Environment-driven mmWave Beamforming for Multi-user Immersive Applications

Ahmad Kamari, Nan Wu, Hemant Kumar Narsani, Bo Han, and Parth Pathak
George Mason University
{akamari,nwu5,hkumar4,bohan,phpathak}@gmu.edu

Abstract

This position paper explores the challenges and opportunities for high-quality immersive volumetric video streaming for multiple users over millimeter-wave (mmWave) WLANs. While most of the previous work has focused on single-user streaming, there is a growing need for multi-user immersive applications such as virtual collaboration, classroom education, teleconferencing, etc. While mmWave wireless links can provide multi-gigabit per second data rates, they suffer from blockages and high beamforming overhead. This paper investigates an environment-driven approach to address the challenges. It presents a comprehensive research agenda that includes developing a collaborative 3D scene reconstruction process, material identification, ray tracing, blockage mitigation, and cross-layer multi-user video rate adaptation. Our preliminary results show the feasibility and identify the limitations of existing solutions. Finally, we discuss the open challenges of implementing a practical system based on the proposed research agenda.

ACM Reference Format:
Ahmad Kamari, Nan Wu, Hemant Kumar Narsani, Bo Han, and Parth Pathak. 2023. Environment-driven mmWave Beamforming for Multi-user Immersive Applications. In *The 1st ACM Workshop on Mobile Immersive Computing, Networking, and Systems (ImmerCom '23), October 6, 2023, Madrid, Spain.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3615452.3617945

1 Introduction

3D scene representations by point cloud or polygon mesh allow users to experience the volumetric video content from any arbitrary viewpoint [57]. This ability enhances the interactive experience from three degrees of freedom (3DoF) provided by 360°videos, which allows viewers to change only the viewing direction, to 6DoF which also enables users to move around the scene. Numerous applications can be enabled through such an immersive experience including telepresence, entertainment, education, and healthcare [15].

Despite the promising aspects of volumetric videos, various challenges exist in realizing their potential in practice. Streaming volumetric videos, especially with high quality, requires high bandwidth which makes it challenging for even a single user [59]. While considering limited bandwidth, prior works try to stream according to only the viewport. For example, ViVo [15], by taking into account the viewpoint, the distance to displayed content, and their occlusion, reduces the effective bandwidth by an average of 40%, which results in a 100 to 200 Mbps data rate for a single user. GROOT [23] accelerates point-cloud operations with GPU-assisted compression to reduce the computation overhead which still requires 100 to 500 Mbps bandwidth for a single user. These works consider only single-user experience which is not the case for many of the applications of volumetric video streaming. In multi-user scenarios, these challenges are more severe, and meeting the bandwidth requirement with the current technologies is extremely challenging.

mmWave WLANs (wireless local area networks), offering multi-gigabit per second data rates, are the key to satisfying the bandwidth requirement of volumetric video streaming to multiple users [59]. However, mmWave performance scales poorly with an increasing number of users because of the high overhead of narrow beam management. The default WLAN 802.11ad protocols sequentially probe the beams for users. On the other hand, by increasing the number of users the probability of inter-user blockage also increases, which forces the process of beam searching to repeat. While an abundance of previous studies proposed different techniques [4, 18, 26, 33] to reduce beam searching overhead, these methods are not suitable for highly dynamic multi-user environments. To address this problem, Zhang et al. [59] exploit 6DoF motion prediction for multiple users to reduce the impact of blockages and the beam searching overhead. SpaceBeam [52] uses LiDAR scanning of the environment to create a reflection profile of the environment to eliminate the beam searching overhead. Although the idea of modeling the reflection environment is promising, using it in practice and particularly implementing it on AR/VR devices with dynamic multi-user movements is a challenging task.

ImmerCom '23 , October 6, 2023, Madrid, Spain

ACM ISBN 979-8-4007-0339-3/23/10.
https://doi.org/10.1145/3615452.3617945

In this paper, we focus on identifying different aspects of designing multi-user volumetric video streaming systems over mmWave WLANs. We specifically focus on environment modeling to reduce the beamforming overhead and mitigate blockages. We describe the related challenges and our contributions below.

• In multi-user scenarios, the main obstacle of a practical mmWave communication is the beam searching overhead. AP's complete knowledge of the environment by modeling different reflection surfaces can reduce the overhead. However, the task of reflection environment reconstruction incorporates multiple challenging steps: (1) How can we collect the RGB-D data in a multi-user scenario considering the dynamic nature of these environments? (2) How can we perform the 3D reconstruction by combining the collected RGB-D data into a single mesh in an efficient yet accurate manner? and (3) How can we identify the material properties of each surface robustly in order to estimate the quality of mmWave reflections?

We first demonstrate the overhead for data collection with different qualities for RGB images. Our results show that single-user data collection is not efficient and can take a long time according to the quality of RGB images. To address the first challenge, we propose a collaborative data collection that can reduce the time significantly and can be done in real-time which can result in an up-to-date 3D map. For the second challenge, we present the outcome of different 3D reconstruction methods on path detection accuracy. Our results show that even the recent methods of reconstruction do not provide the required accuracy (millimeter level) which needs further improvement. For the third challenge, We compare the results of two different large-scale material identification datasets and show that their results are far from being directly useful in a practical system. Therefore, we suggest different potential improvements such as combining the current datasets, using other headset sensors (such as IR), and/or incorporating the mmWave radar in the material identification process.

