Remote Production for Live Holographic Teleportation Applications in 5G Networks

Peng Qian, Vu San Ha Huynh, Ning Wang, *Senior Member, IEEE*, Sweta Anmulwar, De Mi, and Rahim Tafazolli, *Senior Member, IEEE*

***Abstract*—Holographic Teleportation is an emerging media application allowing people or objects to be teleported in a real-time and immersive fashion into the virtual space of the audience side. Compared to the traditional video content, the network requirements for supporting such applications will be much more challenging. In this paper, we present a 5G edge computing framework for enabling remote production functions for live holographic Teleportation applications. The key idea is to offload complex holographic content production functions from end user premises to the 5G mobile edge in order to substantially reduce the cost of running such applications on the user side. We comprehensively evaluated how specific network-oriented and application-oriented factors may affect the performances of remote production operations based on 5G systems. Specifically, we tested the application performance from the following four dimensions: (1) different data rate requirements with multiple content resolution levels, (2) different transport-layer mechanisms over 5G uplink radio, (3) different indoor/outdoor location environments with imperfect 5G connections and (4) different object capturing scenarios including the number of teleported objects and the number of sensor cameras required. Based on these evaluations we derive useful guidelines and policies for future remote production operation for holographic Teleportation through 5G systems.**

***Index Terms*—Holographic Teleportation, extended reality (XR), 5G, remote production, edge computing**

I. Introduction

Holographic teleportation is a type of extended reality (XR) media application allowing end users to view live teleported objects with 6 degrees of freedom (6DoF), including the movement of head (roll, pitch and yaw) and body (forward-backward, left-right, and up-down). As a type of volumetric media application, it is expected that holographic teleportation will generate a profound impact on how people will communicate with each other in the future with fully immersive experiences. This is particularly the case with the development of 5G technologies that offer the necessary network and edge computing capabilities to support the delivery of teleportation services. Today, holographic teleportation is still in its infant stage, and the applications are typically operated in a very limited scale. Same as conventional video production operations, the current practice of running holographic teleportation applications require a dedicated content production server co-located with the sensor cameras on the source side (Fig. 1 (a)). This feature can be a significant obstacle to the wide deployment of such an application to ordinary residential users due to the high cost of maintaining high-power production server in individual households. By leveraging 5G new radio (5GNR) and edge computing capabilities, the functionality of content production can be offloaded to the 5G multi-access edge computing (MEC) node that provides open but secured platform for hosting such complex functions (Fig. 1 (b)).

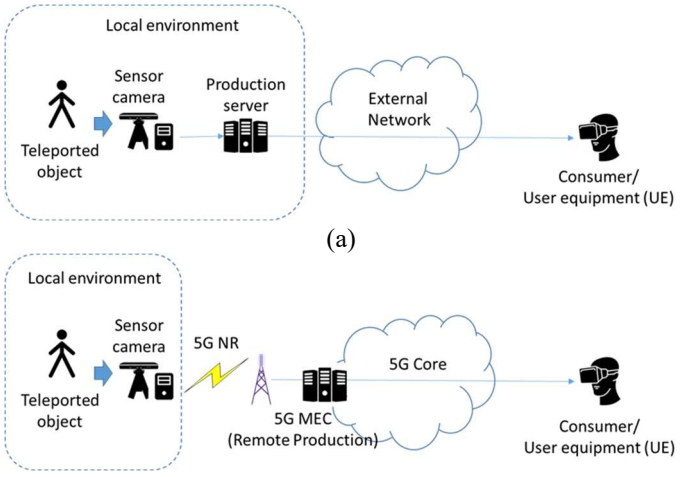


Fig. 1. Concept of 5G remote production for teleportation

In this paper, we present our developed framework for supporting remote production operations of live holographic teleportation services based on 5G MEC. To our best knowledge, this is the first piece of technical work in the literature to leverage 5G capabilities to enable remote production of volumetric media streaming over public networks. The proposed solution is expected to substantially reduce customer's capital investment in running such media applications by shifting the power consuming task of content production from user premises to the 5G mobile edge. In this paper we use the open-source LiveScan3D software [1] as the underlying teleportation platform as a representative application example, based on which the proposed remote production framework was implemented.

In addition to the introduction of the remote production framework for live teleportation, we also comprehensively and systematically analyze a wide range of network and application factors that may affect the quality of remote production operations in the 5G environment. All these experiments were carried out based on the real-life 5G network infrastructure built

at the 5G Innovation Centre (5GIC) hosted at University of Surrey, including 5G RAN, 5G mobile edge and core network components. We consider the following dimensions of analysis: (1) Bandwidth requirements against different content resolution levels, ranging from basic quality to ultra-high definition (UHD) level. We point out through our experiments that today's 5G capabilities are only capable of supporting the low-end of the resolution spectrum. (2) Different transport-layer congestion control algorithms for streaming content frames over the 5G uplink radio. (3) The testing of both outdoor and indoor environment scenarios with varied 5G radio connection qualities. (4) Varied configurations on the object capturing side, including the number of required sensor cameras and the number of human figures to be simultaneously teleported.

It is important to note that holographic teleportation is currently still in its infant stage, and such an application will continue to evolve in the future with much higher network/computing requirements and challenges. This paper aims to conduct a thorough feasibility analysis on what today's 5G network can support for the current early generation of holographic teleportation applications. However, such a study will certainly shed lights on how future 5G beyond network technologies can enable the delivery of next generation teleportation services, and in particular with the identification of a set of key factors that needs to be included in the future feasibility studies. The rest of the paper is organized as follows. In Section II we present our literature review. Section III provides a detailed overview of the holographic teleportation application system with 5G enabled remote production, followed by Section IV that provides in-depth analysis of 5G capabilities for supporting remote production operations. In Section V we conclude the paper.

II. Literature review

A. Application performance in cellular network

Measurement and understanding user-perceived performance of practical cellular network is the premise of deployment and optimization of modern multimedia applications. In 4G era, numerous efforts have been made to examine the achievable throughput and latency in different scenarios [2,3,4,5]. With the help of advanced wireless techniques and evolved network architecture, streaming high-resolution video (e.g., 4k video) at downlink presented an unprecedented success in industry. However, due to the high asymmetry of link conditions in cellular networks, uplink with sluggish performance (e.g., even less than 1 Mbps [4]) limits the application's usage in twofold: First, only application control messages (e.g., HTTP GET) can be observed in industrial network while video streaming is hardly to be supported due to insufficient throughput [5]. Second, with the increasing popularity of cloud computing, mobile users with limited processing ability are able to offload the high computation cost to the cloud where high performance devices reside. However, this is feasible only when the RTT is smaller than the client computes the task locally. In other words, 4G uplink is too poor to get such benefit fully. In contrast, the measured performance in 5G reveals the commercialized cellular performance is now approaching the Gbps era. To name a few, Narayanan, Arvind, et al. measures the uplink performance across three operators, and its throughput varies between 30 Mbps to 60 Mbps [6], Xu, Dongzhu et al. [7] conducted a systematic study in operational 5G network and the uplink throughput is (35-68 Mbps). Tong, JiangYu et al. provided a profound diagnosis of uplink, and the measured throughput ranges from 0.81 Mbps to 217.97 Mbps [8]. Meanwhile, the RTT calculated from the mobile edge to the first hop at the edge is around 20 ms [6,7]. Additionally, such boosted network resources and evolved network infrastructure provide fertile ground for utilizing point-to-multipoint transmission techniques to enhance the performance and efficiency of holographic video delivery. For instance, by comparison of commercial 5G network in different countries and the adoption of broadcast/multicast in 3GPP standardization, Gomez-Barquero, David, et al. [9,10] pointed out that the convergence of broadcast/multicast in 5G New Radio (NR), service-enabled 5G Core, and non-3GPP dual connectivity brings unprecedented advantages for mixed reality video to enhance its service quality, stability and automation ability. These advantages have been demonstrated via a practical 5G testbed with the support of multi-link and multicasting/broadcasting enabled network for immersive video [11]. Similarly, Change Ge, et al. [12] presented that with the assistance of MEC server, high quality video streaming via 5G satellite backhaul can yield satisfied and assured performance.

On the other hand, as transport layer protocols play a critical role in delivering video applications, diagnosing transport layer behavior in mobile has attracted extensive attention in academia. There are three main problems revealed in the traditional mobile network. First, the link utilization of the transport network protocol is low, as applying multiple connections at 5G downlink can triple the throughput of a single connection [5]. Second, the unavoidable RTT and loss variations affect the stability of the transport layer algorithm [5,6] and therefore, application (e.g., panoramic video) uploading throughput shows poor stability [7]. Third, the variations of congestion control algorithms and their parameters (e.g., CUBIC [13] and BBR [14]) exhibit different patterns/behaviors, showing disparities in advantage and disadvantage when facing network variations [15,16]. For instance, the TCP buffer size and unoptimized management policies at mobile phone result in degraded application throughput of mobile network in 5G network [8].

