Development of a Telepresence System for Sharing 3D Physical Objects

Improving visibility of virtual objects by increasing the density of real-time point clouds

○ Daichi Ishikawa[*1], Tomohiro Fukuda[*2] and Nobuyoshi Yabuki[*3]

*1 Graduate Student, Div. of Sustainable Energy and Environmental Engineering, Graduate School of Engineering, Osaka University

*2 Assoc. Prof., Div. of Sustainable Energy and Environmental Engineering, Graduate School of Engineering, Osaka University, Ph.D.

*3 Prof., Div. of Sustainable Energy and Environmental Engineering, Graduate School of Engineering, Osaka University, Ph.D.

Keywords: Telepresence; sharing 3D physical objects; Mixed Reality (MR); real-time point clouds; visualization.

1. Introduction

In the architectural field, consensus building between stakeholders such as architects and clients is necessary to enhance the quality of the decision-making. Communication systems play an important role in providing a platform on which ideas can be shared between several people. Although conducting remote meetings such as video conferencing has many advantages to time, cost, and energy used for traveling when compared with face-to-face meetings, there is a lack of presence felt when communicating with other people. Telepresence allows one to feel as though other users in remote areas were presently at the same location, otherwise known as pseudo-face-to-face communication[1)]. Telepresence has the ability to combine face-to-face meeting with remote meeting.

Virtual Reality (VR) and Mixed Reality (MR) can provide immersive experiences with people wearing VR/MR-Head Mounted Display (HMD). Kantonen *et al*. proposed an MR telepresence system, Augmented Collaboration in Mixed Environment (ACME) that can share 3D virtual objects made in Building Information Modeling (BIM) software between stakeholders[2)]. However, this system requires 3D virtual models to be made in advance, making it difficult to share 3D physical objects whose shapes continuously change in real-time.

Point clouds are a cluster of points with 3D coordinates and RGB color values, etc. captured by a 3D laser scanner. Point clouds enable the surveying of terrain and modeling of 3D physical objects to be done easily and precisely. An RGB-D camera can acquire point clouds in real-time (hereinafter referred to as "real-time point clouds"). Kowalski *et al*. proposed LiveScan3D which automatically reconstructs 3D objects using real-time point clouds[3)]. Fukuda *et al*. developed a telepresence system using MR and real-time point clouds which allowed users to share 3D physical objects[4)]. This system focused only on displaying real-time point clouds. In our previous research, Ishikawa *et al*. proposed an immersive telepresence system to improve operability of real-time point clouds in order to share and manipulate the 3D physical objects more easily in remote areas[5)]. However, the visual quality of the 3D virtual objects using real-time point clouds were insuffficient.

The research presented in this paper aims to improve the aesthetic of 3D virtual objects generated by using real-time point clouds. Although increasing the number of point clouds leads to a better quality in visuals, it also increases the processing speed due to the increasing data size. Thus, we propose a new filtering method for point clouds; point cloud clipping which has ability to improve the point cloud processing speed. The proposed system will then be compared to that of the previous system based on whether aesthetics and point cloud processing speed were enhanced.

2. Proposed Method

Our proposed system is an immersive telepresence system that allows a receiver through an MR-HMD to observe and manipulate 3D virtual objects generated using real-time point cloud data acquired from remote areas (Figure 1).

2.1. GENERATING REAL-TIME POINT CLOUDS

At first, an RGB-D camera simultaneously captures the RGB images and depth data (X, Y and Z direction) in real-time. Using point cloud clipping at this point ensures that the RGB-D camera does not acquire depth data (excluding point in the Z direction) by excluding points which exist further than the distance set by the user. Using the point cloud clipping method before other filtering processes will improve the processing speed.

To generate real-time point clouds, the coordinates of RGB images correspond to that of the depth data coordinates because these coordinates differ when acquired by the RGB-D camera.