• The effectiveness of the reflection model of the environment relies on how fast we can get the paths between the TX and RX from the ray tracer. We find that the current ray-tracing methods are not suitable for real-time usage, and even with high computational overhead, they cannot achieve sub-millisecond speeds (necessary for beamforming). We propose different acceleration methods such as beam and frustum traversal with kd-trees, bounding volume hierarchies and grids, and implementing a fast neural-based ray tracer.

• The ability to predict the paths before starting the communication allows us to detect and even predict the blockages and find different alternative paths between the AP and the clients. Our goal is to use this knowledge to mitigate blockages. Also, the same approach can be applied to deal with data rate fluctuations by finding robust paths between AP and the clients. We propose a cross-layer rate adaptation approach. The 3D material map predictions can be used to set the beams and mitigate the blockage at the physical layer and set the buffer size and adapt the video quality at the application layer according to the quality of the link.

Our ongoing work includes realizing and further optimizing the above ideas and integrating them into a holistic system that can stream high-quality volumetric videos to multiple users over mmWave WLANs.

2 Research Agenda

2.1 Collaborative 3D Scene Reconstruction

AR/VR headsets with various sensors provide the necessary data to accomplish the task of capturing the 3D geometric structure of the surroundings by means of depth sensing. This 3D map along with the 6DoF motion can be leveraged to reduce the overhead of beamforming for mmWave communication. Prior work such as SpaceBeam [52] has shown the use of RGB-D cameras to perform reconstruction of the 3D model of the environment and use it for mmWave beamforming. However, the system operates in an offline manner where a single user is required to scan and reconstruct the environment in advance. Furthermore, it does not operate in real-time and cannot directly account for transient changes (e.g., people moving around) without frequent re-scanning.

When multiple users are consuming the volumetric content, it is possible to let all users participate in the reconstruction process. In these scenarios, all users send their sensor information (e.g., RGB images, depth information, and 6DoF motion) to the AP, and AP combines the information to create a holistic 3D map of the environment. Naive collaborative offline mapping can improve the scanning duration compared to only one user [11, 13]. However, such a process still happens offline and cannot work in a dynamically changing environment. In this research, our goal is to create a real-time framework for collecting and processing sensor data and reconstructing the environment at the AP.

Developing such a framework requires us to address multiple challenges. First, the effect of transient changes is more çç, and depending on the user mobility and the number of users the model can be incomplete. To handle transient changes, one approach is to detect and remove transient objects (e.g., humans) from the RGB-D images and consider only the static environment in the reconstruction process. Model incompleteness comes from the fact that users' FoV might not cover the entire environment because they are focused on the volumetric video content or users occlude the view of each other. To tackle this issue ML approaches can be used to fill the missing pieces of the model [10]. Second, joining the sub-scans of each user to create the full map is not a trivial task, and the AP needs to accurately determine the transformations between each user's scan. There are some recent works [5] for integration of different pieces of the model.

2.2 Material Identification

Material reflection properties is the main criteria for distinguishing different reflection surfaces. Even after having the 3D geometric model of the environment, reflection loss determines where the beams should be guided. There are many different previous studies that try to measure the reflection loss of common materials for both indoor [45, 51] and outdoor [21, 22, 42] environments in mmWave band. The results show a varying behavior for the same material and there are different values for reflection losses. To understand the reflection properties of a surface, we should follow the behavior of a radio signal when the signal reaches it. Upon reaching the surface, the amount of reflected energy depends on the material, its roughness and thickness, the polarisation, and the incident angle. The Fresnel coefficients for reflection and transmission were proposed to describe the effect of polarisation and angle. The Fresnel reflection factor is for smooth surfaces, but in the real world, perfectly smooth surfaces are rare. Therefore, to take the surface roughness into consideration, there is another loss that is calculated using the following equation [52]

$$r/r_y = exp(0.5(\frac{4\pi\Delta_h cos(\theta)}{\lambda})^2) \tag{1}$$

where θ is the incident angle, λ is the wavelength, r_y is the Fresnel reflection coefficient, and Δ_h is the standard deviation of the surface roughness.

Given the variations in the material structure, determining the components of this equation is extremely challenging. Therefore, researchers try different sensors to infer the loss. In this section, we discuss different methods that can be used based to the availability of the sensors on AR/VR headsets:

(1) IR-based material identification. In addition to RGB images, IR sensors available on the headsets can provide extra information about the material properties (e.g., scattering properties) to achieve higher accuracy. For example, a multimodal material segmentation model has been designed that shows the non-RGB imaging modalities including near-infrared images can help in better discriminating different material categories [25]. Also, SpaceBeam [52] uses an IR-based approach to infer the reflection loss by utilizing IR intensity over the surface. Their approach is to estimate the surface roughness and correlate it with the reflection loss. This approach is essentially a material property identifier, rather than material identification, and it can be utilized along with other methods.

(2) RGB images material segmentation. Recent advances in image identification and segmentation and the availability of large-scale datasets such as MINC [7] and DMS [50] make it possible to coarsely predict the material properties of a surface. One important issue with the available datasets is the fact that they are not designed to be leveraged for RF signal propagation. However, with modifications, these approaches can be adapted for the purpose. Material segmentation using RGB images, which can be easily collected using headsets, can give fine-grained classification of surface material in the environments.