B. Holographic Applications

Holographic applications [1,17-22] is the most recent advances in immersive applications that create another dimension for immersive experiences. In order to generate holograms, research literature focuses on two main approaches: image-based [19,21] and volumetric-based solution [1,18,22]. Image-based holograms [19,21] require capturing objects as an array of images at different angles and tilts. This accumulates a massive amount of concurrent images which leads to significant demand in storage and network transmission bandwidth [17, 21]. Different compression mechanisms [21] have been proposed to reduce the overhead of image-based holograms by

leveraging the fact that individual images across the image array represent only minimal differences. Volumetric-based holograms [1,18,22] store objects as a collection of "point clouds" in the three-dimensional space such that each point has its set of X, Y and Z coordinates. The 3D point cloud format allows representing an object as a set of 3D volume pixels or voxels [25]. In order to capture volumetric videos, multiple RGB-D cameras with depth sensors (e.g., Microsoft Azure Kinect [26], Kinect v2 [27], Intel RealSense [28], and other LIDAR scanners [29]) are used to acquire 3D data from different angles which is then merged to produce the entire scene. The whole processing pipeline involves dynamically complex preprocessing and rendering actual images from different viewing angles at the local endpoint, gathering multiple point cloud objects and rendering them simultaneously [1,22,23] where frames from different cameras should be synchronized with unified coordinates and filtered noises in order to build the output hologram. In volumetric-based holograms, each voxel is transmitted only once and the volume is independent of the number of angles or tilts [17,18]. In addition to point clouds, 3D mesh [30] has been studied extensively to represent the volumetric video. However, point cloud introduces better rendering performance, more flexibility and simple representation of 3D objects as 3D mesh algorithms are more complex, employing a set of triangles, quadrangles or general polygons meshes to represent the geometry of 3D models [18,30]. Nevertheless, representing holograms as volumetric media requires complex compression schemes [31] such as octree-based approaches [32] to reduce the volume at the cost of computation. MPEG defines two representations of point cloud compression (PCC), namely video-PCC (V-PCC) [31] and geometry-PCC (G-PCC) [33]. V-PCC decomposes 3D space into two separate video sequences that capture the geometry and texture information while G-PCC encodes the content directly in 3D space by utilizing data structures (e.g., an octree [34]) to describe the point location in 3D space. Most decoding volumetric videos rely on the software where dedicated hardware support for volumetric hologram decoding is still limited [18].

III. HOLOGRAPHIC TELEPORTATION SYSTEM OVERVIEW

In this section, we describe our holographic teleportation system which uses multiple Azure Kinect sensor cameras to capture scenes from multiple viewpoints and stream aggregated 6DoF video content to the user in real time. The rest of the section is organized as follow. Section 3.1 presents the system overview of our holographic teleportation, section 3.2 describes the network video stream specification, frame structure and signaling protocol of the teleportation system, section 3.3 discusses the use cases and applications of the holographic teleportation system in 5G networks.

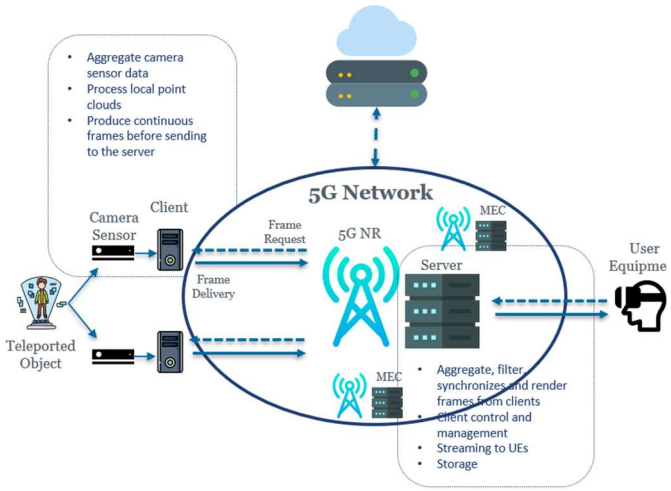


Fig. 2. Holographic Teleportation System Overview.

A. Holographic Teleportation System

We leverage and substantially extend the open-source user-friendly interface holographic application LiveScan3D [1] that allows us to capture the 3D objects using multiple sensor cameras, aggregate and produce 3D hologram at the server before streaming 3D hologram as 6DoF video content to the user. It is worth mentioning that, although we implemented our design based on the LiveScan3D, our holographic teleportation is generically applicable to other holographic-type communication platforms.

Figure 2 shows an overview of our holographic teleportation system. To enable the 6DoF feature, a remote object is surrounded by multiple sensor cameras in order to be captured into a 3D hologram. Each sensor camera is locally connected to a computer acting as a client whose responsibility is to capture sensor camera data, process local point clouds and produce continuous frames before sending them to the server. Each client has a global unique identifier (Guid) as IP addresses are not sufficient in identifying separate clients, especially when clients are in the same subnet or have a common public IP address when operating over the public Internet. In our teleportation system, we use Microsoft Azure Kinect camera as the sensors [26] although our system is portable to work with other types of sensor cameras such as Kinect v2 [27]. Microsoft Azure Kinect sensor camera allows us to enable a variety of frame rates and resolutions from both RGB and depth sensors. To compare with Kinect v2, Azure Kinect at synthetic RGBD resolution of Ultra HD (i.e., 3840x2160) can generate more than three million point clouds per frame assuming one frame is produced every 33.33 milliseconds that is at least three times more produced points than Kinect v2 which results in better quality volumetric [26,27].

The system allows two types of holographic experience: streaming the whole scene or streaming human bodies only. In this article, we mainly focus on the scenario of capturing remote human bodies into the 3D hologram. Note that the human body tracking functionality in Azure Kinect SDK relies on a deep neural network model [26] which requires computational capability from the computer connected to the sensor camera. Nevertheless, the produced frame of human bodies to be

streamed to the remote production server has a substantially smaller size compared to that in streaming whole scenes. Thus, eliminating the volumetric content of the environment allows us to reduce the redundant network traffic and increase the 6DoF video content quality.

With Network Function Virtualization (NFV) technologies, [35,36] the production server function can be flexibly deployed at any location in the network. In this paper, we focus on the scenario of embedding such a power computing function at the 5G mobile edge. From the network point of view, the production server is deployed at the central high-power computing node, aggregating all new frames from multiple groups of clients where each group represents a particular teleported object(s) at a particular location. Due to the differences in frame rate, computational capabilities or network conditions, raw frames captured at the server need to be synchronized as misaligned sensors resulting in a badly rendered image. Local frames of the same object from various clients can be synchronized locally before sending to the server using 3.5-mm synchronization ports [26] that can use to link multiple devices together in daisy-chain or star configuration [26] or in our case, the server can collect all the raw frames and synchronize them [22] before rendering to a 3D object. At the server, we utilize the synchronization function [22] between frames coming from different clients capturing a common object from different angles. These clients, even though are at the same site, may have different hardware and network conditions. Different bandwidth-delay-product (BDP), especially dynamic network latency or packet loss will severely affect data throughput resulting in uncertain time arrival of the frames to the server, out of order frames or frame loss. Even slightly ahead or behind of frames is unacceptable in our holographic teleportation where stringent synchronization between streams is required. One possible solution is in addition to synchronize frames at the server using the timestamp, we put the synchronization function along the end-to-end path, from client to the server and from the server to the user. In the context of 5G networks, frame synchronization at the 5G multi-access edge computing (MEC) [37,38,53] may help to reduce the risk of synchronization disruptions, with the assistance of deterministic content delivery capability of 5GNR In addition to synchronization, the server is also able to filter out noise generated by Azure Kinect sensors to reduce the "flying pixels" effect or compress the frame before transmission [1]. Furthermore, the system utilizes calibration function [1] based on Iterative Closest Point (ICP) algorithm [39] that requests clients to transform its frames from its own coordinate system to the universal coordinate system at the moment of capture in order to correctly render 3D object from different angles based on the position of the sensor cameras.

While the server aggregates, filters, synchronizes and merges the frames from the clients to produce the output frames for live showing, the output frames can be stored as a single volumetric file (e.g., ply or bin file format) for later usage or more importantly to be live streaming to the user equipment (UEs) such as Microsoft HoloLens 2 or any AR/VR [40, 41,42] supported smartphones. The rendered frames of the complete 3D hologram of either single or multiple teleported objects are streamed together to multiple users via multicast network communication. Our holographic teleportation system supports low-latency streaming of 6DoF video content to the user across the globe in real time which allows users, via UEs, not only to view all the teleported objects in a common view field, but also to change the viewing position, manipulate and modify size, position and orientation of each individual object. We propose to offload complex holographic content to the 5G edge computing close to the viewers to improve the end-to-end latency performance by circumventing the long RTT issue of the fixed Internet paths and reduce the network traffic redundancy when streaming to users within the same geographic location [53,54].

B. Network Communication, Frame Structure and Signaling

We utilize pull-based communication pattern between remote production server and clients such that all requests are initiated by the production server. As soon as the clients start a connection with the server, the server sends different types of signals to request new frames from the clients and control the behaviors of connected clients (e.g., transmit whole scene or bodies only, compression level, calibration instructions, etc.) simultaneously. The clients communicate with the server via TCP/IP [15], however, our teleportation system does not limit to only TCP/IP communications but can be extended to other network communication protocols such as QUIC [43].

There are 12 different types of frames for signaling between the clients and the server [22] after connections between clients and server are established: two for the server sending all sensor camera settings and other properties request to the clients (e.g. client's Guid, streaming mode, frame compression level) at the beginning of a session; three for the calibration process where the server instructs clients to transform frames from its own coordinate system to the universal coordinate system before sending; seven for requests, transmissions and acknowledgements of the data exchange process where the clients send either its latest live frames or its pre-recorded frames or no frame signal to the server.

