International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. , No. (Mon-20..)

http://dx.doi.org/10.12785/ijcds/XXXXXX

Improved Real-time 3D Reconstruction Method for Mixed Reality Telepresence

Ajune Wanis Ismail[1], Shafina Abd Karim Ishigaki[1] and Fazliaty Edora Fadzli[1]

[1]*Mixed and Virtual Reality Research Lab, ViCubeLab, Faculty of Computing, Universiti Teknologi Malaysia, 81310 Johor, Malaysia*

Received Mon. 20, Revised Mon. 20, Accepted Mon. 20, Published Mon. 20

Abstract: The advancement of current technologies has allowed long-distance communication between human-to-human to be closer to lifelike encounters. Mixed Reality (MR) telepresence focuses on user involvement within the real and virtual world for telecommunication, which is the current advancement in telepresence applications and is widely recognized. Even so, developing a reliable and affordable three-dimensional (3D) reconstruction of humans in action in real-time for MR telepresence is challenging. Most existing methods have required high computational processes and rely on a large data set to produce a dynamic 3D reconstruction. Therefore, this research introduces an improved real-time three-dimensional reconstruction method that could potentially utilized for MR telepresence. The implementation of double compression techniques on the 3D reconstruction data resulted in substantial enhancements to the real-time 3D reconstruction methodology, hence strengthening the overall framework of the MR telepresence system. An experiment was carried out to evaluate the performance of an improved real-time 3D reconstruction method for MR telepresence. We found the amount of the data transmitted in bytes per frame using the improved method was appropriate for the MR telepresence application. The findings indicate that the real-time performance of 3D reconstruction was accomplished by reducing the data size by twice, resulting in an optimal size that does not impose any limitations on the available bandwidth.

Keywords: Mixed Reality, Telepresence, 3D reconstruction, MR telepresence, real-time, compression

1. INTRODUCTION

The development of telepresence systems which enable remote collaboration of location-wise distant users, has attracted a lot of interest from various communities such as industrial and research groups. Conventionally, telepresence is, at its core, a form of communication that allows users to feel as though they are present in a remote place and interact with the user and objects there in a lifelike way without physically being there [1]. It is frequently used to specify the system which aims to replicate in-person interactions over distance [2]. Therefore, to increase the potential for real-feel interaction between distant users, Mixed Reality (MR) telepresence is introduced as it improves the user's immersive experience of telecommunication.

As described by Milgram and Kishimo [3], MR combines the world of reality and virtual within the same time and location, relying on Virtual Reality (VR) and Augmented Reality (AR) technologies together. The advancement of MR technologies has brought the potential to completely transform the way people interact in a working environment and communicate with distant colleagues [4]. It provides wirelessly broadcast three-dimensional (3D) reconstructions of people or objects by utilizing Red Green Blue - Depth (RGB-D) sensors to Microsoft's HoloLens or head-mounted display (HMD). Tonchev et al. [1] have outlined the specifications and suggested a system design for a particular MR telepresence case study.

The use of MR telepresence has the potential to promote a deeply engaging and immersive encounter, allowing for seamless human-to-human interactions. This technology enables users to communicate and cooperate in a manner that closely resembles being physically present in the same location. The authors of reference [5] introduced a telepresence system in their research, which utilized a life-sized hologram to transmit stereoscopy and continuous motion parallax. Notably, this research achieved these effects without the need for glasses or head tracking. Additionally, the display used in this system was cylindrical and supported a light field, enabling a more immersive experience. Light field displays are capable of rendering three-dimensional (3D) scenes by accurately reproducing the distribution of light in both spatial and angular dimensions surrounding the scene. In addition, a novel telepresence concept known as ”ChameleonMask” is introduced, encompassing a fusion of physical and social telepresence [6]. The ChameleonMask technology facilitates the distant projection of a user's presence onto a physical surrogate, so enabling them to engage with others and surroundings in a manner that

E-mail address: ajune@utm.my, shafina@graduate.utm.my, fazliaty.edora@utm.my

simulates physical presence.

For many immersive telepresence applications, real-time 3D reconstruction is a necessary precondition [7]. In recent years, 3D reconstruction methods have frequently been approached by researchers for data visualization in a variety of disciplines including computer vision, medical imaging, and emerging technologies such as AR, VR, and MR. Hence, this research discusses several studies on different methods of 3D reconstruction for MR telepresence [4], [8]–[10] to analyze and discuss the advantages and disadvantages of each method. Conventionally, the methods for 3D reconstruction in real-time mostly rely on common RGB cameras to gather the images [11]. Other than that, the number of research and projects addressing 3D reconstruction employed in telepresence systems has substantially increased due to the accessibility of cost-effective depth sensing devices such as Microsoft Kinect. However, the 3D reconstruction method is not easy to accomplish [12], especially for MR telepresence, as most works have concentrated on static scenes. MR telepresence requires 3D reconstruction in a dynamic context where geometric colorimetric parameters change over time [13]–[16].

2. RELATED WORKS

Within this research, we constrained this section to discuss 3D reconstruction methods that can reconstruct dynamic context for MR telepresence and compression methods used to reduce the size of 3D reconstruction data. We aim to give an overview of the previous research on related work for the 3D reconstruction method and possible methods used to improve real-time 3D reconstruction within MR telepresence.