(3) mmWave Radar. General category of the material of the surface (e.g., brick, glass, etc.) along with the roughness (Equation 1) can give us an accurate prediction of the RF signal in any frequencies. However, material classification and roughness prediction methods are not accurate [7, 50, 52] and we need live feedback from the same frequency that we are using for the communication to characterize the materials more accurately. For mmWave signals, mmWave radars can provide an in-band response from the environment to be used independently or along with the other sensors' information to improve the material identification accuracy. Although, equipping the headsets with these sensors can be challenging, our research is dedicated to conquering these challenges.

2.3 Robust Real-time Ray-tracing

Scene reconstruction and material identification can only be useful if AP can detect paths to the clients correctly and in real time. There are two approaches to ray-tracing that we plan to explore.

(1) Standard Ray-tracing. Scene reconstruction might not always accurately match the real world and the generated mesh can include noise because of the limited accuracy (±1 cm) that should be considered when AP tries to trace the paths. Signal propagation paths can be determined using standard ray-tracing methods such as shoot-and-bounce ray-tracing [27] which do not expect a 3D model with noisy surfaces. To tackle these issues, SpaceBeam [52] suggested modifications such as path clustering and spurious path pruning. On the other hand, real-time ray-tracing for mmWave becomes infeasible as the complexity of the 3D model increases. Also, it is crucial to have a ray-tracer that operates faster than the actual beamforming time in order for it to be useful. The ray-tracer mentioned in SpaceBeam takes 1.4s when using CPU and it can go down to 100ms on GPU which does not offer any advantage to the mmWave communication specifically in multi-user environments (which is why a precomputed lookup table is used in the SpaceBeam). In order to accelerate this process, the methods in optical ray-tracers such as beam and frustum traversal with the kd-trees and bounding volume hierarchies and grids can be used to reduce the number of rays.

(2) Representing Volumetric Radio Frequency Scene with Neural Networks. Another approach is the adaptation of related work in the optical field such as NeRF [35]

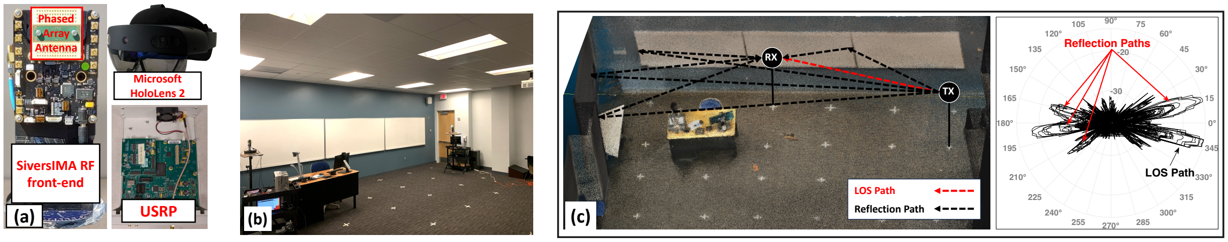

Figure 1: (a) mmWave Devices and HoloLen 2 used in the measurement, (b) measurement environment, (c) an example of collected power angular profile.

which uses neural networks to generate novel views of complex 3D scenes, based on a limited set of 2D images. Inspired by this method, one of the recent works [61] designed a model which can determine the received signal at any position with a modest number of signal measurements. The main challenge in this work is the number of measurements and the duration it takes to train the model for each environment. Our goal is to directly use the RGB images to reduce the number of measurements and use recent advancements such as Instant-NGP [36] to improve the training time.

2.4 Blockage Mitigation and Cross-Layer Rate Adaptation

Blockage mitigation and rate adaptation are critical aspects in the context of multi-user volumetric video streaming over mmWave WLANs. Blockage has a pronounced effect on mmWave links, and it can reduce the data rate significantly or even cause a complete outage (especially human blockage [47]). To overcome this challenge, effective blockage mitigation techniques need to be employed. In these situations, two approaches can be employed to mitigate blockages. The first approach is to use Multi-AP deployment to mitigate the blockages. With the dense deployment of multiple APs, it is possible to perform fast handover between the APs and reduce the chances of blockage through spatial diversity. In this scenario, even if one AP is blocked, there might be a LOS path to another AP, and APs can do dynamic handover between themselves to overcome the blockage. The second approach is to leverage the RGB-D images along with pose information to detect the blockages. Because the main blockages come from inter-user occlusions, a machine-learning model can be trained to detect the persons in the image and AP can steer the beams around it, ensuring continuous connectivity and reducing the impact of signal attenuation.

Additionally, rate adaptation plays a key role in optimizing the QoE. By exploiting the 3D material map to coarsely predict the mmWave channel conditions and adapting the transmission rate accordingly, we can ensure an optimal trade-off between video quality and the available mmWave paths. This adaptive approach allows for seamless streaming by dynamically adjusting the buffer size of the video player, resolution, and compression level to match the varying mmWave paths

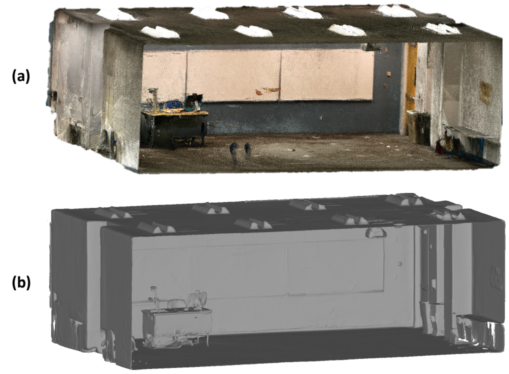

Figure 2: (a) Point cloud generated from the RGB-D images and (b) 3D reconstruction example.

and blockage scenarios, thereby delivering an immersive and uninterrupted volumetric video streaming experience over mmWave WLANs.