Field	Frame Type	Data Length	Compression	Timestamp	QoE/QoS Control Signal
Length (Byte)	1	4	4	4	4

Fig. 3. Header format

As shown in Figure 3, we outline the packet header structure for transmission from clients to server and from server to UEs including frame type (one byte), data length (four bytes), compression (four bytes), timestamp (four bytes) and optional QoE/QoS control signal (four bytes). While the first 13 bytes have been discussed in [22], we add one additional (optional) header field which is used by the server for QoE and QoS control signals. In the real network environment, there is always trade-off challenges between different QoE metrics such as resolution quality, playback disruptions or frame synchronization. Even one single QoE metric (e.g., Frame Per Second (FPS)) is affected by multiple QoS (e.g., bandwidth, latency, packet loss, etc.). To balance the trade-off, the

holographic teleportation system may need to sacrifice one for another (e.g., sacrifice the FPS to reduce playback latency). Based on the current network characteristics and application performance, the QoE and QoS control signal header field in each frame request helps the server to instruct each client to, for example, begin stochastically dropping frames in order to improve the playback experience of the user.

TABLE I
NUMBER OF POINT CLOUDS VS. AZURE KINECT RESOLUTION

Mode	Point Clouds	Bytes Per Frame
Efficiency (NFOV 2x2)	75000	675000
Efficiency (NFOV-Unbinned)	267000	2403000
Efficiency (WFOV-2x2)	240000	2160000
Efficiency (WFOV-Unbinned)	930000	8370000
Quality (HD)	418000	3762000
Quality (FHD)	940000	8460000
Quality (WQHD)	1671000	15039000
Quality (QXGA)	1198000	10782000

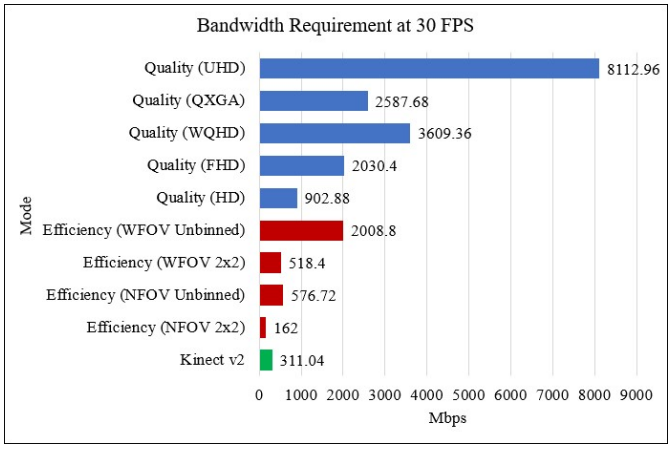


Fig. 4. Bandwidth requirement for different Azure Kinect camera resolution at 30 FPS.

As the client captures and converts RGBD images to points in 3D space, for each point cloud being transmitted, we use nine bytes in the payload: one for each RGB and two for each XYZ dimension. Table 1 and Figure 4 show the number of point clouds in each frame as well as the corresponding bandwidth requirement for different Azure Kinect sensor camera resolutions. In the efficiency mode, the client projects the color resolution into depth space. For example, given the narrow field-of-view unbinned depth mode, each frame contains approximately 267000 point clouds, resulting in 19.2 Mb (or 2.4 MB) per frame. Given the frames produced by the Azure Kinect sensor camera at the maximum rate of 30 Frames Per Second (FPS) as its capacity, the required network bandwidth is 576 Mbps. In the quality mode, the client projects the depth resolution into color resolution space. For example, in order to achieve ultra HD resolution which generates 3.8 million point cloud per frame, the maximum data rate required would be approximately 8 Gbps. However, note that this bandwidth required is theoretically for streaming the whole scene. Since we focus on projecting human bodies only, the bandwidth required would fall significantly depending on how many people being captured by the sensor camera at the same time, their position, postures and their movement relative to the sensor cameras. For streaming one single person, the maximum data rate required in ultra HD resolution at 30 FPS is around 2.9 Gbps. Furthermore, the holographic teleportation system also utilizes Zstandard library [44] to compress data before sending to the client. The data compression level is controlled by the server via the signaling setting frame at the beginning of the session. The data compression applied at the client in real time would further reduce the data rate without losing the frame quality. Note that frames across multiple sensor cameras may include only minimal differences, thus more sophisticated compression schemes can be applied to capture and exploit the spatial-temporal frame coherency and further reduce the volume [17]. As a trade off, it comes at the cost of computation [17]. However, exploring different data compression mechanisms is out of the scope of the paper. In addition to the point cloud data, the holographic teleportation system includes one byte of source id in the payload to indicate frames from different sources or different network locations of the teleported objects. There are also some additional data transmitted for the skeleton and joints information which could be neglected due to its small size compared to the point clouds.

The server sends latest-frame requests to the clients every short adjustable period of time while the clients receive frame requests and respond with a list of its latest produced frames to the server. With network bandwidth limitations and possible high latencies, the environment might result in frames not being transmitted to the server faster than they are captured. On this occasion, new frames will be queued into the client buffer awaiting for the next frame requests to be transmitted to the server. When the buffer is full, the client starts dropping old frames to create more space for newer frames captured. Note that sending frames upon explicit client request have a couple of advantages. On one hand, it enables potential frame-granularity dynamics (e.g., different resolutions per frame) according to varying network conditions and client processing ability. Moreover, blindly pushing frames from multiple cameras via a share 5G link leads to imbalanced throughput, thus the frame requests allow the server to synchronize the frame rate between different sensor cameras. On the other hand, this frame per request behavior comes at the cost of scalability and redundant round trip for every request. In future work, we plan to explore the trade-offs between different data transmission mechanisms. One possible alternative that solves the scalability problem is to keep a unique connection between server-client where the server only needs to send a single request for new frames while the clients actively send new captured frames to the server until receiving further instructions. However, as it is not the focus of our paper, we built our work on top of the existing implementation and system design of the teleportation platform [1,22] which follows frame per request manner.

While the client utilizes one buffer for transmitting frames, there are three core buffering mechanisms running in parallel on the server: receiving buffer, production buffer, and transmitting buffer [22,23], each of them handles specific tasks.

The buffering mechanism in our holographic teleportation helps to deal with network uncertainties especially over the long-distance Internet path and decouple data processing with network operations. The receiving buffer handles all the raw frames received from the clients as well as continuously measure network and buffer status such as throughput, latency, etc. The frames are queued in receiving buffer in FIFO manner and are organized in ascending timestamp order. Production buffer is responsible for synchronization between frames coming from different clients jointly capturing a common real-world 3D object [22,23] from different angles. The synchronization functions rely on the timestamp field in every packet header which in turn relies on the Network Time Protocol (NTP) where both clients and server synchronize their clocks with the NTP server [45] once at the beginning of the transmission. After the frame synchronization, the production buffer handles the rendering process and live showing 3D hologram at the server. Last but not least, the transmitting buffer helps to serve rendered output frames to the UEs as well as continuously observe and record all the statistics of these server-UE connections for control and management.

The connectivity between clients, server and UEs rely on, but not limited to, TCP/IP communications. Our holographic teleportation allows multiple parallel TCP connections per client in order to improve the throughput and playback performance, especially when coping with high RTT [22]. Out of order frames or frame losses at the receiving buffer on the server are controlled based on the frame production timestamp attached in each packet header. When multiple parallel TCP connections are used, the clients check the status of each connection based on round-robin and only write to the available sockets. By decoupling the network communications with data processing in both clients, server and UEs, the holographic teleportation system can optimize its operations' performance in both application and network layer.

C. Holographic-Type Use Cases and Applications in 5G Networks

The holographic Teleportation provides users with 6DoF immersive viewing experience, which is three more degrees of freedom compared to 3DoF VR video including freely moving forward/backward, up/down, or left/right to select the favorite viewing angle of the 3D scene [46]. By enabling viewing from different angles, these volumetric videos can be widely used in many areas such as telecommunication, entertainment, healthcare, education, and training, etc.

Face-to-face communications and interactions among people are not always possible or desirable, taken the pandemic of COVID-19 as an example. Therefore, users may wish to interact with each other in a virtual world rather than a physical environment. Holographic-type applications provide a unique opportunity to implement interactive tele-presence or tele-conference where people from different places (cities, nations, continents, etc.) are projected to a hologram where they not only can see/hear but also can interact with each other such as doing handshaking. Another application is tele-training or tele-learning where trainees/students, while staying remotely, still have the ability to virtually interact with their colleagues/friends, trainers/teachers and dynamically hand on ultra-realistic holographic objects. Tele-surgery is also a standout use case of holographic-type applications where specialists can participate in a surgery remotely through some forms of robotic interfaces and make collaborative decisions with the others in the real world that help to reduce the time, cost and enable them to quickly jump in to solve specific problems when the need arises.

