RESEARCH ARTICLE

Effect of Auto-Erased Sketch Cue in Multiuser Surgical Planning Virtual Reality Collaboration System

SUYEON PARK[1], **(Student Member, IEEE), GAYUN SUH**[2], **SOO-HYUNG KIM**[1], **HYUNG-JEONG YANG**[1], **GUN LEE**[3], **AND SEUNGWON KIM**[1]
[1]Department of Artificial Intelligence Convergence, Chonnam National University, Gwangju 61186, South Korea
[2]School of Artificial Intelligence, Chonnam National University, Gwangju 61186, South Korea
[3]School of Information Technology and Mathematical Sciences, University of South Australia, Adelaide, SA 5001, Australia

Corresponding authors: Gun Lee (gun.lee@unisa.edu.au) and Seungwon Kim (Seungwon.Kim@jun.ac.kr)

ABSTRACT Many researchers have introduced Virtual Reality (VR) systems for various medical applications. One of the typical applications is a medical collaboration system that helps users collaborate in a virtual reality environment while physically staying in different places. In this research, we explore the effect of the using 3-second sketch communication cue in a collaborative VR system for two surgeons planning surgery, when a head-gaze pointer and hand-gesture are available as the baseline communication cues. We conducted a user study comparing four conditions according to the availability of the auto-erased sketch cue by two collaborators: the message provider and receiver. The results demonstrated the benefits of the 3-second sketch communication cue that increased the level of usability and message understanding between two collaborators and decreased the required task load. We additionally found that the sketch cue would be more beneficial if it was erased after finishing each collaborator's communication turn rather than after three seconds. With the lesson learned from the user study, we propose six design principles for the multiuser planning surgery VR system: 1) 3D reconstructed body model should be supported, 2) the cross-sectional view at any angle and position should be supported, 3) visual communication cues, such as sketches, pointers, and hand gestures, are crucial for communication, 4) the erasing sketch should occur when the next sketch starts being drawn 5) the system should support permanent annotations, 6) the system should control the transparency of the hand gestures and head-gaze pointers to mitigate the effect of occlusions.

INDEX TERMS Collaboration, surgical planning, virtual reality (VR), visual communication cue.

I. INTRODUCTION

Virtual reality (VR) is one of the tools for representing a human body in three-dimension (3D). Early researchers [1], [2], [3], [4], [5], [6] in VR medical systems focused on

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaogang Jin.

reconstructing 3D bodies from real-world images that were taken by Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) devices. Zhang et al. [1] and Fiederer et al. [5] reconstructed complex neural structures. Egger et al. [2] developed a system that reconstructed organs and skeletons. Greuter et al. [6] introduced a system that reconstructed a 3D body model based on CT and MRI images. In the early

age, with the 3D reconstructed VR body contents, several researchers simply used it for medical training and simulation [7], [8], [9], [10] by mostly showing the VR body contents to trainers.

Later, researchers started to explore multiuser VR medical applications for collaboration [8], [11], [12], [13], [14], [15], and the communication cues for VR collaboration were their main research topics. Bork et al. [8] explored the use of a head-gaze pointer and colored pins for multiuser communication in VR medical collaboration systems. Yu et al. [11], [12] investigated an interface for precisely sketching the shared 3D body for communication between two surgeons to collaborate.

Multiuser collaboration in VR has been studied not only for medical purposes but also for remote collaboration [8], [11], [12]. In remote collaboration studies, many researchers explored the effect of using visual communication cues such as the pointer, sketch, and hand gestures [16], [17], [18]. The pointer was for pointing information [17], [19], the sketch was for representing a shape by drawing lines [16], [20], [21], and the hand gesture was for hand representations [7], [18], [22]. One of the interesting researches on using the sketch cue was Fussell's study [16] that emphasized the auto-erased functions after three seconds from the time of drawing sketches. This auto-erased sketch was for mitigating the issue of accumulated sketches that hindered collaborative activities and was applied to many remote collaboration studies using sketch cues [22]. Interestingly, this auto-erased sketch has not been applied to recent medical collaboration studies [8], [11], [12].

In this paper, we explored the effect of using the auto-erased sketch cue after 3 seconds for a typical medicine collaboration task: planning surgery together. Since recent collaboration studies [22] support multiple communication cues in addition to the auto-erased sketch cue, our system supported two additional cues: hand gestures and head-gaze pointers. Besides, our system supports rendering a 3D volumetric body, reconstructed with real-world Digital Imaging and Communications in Medicine (DICOM) images taken by CT and MRI and provides a cross-section view of it at any angle and position.

For the user study, we explored the effect of the auto-erased sketch cue by comparing four conditions according to the availability of the auto-erased sketch cue between two collaborators (by turning on and off the 3-second sketch cue for the message provider and/or receiver). The experimental task was for planning liver cancer surgery. To the best of our knowledge, this is the first study exploring the effect of the 3-second sketch cue in combination with hand gestures and gaze pointer for planning a surgery task in VR.

II. RELATED WORK
In this section, we describe the previous studies in medical VR applications first, then in remote collaboration and auto-erased sketch cues.

A. MEDICAL AUGMENTED/VIRTUAL REALITY SYSTEM
The early studies in medical VR applications mostly focused on designing a single user interface including 3D reconstruction of the body with inner body organs, how to improve the user interaction with the dense organs, and how to show cross-section view of the body.

Researchers have tried to integrate real-world DICOM images into a VR system for a 3D display and interface. Reddivari et al. [4] and Ard et al. [23] introduced a function to convert DICOM images into 3D virtual organs and skeletons. Herfarth et al. [24] and Croci et al. [25] investigated the effect of presenting them in a 3D head-mounted display (HMD) view to plan an operation. The 3D display is effective for understanding depth; therefore, the planners can better understand the current state of the organs compared with the 2D display. Further, Lasso et al. [26] introduced a system that a user can navigate the inside of the organs by moving through and viewing inside the 3D organs. Pfeiffer et al. [3] developed a VR framework that automatically creates 3D organs, skeletons, and arteries from DICOM images.

Subsequently, researchers have added various advanced interfaces for visualizing 3D organs and skeletons. The 3D organs and skeletons are dense in a small area, so He et al. [27] developed an interface that spreads the dense organs with the interaction of two hands moving apart. Pfeiffer et al. [3] developed an interface that allows users to control the transparency of the organs and skeletons to easily distinguish them.

Pinter et al. [10] and Luo et al. [28] developed a VR system that supports a cross-sectional view of a 3D reconstructed body. The cross-sectional view shows a plane view of the 3D organs based on the CT or MRI data. They used a virtual panel to determine the position and angle of the cross-section view by positioning it with controllers [10] or a tablet [28]. They also implemented additional functions, such as organ manipulation, drawing sketches, measuring a length, and user navigation to the point indicated by the controller.

Recently, researchers have extended the use of VR medical systems to multiuser systems. Schott et al. [15] and Chheang et al. [14] developed a virtual operating room for educational purposes where the students and a teacher can have a variety of educational medical materials, such as DICOM 2D images, 3D virtual organs, information boards, and recorded videos in a virtual room. In the virtual room, students and teachers can verbally communicate and use the information board (similar to a whiteboard) to share written words.