3 Preliminary Results

In order to identify the challenges in designing volumetric video streaming systems over mmWave WLAN, we perform an experiment using different existing systems. Our focus is on environment modeling to reduce the beamforming overhead and mitigate blockages in mmWave WLANs.

Experimental Setup. We collect data in a university classroom (9.5 m $\times$ 6 m) as shown in Fig 1b. We use Microsoft HoloLens 2 to scan the environment by collecting RGB-D images with different qualities for 3D reconstruction. For 3D environment evaluation including the power angular profile, reflection path assessment and loss prediction, we use a pair of mmWave software radio systems. Each of the mmWave SDRs consists of a phased-array based RF frontend from Sivers Semiconductors [2] and USRP as the baseband processor. The Sivers 60 GHz RF frontend has two phased antenna arrays (one for sending and the other for receiving) with 64 antenna elements each. We use the default codebook from Sivers with 63 beams where the main lobe angle of every beam is approximately 1.5° in Azimuth. We perform a separate measurement campaign to measure the beam patterns of each beam in the codebook. Fig 1a shows the devices used in the experiment. We collect the mmWave channel data (AoA and RSSI for different paths) at 10 TX-RX location pairs. An example of the power angular profile is shown in

Exp.	Capture Time (s)	# of Depth Frames	RGB Resolution	RGB (FPS)	Depth (FPS)	Reconstruction Time (s)
1	362	362	1920×1080	10	1	152
2	528	2640	1280×720	20	5	995

Table 1: Overhead of environment scanning.

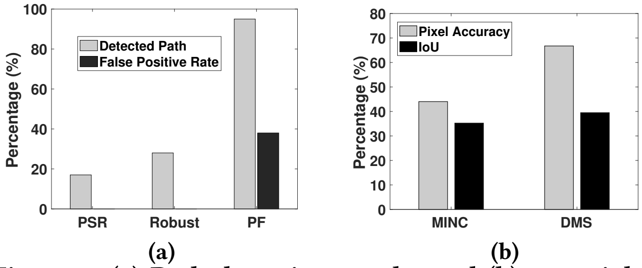

Figure 3: (a) Path detection results and (b) material identification results.

Fig. 1c. To get the precise location of the TX and RX, we use the provided 6DoF pose by the headset as the ground truth.

Environment Scanning Overhead. The process of 3D reconstruction of environment's material map depends on two important factors: (1) the quality of the RGB images which is necessary for the material identification, and (2) the frequency of capturing the depth images which is important in 3D reconstruction. In order to evaluate the overhead of capturing the spatial detail of the environment using HoloLens 2, we perform two data collection experiments. Details of both experiments are listed in Table 1. In the first experiment, we collect the RGB images with high resolution (1920×1080) which can be useful in material identification but it leads to lower RGB and depth capturing frequency (10 FPS for RGB and 1 FPS for depth images). In the second experiment, we set the RGB resolution to medium quality (1280×720) which allows us to increase the capturing frequency to 20 FPS for RGB and 5 FPS for depth images.

The trade-off between the quality of the RGB images and capturing frequency of depth images can cause a large difference in the scanning time and the reconstruction process (an example of a reconstructed scene is shown in Fig 2b). We use the Robust reconstruction method (described below). Having fewer RGB-D images decreases the scanning and reconstruction time (as shown in the table), but it results in missing spatial details which is necessary for mmWave path detection. On the other hand, having medium-quality images can provide enough detail in both material detection and 3D reconstruction.

3D Reconstruction and Path Detection. In this section, we evaluate the effect of 3D reconstruction on path detection accuracy without considering material information. We use the medium quality RGB-D information to perform the 3D reconstruction using three different methods.

(1) PF: The first method is called plane fitting (PF) where after converting the RGB-D data to point cloud (as shown in Fig 2a), we run the RANSAC algorithm [12] to find groups of points that can create a plane. Then, we fit a plane by finding the normal vector and the boundaries. This method is suitable for simpler environments with mostly flat surfaces and can remove small noise in the surface reconstruction.

(2) PSR: The second method uses screened Poisson Surface reconstruction (PSR) [19] to reconstruct mesh from point clouds. This is a classical geometry-based method that results in a complete model albeit while sacrificing accuracy.

(3) Robust: The last model, referred to as Robust, is a high-quality fusion method that uses accurate geometric registration to deal with the accumulated pose estimation errors [9]. This model starts with reconstructing locally smooth scene fragments and deforming these fragments to align them with each other, obtaining high-quality 3D scene models offline.

The resulting 3D models from these 3 methods and the TX and RX locations from the measurement points are then input into the Remcom Wireless InSite [1] channel simulator. Finally, the propagation estimation from the simulations is compared to the measurement results (an example is shown in Fig 1c). Here, we compare the number of detected paths and the false positive for each of the methods. Fig 3a shows the results. Since the experiment environment has mostly flat surfaces as expected by the simulator, the PF method outperforms the other methods, but it also increases the false positive rate significantly. Robust and PSR methods can only detect 28% and 17% of the paths, respectively. This comes from the fact that the model noise causes to have surfaces with high face normal error. The Robust model still performs better than the PSR method which shows the improvement in the reconstruction. Further improvement can be tackling the error in surface normals by considering angle error or taking advantage of the simplicity of the plane fitting in removing face noises.