However, integrating holographic technologies place significant demands on networking infrastructure considering the high bandwidth and stringent ultra-reliable low-latency communication (URLLC) requirements. As immersive technology advances and the content becomes richer, more complex solutions will be required to support new immersive technologies. Thus, enabling immersive interconnected holographic-type applications will require significant advances towards 5G/6G and heterogeneous networks. We propose that our holographic teleportation system can be developed into a type of Virtual Network Function (VNF) [35,36] and deployed in 5G collaborative MECs [37,38] closer to the users in order to improve the end-to-end system performance and network overhead efficiency while reducing the Capital Expenditure (CAPEX). In the next section, we evaluate the application performance from different scenarios across a range of metrics. Based on these evaluations, we derive useful guidelines and policies for future remote production operations for holographic teleportation through 5G systems.

IV. EVALUATION

A. Experiment setup

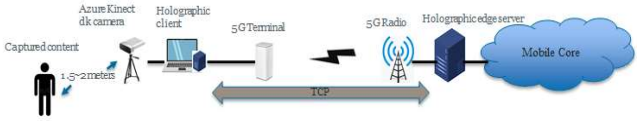


Fig. 5. Setup of holographic application in local 5G test network

Figure 5 depicts the measurement system of holographic application in a local 5G test network. As aforementioned, the deployed holographic server and client are based on an open-source Livescan3d codebase [1]. It is worth noting that in order to maintain its original communication pattern, we focus on the performance features like FPS, server-side perceived throughput per second, and frame delay. The captured content is an adult standing approximately 1.5 to 2 meters in front of the sensor camera, with random gestures and motions. On the display screen at the client-side, facial expression, body motions, and gestures can be correctly identified and restored as a holographic video by the point cloud technique.

The local 5G experiment network is based on the 5G stand-alone (SA) architecture developed at the University of Surrey. Different from "Non-Stand Alone" (NSA) architecture, the 5GNR and its 5G interface are directly connected to the 5G core for both control and user planes, thus can have better performance in terms of capacity, device performance [47]. The receiver's position is well-tuned to have the best throughput, as

the RTT between terminal to edge server ranges from 8.7 ms to 14.5 ms, and the throughput measured by iperf3 [48] varies between 151.2 Mbps to 183.4 Mbps.

In order to record and anatomize the perceived key performance metrics of holographic content in 5G network, we deploy a multi-tiers analysis tool which includes application layer print, packet trace at middlebox (to minimize the application side monitoring pressure), and a system tool 'ss' [49] to export kernel-level TCP parameters in a running TCP connection. Regarding the congestion control algorithm, we deploy an Nginx v17.0 proxy [50] as an NFV between the 5G terminal and sensor, enabling the BBR congestion control algorithm at the edge production server.

B. Evaluation of HD, FHD resolution levels

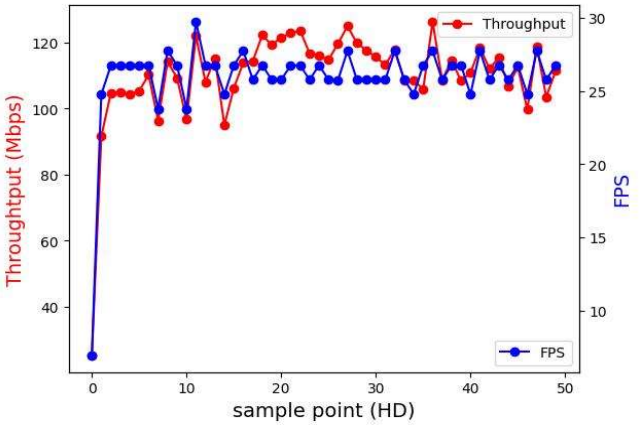


Fig. 6. Throughput and FPS of HD content (50s)

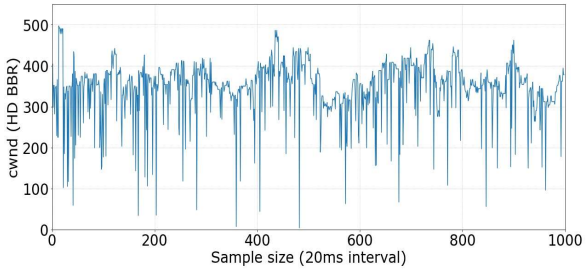


Fig. 7. cwnd print of HD resolution level

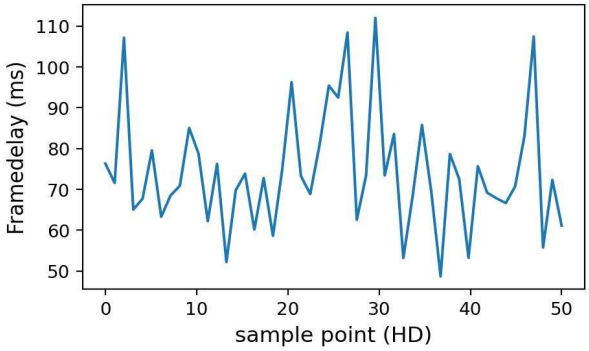


Fig. 8. Frame delay of HD content (50s)

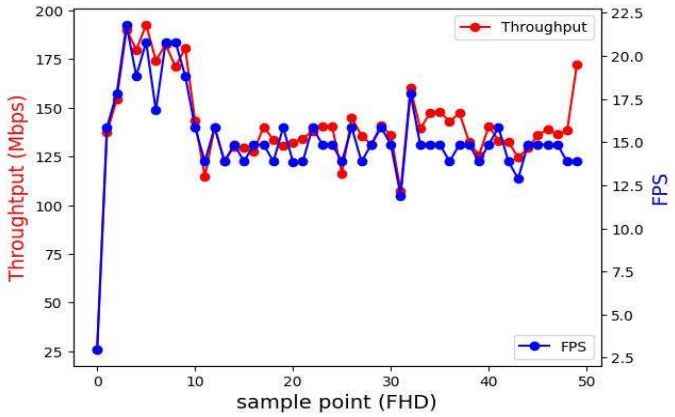


Fig. 9. Throughput and FPS of FHD resolution (50s)

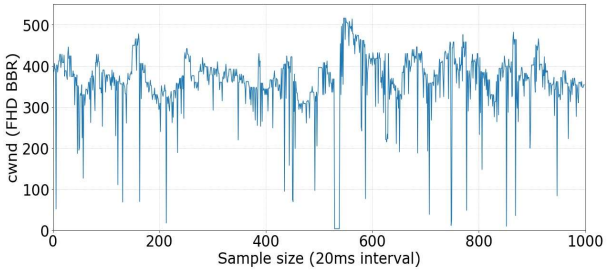


Fig. 10. cwnd print of FHD resolution level

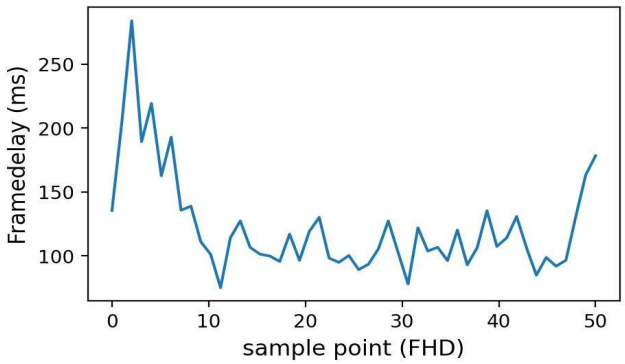


Fig. 11. Frame delay of FHD resolution level

In this sub-section, we present the measurement results of HD and FHD resolution levels in Figs. 6 – 11. First, we present the FPS and throughput of HD and FHD modes, which are two mainstream video resolution levels to allow a user to have satisfactory experiences with the 5G capacity. The enabled congestion control algorithm is BBR. During the 50 s sample period, the HD content can be delivered at above 25 FPS most of the period, with its throughput ranging from 100Mbps to 120 Mbps (see Fig.6). Given the measured throughput by iperf3 at the same location can be 150 Mbps to 180 Mbps, the observed throughput gap implies that streaming HD holographic content through 5G uplink cannot fully utilize the bandwidth due to its low throughput per frame requirement and FPS lock (30 FPS) at the camera side. Additionally, both FPS and throughput present low coefficient of variation (CoV) at around 0.2, indicating that users can perceive high and stable FPS in long-term streaming. This can be attributed to the enhanced state machine of the TCP BBR algorithm. Fig. 7 shows the

corresponding congestion control window (cwnd) sampled every 20 ms during a 20 s recording time. At the transport layer, the observed uncertainty caused by the physical link is that when RTT and bandwidth vary, the estimated min RTT and sending rate calculated by the congestion control algorithm at the sender will become inaccurate, which could jeopardize the application's throughput. However, BBR leverages a comprehensive mechanism to update RTT and pacing rate, as it endeavors to match the latest link condition by adjusting its cwnd per round. This is why the cwnd fluctuates significantly, but the achieved throughput is comparatively stable. Similarly, looking at the FHD holographic content, the FPS stays at 15 FPS, despite a slight surge to 22 FPS at the beginning 10 seconds. Compared to HD resolution, the main limitation of FHD streaming is the network bandwidth. Giving that each frame in FHD will cost 10 to 15 Mbps, it is easy to know that to support a full FPS of FHD the required bandwidth should be doubled.