A. *3D Reconstruction Method for MR Telepresence*

Considering telepresence as an advanced communication approach, MR telepresence has shown to be the upcoming technology trend that facilitates collaborative work. However, 3D reconstruction is a difficult task, nevertheless in MR telepresence [12]. Even after extensive research throughout the years, researchers still encounter difficulties in achieving rapid, reliable, cost-effective, and realistic 3D reconstruction in MR telepresence. We summarize related research work on 3D reconstruction for MR telepresence from recent years in Table I. The 3D reconstruction method is categorized into four categories which are point-based, visual hall, volumetric, and surface reconstruction.

TABLE I. Summary of related works on 3D reconstruction method for MR Telepresence

			3D Reconstruction Method			
Year	Researchers	No of sensors	Point-based	Surface reconstruction	Visual hull	Volumetric
2022	[17]	2	-	-	-	/
2021	[18]	8	-	/	-	-
2021	[1]	3	/	-	-	-
2020	[9]	2	-	/	-	-
2019	[19]	3	-	-	/	-
2019	[13]	2 or more	-	/	-	-
2019	[20]	4	/	-	-	-
2019	[21]	3	-	-	-	/
2018	[22]	2 or more	/	-	-	-
2017	[4]	2	-	/	-	-
2017	[2]	10	-	-	/	-
2016	[23]	8	-	/	-	-

Based on Table I, the point-based approach is when a set of points is gathered as the input data set for the surface reconstruction process [24]. The point-based data may be classified as either structured or unstructured, with structured data containing connection information for the input point, whereas unstructured data does not have the information. Meanwhile, the visual hull approach reconstructs models utilizing the intersection of object silhouettes from several views [25]. Fairchild et al. [2] extract the user's silhouette from a moving background where they perform segmentation by utilizing the IR light spectrum as their 3D reconstruction approach. Nonetheless, as this approach relies on employing silhouettes to extract objects from an image background, it is demanding to do so in a disorganized and open area [13].

Instead, the volumetric reconstruction approach uses voxel grids to represent 3D data. Each volume element may contain data on space occupancy [26], [27], or samples of continuous functions [28]. Joachimczak et al. [4] have presented a system that renders 3D reconstruction directly inside HoloLens which is updated in real-time for the MR telepresence. They accomplished it by sending 3D models with a relatively low polygon count along with highly compressible high-resolution texture data. [22] was designed to be a cost-minimal telepresence system with smooth frame rates where a reconstructed frame is portrayed as a polygonal mesh, and the polygons' textures are created using high-definition data from RGB cameras. Another example is Regenbrecht et al. [21], who proposed a volumetric approach in their study. The capability of these methods lies in their capacity to construct high-quality scene models while simultaneously incorporating noisy input data. Despite that, one significant disadvantage is their excessive memory consumption.

Lastly, surface reconstruction includes the process of obtaining a 3D representation for a physical object from the input data supplied by 3D scanning devices [29]. However, noise reduction, unwanted point removal, data simplification, and clustering of the initial data set to represent the real objects may be required before surface reconstruction. [9] used the RGB-D sensor that captured the human position and transferred it to be presented in an MR collaborative system which enables real-time telepresence. Both local and remote users can perform collaborative tasks.

B. Compression Method for Real-time 3D Reconstruction

The data pertaining to 3D reconstruction is acquired and transmitted instantaneously. To facilitate the efficient transport of data, it is necessary to utilize a specialized file format. The Hi-C matrix, which represents 3D data, is often available in tabular or comma-separated format as the default option [30]. The matrices in question exhibit a significant degree of sparsity, indicating that a large portion of the data contained inside them is represented by zero values. Hence, it is necessary to utilize condensed representations to reduce its overall size. However, it may be argued that the coordinate list, which includes the display of non-empty items together with their matrix coordinates, is the most often employed method for compressing Hi-C matrices. This technique enables efficient matrix updates and facilitates fast random-access times [30]. In practical applications, binary formats can also be employed as a means of reducing the size of files. The BUTLR format, as shown by its binary-indexed encoding, offers an enhanced approach for the efficient retrieval of contact matrices [31].

Other than that, data compression has the potential to drastically reduce data transmission, hence able to minimize the usage of network bandwidth and storage costs [32]. Compression techniques are commonly employed in audio, video, and speech applications to reduce data size by reducing redundancies. Lossy and lossless compression techniques are the two types of compression algorithms.

Lossy algorithms, as their name suggests, may achieve significant compression ratios due to their ability to forgo the need for an exact replication of the original signal during the decompression process. The popularity of audio-video applications has increased because human perception does not necessitate faultless reconstruction. Furthermore, lossy compression is predominantly employed as a technique to effectively compress sensor data, specifically smart meter data, with the primary objective of reducing costs associated with the constrained resources of bandwidth, energy, and storage inside the transmission realm of the Internet of Things (IoT) environment [33]. Various lossy compression techniques are employed in data processing, such as discrete wavelet transformation (DWT), singular value decomposition (SVD), principal component analysis (PCA), and compressive sensing (CS) [34]–[37]. Furthermore, a recent study has examined a compression technique that utilizes Principal Component Analysis (PCA) and Compressive Sensing (CS) to capture the interdependence of several variables provided by a smart meter [32]. Lossy compression techniques are commonly employed in situations when errors may be tolerated, such as when a high compression ratio is required or in wireless communication environments [32].

