UAV to fly to the capture location is τ_{cap}. Thus, the RF link transmission rate R_{cap}^{RF} can be obtained by setting constraint (8) as an equality and rewriting it as:

$$R_{cap}^{RF} = (1 - \tau_{cap} / \Delta T) C_{RF}. \tag{17}$$

This selection of $\mathbf{X}_{opt}$, $\mathbf{Y}_{opt}$, $\boldsymbol{\tau}_{opt}$, ϕ_{opt}, R_{opt}^{FSO}, and R_{cap}^{RF} comprises the solution to the problem in (1). A formal description of the optimum routing path selection and UAV-to-location assignment method is provided in Algorithm 1, which is activated whenever the user moves to a new location.

Algorithm 1 Routing path selection and UAV placement

1: Given: G_{remote}, L_c, L_d, K.
2: Determine ϕ^k, $k \in \{1, 2, ..., K\}$.
3: **for** k = 1 to K **do**
4: Determine $\boldsymbol{\tau}_{\phi^k}^*$, $\mathbf{Y}_{\phi^k}^*$, and $\mathbf{X}_{\phi^\mathbf{k}}^*$ from (14).
5: Determine $R_{\phi^k}^*$ from (15).
6: **end for**
7: Select $\mathbf{X}_{opt}$, $\mathbf{Y}_{opt}$, $\boldsymbol{\tau}_{opt}$, ϕ_{opt}, and R_{opt}^{FSO} from (16).
8: Determine R_{cap}^{RF} from (17).
9: Send control information to the UAVs with role, location assignments, and transmission rates.

VI. Experimental Results and Analysis

We evaluate the performance of our proposed dual-path scalable 360° video aerial streaming system through extensive experimental analysis. We assess the delivered immersion fidelity enabled by our system in terms of the expected PSNR (Peak Signal to Noise Ratio) of the user viewport during a VR session. We formulate the latter as $PSNR = 10\log_{10}(255^2/D_{opt})$, where D_{opt} represents the expected reconstruction error of the delivered viewport content over the duration of the session. We use the popular 360° video *Angel Falls* [10] as the remote scene location viewpoints the user can navigate. It features 12K spatial resolution and temporal frame rate of 120fps. We set the minimum and maximum encoding rates for the entire 360° content as 61.56 Mbps and 2.4 Gbps, respectively. We construct the base and enhancement layers streamed in our system, by encoding this content as described in Section III-D, using R_{cap}^{RF} and R_{opt}^{FSO} as the encoding data rates of the two layers, respectively.

We obtain R_{opt}^{FSO} and R_{cap}^{RF} from (16) and (17) resepctively, where $C_{FSO} = B_{FSO}\log_2(1 + (R_{PD}P_r)^2/N_0B_{FSO})$ and $C_{RF} = B_{RF}\log_2(1 + P_U/(B_{RF}N_0 + \mathcal{L}_{us}))$. Here, $B_{FSO} = 100$ MHz and $B_{RF} = 500$ MHz are the FSO and RF channel bandwidths, respectively, $R_{PD} = 0.53$ A/W is the FSO receiver responsivity, $N_0 = 10^{-21}$ A^2/Hz is the noise power spectral density, $P_U = 50$ mW is the RF signal transmit power, and $\mathcal{L}_{us}$ represents the pathloss between a UAV and the server.

We consider a remote scene grid with 4×4 location points as shown in Fig. 3. We set the VR session duration to be 600 seconds. In line with video compression practices that commonly use a GOP duration of one second, we set $\Delta T = 1$ second. We also implement a state-of-the-art reference method [2] in our setting that uses the RF downlink channel to stream the 360° content. In this method, a UAV is assigned the task of either capturing and streaming the content directly to the server or just relaying it from the capture UAV to the server, depending on the distance between the capture location and the server. We compare our system's performance with this reference method.

A. Optimum solution at low complexity

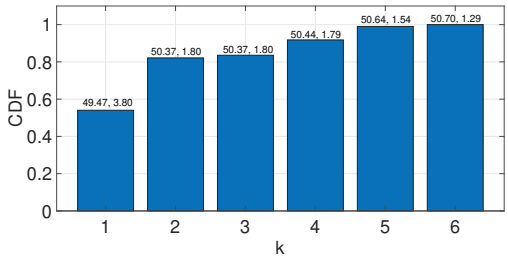

Fig. 5: Cumulative probability of the optimum routing path being within the first k-shortest paths identified by Algorithm 1 ($N_d = 7$, $d_{link} = 10$ m).

We explore the accuracy and low complexity nature of our optimization framework in this section. Fig. 5 shows the cumulative probability of the optimum routing path identified

by our proposed framework to be lying within the first k-shortest paths ($k = 1, 2, 3, 4, 5, 6$) We can observe that in more than half of the cases the shortest path is identified as the optimum path. In 92% of the cases the optimum path lies within the first 4 shortest paths. And in all cases the proposed framework identifies the optimum path within the first 6 shortest paths, thus K can be limited to 6. Hence, our optimization framework accurately captures the optimal solution at low complexity.

We also show the average value and standard deviation (σ_{PSNR}) of the delivered immersion fidelity as a 2-tuple on top of each bar in Fig. 5. We can observe that as k increases, the immersion fidelity reaches a plateau and small further gain is achieved. We can also see that the standard deviation of the delivered immersion fidelity decreases with an increase in k, e.g., for $k = 1$, $\sigma_{PSNR} = 3.80$ and for $k = 6$, $\sigma_{PSNR} = 1.29$.