Another critical metric of holographic content is frame delay. The frame delay recorded at the edge server side is defined as the time difference between the sensor generating an instant frame upon a frame request and edge server fully receiving and rendering this frame. Figure 8 and Figure 11 depict the frame delay of HD and FHD resolution levels during a 50 s sampling period. The resolution level HD has a median frame delay at 72 ms, with its standard deviation presenting at 14.58 ms. Different from FPS and throughput, the frame delay was measured by each frame, then it can capture the network's variation at more fine-grained level (e.g., every 30 ms) therefore, the upper bound value is even two times higher than its lowest value. For FHD level, it also suddenly ramped up to around 300 ms and then kept stable at 100 ms. This is because the measured RTT variation at the transport layer surges to 41 ms due to the 5G link fluctuation and then drops to 16 ms after the beginning 10 seconds (by examining the ss log file). Another key factor that leads to this frame delay variation is the different frame sizes. For instance, in the FHD scenario, the smallest frame size is 612 KB, and the largest one is 973 KB. As a larger frame will require higher transmission and computation time, the downloading time of each frame varies from 74 ms to 124 ms after the beginning 10 seconds.

Moreover, as mentioned before, in this test framework, we adopt explicit request messages from the edge server to the sensor to request each frame. By examining the time-to-first-byte delay in FHD and HD's data set in outdoor scenario, these two resolution levels share a similar median time-to-first-byte delay at around 12.1 ms, which is the measured round-trip time in 5G link. In other words, if the holographic sensor actively and properly pushes each frame, this time-to-first byte delay can be further improved. However, explicit frame request can allow a client to embed real-time information and schedule the frames between different cameras. In this sense, we plan to investigate the server push mode in different network conditions and multiple cameras scenario in our future work.

C. *Effect of different resolution levels*

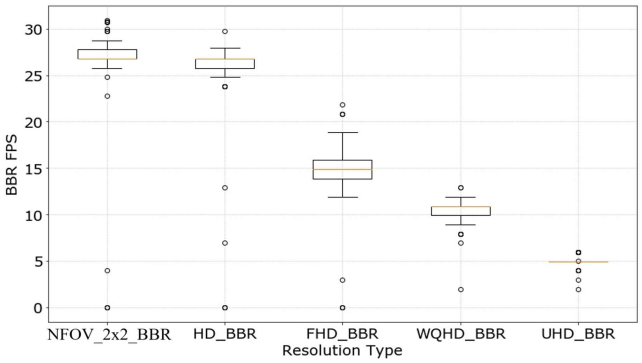


Fig. 12. FPS comparison between different resolution levels

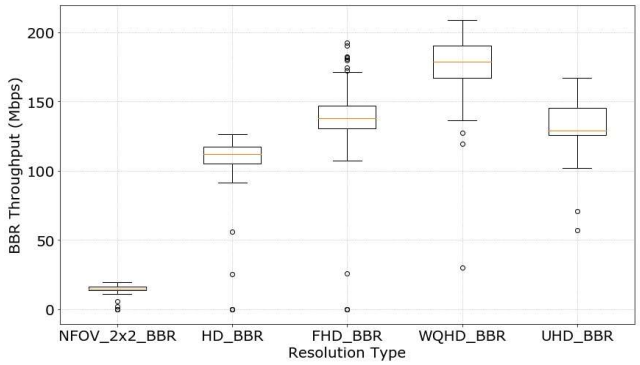


Fig. 13. Throughput comparison between different resolution levels

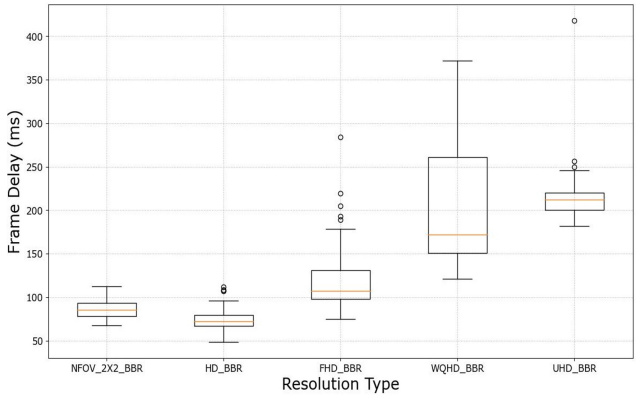


Fig. 14. Frame delay comparison of different resolution levels

Now we compare the performance between different resolution levels (see Fig. 12-14). In general, the results fall into three groups: 1) NFOV_2x2 and HD, due to its low bandwidth requirement, the uplink bandwidth is sufficient to support full FPS for NFOV_2x2 and HD, and their frame delay are low and more robust (e.g., less than 100 ms). This is because the redundant network resources can be utilized once network fluctuation happens (see Fig 14). 2) FHD and WQHD. The achievable FPS performances are limited by the available uplink bandwidth, especially for WQHD its FPS drops to 10 FPS although it can almost saturate the available bandwidth. Additionally, resolution levels whose FPS is limited by the 5G uplink bandwidth are more vulnerable to the network fluctuations, as any bandwidth and RTT rise or drop will be directly embodied in the perceived FPS and throughput, and also incur high frame delay variation. 3) UHD. Its throughput per frame is around 25 MB, which is triple of FHD. Its total FPS

performance is limited by the client-side processing ability, presenting only at 5 and therefore, its throughput is even lower than FHD. Regarding its frame delay performance, it presents the highest value among all resolution levels (e.g., 215 ms on average).

D. Analysis of different congestion control algorithms

Table. 1. FPS and Throughput comparison between BBR and CUBIC

	FPS (BBR-CUBIC)	Throughput (BBR-CUBIC , Mbps)
NFOV_2x2	-1.00	-1.76
HD	9.55	42.41
FHD	6.02	57.77
WQHD	5.75	98.88
UHD	1.89	50.07

We have also evaluated the performance of the congestion control algorithms by tweaking the congestion control algorithm at the sensor side and comparing the performance between CUBIC and BBR. Table 1 shows the FPS and throughput difference between BBR and CUBIC. Obviously, BBR outperforms CUBIC for all resolution levels except for NFOV_2x2, for which these two algorithms have equal performance due to NFOV_2x2's low throughput requirement. For all other resolution levels, CUBIC's FPS falls below 15 FPS, and therefore it is unable to support an acceptable performance of holographic content in 5G uplink.

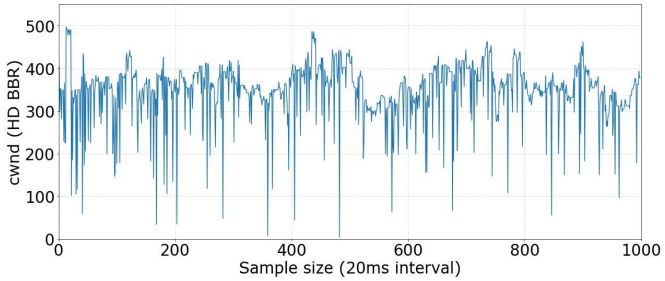


Fig. 15. cwnd print (HD, BBR)

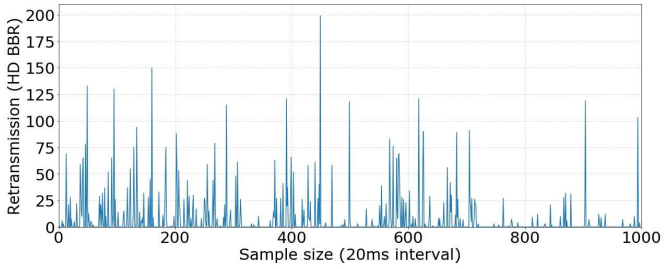


Fig. 16. Number of packet retransmission (HD, BBR)

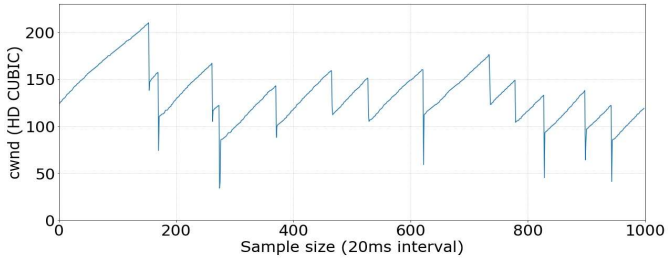


Fig. 17. cwnd print (HD, CUBIC)

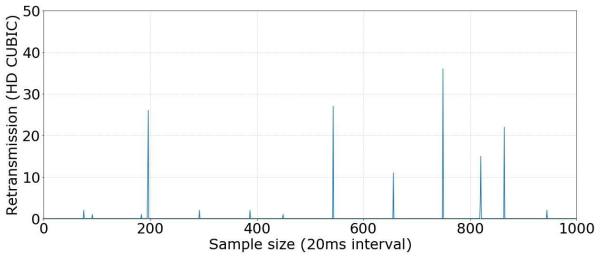


Fig. 18. Number of packet retransmission (HD, BBR)

We present a more detailed comparison of cwnd values in Fig. 15 to 18. Figs. 15 and 16 show a detailed BBR congestion window print in a 20s period. If we assume in this experiment BBR and CUBIC experienced similar network fluctuation on 5G uplink, BBR's cwnd is more oscillating, and in contrast, there are many jagged lines in CUBIC's cwnd pattern. This is because on a time-varying link, BBR's mechanism of the 4-state machine (startup, drain, probe_BW and probe_RTT) can proactively tune its cwnd to match the fluctuating link conditions. For instance, once the link RTT drops, using the same cwnd to send data will lead to buffer bloat or even packet loss. To overcome this issue, BBR leverages the per-round estimation of sending rate and minimal RTT to avoid sending excessive data. Additionally, the RTT from the client to the network edge is around 10 ms, which is approximately one-third of the application's requesting interval (33.3 ms per frame). Therefore, during the transmission time of a single frame, the BBR algorithm can sense and adapt its cwnd approximately in three rounds, to match the potential network fluctuation and maintain the robustness of frame delivery. Consequently, the application layer performance (see Fig. 12-14.) is relatively stable, although the cwnd and packet retransmission present high variation. In contrast, CUBIC relies on an explicit loss as the congestion signal, thus it is unable to timely react to such fluctuation conditions. Moreover, as CUBIC's cwnd will stay at a limited rate once it enters the TCP-friendly region during the congestion avoidance phase, thus the unavoidable and consecutive retransmission events and short RTT on 5G uplink will limit cwnd at a low level for a long period. Looking into the observed retransmission packet number, BBR experienced more retransmissions and bursts than CUBIC. This is because when encountering packet retransmission, BBR's policy is to use its state machine to tame such retransmission while maintaining a reasonable rate, but CUBIC just blindly reduces its cwnd in a multiplicative manner instead of estimating whether link congestion is truly happening.