One of the important research topics for the multiuser medical system is improving the quality of communication for collaboration in VR. Bork et al. [8] also introduced a medical educational system and added visual communication cues, such as a head-gaze pointer, and colored pins for better user communication. Yu et al. [11], [12] used the sketch communication cue for collaboration between surgeons in VR medical systems. They introduced the Magnoramas system [11] that enhances the drawing accuracy of a sketch cue by

bringing the target object (e.g., the skull) close to the user and allowing the user to manipulate it for better and more precise drawing sketches. This system was extended to a co-located collaboration system [12] to augmented reality (AR) that solved the problem of a too-small operational area for multiple surgeons by providing a digital copy of a 3D region of interest and allowing multiple surgeons to have enough space and share sketches on copies. However, their studies [11] primarily focused on accurately drawing sketches but did not explore previously introduced interaction techniques, such as automatically erasing sketches after three seconds from the time they were drawn [16], that still needs further investigation in the VR planning surgery task.

B. AUGMENTED/VIRTUAL REALITY COLLABORATION SYSTEM

Collaboration is a process of joint and interdependent activities to achieve a common goal [21]. Collaborators try to align their activities, and the critical factor for aligning them is a high level of awareness of what is occurring in the task space and understanding others' activities [29].

To provide a better understanding of others' activities, researchers focused on communication cues between collaborators. The major communication cues that they explored are virtual hand interactions, sketches, and pointer cues, to supplement voice communication. Early researchers used a pointer and drew sketches over a 2D screen view [16]. Later, with advanced computer vision technology, Chen et al. [30] and Kunz et al. [31] extracted hands from live camera images and overlaid them onto the 2D shared view, supporting hand-gesture communication. Later, as the collaboration task space was captured and shared in 3D, the visual communication cues were also in 3D [32]. Moreover, the display system became portable with an HMD; thus, the interface for visual communication cues also became portable or used bare-hand interaction with the hand-tracking technology [22]. Recently, with the advances in eye tracking technology, several researchers have started to use a gaze pointer as a visual communication cue. Piumsomboon et al. [33] and Lee et al. [34] used eye gaze to represent collaborators' points of interest (POIs) in the task space and explored how the increased awareness of collaborators' POIs influenced the collaboration. Jing et al. extended the use of the eye-gaze pointer by integrating speech recognition with the gaze pointer [35] and suggesting a visual representation of linking multiuser gaze pointers [19].

One interesting recent trend in the study of visual communication cues is combining or integrating multiple visual cues to better use the collaboration system. Huang et al. [36] introduced a system supporting sketch and hand-gesture cues, and Kim et al. [22] discussed using three typical visual cues, virtual hands, sketches, and pointers. Their studies reported that each visual communication cue has benefits, as summarized in Table 1. The hand-gesture cue was more useful with a task requiring sophisticated hand operation [37],

the sketch cue was better with a task requiring representing lines [18], and pointer cues with a task requiring quick participation [17].

TABLE 1. Benefits of visual communication cues.

	Benefits
Hand Gestures	Most natural and intuitive [37], Displays shapes with diverse finger poses [18], Has diverse symbolic meaning [22]
Pointers	Easily illustrates points of interest and simple positions [19], Fast reactions to collaboration activity [17]
Sketches	Displays specific shapes or lines [16], written words [20], and accumulated information with nonvolatility [21]

C. AUTO-ERASED SKETCH CUE FOR COLLABORATION

Most collaboration includes a sequence of communication, and the information contained at a certain step of communication is meaningful in the combination of the previous and next steps of the communication. In this sense, old information that was already used for previous collaborative activities is not necessary anymore as collaborators already understood and used it. The hand gesture and pointer communication cues are only visible when the collaborators are currently using them, so it is impossible to retain old information by the hand gesture and pointer cues. However, the sketch cue remained like an annotation as drawn, and this causes an issue of retaining old information in the task space and hinders the clear view of the task space.

To solve this issue, Fussell et al. [16] introduced the auto-erased sketch cue after three seconds from the time the sketch was drawn. The main purpose of the automatically erased sketches was to mitigate the user's misunderstanding of the task space by the accumulated and old sketch cues. In her user study [16], the results showed that 3-second sketch cues improved usability and achieved better communication between collaborators while reducing mental effort for collaboration.

Manually erasing the sketch function might not be a good solution according to Kim's early study [17]. In his study, he found that the sketch providers (or drawers) did not frequently erase their sketches because the sketch providers thought the old sketch did not hinder understanding task space while the sketch receiver did not have the confidence to erase the provider's sketch or did not know whether it was okay to erase provider's sketch or not. In short, the manually erasing sketches were not performed as many as the automatically erasing sketch function as the message receiver wanted. Therefore, he implemented the auto-erased function in his later remote collaboration study with the sketch cue [22].

Fussell et al. [16] implemented automatically erasing the sketch after three seconds, and Kim et al. [22] erased the sketch after one second.

D. SUMMARY AND HYPOTHESES

Recent studies with a VR medical collaboration system do not fully adopt recent design recommendations for VR remote collaboration systems that support the 3-seconds sketch cue and multiple visual communication cues together. Thus, we apply this design recommendation, mainly for planning surgery tasks, and explore whether the recommendation is still significant or not.

To solve this question, we compared four conditions in this user study with two independent variables: whether the sketch cue was supported or not and which collaborator (the message provider and receiver) uses the sketch cue (to observe how it differently influences based on the collaborators' role). We refer to this 3-second sketch cue as the 3-s sketch cue.

Based on the previous studies, we pose the following hypotheses in planning surgery tasks:

1) The 3-s sketch cue improves system usability.
2) The 3-s sketch cue supports better communication between collaborators.
3) Collaborators have a higher level of co-presence with the 3-s sketch cue.
4) Collaborators spend less mental effort with the 3-s sketch cue.

III. SYSTEM DESIGN

This section describes system requirements and how to develop a prototype system. Section III-A outlines the requirements for the system, while Section III-B introduces our prototype. Finally, in Section III-C, we describe the setup used for development and user studies.

A. SYSTEM REQUIREMENTS

To develop a practical VR surgery planning system for surgeons, we interviewed three surgeons and defined three main requirements:

1) The system should display the cross section to help surgeons find the best incision line.
2) The system should provide a method to represent the incision lines clearly and easily.
3) It should support pointing activities to specify and indicate the POI.

To facilitate the requirements, our prototype provides several functions. First, the prototype supports the volumetric rendering of the DICOM data and generates a 3D body and a cross-sectional view. Second, we implemented a 3-s sketch cue to clearly represent incision lines. Third, we integrated the head-gaze pointer to present the collaborators' POIs. Additionally, it is a multiuser collaboration system; therefore, we added the functions of synchronizing user activities with avatars and visual cues. Furthermore, the proposed system supports recording and saving the content and current scene. Section III-B discusses more details.