In contrast, lossless compression results in reduced compression ratios but enables perfect signal reconstruction [38]. In their research, researchers [33] have introduced a lossless compression technique that effectively compresses waveforms by utilizing the Gaussian approximation with a limited number of approximations. In the present scenario, [4] and [39] are employing the Zstandard lossless compression technique to minimize delay and optimize throughput for reconstructions. One further lossless compression technique [40] that is suggested involves the utilization of the Lempel-Ziv-Free (LZF) compression algorithm, which is effective in reducing both the time required for reading and writing operations. The researchers utilized the LZF compression technique to compress and subsequently decompress the data upon its arrival at the distant site, as documented in the study conducted by reference [17]. The LZF compression approach was also endorsed by [30] as it requires minimal code space and working memory due to its reduced utilization of code space and working memory. Hence, the present research employs the LZF compression technique to compress and decompress the data pertaining to the 3D reconstruction process. Further elaboration on the subject matter will be provided in the subsequent section.

3. PROPOSED METHOD

In this section, we delve into the framework we have proposed to enhance the real-time 3D reconstruction method for MR telepresence. This research did not claim the method as an optimization method. Apply vertex compression to reduce the memory footprint and bandwidth requirements of transmitting vertex data over network connections. As illustrated in Figure 1, our proposed framework comprises three primary processes: real-time 3D reconstruction at the sender, double compression before and after data transmission, and MR telepresence at the receiver site. The following subsection will provide a detailed explanation of each of these processes.

A. Real-time 3D Reconstruction

We begin the real-time 3D reconstruction process by gathering in-depth information from the real-world object and translating it into a numerical array that the computer may manipulate to build 3D data. Traditional cameras can only capture Two-dimensional (2D) RGB data, such as images and films, and cannot collect the necessary depth information for 3D reconstruction. We use the Microsoft Kinect sensor and its time-of-flight (ToF) technology to overcome this gap. ToF technology detects light actively and calculates the phase difference between incident and reflected light at each pixel, allowing us to calculate exact depth distances [41]. We may map pixel points in Euclidean space by assigning them x, y, and z values, resulting in the development of a point cloud data representation. This

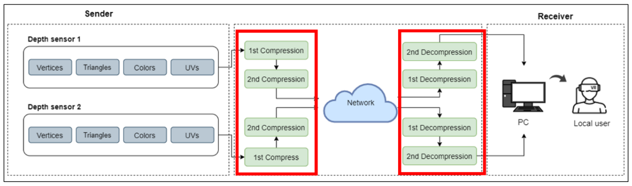

Figure 1. Framework proposed a method using double compression to improve the real-time 3D Reconstruction

point cloud data is used to visualize the 3D model output, which then becomes the input data for later phases. In our experiment, we employed two depth sensors. Each depth sensor captures point cloud data comprising vertices, triangles, color, and UVs (“U” and “V” denote the axes of the 2D texture), as depicted in Figure 1. Once we successfully acquire point cloud data from both depth sensors, we subject this data to a double compression process before transmitting it to the network.

B. *Double Compression*

To optimize network capacity utilization for data transfer, it is often essential to employ data compression techniques to reduce the size of the data while minimizing or eliminating any loss of content. The method to improve the 3D reconstruction data is by using a double compression approach. Double compression is applying the compression algorithm twice on the 3D data to achieve further compression [42]. The aim to use double compression is to achieve even higher compression ratios compared to a single pass as [43][43] has proved double compression on 3D data resulting higher compression ratio. The compression used to compress the data is Lempel-Ziv-Free (LZF). LZF is a fast compression that is ideal for reducing space without sacrificing speed. It is also suitable for repeating data. The process of compression is shown in Figure 1. As presented in Figure 1 after acquiring the point cloud data that consists of vertices, triangles, colors, and UVs from each depth sensor. Each data will undergo compression separately. The initial compression pass significantly reduces data size, achieving remarkable compression efficiency. Subsequently, we perform a second compression to further refine the data’s compression size. We did not perform any optimization techniques such as efficient culling techniques, for example; frustum culling and occlusion culling, to discard or minimize processing of off-screen or occluded vertices. We implement the streaming techniques to load and render vertex data progressively, prioritizing high-detail areas or objects within the mixed reality environment.

Our compression process involves the use of a codec compression tool. This compressor/decompressor tool operates seamlessly over the network, effectively compressing and transmitting data from one end, and then performing the reverse operation to decompress the data to its original form at the other end. This codec compression methodology is a crucial element of our approach, ensuring efficient data transmission and preservation of data integrity throughout the process. The codec compression flow is shown in Figure 2.

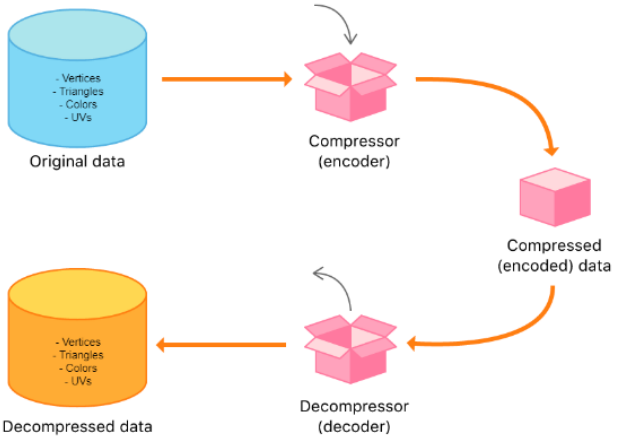

Figure 2. Codec compression process.

After successfully compressing the 3D data, as facilitated by the codec compression process, the next step is to transmit this compressed data through the network. This compression strategy serves important dual purposes. Firstly, it helps minimize the data size before sending it over the network, ensuring more efficient utilization of available bandwidth. This, in turn, contributes to optimizing network resources and enhancing overall network performance. Secondly, this compression-driven reduction in data size also aids in reducing transfer times, particularly in real-time scenarios.