2.2. POINT CLOUD PROCESSING

To enable a user wearing MR-HMD the ability to manipulate 3D objects individually, we need to classify real-time point clouds into clusters using point cloud segmentation. In the proposed

日本建築学会情報システム技術委員会
第42回情報･システム･利用･技術シンポジウム論文集，216-219，2019年12月，東京
Proceedings of the 42nd Symposium on Computer Technology of Information,
Systems and Applications, AIJ, 216-219, Dec., 2019, Tokyo

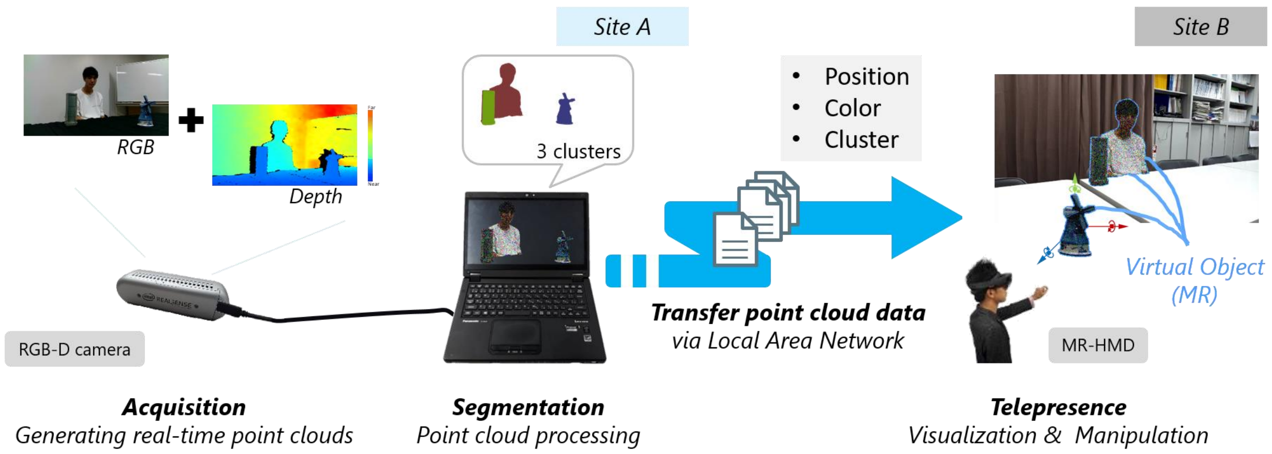


Figure 1. The proposed system architecture

system, point cloud segmentation is done using the Euclidean Cluster Extraction[6)] method which classifies point clouds based on the distances between each point. Before point cloud segmentation, some point cloud processing is required as follows:

- **Filtering**: removing unnecessary points from the point cloud data depending on the threshold set by the user in X, Y, and Z direction to accelerate point cloud processing speed. It is unnecessary to consider in the Z direction because unnecessary points have already removed by point cloud clipping.
- **Extracting and removal planes**: extraction and removal of some planes such as desk or floor surfaces using Random Sample Consensus (RANSAC). This process is necessary for the Euclidean Cluster Extraction method to prevent point clouds from being recognized as the same cluster through some planes.
- **Reducing noise**: eliminating point clouds which exceed the threshold set by the user to achieve better visuals of 3D virtual objects represented as real-time point clouds.

2.3. TRANSFERRING POINT CLOUD DATA

Point cloud data with XYZ coordinates, RGB color values, and cluster information of real-time point clouds are transmitted to an MR-HMD device in a remote area connected via the same Local Area Network (LAN) by using a TCP socket.

The proposed system saves the processed data as a PLY file, which is a file format used for point cloud transmittance and then plays the recorded video output continuously on a PC display.

2.4. TELEPRESENCE ON MR-HMD

A receiver wearing MR-HMD is able to observe, in the real world, 3D virtual objects generated by point cloud data through mixed reality in real-time. The user is able to observe real-time point clouds in 3D and can then manipulate them using a pinch hand gesture.