B. PROTOTYPE DEVELOPMENT

We used an open-source library[1] that generated a 3D body model based on DICOM images (Fig. 1(a)). With the 3D body model, the proposed prototype supports two types of cross-sectional views. The first is the slice view, presenting the remaining part of the body after making an unnecessary body part transparent (Fig. 1(b)), and the second is a cross-sectional view depicting CT or MRI images at the selected angle and position (Fig. 1(c)).

For the two views, we designed two 2D panels. The transparent 2D panel with a green boundary was for the slice view, and the nontransparent 2D panel with a black background was for the cross-sectional view. Collaborators can select the appropriate view by placing the 2D panels at the angles and positions with simple hand holding, translating, and releasing interactions. Additionally, collaborators can manipulate any object in the VR environment with hand holding, translating, and releasing interactions. The collaborators can also scale any object up or down with two hands moving apart or close together when both hands hold the object.

The sketch cue draws lines with the line render object of the Unity Engine and forms a line(s) with the series of the position of the index fingertip. The sketch starts being drawn when a collaborator has the finger pose as Figure 1d and 1h: stretching the thumb and index finger while closing the others. Regarding the drawing position, our system followed the suggestion by Yu et al. [11] rather than Kim's [22]; drawing a sketch at the index fingertip (Fig. 1(d)) rather than where the index finger points at in distance. The position of the index fingertip is saved and accumulated in the line render object while the collaborator is moving the index finger, and the line render object connects multiple positions of the index fingertip to form a sketch. During the finger drawing sketch, the collaborator should keep the finger pose, and changing the finger pose stop drawing the sketch (i.e., closing thumb and index finger) The sketches are automatically erased 3 seconds after drawn.

For presenting the collaborator's POI or specifying a distance object, we implemented a small, green, sphere-shaped head-gaze pointer. The position of it is the first collision point of the ray at the front head direction of an HMD. The starting position of the ray is the center of the viewpoint as several previous studies did [33], [34], [35]. For hand-gesture cues, we adopted direct hand-gesture interaction that synchronizes real-world hand gestures into virtual hands and supports hand-object holding, translation, and releasing interactions.

We implemented a network using the Photon Fusion framework to synchronize the virtual medical environment for multiple collaborators [38]. This framework synchronizes the position and orientation of the avatar, hand model, and other virtual objects. Voice communication was available with the Photon Voice Network [39].
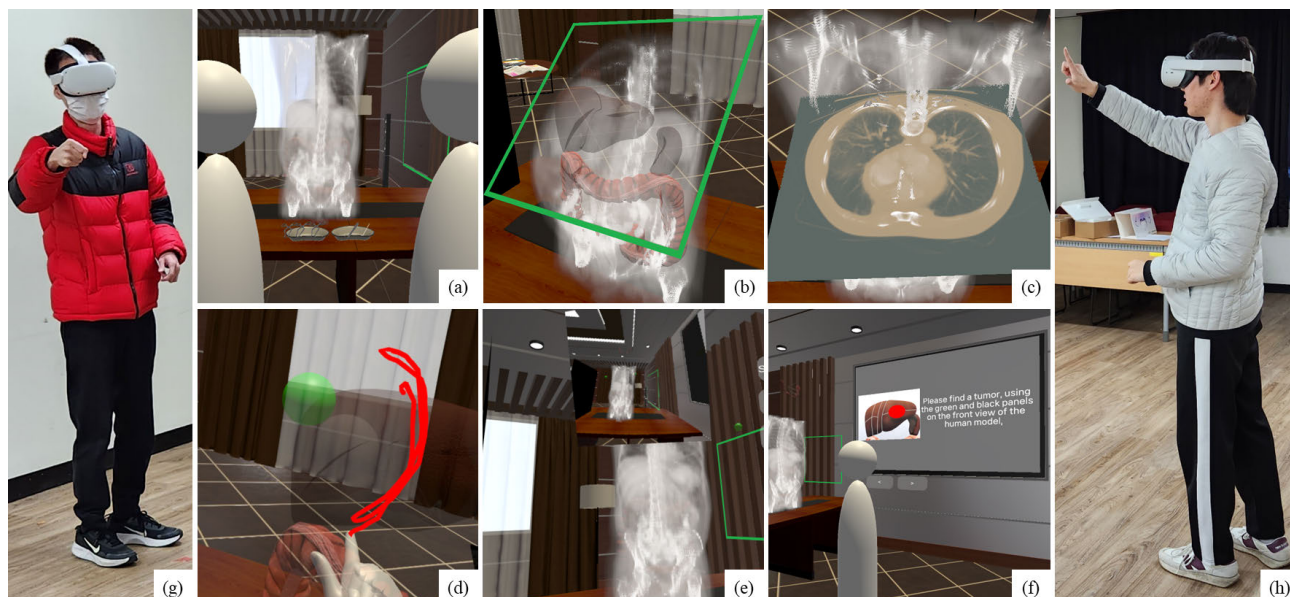
---

[1]https://github.com/mlavik1/UnityVolumeRendering

FIGURE 1. System overview: (a) a 3D body model from DICOM image reconstruction investigated by two users represented as avatars, (b) slice view and panel, (c) cross-sectional view and panel, (d) user sketch (red lines) on the liver and a green ball representing head gaze pointer, (e) window displaying the other user's first-person view, and (f) Instruction window displaying the surgical step to the provider.

Moreover, we implemented several other functions. One is sharing the other collaborator's first-person view with a window (Fig. 1(e)). To implement it, each collaborator's system has additional camera objects for taking other's view while sharing the position of each collaborator's viewpoint between two systems. When pressing the window with a hand, the system maximizes the window and supports to have other view rather than his/her own view.

Furthermore, the proposed prototype includes an instruction window that provides information about the surgical step (Fig. 1(f)). The arrow button interface is provided to support moving to the previous and next steps of the instructions. Additionally, the prototype supports functions for the capturing scene and screen. The scene capture function saves the position and orientation of every virtual object (including the reconstructed body, 2D panels, and visual communication cues) at the moment of saving. A screen capture function saves the image of the current user view to a desktop folder, allowing the collaborators to check the images without loading the VR medical prototype with an HMD, but using a desktop computer instead.

C. DEVICES

We used the Unity Game Engine (*2021.3.14f), Oculus Integration 42.0, and OVR Plugin Legacy for implementing the prototype. The prototype system is built on Oculus Quest 2 VR HMDs with a resolution of 1832×1920 pixels per eye and a 90 Hz refresh rate. Two HMDs (one for each user) were tethered to two desktops. One desktop is equipped with an Intel Core i7-9700K CPU at 3.60 GHz, 16 GB of RAM, an NVIDIA GeForce RTX 2080 Ti, and Windows 10, and the other one with Intel Core i9-12900K at 3.2 GHz, 80 GB of RAM, an NVIDIA GeForce RTX 3090, and Windows 10.

IV. USER STUDY

We designed and conducted a user study to investigate using the 3-s sketch cue for a VR surgery planning system.