In MR telepresence applications, where timely data transmission is crucial for a seamless user experience, this approach helps minimize potential delays and plays a pivotal role in achieving real-time performance. When the compressed data arrives within the network, it promptly becomes accessible for further processing and analysis. This seamless integration of data optimization and real-time transmission enhances the overall efficiency and responsiveness of the system, contributing to a smoother and more responsive user experience. As soon as the compressed data reaches the receiver site, it undergoes a double decompression process to restore it to its original size before compression. Subsequently, the 3D reconstruction data collected from multiple depth sensors are seamlessly integrated at the receiver site, poised for presentation in the context of MR telepresence.

C. *MR Telepresence*

The next stage involves preparing the necessary experimental setups for MR telepresence. These setups encompass both the local and remote user workspaces and the networking infrastructure, which relies on service providers.

Figure 3(a) illustrates the configuration of the experimental setup for remote users, while Figure 3(b) presents

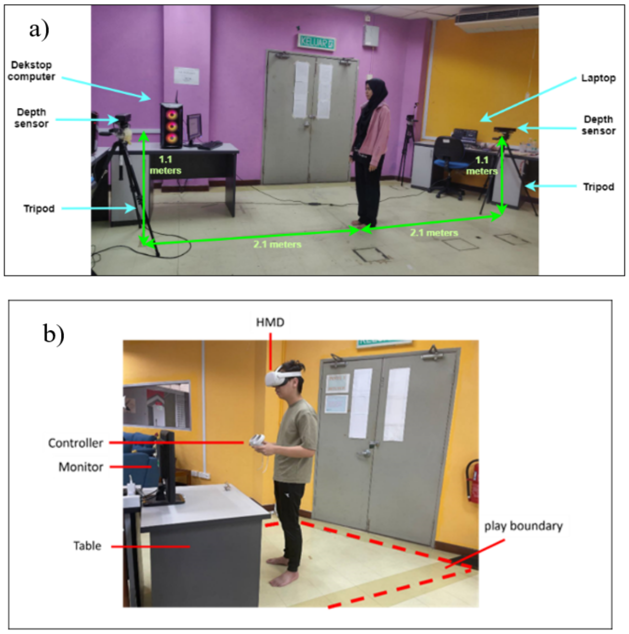

Figure 3. The experimental setup for (a) remote user and (b) local user

the setup for local users. The 3D reconstruction process occurs within the remote user workspace as shown in Figure 3(a) before the data is transmitted to the local user workspace in Figure 3(b). As seen in Figure 3(a), the setup requires two Kinect sensors mounted on tripods and two separate computers to operate the system. This arrangement is necessary because Kinect sensors are unable to support more than one device on a single computer. Meanwhile, the experimental setup of the local user in Figure 3(b) consists of an HMD device, controller, monitor, and a designated play area boundary designed for an immersive MR telepresence experience.

The visual scale of the 3D reconstruction is heavily determined by the distance between the sensors and the subject, which is critical for correctly capturing the remote user's full body. The scope of the reconstruction varies depending on the distance, ranging from half-body to full-body depiction. Based on previous research [17], we recommend a depth sensor height of 1.1 meters and a distance of 2.1 meters between the remote user and the depth sensors. Multiple Kinect sensors are carefully placed to record the remote user from both the front and the back, ensuring a full-body 3D reconstruction.

The remote user is then invited to join the telepresence experience at a remote location. A stable network connection is required for telepresence to work properly. To build stable network connections and enable real-time communication in this research, we use a data streaming library as our network service provider. This service provides global coverage, making a large range of services available from anywhere on the planet. Figure 4 depicts the networking structure. Users must get tokens from the server for validation to begin using the services. The server is activated and ready to use once it has been confirmed. To connect to the same shared area, both local and remote users must enter the same room number, promoting contact and collaboration.

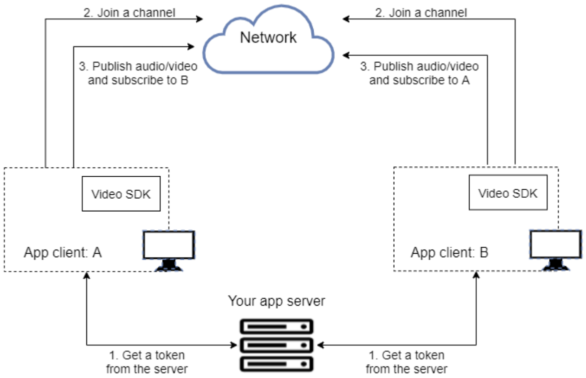

Figure 4. The data streaming networking structure

4. EXPERIMENT OVERVIEW

This section presents a comprehensive description of the experimental setting employed to evaluate the efficacy of the improved real-time 3D reconstruction method for MR telepresence. The test application employed in this research is described and the findings are analyzed, with a comparative analysis of other relevant studies.

Our experiment will evaluate the improved real-time 3D reconstruction approach for MR telepresence, as described by Fadzli et al. [17]. The distant user's location is where the real-time 3D reconstruction process takes place, where depth sensors collect and rebuild the remote user before delivering the compressed data to the local user. An appropriate setup at the distant user's location, as explained in earlier sections, is required to successfully create a 3D reconstruction of a moving human.