3. Experiment and Results

3.1. EXPERIMENTAL CONDITION

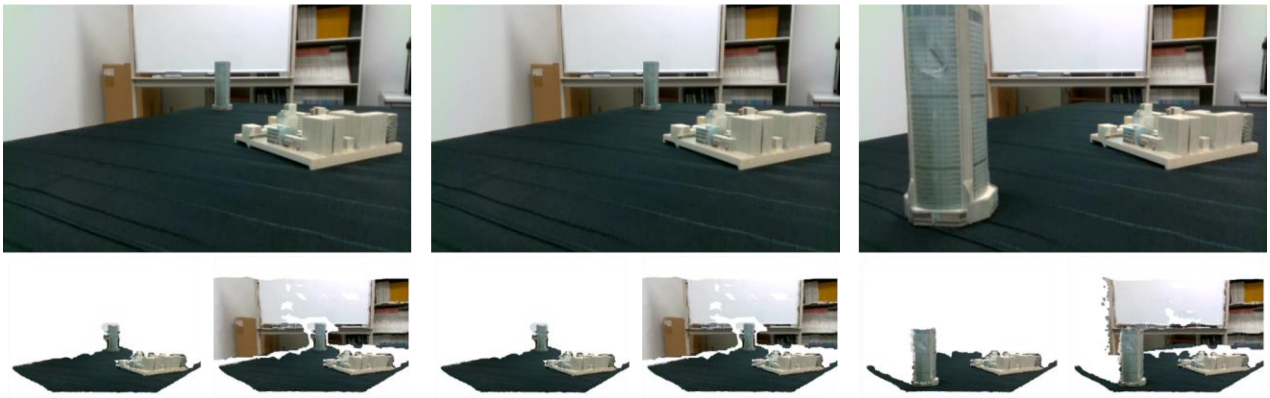
The proposed system requires some devices as shown in Table 1. Although RealSense D435i enables capturing depth data in high resolution (1280×720), we set the resolution as 640×360 in the proposed system to accelerate point cloud processing speed. We also set the threshold of point cloud clipping (Z direction) as 0−1.5 m and that of filtering (X and Y direction) as -1.0−1.0 m. Point cloud processing and transferring point cloud data are conducted by the laptop PC. HoloLens receives point cloud data and displays virtual objects generated by real-time point clouds.

Table 1. Required devices of the proposed system

Name	Use	Feature
RealSense D435i	RGB-D camera	Depth Technology: Active IR Stereo Minimum Depth Distance: 0.105 m
Let's note MX4	Laptop PC	RAM: 8.0 GB CPU: Intel® Core™ i5-5300U GPU: Intel® HD Graphics 5500
HoloLens	MR-HMD	Wi-Fi 802.11ac

3.2. POINT CLOUD CLIPPING METHOD

In the proposed system, we modified the filtering process by adding a point cloud clipping method in order to accelerate the point cloud processing speed. To evaluate the feasibility of the proposed system, the processing speed of the initially proposed system is compared with the improved system. We measured relations between the number of output points and the processing speed three times when generating real-time point clouds based on two static 3D physical models on a desk while changing the position of them (Figure 2).



(a) (b) (c)

Figure 2. Input RGB image (top), clipped input points (bottom left), unclipped input points (bottom right), respectively

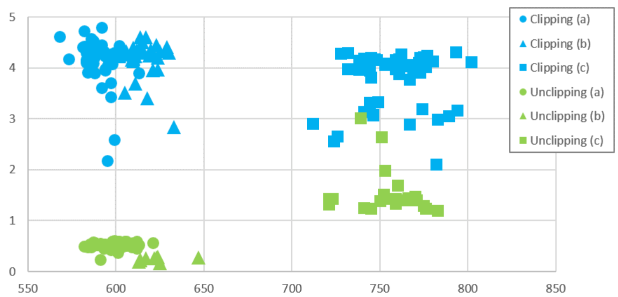


Figure 3. The relation between the number of output points and point cloud processing speed

In Figure 3, the green dots indicate the result of the relation between the number of output points and the processing point cloud speed when not using the point cloud clipping method, while the blue dots show the improved system using the clipping method. Regarding the shape of dots, circles, triangles, and squares indicate in case (a), (b), and (c), respectively. In all cases, necessary objects are represented. Consequently, using the point cloud clipping method was successful in reducting point cloud processing speed.