A. TASK

For practical use, we applied a real-world scenario for the experimental task. We chose the liver resection task because it has a typical sequence of steps with a confirmed scenario for treating liver cancer, which has the sixth highest incidence rate globally. The real-world surgery task includes dozens of detailed operations, so taking every detail in the VR planning of a surgery task is practically impossible.

Thus, we discussed this with surgeons and applied the five major steps for such a task, as follows:

1) Check the position of the tumor with the cross-sectional and slicing views in the 3D volume rendered body model (Fig. 2(a));
2) Share and confirm the incision lines of the abdomen (Fig. 2(b)) between collaborators;
3) Specify two or three skin-pinching points for holding the opened abdomen with forceps (in this study, the collaborators simply leave the forceps at the red points Fig. 2(c));
4) Share and confirm an incision line of the liver to cut off the tumor (Fig. 2(d));
5) Remove the forceps.

In the real world, there is another step for closing and suturing the abdomen, but we did not consider this step because the surgeons told us it is not significant in planning surgery.

Additionally, to complete the task within an acceptable experimental time, we designed it to be one collaborator providing an idea of operation using visual communication cues, and the other collaborator confirming it by duplicating the same visual cue activities to demonstrate how well the collaborator understands the idea. For example, if one collaborator explains operation activities with a sketch drawing activity, the other collaborator confirms his/her understanding of it by drawing the sketch again at the same position of the reconstructed body.
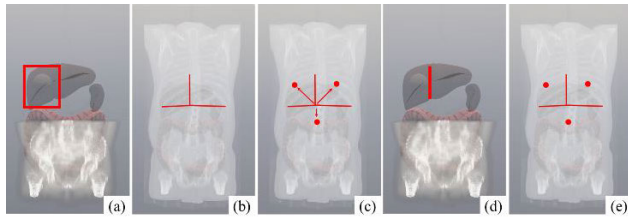

FIGURE 2. Task steps: (a) check position, (b) confirm incision line on the abdomen, (c) skin-pinching point, (d) incision line on the liver, and (e) remove forceps.
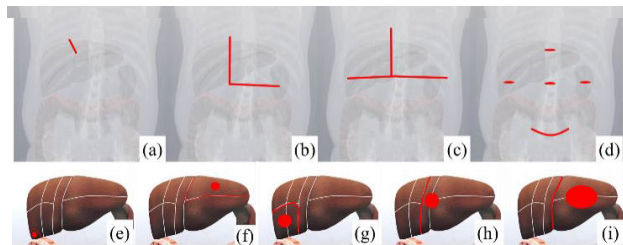

FIGURE 3. Abdomen incision types and liver incision types. (a) minimal incision, (b) j-section window incision, (c) inverted T incisions, (d) laparoscopic, (e) wedge resection, (f) segmentectomy, (g) two-segmentary, (h) tri-segmentectomy, (i) lobectomy.

We employed the generally accepted incision methods for a liver cancer. Hence, the types of abdomen incisions were the minimal, J-section window, inverted T, and laparoscopic incisions (Fig. 3(a–d)), and the types of liver incisions included the wedge resection, segmentectomy, two-segmentary, tri-segmentectomy, and lobectomy (Fig. 3(e–i)). The minimum incision type opens the abdomen around the lesion with a small straight line. The J-section window incision is larger than the minimum incision and creates a J-shaped opening in the abdomen, revealing a liver lobe. An inverse T-shaped incision reveals almost the entire liver. The tumor size determines the liver incision method.

The liver is divided into eight sections, and if a tumor is smaller than half of one section and is inside of it, the tumor can be removed using two incision lines, making a wedge resection. If a tumor is larger than half of one section and is still inside of the section, the section is removed through a segmentectomy incision. When the tumor has invaded one section from another, removing two or three segments is

called a two-segmentectomy or tri-segmentectomy. When one side lobe of the liver is removed, it is called a lobectomy.

We designed nine tasks involving various positions and directions of the incision lines and tumor locations. The experiment included four conditions with two rounds, with a pair of participants switching roles: eight tasks were for the experimental sessions, and one was for a train. During the experiment, the tasks were randomly assigned to the participants. The experiment was held in a room with a 3.5 m × 3.5 m empty space.

TABLE 2. Open-ended questions after a round of four sessions with four conditions.

	Questions
Q1.	Please describe what you like best about this system and the reason.
Q2.	Please describe what you dislike about this system and the reason.
Q3.	Please describe how you felt about the head gaze.

B. CONDITIONS AND PROCEDURE

We assessed four conditions in this user study: None, RSketch, PSketch, and Both. In the "None" condition, the 3-s sketch cue was not provided to either the receiver or provider, but they still could use baseline visual cues, including hand gestures and a head-gaze pointer. In the RSketch and PSketch conditions, the receiver or provider, correspondingly, could sketch and use other baseline visual cues. The "Both" conditions allowed both collaborators to sketch and use baseline visual cues. The experiment was in a within-subject design.

After a pair of participants entered a room, we explained the purpose and procedure of the user study. Once they agreed to the study, they completed a consent form and demographic questionnaire asking about gender, age, and prior experience with VR systems, and we randomly assigned the roles to them. Then, they wore the Oculus Quest 2 and were asked to use sketch and two baseline cues (hand gestures and head-gaze pointers) to communicate each other for manipulating virtual objects. After they felt comfortable using the visual cues and manipulating objects, they conducted a training session to familiarize themselves with the task. The training session included two steps: becoming acquainted with the task and performing the collaborative task. In the acquainted step, an experimenter explained the surgery steps and showed where the tumor and incision lines were on an instruction paper to the provider participant. In the performing task step, the provider explained the surgery steps, and the receiver confirmed the surgery activities with visual cues, as described in Section IV-A. During this performing task step, the provider can refer to the instructions displayed on a board on the right side of the 3D body model, but the receiver cannot see it.

After the training task, participants performed eight experimental sessions with the given conditions and roles. They
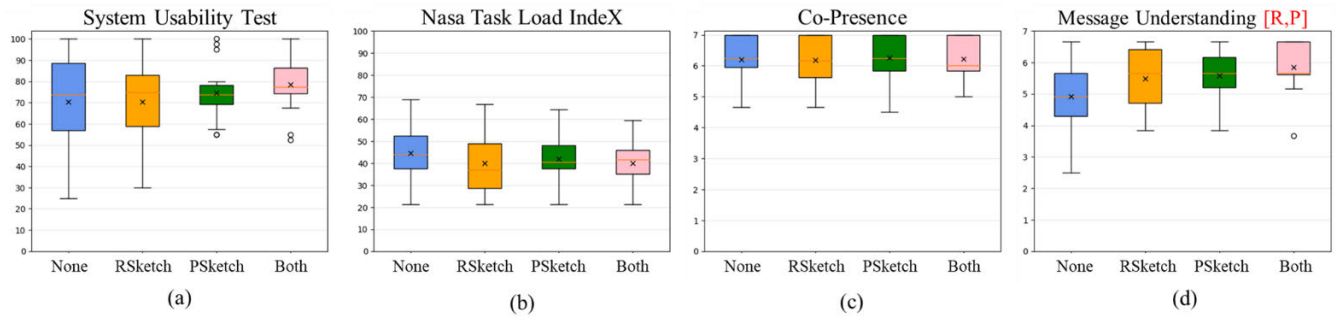
FIGURE 4. Questionnaire results for providers: (a) SUS, (b) NASA TLX, (c) Co-Presence, (d) Message Understanding. (x: mean; R and P: significant effects of the receiver or provider using 3-s sketches).