Material Identification. As mentioned in Sec. 2, there are multiple large-scale datasets for material identification. However, using them in practice in the context of signal propagation can be challenging. In this section, we use pre-trained models on MINC [7] and DMS [50] datasets and evaluate the material identification on 25 RGB images from the collected medium-quality data by the headset. Our goal is to see the performance of existing models in identifying common materials in an indoor classroom environment.

We carefully annotate the RGB images according to the categories in MINC and DMS datasets separately. We run both models and compare their accuracy based on two metrics. Pixel accuracy is the number of correctly identified pixels, and intersection over union (IoU) is the overlap of the ground truth and prediction region over the union of them. The results are shown in Fig. 3b. Results indicate that the model trained on DMS with pixel accuracy=66.74% and IoU=39.53% outperforms the MINC with pixel accuracy=44.03% and IoU=35.26% in both metrics. Despite the fact that the DMS dataset was collected over fewer images, the model is trained on polygon labels which is different from the MINC which is trained on points. This aspect causes to have optimized boundaries and higher accuracy because of less noise. Another key aspect that makes DMS better is the diversity of material categories which is important in signal propagation. For example, because the MINC dataset does not include whiteboard and ceiling tile categories, they are usually detected as painted and tile categories that have different signal propagation properties. On the other hand, in DMS the accuracy in some of the categories is not high and can cause misidentifying some categories of materials. For example, our results show that DMS has difficulty in identifying whiteboards and plastic. This can happen because of two reasons: (1) the number of samples and their scene diversity (different lighting, angle, etc.) for some of the categories, and (2) the similarity between materials which is one of the most important difficulties in material segmentation tasks.

In general, although these two datasets are the largest material collections, they cannot be directly used for the task of material identification in the context of signal propagation. Improving the segmentation models, combining both datasets and modifying the categories, and adding more samples while making the categories more specific to RF propagation (as opposed to visual classification) are important outstanding problems.

4 Related Work

Multi-user AR, VR, and 360° Video Streaming. Recent studies investigated the support of multiple users in the fields of AR [31, 43, 60], VR [24, 28, 34], and 360° video streaming [6, 8, 41]. For example, SPAR [43] targets minimizing the spatial inconsistencies of visual content and reducing initialization latencies, by taking the positions of virtual objects into consideration. Coterie [34] reduces bandwidth demands by leveraging the consistency in background content across consecutive frames. Bao et al. [6] and M5 [59] leverage multicast to deliver shared content across multiple users. Our work leverages mmWave networks for volumetric video streaming, which are highly susceptible to inter-user blockages. To address this, we propose a novel approach that involves collaborative 3D scene reconstruction among multiple users. By detecting multiple paths to the AP with the 3D scenes and the pose of headsets, our approach can effectively mitigate blockage effects.

Volumetric Video Streaming. Existing work [15, 23, 30, 57, 59] centered on addressing the computation and bandwidth-intensive nature of volumetric video streaming. For example, early work reduced the computation overhead and bandwidth requirement of mobile volumetric video streaming by leveraging visibility-aware optimizations (e.g., ViVo [15]) and accelerating point-cloud decompression with GPU-assisted compression scheme (e.g., GROOT [23]). Recent efforts include Vues [30] that improves the quality of experience (QoE) by transcoding a point cloud frame into multiple 2D images, YuZu [57] that enhances the volumetric video streaming with super-resolution, and M5 [59] that performs 6DoF motion prediction for adapting mmWave beams and dynamically prefetches content to mitigate the impact of blockage for multi-user streaming. MetaStream [14] expands these video-on-demand works to a practical live volumetric content capture, creation, delivery, and rendering system. Different from these works, we focus on supporting the volumetric video streaming with mmWave networks on AR/VR headsets and propose a research agenda that mitigates the blockage effects for providing a stable and high QoE during video streaming.

60 GHz WLANs. The performance evaluation of mmWave channels is carried out in 5G communications and other networks [44, 46, 48, 58, 63]. The blockage loss [29], avoidance [32], prediction, and mitigation [3, 20, 39, 53, 55] have been thoroughly studied. Several mmWave network models for mobility management [16, 17, 37, 49], interference cancellation, reduction and avoidance [38, 40, 54], and beam selection and management [18, 26, 56, 62] have been designed. However, the aforementioned solutions reflect only on physical layer knowledge for improving link resiliency and do not leverage the information that is available through AR/VR headsets that includes RGB and depth information.

5 Conclusion

This position paper identifies several key areas that pose ongoing challenges for the development of practical mmWave communication for multi-user volumetric video streaming. We propose a holistic research agenda and identify the existing challenges by conducting preliminary experiments utilizing a mmWave testbed and Microsoft HoloLens 2. We show that while collaborative 3D scene reconstruction holds the potential to create environment-driven mmWave beamforming solutions, current approaches fall short and novel solutions that encompass material identification and efficient ray tracing are needed for blockage mitigation and alleviation of beamforming overhead.

Acknowledgment

We thank the anonymous reviewers for their insightful comments. This work was supported in part by the U.S. NSF under Grants 2212296 and 2235049.

References

[1] 2021. Remcom Wireless InSite 3D Wireless Prediction Software. . https://www.remcom.com/wireless-insite-em-propagation-software.

[2] [n. d.. The Evaluation Kits (EVKs) EVK06002. https://www.sivers-semiconductors.com/sivers-wireless/evaluation-kits/evaluation-kit-evk06002/.