B. Impact of the remote scene grid size

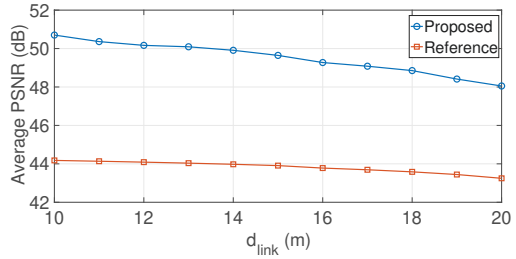

Fig. 6: Immersion fidelity for different remote scene grid size.

In Fig. 6, we examine the impact of the remote scene grid (G_{remote}) size on the delivered immersion fidelity for our proposed method and the reference method, when $N_D = 7$. We vary the size of G_{remote} (Fig. 3) by changing the length d_{link} between two adjacent points on the grid. We can observe that for our method the immersion fidelity monotonically decreases with an increase in d_{link}, as expected. In particular, as the size of G_{remote} increases the time required for the UAVs to form the routing path by moving to their assigned grid locations on the path increases, which results in smaller FSO transmission rates and hence decreased immersion fidelity. For the reference method, as the size of G_{remote} increases the capture UAV requires more time to move to its assigned location, resulting in smaller RF transmission rate and hence reduced immersion fidelity.

We can also see that the reference method achieves immersion fidelity of 43.25 dB - 44.18 dB for different values of d_{link}. On the other hand, our system enables streaming a 360° viewpoint with high immersion fidelity (48.05 dB - 50.7 dB). This gain in performance is enabled by our system by using parallel FSO and RF links, and creating an optimum multi-hop FSO routing path and optimum placement of UAVs along the path, thus ensuring maximum streaming rate selection.

C. Impact of the number of UAVs

In Fig. 7, we show the performance of our proposed method and the reference method for a different number of available UAVs (N_D), when d_{link}=10 m. As we increase the number of UAVs, for our system the probability of FSO

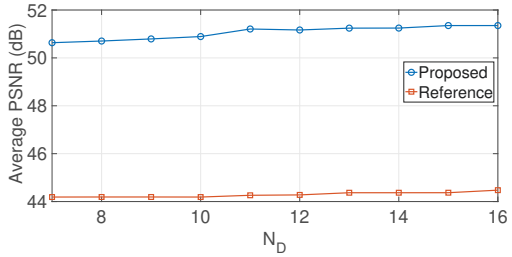

Fig. 7: Achieved performance with different number of UAVs.

links being established and delivering the captured 360° video viewpoint within ΔT increases. Thus, as a result, the delivered immersion fidelity improves with an increase in N_D. With only 7 UAVs, a very high fidelity immersion of 50.75 dB can be achieved and increasing the number of UAVs to 16 results in a very small additional gain of less than 1 dB. Thus, it is sufficient to deploy 7 UAVs to ensure a high quality experience for the user. Similarly, for the reference method the increase in immersion fidelity with an increase in N_D is negligible. We can also see that our proposed system enables an up to 6 dB gain in immersion fidelity compared to the reference method. The gain in performance is enabled by the different aspects of our system as discussed in Section VI-B.

VII. Conclusion

We investigated a novel RF-FSO dual-path UAV network for high fidelity aerial scalable 360° video streaming. The dual-path connectivity aspect of our system helps enable both improved reliability and high immersion fidelity. We formulated an optimization problem to maximize the delivered immersion fidelity which depends on the FSO and RF link rates, effective routing path selection, and fast deployment of the UAVs along the path. The problem is mixed integer programming and we formulated an optimization framework that accurately captures the optimal solution at lower complexity. Our experimental results show an up to 6 dB gain in delivered immersion fidelity over a state-of-the art method. We also demonstrate that our system can enable for the first time high fidelity aerial 360° video streaming at 12K resolution and 120 fps, thereby helping enable future societal VR applications.

References

[1] J. Chakareski, "UAV-IoT for next generation virtual reality," *IEEE Transactions on Image Processing*, vol. 28, no. 12, pp. 5977–5990, 2019.

[2] X. Wang, A. Chowdhery, and M. Chiang, "Networked drone cameras for sports streaming," in *Proc. IEEE ICDCS*, 2017, pp. 308–318.

[3] G. A. Hollinger and G. S. Sukhatme, "Sampling-based robotic information gathering algorithms," *IJRR*, vol. 33, no. 9, pp. 1271–1287, 2014.

[4] X. Zhang and L. Duan, "Fast deployment of UAV networks for optimal wireless coverage," *IEEE TMC*, vol. 18, no. 3, pp. 588–601, 2018.

[5] M. Khan and J. Chakareski, "Neighbor discovery in a free-space-optical UAV network," in *Proc. IEEE GLOBECOM*, 2019, pp. 1–6.

[6] J. M. Boyce, Y. Ye, J. Chen, and A. K. Ramasubramonian, "Overview of SHVC: Scalable extensions of the high efficiency video coding standard," *IEEE TCSVT*, vol. 26, no. 1, pp. 20–34, 2015.

[7] V. Vazirani, *Approximation Algorithms*, 2nd ed. Berlin Heidelberg: Springer-Verlag, 2003.

[8] J. Y. Yen, "Finding the k shortest loopless paths in a network," *Management Science*, vol. 17, no. 11, pp. 712–716, 1971.

[9] A. P. Punnen and K. Nair, "Improved complexity bound for the maximum cardinality bottleneck bipartite matching problem," *Discrete Applied Mathematics*, vol. 55, no. 1, pp. 91–93, 1994.

[10] "Angel Falls." https://www.youtube.com/watch?v=L_tqK4eqelA.