FIGURE 5. Questionnaire results for receivers: (a) SUS, (B) NASA TLX, (c) Co-Presence, (d) Message understanding. (x: mean; R and P: significant effects of the receiver or provider using 3-s sketches).
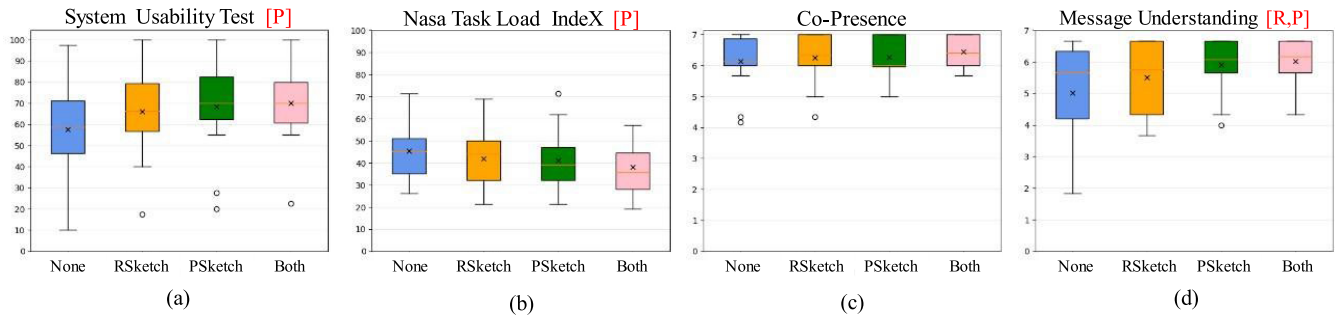
were acquainted with performing the task steps in every experimental session, similar to the training session. After completing the task with a given condition, they answered the questions in the System Usability Survey (SUS) [40], NASA Task Load Index (NASA-TLX) [41], Co-Presence [42], and Message Understanding [42] questionnaires. After completing four experimental sessions with the four conditions, they answered open-ended questions (Table 2) and ranked the conditions according to their preference. The participants switched their roles and performed another round of four sessions with the four conditions. During experiments, we recorded the screen view and collected the task completion time and number of sketch strokes. The experiment took about 1 h and 30 min to 2 h, and participants received a gift certificate worth about ten dollars as a reward.

C. PARTICIPANTS

We recruited 20 participants in pairs, and each pair knew each other as friends because normally the surgeon usually plans a surgery with the colleagues, they know well. There were 11 females and nine males ranging from 20 to 28 years old ($M = 23.12$; $SD = 2.87$). Seventeen participants were moderately familiar with the VR system; the other three had never played VR. We excluded the data from two pairs of participants who did not focus on completing the task but playing together in the VR environment.

V. RESULT

According to the Shapiro–Wilk test, the objective data (task completion time) were not normally distributed for each condition, and subjective questionnaire data were in ordinal scales. Thus, we ran the aligned rank transform as proposed by Wobbrock et al. [43]. We used a two-way repeated measures analysis of variance ($\alpha = .05$). For the number of the sketch strokes, the availability of the sketch cue was different according to the condition; thus, we only compared the provider's number of strokes between PSketch and Both conditions and the receiver's number of strokes between RSketch and Both conditions using the paired t-tests ($\alpha = .05$).

A. QUESTIONNAIRE RESULTS

Figure 4 and 5 summarizes the results from the questionnaires for providers and receivers, correspondingly.

1) CO-PRESENCE

Participants perceived no effect of conditions on the co-presence. The provider did not feel a significant difference in co-presence regardless of the provider ($F(1,15) = 0.01$, $p = .920$) or receiver ($F(1,15) = 0.359$, $p = .558$) using the sketch cue or not. There was no significant interaction effect ($F(1,15) = 0.020, p = .890$) between the use of sketch cues by the two collaborators. Additionally, the receiver did not feel a significant difference in co-presence according to sketch-cue

availability (provider: F(1,15) = 0.947, p =.346; receiver: F(1,15) = 3.902, p =.067), and no interaction effect was found (F(1,15) = 0.190, p =.669).

2) SYSTEM USABILITY SURVEY

The provider did not feel a significant difference in usability according to the use of the sketch cue by the provider (F(1,15) = 3.299, p =.089) and receiver (F(1,15) = 1.279, p =.276). In addition, there was no interaction effect (F(1,15) = 4.027, p =.063). The receiver felt a higher level of usability when the provider used the sketch cue (F(1,15) = 6.608, p =.021, η =.306) compared to not using it, but the results did not indicate a significant difference regarding whether the receiver used the sketch cue or not (F(1,15) = 2.594, p =.128). Moreover, no interaction effect was revealed (F(1,15) = 3.951, p =.065).

3) NASA TASK LOAD INDEX

The results for NASA-TLX were similar to those for SUS. The provider did not feel a significant difference according to the availability of the sketch cue (provider usage: F(1,15) = 1.220, p =.287; receiver usage: F(1,15) = 1.878, p =.191), and no interaction effect was found (F(1,15) = 1.870, p =.191). However, the receiver felt a lower level of task load when the provider used the sketch cue (F(1,15) = 7.259, p =.017, η =.326) compared to not using it. The receiver did not feel a significantly different level of task load according to the receiver's sketch usage (F(1,15) = 2.167, p =.162), and no interaction effect was found (F(1,15) = 0.006, p =.940).

4) MESSAGE UNDERSTANDING

In contrast to the previous results, participants reported a significant effect of the provider and receiver using the 3-s sketch cue. The provider reported a better understanding of messages when the provider and receiver used the sketch cue compared to when they did not (provider: F(1,15) = 21.510, p <.000, η =.589; receiver: F(1,15) = 5.896, p =.028, η =.282). In addition, the receiver reported they felt they better understood the messages when the provider and receiver used the sketch cue (provider: F(1,15) = 12.522, p =.003, η =.455; receiver: F(1,15) = 8.100, p =.012, η =.351). However, no interaction effect was found between the two factors (provider: F(1,15) = 1.597, p =.226; receiver: F(1,15) = 1.088, p =.314).

5) PREFERENCE

Table 3 summarizes the results of preferences. The sketch usage by providers had a significant effect on the preference of both providers (F(1,15) = 43.391, p <.001, η =.743) and receivers (F(1,15) = 55.477, p <.001, η =.787) preferring the conditions where the sketch cues are available. The sketch usage by a receiver also had an effect on the preferences of the providers (F(1,15) = 63.505, p <.001, η =.809 and receivers F(1,15) = 27.262, p <.001, η =.645) preferring having sketch cues. There was no significant interaction effect on the preference of both providers (F(1,15) = 2.459, p =.138) and receivers (F(1,15) = 2.288, p =.151).