To begin the evaluation, participants are welcomed to the remote user workspace area and placed between the Kinect sensors. They are then taught to make little movements to examine the 3D reconstruction output. The evaluation session lasts 5 minutes, during which we collect all relevant data. Figure 5 depicts the evaluation process. The data obtained includes quantitative measurements such as processing time, and 3D reconstruction processing rate, and statistics such as the number of vertices, triangles, color, UVs, and frames per second.

Meanwhile, acquiring access to the MR room entails completing the processes specified in Figure 6, which serves as a flowchart guiding both local and remote users through the essential sequence of steps and decisions. The

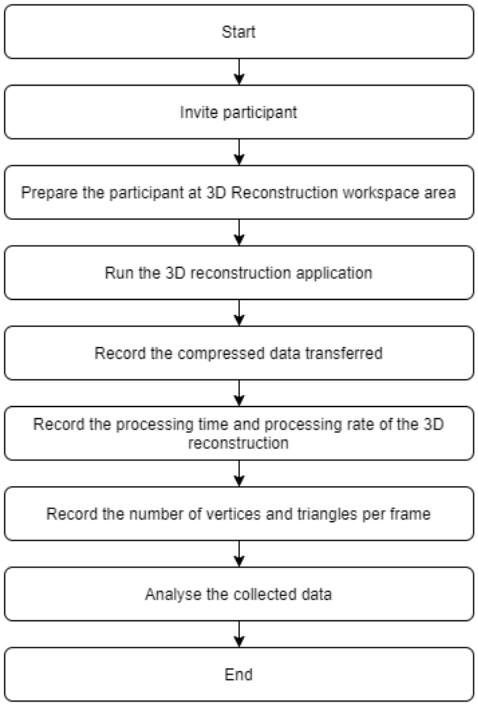

Figure 5. The procedure of the experiment

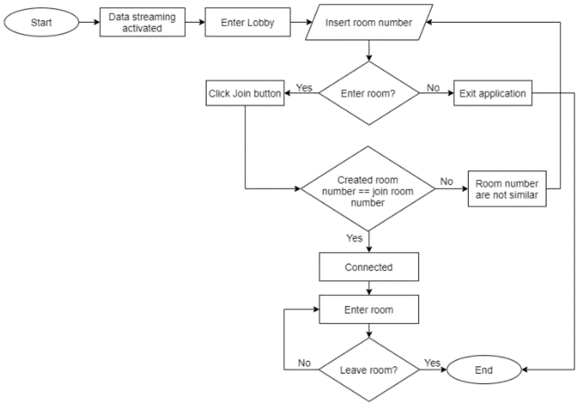

Figure 6. Flowchart to enter the MR room

complete method is depicted graphically in this diagram. The flowchart starts with the data streaming network being activated. Once the network is operational, both users enter the lobby and enter the assigned room number. Both the local and remote users must use the same room number to ensure a successful connection in the same room. Both users are instantly linked and provided access to the shared room once the room number is checked as matching. They can stay in the room until they decide to leave.

5. RESULTS AND DISCUSSION

The experiment recorded the 3D reconstruction data consisting of vertices, triangles, colors, and UVs to determine the size of data that may be optimized. Table II displays the amount of data transferred every frame in bytes for raw data without compression, single compression, and double compression for 3D reconstruction data.

TABLE II. Results of compression on 3D reconstruction data

	Vertices	Triangles	Colors	UVs
Raw data	56090	96660	868352	37396
1st Compression	52112	52196	508050	21863
2nd Compression	53252	28137	503678	11200

The size of the 3D reconstruction data is measured before and after the single and double compression. Table II shows the results collected for the transmitted data in bytes per frame. Based on Table II, triangles show the most significant difference in the first compression, followed by the colors, UVs, and vertices. After the second compression, triangles are still able to reduce a large amount of the data, followed by the colors and UVs. However, the bytes size for the vertices has increased after the second compression. This might happen due to the vertices having exceeded the limit for the compression. To overcome these issues, further study is required to improve the compression for vertices.

The compression results of the 3D reconstruction data were also compared with [39]. The results have been tabulated as shown in Table Table III. From the comparison table, the double compression using the Lempel-Ziv-Free (LZF) method could reduce the byte size of various types of data such as vertices, triangles, and UVs, compared to the Z standard method implemented by [39]. The results for client-to-server show how the accumulated byte size to be sent to the server using the Z standard produces a smaller size compared to this research. However, the excessive reduction in the [39] experiment resulted in a significant reduction in the resolution of the output from 1920 x 1080 to 640 x 480. Hence, balancing the trade-off of quality and performance is necessary when implementing data optimization for real-time 3D reconstruction.

The results of the proposed 3D reconstruction method are presented in Figure 6, in which the local user can see and interact with the 3D reconstruction of the remote user through the network in MR telepresence. To achieve MR telepresence, we display the results via an HMD device. As described in Figure 7(a), the remote user appears in a fully virtual world as if it is in the same space as the local user. Meanwhile, Figure 7(b)(c) shows the results from the side and behind the remote user. The combination of a remote user's 3D reconstruction and a virtual environment has resulted in MR telepresence.