E. Varying 5G terminal's locations

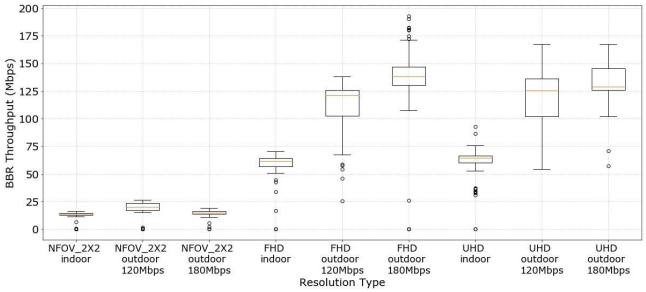


Fig. 19. Throughput comparison of three locations

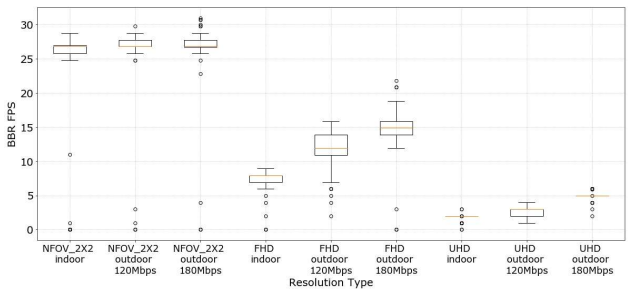


Fig. 20. FPS comparison of three locations

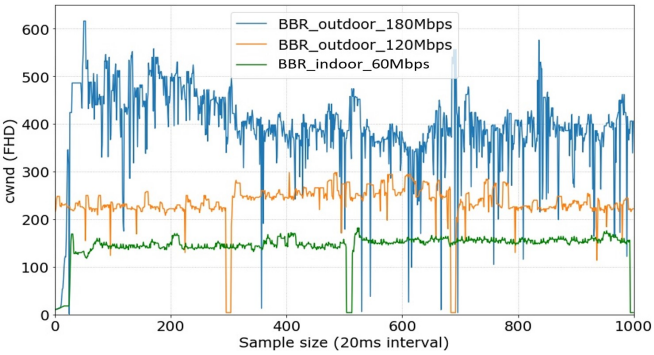


Fig. 21. cwnd comparison in three different locations (FHD, BBR)

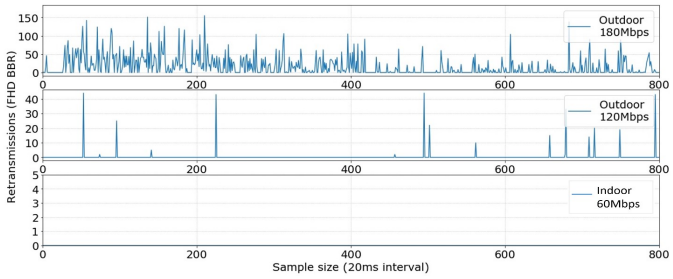


Fig. 22. number of retransmission packets in different locations (FHD, BBR)

After examining the performance of different resolution levels, we turn our focus on the impact of different locations of where the holographic clients are connected to 5G radio. In addition to the best location which has well-tuned line-of-sight UE-panel orientation and no impact from rain and moisture [6], we select alternative outdoor and indoor positions that the client cannot have the ideal radio performance in a 5G cell due to environmental obstruction. In detail, at the alternative outdoor position, the 5G terminal connected to the holographic client is selected to be partially covered behind a tree and turned to a different angle to the radio panel. At the indoor position, the 5G terminal is placed in a meeting room blocked by the building wall. The throughput measured at these two locations are 120 Mbps and 60 Mbps, respectively. Figs. 19 and 20 depict the throughput and FPS comparison between NFOV_2x2, FHD and UHD in three locations. A common observation is that due to its low throughput requirement, NFOV_2x2 can always have stable and satisfying FPS (e.g., above 25 FPS) across different locations, implying configuration of holographic content in low-resolution mode can be well supported in 5G uplink without careful consideration of terminal's position. In contrast, outdoor 180 Mbps is the only position that can provide sufficient bandwidth for FHD to achieve 15 FPS, and apparently, UHD can hardly be streamed in a 5G uplink.

We also print the cwnd and number of retransmission packets of BBR in these three locations (see Fig. 21 and Fig. 22). Surprisingly, unlike the location outdoor 180 Mbps that packet retransmission frequently happened, the FHD video uploading at outdoor 120 Mbps position only experienced several retransmission spikes. Moreover, there was no retransmission observed at the indoor position, and therefore its performance presents better stability than outdoor positions. We attribute this drop of retransmission to its limit bandwidth which cannot incur high jitter on the per-round statistics of the BBR. In other words, taming uncertainty in high bandwidth high fluctuation scenarios will be the key challenge for remote production in the 5G era.

F. Testing with different teleportation objects

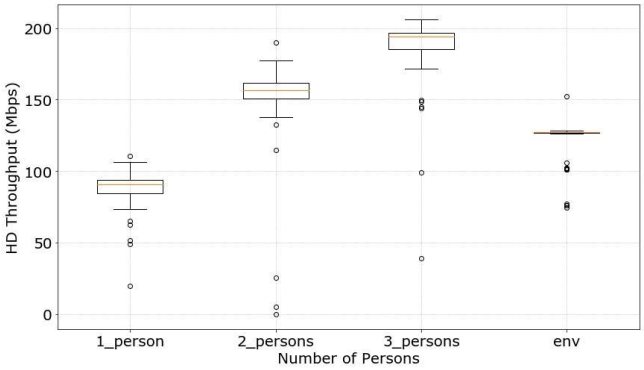


Fig. 23. Throughput when the captured content changes

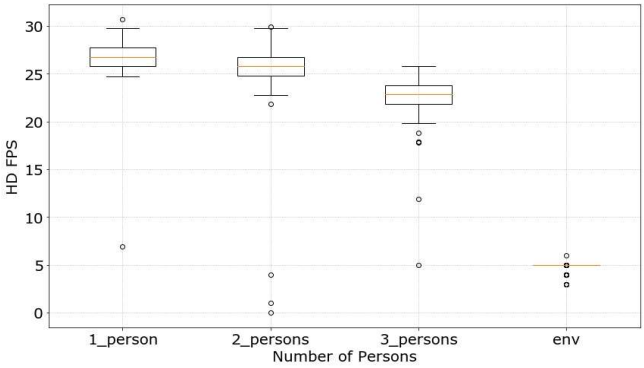


Fig. 24. FPS when the captured content changes

It is also possible to teleport multiple objects at the same location with a single sensor camera, and we now analyze how the number of teleported objects at one location affects the data rate requirements with regard to the 5G uplink capacity for remote production. As the results above show, resolution modes not lower than FHD with single person captured have already fully utilized the uplink radio. We set the resolution level to HD and varied the number of persons to be teleported from 1 to 3. According to its throughput and FPS figures (see Figs. 23 and 24), each person accounted for approximately 60~70 Mbps, and the overall FPS can stay above 22 FPS even when 3 persons are standing together. In contrast, if we change the setting of the sensor camera to whole environment mode (e.g., capturing an office area), its FPS drops to 5 due to limited client-side

processing ability. Regarding the performance robustness, capturing the whole environment shows a much smaller variation than capturing multiple persons. This is because in an interactive event, gesture or motion may lead to body overlapping, thus the number of captured points will dynamically drop accordingly.