B. QUANTITATIVE RESULTS

Figure 6 summarizes the quantitative results from objective measures.

1) NUMBER OF SKETCH STROKES

We compared the number of sketch strokes made by providers between PSketch and Both conditions, and by receivers between RSketch and Both conditions. The comparison exhibited no significant difference in the number of provider's strokes between the PSketch and Both conditions (t(15) = –0.009, p =.993). Similarly, the number of receiver's strokes was not significantly different between RSketch and Both conditions (t(15) = –0.055, p =.957). Both the receiver and provider made a similar number of sketch strokes in Both conditions compared to RSketch (t(15) = 0.00, p = 1) and PSketch (t(15) = 0.35, p =.972) conditions, respectively.

2) COMPLETION TIME

There was no significant difference in the task completion time according to the use of the sketch cue by the provider (F(1,15) = 0.363, p =.567) or receiver (F(1,15) = 0.004, p =.950). Additionally, the interaction effect was not revealed (F(1,15) = 1.653, p =.218).
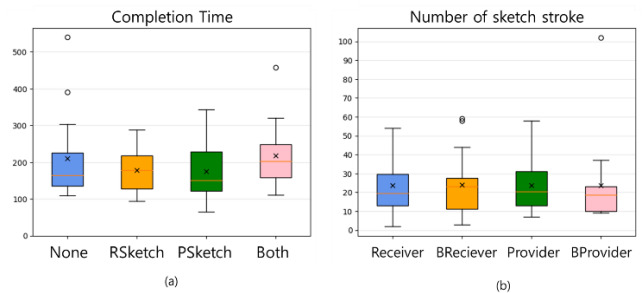

FIGURE 6. Result of Quantitative results. (a) Completion time. (b) Number of sketch stroke (x: mean).

C. RESULTS OF OPEN-ENDED QEUSTIONS AND OBSERVATIONS

This section summarizes the participants' answers to open-ended questions in Table 2.

TABLE 3. Participant ranks for the conditions.

Condition	Receiver [R, P]					Provider [R, P]				
	1st	2nd	3rd	4th	Mean (SD)	1st	2nd	3rd	4th	Mean (SD)
None	2	0	0	14	3.81 (0.75)	1	0	0	15	3.63 (1.02)
Receiver	0	4	12	0	2.56 (0.51)	0	7	9	0	2.75 (0.45)
Provider	2	11	3	0	2.25 (0.89)	2	8	6	0	2.06 (0.87)
Both	12	1	1	2	1.38 (0.89)	13	1	1	1	1.56 (1.09)

1) SKETCH CUE

Participants commented that the sketch cue was crucial for the VR planning surgery task. Nine participants (P1, P2, P4, P7, P8, P11, P13, P14, and P15) reported that visually displaying the incision lines was intuitive and easy to understand. For example, P2 stated, "*The communication was effective with the sketch because it was possible to visually share the exact incision lines,*" and P14 stated, "*The sketch was convenient to directly indicate the region that I want to express.*" The benefits of presenting the incision line were also revealed when the receiver confirmed it. For example, P7 stated, "*It was easy to understand whether the receiver understood my messages or not by the confirming sketches,*" and P11 stated, "*When the sketch cue was not available to the receiver, the receiver showed the incision line with hands, and it was unclear.*"

However, the 3-seconds sketch cue was not best design for the sketch cue in VR planning surgery task even if it had a benefit in collaborators' communication. Some participants (P3, P5, P12, P13, and P14) reported that holding a sketch only for 3 seconds was inconvenience. We observed that the provider drew over the incision lines several times with one stroke for keeping the sketches until the provider felt the receiver understood the sketches. P3, P5, and P12 reported that these continuous drawings were inconvenient and made a thick line and displayed unclear incision line (according to the comment from P13). They suggested erasing the sketch when drawing the next sketch, so the current sketch could be remained until the next phase of communication with a new sketch starts. Moreover, this continuous drawing by the provider also affected the receiver, so the receiver also drew over the line back and forth continuously by following the provider.

Additionally, we observed that participants sometimes verbally explained the incision line after drawing the sketches, which consumed time and might have influenced the task completion time. Specifically, the task completion time increased slightly but not significantly when both participants' sketches were available because, regardless of the reduced message understanding time, participants had to wait for 3 s each time until the sketches are erased after using them to explain next surgical steps.

2) 3D BODY AND CROSS-SECTIONAL VIEW

Many participants were interested in the cross-sectional view. They especially liked that they could observe the CT or MRI scan images in VR using a simple panel interaction. For example, P10 stated, "*It is impressive that I can see both inside and outside of the body just by making simple interactions,*" and P16 stated, "*It was visually intuitive to see the precise location of the scanned images in three dimensions.*" One of the interesting observations was collaborative activities to determine the proper angles and positions of the cross-sectional views. The system supported one cross panel, and the provider usually held and controlled it to view certain angles and positions in the cross-sectional view. Hence, to control the cross-sectional view, some receivers held the 3D body and moved it to set the proper position of the cross-sectional view. For instance, P8 stated, "*Since the provider used the cross-section panel, I simply took the 3D body and positioned it for having proper angle view of the cross section.*" The receiver hardly held the cross-section panel once after the provider started using it.

Additionally, three participants, P3, P6, and P12, complained about not having any annotation attached to the 3D body. For example, P6 stated, "*I wanted to place an annotation on the body, but there was no way to do that.*" The system's three visual cues did not remain or last so could not be used like an annotation or note.

3) HEAD-GAZE POINTERS AND HAND GESTURES

Most participants used hand gestures to represent the incision lines when the sketch cue was unavailable in the None condition. However, it was insufficient to represent the length of the incision line. For example, P4 stated, "*I positioned and shaped a hand to represent an incision line, but it was not good enough to show [the] exact line position and length.*" With the head-gaze pointer, participants could know where the other participant looked at and confirm whether the other participant focused on the collaborative activities. For instance, P9 stated, "*It was nice to know where the other person was looking and whether he/she focused on my explanation or not.*" However, two participants (P2 and P5) complained about the hand and gaze pointer obscuring the view and suggested making the gaze pointer transparent when both participants looked at the same objects or area.

VI. DISCUSSION

This section discusses whether the results of the user experiment align with the hypotheses and what implications they hold. The effect of the sketch cue was evident on the level of understanding of the message in communication between the provider and receiver, and the second hypothesis (3-s sketch cues support better communication between collaborators) is validated.

For the level of co-presence, no significant effect of using the sketch cue was found within the conditions. Therefore, the third hypothesis (collaborators have a higher level of co-presence with the sketch cue) is not supported. For this result, we contemplated the base visual communication cues, including hand gestures, gaze pointer, and avatar, were always available in all four conditions. They already visually represented participants' activities to themselves and their counterparts (regardless of effectiveness of communication); thus, the availability of the sketch cue might not have significantly influenced the participant's feeling of co-presence.