For this research, the results are compared to Fadzli et al. [17], since their technique was identified as the most current and closely aligned with our own investigation. According to their assessment [17], the researchers conducted research on real-time 3D reconstruction for holographic telepresence. They evaluated the performance of 3D reconstruction by

TABLE III. Results of compression on 3D reconstruction data compared with previous work

Features		Our proposed method	Cordova-Esparza et al. [39]
Method		LZF	Z standard
Lossless / Lossy		Lossless	Lossless
Colors	Raw	868352 bytes	6220800 bytes
	Compression	503678 bytes	76800 bytes
Depth Image	Raw	-	434176 bytes
	Compression	-	34734 bytes
Vertices	Raw	56090 bytes	-
	Compression	53252 bytes	-
Triangles	Raw	96660 bytes	-
	Compression	28137 bytes	-
UVs	Raw	37396 bytes	-
	Compression	11200 bytes	-

assessing many factors, including processing time (in milliseconds), processing rate (in frames per second), average number of vertices, and average number of triangles. The outcome of the research may vary due to the use of different display devices. The display device suggested in reference [17] is based on projectors, but this research employs the Oculus Quest Head-Mounted Display (HMD). Both devices possess distinct consoles that have the potential to provide varying numerical outputs.

During the assessment, many metrics were measured upon the participant's entry into the MR telepresence, including the processing time, processing rate, average number of vertices, and average number of triangles. The findings have been collected and organized in Table IV. According to the findings presented in Table IV, it can be observed that the optimized approach exhibits a longer processing time in comparison to the results reported by Fadzli et al. [17]. The procedure yields a reduced processing rate in comparison to the findings of the prior study. This phenomenon can likely be attributed to the integration of the Oculus Quest device with the MR system, which necessitates increased computational requirements. Nevertheless, there is no statistically significant disparity between the mean number of vertices and the mean number of triangles when compared to the findings of the prior investigation.

TABLE IV. Amount of data transmitted in bytes per frame

	Processing Time (ms)	Processing Rate (fps)	Average Number of Vertices	Average Number of triangles
This research	171	30	2930	14172
Curless et al [28]	78	78	2274	12303

Figure 7. Results of the proposed 3D reconstruction method in MR telepresence; (a) front view; (b) side view; (c) back view

6. CONCLUSION

This research achieved an improved real-time 3D reconstruction method, which addressed key issues in data acquisition and transmission as well as integration with MR telepresence. Another highlighted issue in real-time 3D reconstruction pertains to data transmission, which is critical since the 3D reconstruction data relies on the network to transmit data to the receiver. However, due to the large processing data involved in 3D reconstruction, a high bandwidth network is required for data transmission. Bandwidth constraints often result in a drop in transmission time, and subsequently, real-time performance cannot be achieved. Through the study of research findings and previous work in the 3D reconstruction field, key issues in real-time 3D reconstruction were identified, and solutions were developed. To address these issues, a compression method was introduced to optimize 3D reconstruction data during data transmission.

Based on the review of the relevant literature, it is evident that the real-time 3D reconstruction method for MR telepresence has the potential to facilitate long-distance communication in a significant way. The study presented and analysed prior work on MR telepresence employing various 3D reconstruction techniques. This paper also compares the extant method for 3D reconstruction used for MR telepresence and the compression method used to reduce the size of the 3D reconstruction data, which can enhance the real-time performance of 3D reconstruction. The research intended to enhance the real-time 3D reconstruction method for MR telepresence using double compression. The proposed method entailed capturing a remote user with two depth sensors and generating depth data consisting of vertices, triangles, colours, and UVs. The depth data is subjected to surface reconstruction in order to generate mesh. Real-time 3D reconstruction of the remote user is displayed within the MR environment enabled on the HMD worn by the local user as a result of the process.

Limitations such as networking infrastructure, number of sensors, and the play area boundary setting. There were bandwidth and latency constraints in networking since the real-time 3D reconstruction methods often require high-speed data transmission for streaming sensor data, rendering virtual environments, and synchronizing MR experiences over network connections. Limited network bandwidth and high latency can introduce delays in data transmission, resulting in lag or jitter in MR telepresence applications. This latency can degrade the real-time responsiveness and immersion of the experience. Unreliable or unstable network connections can lead to packet loss, dropped frames, or connection interruptions, impacting the quality and continuity of MR telepresence sessions.

Numbers of sensors, we used calibration techniques to align and synchronize data from multiple sensors, such as depth cameras, RGB cameras, and inertial sensors, to establish a common reference frame. Employ sensor fusion algorithms to integrate data from different sensors and modalities, compensating for sensor noise, drift, and inconsistencies. Defining play area boundaries limits users' freedom of movement and immersion within the virtual environment. Adapting play area boundaries to changing environmental conditions or user preferences requires dynamic calibration and adjustment mechanisms. The consequences of inaccurately determining the visual scale can impact the quality of MR telepresence and overall user satisfaction.

As we look to the future, we will expand the scope of this research by including a larger number of users in the sample. This will enable us to evaluate the scalability and robustness of our method. In addition, we intend to undertake user usability testing in the context of MR telepresence to evaluate the social presence and overall experience of users. Besides that, our future efforts will concentrate on enhancing the precision and quality of data acquired from 3D reconstruction, thereby facilitating a more exhaustive evaluation. To further improve the MR telepresence experience, we will also focus on addressing latency and other network requirements.

In conclusion, our proposed method, which utilizes an HMD as a display, has demonstrated its efficacy in real-time 3D reconstruction for MR telepresence. The journey continues as we endeavor to refine and expand this technology, creating new opportunities for immersive remote interactions and applications.

7. Acknowledgement

We would like to extend our heartfelt gratitude to the Ministry of Higher Education for the financial support under the Fundamental Research Grant Scheme (FRGS), UTM, FRGS/1/2023/ICT10/UTM/02/2. Thanks to ViCubeLab at Universiti Teknologi Malaysia for their invaluable support and facilities.