3.3. THE VISIBILITY OF 3D VIRTUAL OBJECTS

This section evaluates whether the aesthetics of the 3D virtual objects is enhanced by using the proposed system. As shown in Figure 4, an experiment is conducted via Local Area Network (LAN) on two sites separated by a wall. In site A, real-time point clouds that represent a sender and a static physical 3D building model are generated by acquiring RGB images and depth data using RealSense D435i (Figure 5a). In site B, a receiver wearing HoloLens receives the processed point cloud data transferred from site A and observes the state of the sender and 3D building model from site A through real-time point clouds in mixed reality (Figure 5b). The receiver is also able to manipulate 3D virtual objects generated by real-time point clouds using a hand pinch gesture.

Figure 6 shows screenshots of the view of the receiver captured by the HoloLens. The receiver in site B was able to observe the sender and the 3D building model in remote area site B in 3D using real-time point clouds.

Next, comparison of the visibility of our proposed system is made with that of the previous research[5)] (Figure 7). Visuals in the improved system in Figure 6 is compared to that of visuals created in the previous research in Figure 7. It is evident that the visibility of objects represented by real-time point clouds was improved by increasing the number of output point clouds.

This experiment was done using a laptop PC for point cloud processing. If a PC with better specifications is used instead of the laptop PC, the point cloud processing speed in our proposed system is almost same as that of previous research. Priority of desired outcome should be made regarding aesthetics of virtual objects, faster the processing speed or mobility depending on the situation when using this system.

4. Discussion

As described in sections 3.2, we could improve the visibility of 3D virtual object by increasing the density of real-time point clouds in the proposed system　Improving the visibility leads to better understandings and smooth consensus building between stakeholders. In addition, we confirmed the point cloud clipping method could reduce the point cloud processing speed.

However, our proposed system still has some challenges.

The first is the lack of bidirectional connection. Currently only the sender is able to transmit point cloud data and the receiver can only view and manipulate the objects. A desired outcome for future works is the sender's ability to view and manipulate point cloud data as well; making the system bidirectional.

The second is the limitation of environment. In order to use the system on a greater scale, adapting this system into the internet environment is essential.

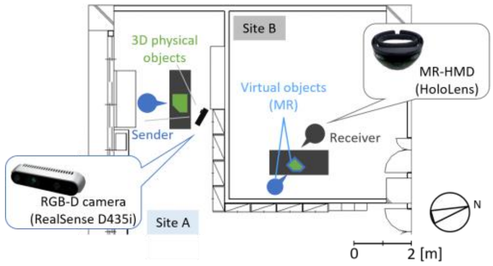
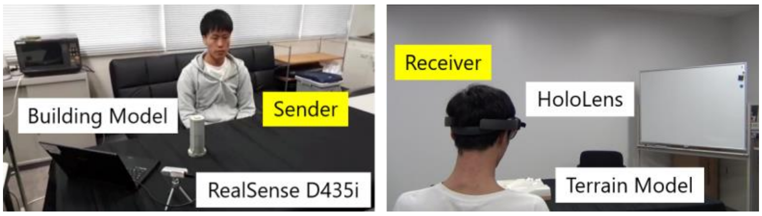


Figure 4. Floor plan showing two sites



(a) Sender (site A) (b) Receiver (site B)