F. Testing with different numbers of sensor cameras

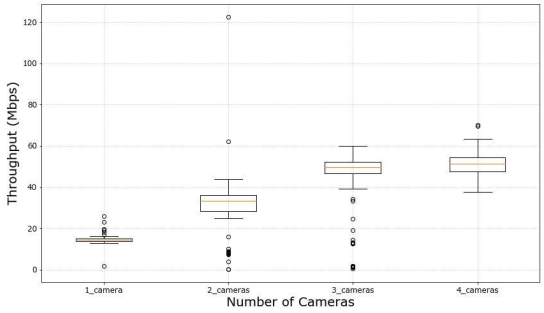


Fig. 25. Throughput with different numbers of cameras

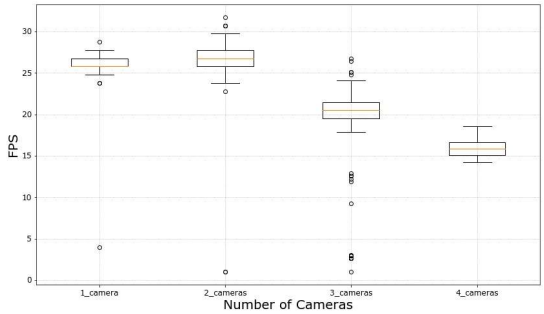


Fig. 26. FPS with different numbers of cameras

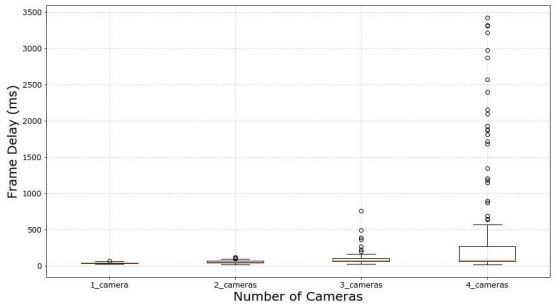


Fig. 27. Frame delay with different numbers of cameras

As an immersive content application allowing 6DoF flexibility of object viewing, different numbers of sensor cameras can be configured on per-site basis. As a result, different levels of immersive experiences can be achieved with specific number of cameras. Given the high bandwidth requirements, we range the number of cameras per site between 1 and 4. Increasing the number of cameras and locating them at different sides of the captured target can generate more raw frames to be streamed to the production server side. By receiving these frames remotely and synchronously, the production server at the edge can process and integrate these frames into an enhanced frame with densified points, mitigating the potential quality degradation caused by unoptimized environment light and capturing angle. We assess its performance as throughput, average FPS (total FPS received divided by the camera number), and the average frame delay. Figs 25, 26 and 27 depict the throughput, FPS and frame delay when the camera number increases from 1 to 4 and fixing the resolution at NFOV_2x2.

It can be observed that each additional camera contributes to approximately 18 Mbps when the additional camera number varies from 1 to 3, but its average FPS drops from 25 to 20 due to resource competition at the 5G uplink. Accordingly, the average frame delay increases from around 42.4 ms to 98.8 ms. Meanwhile, all boxplot pictures at camera number 3 show an increased variance. Given the aggregated throughput (e.g., 50 Mbps see Fig. 25) is still far less from the achievable bandwidth at uplink (e.g., 120 Mbps), we attribute this fluctuation to the contention between multiple connections from different clients sending high FPS concurrently over 5G radio. Additionally, this contention becomes more severe if the four camera work simultaneously, as processing quadruple frames at the same time will lead to a high computation delay at the server side, and that could delay the normal request/respond pattern on part of the connections. Consequently, the throughput in 4 cameras mode does not increase further, the FPS also drops to 16 FPS, and the frame delay may occasionally rise to 266.5 ms (75th percentile). However, compared to other evaluation results like Panoramic video in real-time 4G videotelephony [51] which has 460 ms frame relay, multiple cameras work in NFOV_2x2 mode in 5G uplink can still support a much better frame delay level (e.g., 160.4 ms in 95th percentile, 3 cameras). Furthermore, as applying coordinated multiple connections can increase bandwidth utilization [6] and mitigate the cwnd degradation caused by network fluctuations [52], further improvement in frame delay, FPS is foreseeable if the multiple connections can be well scheduled and managed.

V. CONCLUSION

With the help of the enhanced network capability and evolved mobile edge architecture in 5G, XR-based holographic teleportation becomes a promising media service type. In this paper, a novel 5G MEC framework for supporting remote production operations of live holographic teleportation services is presented and evaluated. By offloading complex media processing task to remote edge server that has powerful processing ability, local 5G client is able to focus on collecting data from camera and leverage the 5G uplink to transfer the raw data to network edge, making the capital investment to be significantly reduced.

Extensive evaluations have been conducted in a real-life 5G test network based on carefully identified scenarios, also with detailed analysis about its achievable throughput, FPS, impact of transport layer protocol, radio condition and the application settings in different scenarios. The results indicate that low resolution levels (e.g., HD) can be well supported with satisfied performance via 5G radio, with a stable FPS and throughput, and low frame delay by the virtue of TCP BRR algorithm. Additionally, by understanding the performance of different

person number, camera number and client's indoor/outdoor location, diversified holographic application scenarios can be designed in the prosperous 5G industry.

VI. ACKNOWLEGEMENTS

This work is funded by UK EPSRC NGCDI project(EP/R004935/1). The authors also acknowledge the support from 5GIC industry members (5GIC) (http://www.surrey.ac.uk/5gic) for this work, and also the colleagues in the HOLOTWIN Project (D01-285/06.10.2020) funded by the Ministry of Education and Science of Bulgaria.

REFERENCES

[1] Kowalski, M.; Naruniec, J.; Daniluk, M.: "LiveScan3D: A Fast and Inexpensive 3D Data Acquisition System for Multiple Kinect v2 Sensors". in 3D Vision (3DV), 2015 International Conference on, Lyon, France, 2015

[2] Huang, Junxian, et al. "An in-depth study of LTE: Effect of network protocol and application behavior on performance." ACM SIGCOMM Computer Communication Review 43.4 (2013): 363-374.

[3] Hyun, Jonghwan, et al. "Measurement and analysis of application-level crowd-sourced lte and lte-a networks." 2016 IEEE NetSoft Conference and Workshops (NetSoft). IEEE, 2016.

[4] Guo, Yihua, et al. "Understanding on-device bufferbloat for cellular upload." Proceedings of the 2016 Internet Measurement Conference. 2016.

[5] Li, Feng, et al. "Who is the King of the Hill? Traffic Analysis over a 4G Network." 2018 IEEE International Conference on Communications (ICC). IEEE, 2018.

[6] Narayanan, Arvind, et al. "A first look at commercial 5G performance on smartphones." Proceedings of The Web Conference 2020.

[7] Xu, Dongzhu, et al. "Understanding operational 5g: A first measurement study on its coverage, performance and energy consumption." Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the applications, technologies, architectures, and protocols for computer communication. 2020.

[8] Liu, Tong, Jiangyu Pan, and Ye Tian. "Detect the Bottleneck of Commercial 5G in China." 2020 IEEE 6th International Conference on Computer and Communications (ICCC). IEEE, 2020.

[9] Gomez-Barquero, David, et al. "IEEE transactions on broadcasting special issue on: 5G for broadband multimedia systems and broadcasting." IEEE Transactions on Broadcasting 65.2 (2019): 351-355.

[10] Gomez-Barquero, David, et al. "IEEE transactions on broadcasting special issue on: Convergence of broadcast and broadband in the 5G era." IEEE Transactions on Broadcasting 66.2 (2020): 383-389.

[11] Mi, De, et al. "Demonstrating immersive media delivery on 5G broadcast and multicast testing networks." IEEE Transactions on Broadcasting 66.2 (2020): 555-570.

[12] Ge, Chang, et al. "QoE-assured live streaming via satellite backhaul in 5G networks." IEEE Transactions on Broadcasting 65.2 (2019): 381-391.

[13] Ha, Sangtae, Injong Rhee, and Lisong Xu. "CUBIC: a new TCP-friendly high-speed TCP variant." ACM SIGOPS operating systems review 42.5 (2008): 64-74.

[14] Neal Cardwell, Yuchung Cheng, C. Stephen Gunn, Soheil Hassas Yeganeh, and Van Jacobson. 2017. BBR: congestion-based congestion control. Commun. ACM 60, 2 (February 2017), 58–66.

[15] Scholz, Dominik, et al. "Towards a deeper understanding of TCP BBR congestion control." 2018 IFIP networking conference (IFIP networking) and workshops. IEEE, 2018.

[16] Atxutegi, E., Liberal, F., Haile, H. K., Grinnemo, K. J., Brunstrom, A., & Arvidsson, A. (2018). On the use of TCP BBR in cellular networks. IEEE Communications Magazine, 56(3), 172-179.

[17] A. Clemm, M.T. Vega, H.K. Ravuri, T. Wauters, F.D. Turck, "Towards truly immersive holographic-type communications: challenges and solutions". IEEE Communications Magazine. 2019.

[18] F. Qian, B. Han, J. Pair, and V. Gopalakrishnan. "Toward Practical Volumetric Video Streaming on Commodity Smartphones". In Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications. ACM, USA, pp. 135–140. 2019.

[19] J. Karafin and B. Bevensee, "25-2: On the Support of Light Field and Holographic Video Display Technology," SID Symp. Dig. Tech. Pap., vol. 49, no. 1, p. 318–21. 2018.

[20] J. V. D. Hooft, M. T. Vega, T. Wauters, H. K. Ravuri, C. Timmerer, H. Hellwagner , F. D. Turck. "Towards 6DoF Virtual Reality Video Streaming: Status and Challenges". IEEE COMSOC. 2019.

[21] P. A. Kara et al., "Evaluation of the Concept of Dynamic Adaptive Streaming of Light Field Video," IEEE TBC, vol. 64, no. 2, pp. 407–21. 2018.