References

[1] K. Tonchev, I. Bozhilov, R. Petkova, V. Poulkov, A. Manolova, and P. Lindgren, "Implementation requirements and system architecture for mixed reality telepresence application scenario," in *2021 24th International Symposium on Wireless Personal Multimedia Communications (WPMC)*. IEEE, 2021, pp. 1–6.

[2] A. J. Fairchild, S. P. Campion, A. S. García, R. Wolff, T. Fernando, and D. J. Roberts, "A mixed reality telepresence system for collaborative space operation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 4, pp. 814–827, 2016.

[3] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays unconscious computing view project using stereoscopic video for defence teleoperation view project a taxonomy of mixed reality visual displays," *IEICE Trans Inf Syst*, vol. 12, 1994.

[4] M. Joachimczak, J. Liu, and H. Ando, "Real-time mixed-reality telepresence via 3d reconstruction with hololens and commodity depth sensors," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 2017, pp. 514–515.

[5] D. Gotsch, X. Zhang, T. Merritt, and R. Vertegaal, "Telehuman2: A cylindrical light field teleconferencing system for life-size 3d human telepresence." in *CHI*, vol. 18, 2018, p. 552.

[6] K. Misawa and J. Rekimoto, "Chameleonmask: Embodied physical and social telepresence using human surrogates," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, 2015, pp. 401–411.

[7] P. Stotko, S. Krumpen, M. B. Hullin, M. Weinmann, and R. Klein, "Slamcast: Large-scale, real-time 3d reconstruction and streaming for immersive multi-client live telepresence," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 5, pp. 2102–2112, 2019.

[8] W. Lin, C. Zheng, J.-H. Yong, and F. Xu, "Occlusionfusion: Occlusion-aware motion estimation for real-time dynamic 3d reconstruction," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 1736–1745.

[9] F. E. Fadzli and A. W. Ismail, "A robust real-time 3d reconstruction method for mixed reality telepresence," *International Journal of Innovative Computing*, vol. 10, no. 2, 2020.

[10] S. S. Joseph and D. Aju, "A comparative survey on three-dimensional reconstruction of medical modalities based on various approaches," in *Information Systems Design and Intelligent Applications: Proceedings of Fifth International Conference INDIA 2018 Volume 1*. Springer, 2019, pp. 223–233.

[11] Q. Jia, L. Chang, B. Qiang, S. Zhang, W. Xie, X. Yang, Y. Sun, and M. Yang, "Real-time 3d reconstruction method based on monocular vision," *Sensors*, vol. 21, no. 17, p. 5909, 2021.

[12] F. E. Fadzli, A. W. Ismail, M. Y. F. Aladin, and N. Z. S. Othman, "A review of mixed reality telepresence," in *IOP conference series: materials science and engineering*, vol. 864, no. 1. IOP Publishing, 2020, p. 012081.

[13] S. Meerits, V. Nozick, and H. Saito, “Real-time scene reconstruction and triangle mesh generation using multiple rgb-d cameras,” *Journal of Real-Time Image Processing*, vol. 16, pp. 2247–2259, 2019.

[14] A. Collet, M. Chuang, P. Sweeney, D. Gillett, D. Evseev, D. Calabrese, H. Hoppe, A. Kirk, and S. Sullivan, “High-quality streamable free-viewpoint video,” *ACM Transactions on Graphics (ToG)*, vol. 34, no. 4, pp. 1–13, 2015.

[15] A. Maimone and H. Fuchs, “Encumbrance-free telepresence system with real-time 3d capture and display using commodity depth cameras,” in *2011 10th IEEE International Symposium on Mixed and Augmented Reality*. IEEE, 2011, pp. 137–146.

[16] C. Plüss, N. Ranieri, J.-C. Bazin, T. Martin, P.-Y. Laffont, T. Popa, and M. Gross, “An immersive bidirectional system for life-size 3d communication,” in *Proceedings of the 29th International Conference on Computer Animation and Social Agents*, 2016, pp. 89–96.

[17] F. E. Fadzli, A. W. Ismail, S. Abd Karim Ishigaki, M. N. A. Nor’a, and M. Y. F. Aladin, “Real-time 3d reconstruction method for holographic telepresence,” *Applied Sciences*, vol. 12, no. 8, p. 4009, 2022.

[18] Y. Luo, Z. Mi, and W. Tao, “Deepdt: Learning geometry from delaunay triangulation for surface reconstruction,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 3, 2021, pp. 2277–2285.

[19] V. Parikh and M. Khara, “A mixed reality workspace using telepresence system,” in *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)*. Springer, 2019, pp. 803–813.

[20] B. Thoravi Kumaravel, F. Anderson, G. Fitzmaurice, B. Hartmann, and T. Grossman, “Loki: Facilitating remote instruction of physical tasks using bi-directional mixed-reality telepresence,” in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, 2019, pp. 161–174.

[21] H. Regenbrecht, J.-W. Park, C. Ott, S. Mills, M. Cook, and T. Langlotz, “Preaching voxels: an alternative approach to mixed reality,” *Frontiers in ICT*, vol. 6, p. 7, 2019.

[22] J. Kolkmeier, E. Harmsen, S. Giesselink, D. Reidsma, M. Theune, and D. Heylen, “With a little help from a holographic friend: The openimpress mixed reality telepresence toolkit for remote collaboration systems,” in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, 2018, pp. 1–11.