Figure 5. System users

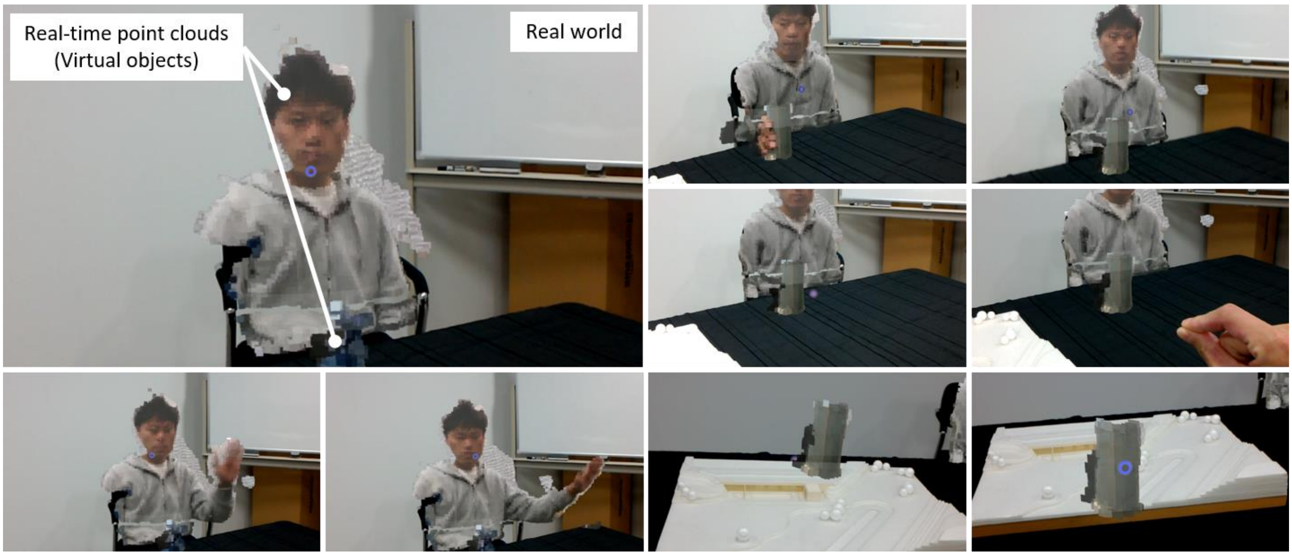


Figure 6. The receiver's view through HoloLens

Figure 7. Insufficient views of real-time point clouds in previous research

5. Conclusion and Future works

We proposed an immersive telepresence system that has improved aesthetics for sharing objects between users generated using real-time point clouds. Although latency is slower due to the increased number of point clouds, we confirmed that including the clipping method during the filtering process has the ability to accelerate the point cloud processing speed.

Future works include implementation of bidirectional connection between a sender and a receiver and adapting the proposed system into the internet environment to realize more effective telepresence.

References

1) Buxton, W. (1992). Telepresence: Integrating shared task and person spaces, In Proceedings of graphics interface, Vol. 92, No. 1992, pp. 123-129.
2) Kantonen, T., Woodward, C., and Katz, N. (2010) Mixed reality in virtual world teleconferencing, In Virtual Reality Conference (VR), 2010 IEEE, pp. 179-182.
3) Kowalski, M., Naruniec, J., and Daniluk, M. (2015). Livescan3d: A fast and inexpensive 3d data acquisition system for multiple kinect v2 sensors. In 2015 International Conference on 3D Vision pp. 318-325.
4) Fukuda, T., Zhu, Y., and Yabuki, N. (2018) Point Cloud Stream on Spatial Mixed Reality Toward Telepresence in Architectural Field, Proceedings of the 36th eCAADe Conference (Education and Research in Computer Aided Architectural Design in Europe), Vol. 2, pp.737-744.
5) Ishikawa, D., Fukuda, T., and Yabuki, N. (2019). Development of a Telepresence System for Sharing 3D Physical Objects Using Real-time Point Clouds. AIJ HOKURIKU 2019, Summaries of Technical Papers of Annual Meeting, Information Systems Technology, 13-14.
6) Rusu, R. B. (2010). Semantic 3d object maps for everyday manipulation in human living environments. KI-Künstliche Intelligenz, 24(4), 345-348.