[3] Ahmed Alkhateeb, Iz Beltagy, and Sam Alex. 2018. Machine learning for reliable mmwave systems: Blockage prediction and proactive handoff. In *2018 IEEE Global conference on signal and information processing (GlobalSIP)*. IEEE, 1055–1059.

[4] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath. 2014. Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems. *IEEE Journal of Selected Topics in Signal Processing* 8, 5 (2014), 831–846.

[5] Xuyang Bai, Zixin Luo, Lei Zhou, Hongkai Chen, Lei Li, Zeyu Hu, Hongbo Fu, and Chiew-Lan Tai. 2021. PointDSC: Robust Point Cloud Registration Using Deep Spatial Consistency. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 15859–15869.

[6] Yanan Bao, Tianxiao Zhang, Amit Pande, Huasen Wu, and Xin Liu. 2017. Motion-Prediction-Based Multicast for 360-Degree Video Transmissions. In *Proceedings of IEEE International Conference on Sensing, Communication, and Networking (SECON)*.

[7] Sean Bell, Paul Upchurch, Noah Snavely, and Kavita Bala. 2015. Material Recognition in the Wild with the Materials in Context Database. *Computer Vision and Pattern Recognition (CVPR)* (2015).

[8] Jiangong Chen, Xudong Qin, Guangyu Zhu, Bo Ji, and Bin Li. 2021. Motion-Prediction-based Wireless Scheduling for Multi-User Panoramic Video Streaming. In *Proceedings of IEEE INFOCOM*.

[9] Sungjoon Choi, Qian-Yi Zhou, and Vladlen Koltun. 2015. Robust Reconstruction of Indoor Scenes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[10] Angela Dai, Daniel Ritchie, Martin Bokeloh, Scott Reed, Jürgen Sturm, and Matthias Nießner. 2018. ScanComplete: Large-Scale Scene Completion and Semantic Segmentation for 3D Scans. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[11] Siyan Dong, Kai Xu, Qiang Zhou, Andrea Tagliasacchi, Shiqing Xin, Matthias Nießner, and Baoquan Chen. 2019. Multi-Robot Collaborative Dense Scene Reconstruction. *ACM Trans. Graph.* 38, 4, Article 84 (jul 2019), 16 pages. https://doi.org/10.1145/3306346.3322942

[12] Martin A. Fischler and Robert C. Bolles. 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM* 24, 6 (jun 1981), 381–395. https://doi.org/10.1145/358669.358692

[13] Stuart Golodetz, Tommaso Cavallari, Nicholas A. Lord, Victor A. Prisacariu, David W. Murray, and Philip H. S. Torr. 2018. Collaborative Large-Scale Dense 3D Reconstruction with Online Inter-Agent Pose Optimisation. *IEEE Transactions on Visualization and Computer Graphics* 24, 11 (2018), 2895–2905. https://doi.org/10.1109/TVCG.2018.2868533

[14] Yongjie Guan, Xueyu Hou, Nan Wu, Bo Han, and Tao Han. 2023. MetaStream: Live Volumetric Content Capture, Creation, Delivery, and Rendering in Real Time. In *Proceedings of ACM MobiCom*.

[15] Bo Han, Yu Liu, and Feng Qian. 2020. ViVo: Visibility-Aware Mobile Volumetric Video Streaming. In *Proceedings of ACM MobiCom*.

[16] Tarek Hassan and Munjure Mowla. 2022. A Novel MAC Scheduling Approach for Mobility based 5G Millimeter Wave Networks. In *2022 4th International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*. IEEE, 1–4.

[17] Yubing Jian, Yuchen Liu, Shyam Krishnan Venkateswaran, Douglas M Blough, and Raghupathy Sivakumar. 2020. A quantitative exploration of access point mobility for mmwave wifi networks. In *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 1–7.

[18] Honghao Ju, Yan Long, Xuming Fang, and Rong He. 2020. Systematic beam management in mmWave network: tradeoff among user mobility, link outage, and interference control. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*. IEEE, 1–5.

[19] Michael Kazhdan and Hugues Hoppe. 2013. Screened Poisson Surface Reconstruction. *ACM Trans. Graph.* 32, 3, Article 29 (jul 2013), 13 pages. https://doi.org/10.1145/2487228.2487237

[20] Dileep Kumar, Jarkko Kaleva, and Antti Tölli. 2021. Blockage-aware reliable mmWave access via coordinated multi-point connectivity. *IEEE Transactions on Wireless Communications* 20, 7 (2021), 4238–4252.

[21] O. Landron, M.J. Feuerstein, and T.S. Rappaport. 1996. A comparison of theoretical and empirical reflection coefficients for typical exterior wall surfaces in a mobile radio environment. *IEEE Transactions on Antennas and Propagation* 44, 3 (1996), 341–351. https://doi.org/10.1109/8.486303

[22] B. Langen, G. Lober, and W. Herzig. 1994. Reflection and transmission behaviour of building materials at 60 GHz. In *5th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Wireless Networks - Catching the Mobile Future.*, Vol. 2. 505–509 vol.2. https://doi.org/10.1109/WNCMF.1994.529141

[23] Kyungjin Lee, Juheon Yi, Youngki Lee, Sunghyun Choi, and Young Min Kim. 2020. GROOT: A Real-Time Streaming System of High-Fidelity Volumetric Videos. In *Proceedings of ACM MobiCom*.

[24] Yong Li and Wei Gao. 2018. MUVR: Supporting Multi-user Mobile Virtual Reality with Resource Constrained Edge Cloud. In *Proceedings of IEEE/ACM Symposium on Edge Computing (SEC)*.