Moreover, the results revealed that the receiver felt a higher level of usability and a lower level of required mental effort when the provider used the sketch cue, but the provider did not feel a significant effect of receiver's sketches on the level of usability and required mental effort. Thus, Hypotheses

1 and 4 (H1: 3-s sketch cues improve system usability, and H4: collaborators spend less mental effort with sketch cues) are partially supported. The different results between the providers and receivers on the usability and mental effort might come from the different levels of information they had. In this user study, the provider had information before starting the task, whereas the receiver did not. Thus, the receiver might need clear and precise representation of the transferred information from the provider to fully understand it from none, and the sketch was suitable for it. In contrast, receiver's sketch might be less important to the provider as the provider was able to understand receiver's hand expression when already knowing the correct answer.

One interesting finding was that the 3-seconds sketch was not the best design of the sketch cue for the VR planning surgery task (even if the sketch cue still played important communication cue). Some participants in our study complained the erasing function after three seconds but it was not reported in the studies of Fussell et al. [16] and Kim et al. [22]. This difference may be because of the task type in the user study. Fussell et al. [16] and Kim et al. [22] prepared object assembly tasks, such as Lego blocks and Tangrams, and simple position and orientation information was required. For the position and orientation information of Lego blocks, 3 seconds visibility of the sketch cue might be sufficient to understand and perform assembly. However, we prepared a complex surgery-planning task, including explanation of the surgery steps and the receiver's remembering incision lines (to make sure both collaborators correctly understand the surgery plan), hence the participants preferred the sketches to be displayed until the time they ensured whether the counterpart well understand and remember the sketch communication.

This study has a few limitations. First, the proposed prototype does not cover all tasks for planning surgery; thus, it may differ from real-world one. However, we designed the five steps in our experimental task according to surgeons' opinions in the field. Additionally, this system is limited to a two-user collaboration, but it could easily be extended to a system with more than two users. Moreover, the participants in this study were not surgeons who could be target users. The effects, however, may be marginal because the sketch cue is a common visual communication cue in daily life; therefore, participants may have similar attitudes toward the sketch cue regardless of their profession. In addition, the interface design was proposed by surgeons.

We make the following design recommendations for planning surgery VR application from the surgeons' comments, literature reviews, and our study results. First, a 3D reconstructed body model should be supported based on real-world data, such as CT and MRI images. Second, the cross-sectional view at any angle and position of the body should be supported for examining inside the body and organs. Third, visual communication cues, such as sketches, pointers, and hand gestures, are crucial for communication. Fourth, the erasing sketch should occur after ensuring that the counterpart

well understand and remember the sketch communication. Since collaborators might continue next communication after ensuring it, we recommend erasing a sketch when the next sketch starts being drawn. Fifth, the system should support permanent annotations (like a colored pins by Bork et al. [8]), especially attached to the 3D body, together with the volatile visual communication cues, such as hand gestures, erased sketches, and head-gaze pointers. Sixth, when both collaborators focus on the same object or area, the system should control the transparency of the hand gestures and head-gaze pointers to mitigate the effect of occlusions.

VII. CONCLUSION

In this paper, we explored the effect of the 3-second sketch cue for multiuser surgery planning tasks in VR, where head-gaze pointers and hand-gesture visual communication cues were available as baseline cues. In the user study with the scenario where one collaborator provided a surgical planning idea to another, we found that the 3-s sketch cue supports better communication for collaboration. For better design of the sketch cue, erasing sketch after drawing next was suggested. Additionally, we proposed five more suggestions to design the planning surgery VR application.

In the future, we aim to extend this study to an approach with more than two collaborators and explore medical training tasks.

REFERENCES

[1] S. Zhang, C. Demiralp, D. F. Keefe, M. DaSilva, D. H. Laidlaw, B. D. Greenberg, P. J. Basser, C. Pierpaoli, E. A. Chiocca, and T. S. Deisboeck, "An immersive virtual environment for DT-MRI volume visualization applications: A case study," in *Proc. Visualization*, Oct. 2001, pp. 437–584, doi: 10.1109/VISUAL.2001.964545.

[2] J. Egger, M. Gall, J. Wallner, P. Boechat, A. Hann, X. Li, X. Chen, and D. Schmalstieg, "HTC vive MeVisLab integration via OpenVR for medical applications," *PLoS ONE*, vol. 12, no. 3, Mar. 2017, Art. no. e0173972, doi: 10.1371/journal.pone.0173972.

[3] M. Pfeiffer, H. Kenngott, A. Preukschas, M. Huber, L. Bettscheider, B. Müller-Stich, and S. Speidel, "IMHOTEP: Virtual reality framework for surgical applications," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 5, pp. 741–748, May 2018, doi: 10.1007/s11548-018-1730-x.

[4] S. Reddivari, J. Smith, and J. Pabalate, "VRvisu: A tool for virtual reality based visualization of medical data," in *Proc. IEEE/ACM Int. Conf. Connected Health, Appl., Syst. Eng. Technol. (CHASE)*, Jul. 2017, pp. 280–281, doi: 10.1109/CHASE.2017.102.

[5] L. D. J. Fiederer, H. Alwanni, M. Völker, O. Schnell, J. Beck, and T. Ball, "A research framework for virtual-reality neurosurgery based on opensource tools," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Mar. 2019, pp. 922–924, doi: 10.1109/VR.2019.8797710.

[6] L. Greuter, A. De Rosa, P. Cattin, D. M. Croci, J. Soleman, and R. Guzman, "Randomized study comparing 3D virtual reality and conventional 2D on-screen teaching of cerebrovascular anatomy," *Neurosurgical Focus*, vol. 51, no. 2, p. E18, Aug. 2021, doi: 10.3171/2021.5.FOCUS21212.

[7] T. Huber, M. Paschold, C. Hansen, T. Wunderling, H. Lang, and W. Kneist, "New dimensions in surgical training: Immersive virtual reality laparoscopic simulation exhilarates surgical staff," *Surgical Endoscopy*, vol. 31, no. 11, pp. 4472–4477, Nov. 2017, doi: 10.1007/s00464-017-5500-6.

[8] F. Bork, A. Lehner, D. Kugelmann, U. Eck, J. Waschke, and N. Navab, "VesARlius: An augmented reality system for large-group co-located anatomy learning," in *Proc. IEEE Int. Symp. Mixed Augmented Reality Adjunct (ISMAR-Adjunct)*, Oct. 2019, pp. 122–123, doi: 10.1109/ISMAR-Adjunct.2019.00-66.

[9] R. Kurul, M. N. Ögün, A. N. Narin, Ş. Avci, and B. Yazgan, "An alternative method for anatomy training: Immersive virtual reality," *Anatomical Sci. Educ.*, vol. 13, no. 5, pp. 648–656, Sep. 2020, doi: 10.1002/ase.1959.

[10] C. Pinter, A. Lasso, S. Choueib, M. Asselin, J.-C. Fillion-Robin, J.-B. Vimort, K. Martin, M. A. Jolley, and G. Fichtinger, "SlicerVR for medical intervention training and planning in immersive virtual reality," *IEEE Trans. Med. Robot. Bionics*, vol. 2, no. 2, pp. 108–117, May 2020, doi: 10.1109/TMRB.2020.2983199.