[22] Selinis, I., Wang, N., Da, B., Yu, D., & Tafazolli, R. "On the Internet-scale streaming of holographic-type content with assured user quality of experiences". In 2020 IFIP networking conference (networking) (pp. 136-144). IEEE. 2020.

[23] Anmulwar S., Wang N., Pack A., Huynh V. S. H, Yangy J., Tafazolli R. "Frame Synchronisation for Multi-Source Holograhphic Teleportation Applications - An Edge Computing Based Approach". IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE PIMRC 2021). 2021.

[24] T. Taleb; Z. Nadir; H. Flinck; J. Song. "Extremely Interactive and Low-Latency Services in 5G and Beyond Mobile Systems". IEEE Communications Standards Magazine issue 2, vol. 5, pp. 114-119. 2021.

[25] Favalora, G. E. "Volumetric 3D displays and application infrastructure". Computer 38, 8, 37–44. 2005.

[26] Microsoft Azure Kinect DK. [online] Available at: https://docs.microsoft.com/en-us/azure/Kinect-dk/ [Accessed 20 August 2021].

[27] Kinect v2. Available: https://developer.microsoft.com/en-us/windows/kinect/. Last accessed 20th August 2021.

[28] Intel RealSense. [online] Available at: https://www.intel.co.uk/content/www/uk/en/architecture-and-technology/realsense-overview.html.

[29] H. Qiu, F. Ahmad, F. Bai, M. Gruteser, and R. Govindan. "Augmented Vehicular Reality". In Proceedings of MobiSys, 2018.

[30] A. Maglo, G. Lavoue, F. Dupont, and C. Hudelot. "3D Mesh Compression: Survey", Comparisons, and Emerging Trends. ACM Computing Surveys, 47(3), 2015.

[31] S. Schwarz et al. "Emerging MPEG Standards for Point Cloud Compression," IEEE JETCAS, vol. 9, no. 1, pp. 133–48. 2019.

[32] R. Schnabel and R. Klein. "Octree-based Point-Cloud Compression". In Euro. Symp. on Point-Based Graphics, 2006.

[33] D. Graziosi, O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki, and A. Tabatabai, "An overview of ongoing point cloud compression standardization activities: video-based (V-PCC) and geometry-based (G-PCC)," APSIPA Transactions on Signal and Information Processing, vol. 9, 2020.

[34] Meagher D. "Geometric modeling using octree encoding". Comput. Graph. Image Process., 19, pp. 129–147. 1981.

[35] S. Abdelwahab, B. Hamdaoui, M. Guizani, and T. Znati, "Network function virtualization in 5G," IEEE Commun. Mag., vol. 54, no. 4, pp. 84–91. 2016.

[36] S. Sun, M. Kadoch, L. Gong, and B. Rong, "Integrating network function virtualization with SDR and SDN for 4G/5G networks," IEEE Netw., vol. 29, no. 3, pp. 54–59. 2015.

[37] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing—A key technology towards 5G," White Paper, ETSI. 2015.

[38] T. X. Tran, A. Hajisami, P. Pandey, and D. Pompili, "Collaborative mobile edge computing in 5G networks: New paradigms, scenarios, and challenges," IEEE Commun. Mag., vol. 55, no. 4, pp. 54–61. 2017

[39] P. J. Besl and N. D. McKay. "A method for registration of 3-d shapes." IEEE Trans. Pattern Anal. Mach. Intell., 14(2):239– 256, Feb. 1992

[40] Elmqaddem N. "Augmented reality and virtual reality in education. Myth or reality?". Int J Emerging Technol Learn 14. pp.234–242. 2019

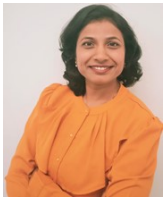
[41] Hu, M., Luo, X., Chen, J., Lee, Y. C., Zhou, Y., and Wu, D. "Virtual reality: A survey of enabling technologies and its applications in IoT". Journal of Network and Computer Applications, 102970. 2021.

[42] Y.-G. Kim and W.-J. Kim, ‘‘Implementation of augmented reality system for smartphone advertisements,’’ Int. J. Multimedia Ubiquitous Eng., vol. 9, no. 2, pp. 385–392, 2014.
[43] A. Langley, A. Riddoch, A. Wilk, A. Vicente, C. Krasic, D. Zhang, F. Yang, F. Kouranov, I. Swett, J. Iyengar, et al. “The QUIC Transport Protocol: Design and Internet-Scale Deployment”. In Proceedings of the Conference of the ACM Special Interest Group on Data Communication. ACM, 183–196. 2017.
[44] ZSTD Compression Library. [online] Available at: https://facebook.github.io/zstd/ [Accessed 20 August 2021].
[45] D. Mills, “Internet time synchronization: The network time protocol,” IEEE Trans. Commun., vol. 39, pp. 1482-1493, Oct. 1991
[46] S. Subramanyam, J. Li, I. Viola, and P. Cesar, “Comparing the Quality of Highly Realistic Digital Humans in 3DoF and 6DoF: A Volumetric Video Case Study,” in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 127–136. 2020.
[47] Liu, Guangyi, et al. "5G deployment: Standalone vs. Non-standalone from the operator perspective." IEEE Communications Magazine 58.11 (2020): 83-89.
[48] "iperf3", [online] Available at: https://iperf.fr/iperf-download.php
[49] "ss tool". [online]Available at: https://man7.org/linux/man-pages/man8/ss.8.html
[50] "Nginx". [online] Available at: https://www.nginx.com/
[51] Xie, Xiufeng, and Xinyu Zhang. "Poi360: Panoramic mobile video telephony over lte cellular networks." Proceedings of the 13th International Conference on emerging Networking EXperiments and Technologies. 2017.
[52] Qian, Peng, Ning Wang, and Rahim Tafazolli. "Achieving robust mobile web content delivery performance based on multiple coordinated QUIC connections." IEEE Access 6 (2018): 11313-11328.
[53] M. Radenkovic and V. S. H. Huynh, “Cognitive caching at the edges for mobile social community networks: A multi-agent deep reinforcement learning approach,” IEEE Access, vol. 8, pp. 179561–179574, 2020.
[54] V. S. H. Huynh and M. Radenkovic. “Leveraging Social Behaviour of Users Mobility and Interests for Improving QoS and Energy Efficiency of Content Services in Mobile Community Edge-clouds.”. CLOSER, p. 382-389. 2020.

PENG QIAN received the B.Eng. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2008, and the M.Sc. degree (Hons.) in 2013 and PhD degree in 2018 from the University of Surrey. He is currently a Research Fellow in the Institute for Communication Systems, Home of 5G Innovation Centre (5GIC) and 6G Innovation Centre (6GIC), University of Surrey. He was involved in the field of the telecommunication industry from 2008 to 2012. His research interests include web and multimedia content acceleration, mobile edge computing, and next-generation Internet and routing protocols.

VU SAN HA HUYNH received the B.Sc. (Hons) degree in computer science from the University of Nottingham in 2016 and the PhD degree from the University of Nottingham in 2021. He has authored in venues including IEEE Access, ACM MobiCom, CLOSER, PECCS, IEEE WiMob, IEEE FMEC and IEEE IWCMC. His research interests include intelligence in the network transport layer, mobile edge cloud and fog computing, distributed data caching services, mobile heterogeneous opportunistic networks, large-scale internet service architectures, Internet of Things and network science.

Ning Wang (SM'17) obtained his PhD degree in Electronic Engineering from the Centre for Communication Systems Research at University of Surrey where he has been working as a professor. Professor Wang is currently leading a research team focusing on 5G and beyond networking and applications. He has published over 150 research papers at prestigious international conferences and journals. His main research areas include future networks, multimedia networking, network management and control and QoS/QoE assurances.



Sweta Anmulwar Sweta Anmulwar is a Research SDN Software Developer and a PhD researcher working on Frame Synchronisation for Multi-source Holographic Teleportation Applications at University of Surrey, UK. Sweta has an experience in Computer Networking, Software development and Wireless communication. She is a key member in design and development of various software development projects like SDN Online Lab, Smart city, Traffic Generators. She has conducted various tutorials on SDN at IEEE conferences and faculty development programme on “SDN and OpenFlow”.

De Mi (M’17) received the B.Eng. degree from the Bei jing Institute of Technology, China, the M.Sc. degree from the Imperial College London and the Ph.D. degree from the University of Surrey, U.K.. He is currently with the Institute for Communication Systems, Home of 5G Innovation Centre (5GIC) and 6G Innovation Centre (6GIC), University of Surrey. He is also a U.K. Engineering and Physical Sciences Research Council (EPSRC) Ph.D. Academic Supervisor, and a Visiting Scholar with the Shanghai Jiao Tong University, China. His research interests include B5G/6G radio access techniques, radio access network architecture and protocols, hypercomplex signal processing, and the next generation broadcast and multicast communications.

Regius Rahim Tafazolli, Fellow of the Royal Academy of Engineering, IET, WWRF and Regius Professor of Electronic Engineering; Professor of Mobile and Satellite Communications; Founder and Director of 5GIC, 6GIC and ICS (Institute for Communication Systems) at the University of Surrey. He has over 30 years of experience in digital communications research and teaching. He has authored and co-authored more than 1000 research publications and is regularly invited to deliver keynote talks and distinguished lectures to international conferences and workshops. He was advisor to the Mayor of London (Boris Johnson) on London 2050 Infrastructure.