[25] Yupeng Liang, Ryosuke Wakaki, Shohei Nobuhara, and Ko Nishino. 2022. Multimodal Material Segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 19800–19808.

[26] Chia-Hung Lin, Wei-Cheng Kao, Shi-Qing Zhan, and Ta-Sung Lee. 2019. BsNet: A deep learning-based beam selection method for mmWave communications. In *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*. IEEE, 1–6.

[27] H. Ling, R.-C. Chou, and S.-W. Lee. 1989. Shooting and bouncing rays: calculating the RCS of an arbitrarily shaped cavity. *IEEE Transactions on Antennas and Propagation* 37, 2 (1989), 194–205. https://doi.org/10.1109/8.18706

[28] Xing Liu, Christina Vlachou, Feng Qian, Chendong Wang, and Kyu-Han Kim. 2020. Firefly: Untethered Multi-user VR for Commodity Mobile Devices. In *Proceedings of USENIX Conference on Usenix Annual Technical Conference (ATC)*.

[29] Ximan Liu, Yuxiang Zhang, Tao Jiang, Li Yu, Jianhua Zhang, and Liang Xia. 2022. Multi-Person Blockage Loss Modeling at Millimeter-Wave Band. In *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)*. IEEE, 1–6.

[30] Yu Liu, Bo Han, Feng Qian, Arvind Narayanan, and Zhi-Li Zhang. 2022. Vues: Practical Volumetric Video Streaming through Multiview Transcoding. In *Proceedings of ACM MobiCom*.

[31] Zida Liu, Guohao Lan, Jovan Stojkovic, Yunfan Zhang, Carlee Joe-Wong, and Maria Gorlatova. 2020. CollabAR: Edge-assisted Collaborative Image Recognition for Mobile Augmented Reality. In *Proceedings of ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*.

[32] Jie Lv, Xinxin He, and Tao Luo. 2021. Blockage avoidance based sensor data dissemination in multi-hop mmwave vehicular networks. *IEEE Transactions on Vehicular Technology* 70, 9 (2021), 8898–8911.

[33] Z. Marzi, D. Ramasamy, and U. Madhow. 2016. Compressive Channel Estimation and Tracking for Large Arrays in mm-Wave Picocells. *IEEE Journal of Selected Topics in Signal Processing* 10, 3 (April 2016), 514–527. https://doi.org/10.1109/JSTSP.2016.2520899

[34] Jiayi Meng, Sibendu Paul, and Y Charlie Hu. 2020. Coterie: Exploiting Frame Similarity to Enable High-quality Multiplayer VR on Commodity Mobile Devices. In *Proceedings of International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*.

[35] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In *ECCV*.

[36] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. 2022. Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. *ACM Trans. Graph.* 41, 4, Article 102 (July 2022), 15 pages. https://doi.org/10.1145/3528223.3530127

[37] Tariq Mumtaz, Shahabuddin Muhammad, Nazeeruddin Mohammad, Muhammad Imran Aslam, and Irfan Ahmed. 2018. Modeling and evaluation of mobility management in mmWave cellular networks. In *2018 IEEE 21st International Multi-Topic Conference (INMIC)*. Ieee, 1–6.

[38] Hemant Kumar Narsani, Prasanna Raut, Kapal Dev, Keshav Singh, and Chih-Peng Li. 2021. Interference limited network for factory automation with multiple packets transmissions. In *2021 IEEE 18th annual consumer communications & networking conference (CCNC)*. IEEE, 1–6.

[39] Mitsumasa Nitta, Hirofumi Suganuma, and Fumiaki Maehara. 2021. Blockage mitigation method employing antenna height and user position in millimeter-wave communications. In *2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*. IEEE, 1–2.

[40] Adewale K Oladeinde, Ehsan Aryafar, and Branimir Pejcinovic. 2021. EBG-based self-interference cancellation to enable mmwave full-duplex wireless. In *2021 IEEE Texas Symposium on Wireless and Microwave Circuits and Systems (WMCS)*. IEEE, 1–6.

[41] Cristina Perfecto, Mohammed S Elbamby, Javier Del Ser, and Mehdi Bennis. 2020. Taming the Latency in Multi-User VR 360°: A QoE-Aware Deep Learning-Aided Multicast Framework. *IEEE Transactions on Communications* 68, 4 (2020), 2491–2508.

[42] Telecom Infra Project. 2019. Analysis of 28GHz and 60GHz Channel Measure- ments in an Indoor Environment. (2019).

[43] Xukan Ran, Carter Slocum, Yi-Zhen Tsai, Kittipat Apicharttrisorn, Maria Gorlatova, and Jiasi Chen. 2020. Multi-user Augmented Reality with Communication Efficient and Spatially Consistent Virtual Objects. In *Proceedings of International Conference on emerging Networking EXperiments and Technologies (CoNEXT)*.

[44] Theodore S Rappaport, Shu Sun, Rimma Mayzus, Hang Zhao, Yaniv Azar, Kevin Wang, George N Wong, Jocelyn K Schulz, Mathew Samimi, and Felix Gutierrez. 2013. Millimeter wave mobile communications for 5G cellular: It will work! *IEEE access* 1 (2013), 335–349.