[23] S. Orts-Escolano, C. Rhemann, S. Fanello, W. Chang, A. Kowdle, Y. Degtyarev, D. Kim, P. L. Davidson, S. Khamis, M. Dou *et al.*, “Holoportation: Virtual 3d teleportation in real-time,” in *Proceedings of the 29th annual symposium on user interface software and technology*, 2016, pp. 741–754.

[24] S. P. Lim and H. Haron, “Surface reconstruction techniques: a review,” *Artificial Intelligence Review*, vol. 42, pp. 59–78, 2014.

[25] A. Laurentini, “The visual hull concept for silhouette-based image understanding,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 16, no. 2, pp. 150–162, 1994.

[26] C.-H. Chien, Y. Sim, and J. Aggarwal, “Generation of volume/surface octree from range data,” in *Proceedings CVPR’88: The Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 1988, pp. 254–255.

[27] C. Connolly, “Cumulative generation of octree models from range data,” in *Proceedings. 1984 IEEE International Conference on Robotics and Automation*, vol. 1. IEEE, 1984, pp. 25–32.

[28] B. Curless and M. Levoy, “A volumetric method for building complex models from range images,” in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, 1996, pp. 303–312.

[29] A. Khatamian and H. R. Arabnia, “Survey on 3d surface reconstruction.” *Journal of Information Processing Systems*, vol. 12, no. 3, 2016.

[30] J. Waldispühl, E. Zhang, A. Butyaev, E. Nazarova, and Y. Cyr, “Storage, visualization, and navigation of 3d genomics data,” *Methods*, vol. 142, pp. 74–80, 2018.

[31] C. Genomics, “Binary upper triangular matrix (butlr) file format,” 2018. [Online]. Available: http://promoter.bx.psu.edu/hi-c//butlr.html(accessedFeb.26,2023)

[32] M. R. Chowdhury, S. Tripathi, and S. De, “Adaptive multivariate data compression in smart metering internet of things,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1287–1297, 2020.

[33] A. Abuadbba, I. Khalil, and X. Yu, “Gaussian approximation-based lossless compression of smart meter readings,” *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5047–5056, 2017.

[34] S. Das and P. N. Rao, “Principal component analysis based compression scheme for power system steady state operational data,” in *ISGT2011-India*. IEEE, 2011, pp. 95–100.

[35] J. C. S. de Souza, T. M. L. Assis, and B. C. Pal, “Data compression in smart distribution systems via singular value decomposition,” *IEEE transactions on smart grid*, vol. 8, no. 1, pp. 275–284, 2015.

[36] Y. Lee, E. Hwang, and J. Choi, “A unified approach for compression and authentication of smart meter reading in ami,” *IEEE access*, vol. 7, pp. 34 383–34 394, 2019.

[37] M. P. Tcheou, L. Lovisolo, M. V. Ribeiro, E. A. Da Silva, M. A. Rodrigues, J. M. Romano, and P. S. Diniz, “The compression of electric signal waveforms for smart grids: State of the art and future trends,” *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 291–302, 2013.

[38] A. Tiwari and T. H. Falk, “Lossless electrocardiogram signal compression: A review of existing methods,” *Biomedical Signal Processing and Control*, vol. 51, pp. 338–346, 2019.

[39] D.-M. Córdova-Esparza, J. R. Terven, H. Jiménez-Hernández, A. Herrera-Navarro, A. Vázquez-Cervantes, and J.-M. García-Huerta, “Low-bandwidth 3d visual telepresence system,” *Multimedia Tools and Applications*, vol. 78, pp. 21 273–21 290, 2019.

[40] T. A. Welch, “A technique for high-performance data compression,” *Computer*, vol. 17, no. 06, pp. 8–19, 1984.

[41] Y. He and S. Chen, “Recent advances in 3d data acquisition and processing by time-of-flight camera,” *IEEE Access*, vol. 7, pp. 12 495–12 510, 2019.

[42] T. H. Thai and R. Cogranne, “Estimation of primary quantization steps in double-compressed jpeg images using a statistical model of

discrete cosine transform," *IEEE Access*, vol. 7, pp. 76 203–76 216, 2019.

[43] R. Starosolski, "Employing new hybrid adaptive wavelet-based transform and histogram packing to improve jp3d compression of volumetric medical images," *Entropy*, vol. 22, no. 12, p. 1385, 2020.

Ajune Wanis Ismail is a senior lecturer at Johor Universiti Teknologi Malaysia (UTM). She earned her B.Sc., M.Sc., and Dr. Sc. degrees from UTM. She earned a B.Sc. in computer graphics and computer vision and began research on Augmented Reality, which is now her primary research focus. In three semesters, she completed her M.Sc research mode degree, working on vision-based tracking in Augmented Reality.

Shafina Abd Karim Ishigaki graduated from Universiti Teknologi Malaysia (UTM) in 2021 with a B.S. in Computer Science (Graphics and Multimedia Software). She has completed her UTM's master's program under the supervision of Dr. Ajune Wanis. Her life-size telepresence employing holographic display is the topic of her B.S. final project. For her master's degree, she keeps pursuing her interest in MR telepresence research.

Fazliaty Edora Fadzli is a lecturer at Universiti Teknologi Malaysia (UTM). In 2023, she obtained her PhD in Computer Science degree from UTM by enhancing a life-size holographic telepresence framework with real-time 3D reconstruction for dynamic scene. She was offered a fast-track offer letter to a full-time study after receiving B.S. degree in Computer Science (Graphics and Multimedia) from UTM in 2019 with first class honor.