[45] K. Sato, T. Manabe, T. Ihara, H. Saito, S. Ito, T. Tanaka, K. Sugai, N. Ohmi, Y. Murakami, M. Shibayama, Y. Konishi, and T. Kimura. 1997. Measurements of reflection and transmission characteristics of interior structures of office building in the 60-GHz band. *IEEE Transactions on Antennas and Propagation* 45, 12 (1997), 1783–1792. https://doi.org/10.1109/8.650196

[46] Min Soo Sim, Shubham Ahuja, Wooseok Nam, Alexander Dorosenco, Zhifei Fan, and Tao Luo. 2022. 60 GHz mmWave Full-Duplex Transceiver Study and Over-the-Air Link Verification. In *GLOBECOM 2022-2022 IEEE Global Communications Conference*. IEEE, 3797–3802.

[47] Sanjib Sur, Vignesh Venkateswaran, Xinyu Zhang, and Parmesh Ramanathan. 2015. 60 GHz Indoor Networking through Flexible Beams: A Link-Level Profiling. *SIGMETRICS Perform. Eval. Rev.* 43, 1 (jun 2015), 71–84. https://doi.org/10.1145/2796314.2745858

[48] Sanjib Sur, Vignesh Venkateswaran, Xinyu Zhang, and Parmesh Ramanathan. 2015. 60 GHz indoor networking through flexible beams: A link-level profiling. In *Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*. 71–84.

[49] Aysenur Turkmen, Shuja Ansari, Paulo Valente Klaine, Lei Zhang, and Muhammad Ali Imran. 2021. Indoor mobility prediction for mmWave communications using Markov Chain. In *2021 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 1–5.

[50] Paul Upchurch and Ransen Niu. 2022. A Dense Material Segmentation Dataset for Indoor and Outdoor Scene Parsing. In *European Conference on Computer Vision (ECCV)*.

[51] Usman Tahir Virk, Sinh Le Hong Nguyen, Katsuyuki Haneda, and Jean-Frederic Wagen. 2018. On-Site Permittivity Estimation at 60 GHz Through Reflecting Surface Identification in the Point Cloud. *IEEE Transactions on Antennas and Propagation* 66, 7 (2018), 3599–3609. https://doi.org/10.1109/TAP.2018.2829798

[52] Timothy Woodford, Xinyu Zhang, Eugene Chai, Karthikeyan Sundaresan, and Amir Khojastepour. 2021. SpaceBeam: LiDAR-Driven One-Shot MmWave Beam Management. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services* (Virtual Event, Wisconsin) *(MobiSys '21)*. Association for Computing Machinery, New York, NY, USA, 389–401. https://doi.org/10.1145/3458864.3466864

[53] Shunyao Wu, Muhammad Alrabeiah, Andrew Hredzak, Chaitali Chakrabarti, and Ahmed Alkhateeb. 2022. Deep Learning for Moving Blockage Prediction using Real mmWave Measurements. In *ICC 2022-IEEE International Conference on Communications*. IEEE, 3753–3758.

[54] Yue Yin, Haoze Chen, Zongdian Li, Ryuichi Fukatsu, Tao Yu, and Kei Sakaguchi. 2020. Design of antenna configuration for interference control in mmwave V2V communication systems. In *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*. IEEE, 1–5.

[55] Li Yu, Jianhua Zhang, Yuxiang Zhang, Xiaogang Li, and Guangyi Liu. 2022. Long-range blockage prediction based on diffraction fringe characteristics for mmWave communications. *IEEE Communications Letters* 26, 7 (2022), 1683–1687.

[56] Matteo Zecchin, Mahdi Boloursaz Mashhadi, Mikolaj Jankowski, Deniz Gündüz, Marios Kountouris, and David Gesbert. 2022. LIDAR and position-aided mmWave beam selection with non-local CNNs and curriculum training. *IEEE Transactions on Vehicular Technology* 71, 3 (2022), 2979–2990.

[57] Anlan Zhang, Chendong Wang, Bo Han, and Feng Qian. 2022. YuZu: Neural-enhanced Volumetric Video Streaming. In *Proceedings of USENIX NSDI*.

[58] Ding Zhang, Panneer Selvam Santhalingam, Parth Pathak, and Zizhan Zheng. 2019. Characterizing interference mitigation techniques in dense 60 ghz mmwave wlans. In *2019 28th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 1–9.

[59] Ding Zhang, Puqi Zhou, Bo Han, and Parth Pathak. 2022. M5: Facilitating Multi-User Volumetric Content Delivery with Multi-Lobe Multicast over mmWave. In *Proceedings of ACM SenSys*.

[60] Wenxiao Zhang, Bo Han, Pan Hui, Vijay Gopalakrishnan, Eric Zavesky, and Feng Qian. 2018. CARS: Collaborative Augmented Reality for Socialization. In *Proceedings of International Workshop on Mobile Computing Systems & Applications (HotMobile)*.

[61] Xiaopeng Zhao, Zhenlin An, Qingrui Pan, and Lei Yang. 2023. NeRF2: Neural Radio-Frequency Radiance Fields. *arXiv preprint arXiv:2305.06118* (2023).

[62] Yadan Zheng, Shanzhi Chen, and Rui Zhao. 2021. A deep learning-based mmWave beam selection framework by using lidar data. In *2021 33rd Chinese Control and Decision Conference (CCDC)*. IEEE, 915–920.

[63] Yibo Zhu, Zengbin Zhang, Zhinus Marzi, Chris Nelson, Upamanyu Madhow, Ben Y Zhao, and Haitao Zheng. 2014. Demystifying 60GHz outdoor picocells. In *Proceedings of the 20th annual international conference on Mobile computing and networking*. 5–16.