[11] K. Yu, A. Winkler, F. Pankratz, M. Lazarovici, D. Wilhelm, U. Eck, D. Roth, and N. Navab, "Magnoramas: Magnifying dioramas for precise annotations in asymmetric 3D teleconsultation," in *Proc. IEEE Virtual Reality 3D User Interfaces (VR)*, Mar. 2021, pp. 392–401, doi: 10.1109/VR50410.2021.00062.

[12] K. Yu, U. Eck, F. Pankratz, M. Lazarovici, D. Wilhelm, and N. Navab, "Duplicated reality for co-located augmented reality collaboration," *IEEE Trans. Vis. Comput. Graphics*, vol. 28, no. 5, pp. 2190–2200, May 2022, doi: 10.1109/TVCG.2022.3150520.

[13] M. Vincze, M. D. Kucarov, B. Biricz, A. Benhamida, M. Ogbolu, M. Kozlovszky, V. Jònàs, and R. Paulik, "Real-time multi-user 3D visualization software in medicine," in *Proc. IEEE 22nd Int. Symp. Comput. Intell. Informat. 8th IEEE Int. Conf. Recent Achievements Mechatronics, Autom., Comput. Sci. Robot. (CINTI-MACRo)*, Nov. 2022, pp. 45–50, doi: 10.1109/CINTI-MACRo57952.2022.10029449.

[14] V. Chheang, D. Schott, P. Saalfeld, L. Vradelis, T. Huber, F. Huettl, H. Lang, B. Preim, and C. Hansen, "Towards virtual teaching hospitals for advanced surgical training," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Abstr. Workshops (VRW)*, Mar. 2022, pp. 410–414, doi: 10.1109/VRW55335.2022.00089.

[15] D. Schott, P. Saalfeld, G. Schmidt, F. Joeres, C. Boedecker, F. Huettl, H. Lang, T. Huber, B. Preim, and C. Hansen, "A VR/AR environment for multi-user liver anatomy education," in *Proc. IEEE Virtual Reality 3D User Interfaces (VR)*, Mar. 2021, pp. 296–305, doi: 10.1109/VR50410.2021.00052.

[16] S. R. Fussell, L. D. Setlock, J. Yang, J. Ou, E. Mauer, and A. D. I. Kramer, "Gestures over video streams to support remote collaboration on physical tasks," *Hum.-Computer Interact.*, vol. 19, no. 3, pp. 273–309, Sep. 2004, doi: 10.1207/s15327051hci1903_3.

[17] S. Kim, G. Lee, N. Sakata, and M. Billinghurst, "Improving co-presence with augmented visual communication cues for sharing experience through video conference," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Sep. 2014, pp. 83–92, doi: 10.1109/ISMAR.2014.6948412.

[18] S. Kim, M. Billinghurst, and K. Kim, "Multimodal interfaces and communication cues for remote collaboration," *J. Multimodal User Interfaces*, vol. 14, no. 4, pp. 313–319, Dec. 2020, doi: 10.1007/s12193-020-00346-8.

[19] A. Jing, K. Gupta, J. McDade, G. A. Lee, and M. Billinghurst, "Comparing gaze-supported modalities with empathic mixed reality interfaces in remote collaboration," in *Proc. IEEE Int. Symp. Mixed Augmented Reality (ISMAR)*, Oct. 2022, pp. 837–846, doi: 10.1109/ISMAR55827.2022.00102.

[20] D. Datcu, M. Cidota, H. Lukosch, and S. Lukosch, "On the usability of augmented reality for information exchange in teams from the security domain," in *Proc. IEEE Joint Intell. Secur. Informat. Conf.*, Sep. 2014, pp. 160–167, doi: 10.1109/JISIC.2014.32.

[21] J. Hauber, "Understanding remote collaboration in video collaborative virtual environments," Ph.D. dissertation, Univ. Canterbury Press, Christchurch New Zealand, 2008.

[22] S. Kim, G. Lee, W. Huang, H. Kim, W. Woo, and M. Billinghurst, "Evaluating the combination of visual communication cues for HMD-based mixed reality remote collaboration," Presented at the CHI Conf. Hum. Factors Comput. Syst., Glasgow, U.K., May 2019, doi: 10.1145/3290605.3300403.

[23] T. Ard, D. M. Krum, T. Phan, D. Duncan, R. Essex, M. Bolas, and A. Toga, "NIVR: Neuro imaging in virtual reality," in *Proc. IEEE Virtual Reality (VR)*, Mar. 2017, pp. 465–466, doi: 10.1109/VR.2017.7892381.

[24] C. Herfarth, W. Lamadé, L. Fischer, P. Chiu, C. Cardenas, M. Thorn, M. Vetter, L. Grenacher, and H. P. Meinzer, "The effect of virtual reality and training on liver operation planning," *Swiss Surg.*, vol. 8, no. 2, pp. 67–73, Apr. 2002, doi: 10.1024/1023-9332.8.2.67.

[25] D. M. Croci, R. Guzman, C. Netzer, L. Mariani, S. Schaeren, P. C. Cattin, and G. F. Jost, "Novel patient-specific 3D-virtual reality visualisation software (SpectoVR) for the planning of spine surgery: A case series of eight patients," *BMJ Innov.*, vol. 6, no. 4, pp. 215–219, Oct. 2020.

[26] A. Lasso, H. H. Nam, P. V. Dinh, C. Pinter, J.-C. Fillion-Robin, S. Pieper, S. Jhaveri, J.-B. Vimort, K. Martin, M. Asselin, F. X. McGowan, R. Kikinis, G. Fichtinger, and M. A. Jolley, "Interaction with volume-rendered three-dimensional echocardiographic images in virtual reality," *J. Amer. Soc. Echocardiogr.*, vol. 31, no. 10, pp. 1158–1160, Oct. 2018, doi: 10.1016/j.echo.2018.06.011.

[27] L. He, A. Guayaquil-Sosa, and T. McGraw, "Medical image atlas interaction in virtual reality," in *Proc. IEEE Vis. Immersive Anal. Workshop*, 2017. [Online]. Available: http://immersiveanalytics.net

[28] W. Luo, E. Goebel, P. Reipschläger, M. O. Ellenberg, and R. Dachselt, "Exploring and slicing volumetric medical data in augmented reality using a spatially-aware mobile device," in *Proc. IEEE Int. Symp. Mixed Augmented Reality Adjunct (ISMAR-Adjunct)*, Oct. 2021, pp. 334–339, doi: 10.1109/ISMAR-Adjunct54149.2021.00076.

[29] S. Kim, M. Billinghurst, and G. Lee, "The effect of collaboration styles and view independence on video-mediated remote collaboration," *Comput. Supported Cooperat. Work*, vol. 27, nos. 3–6, pp. 569–607, Dec. 2018, doi: 10.1007/s10606-018-9324-2.

[30] S. Chen, M. Chen, A. Kunz, A. E. Yantaç, M. Bergmark, A. Sundin, and M. Fjeld, "SEMarbeta: Mobile sketch-gesture-video remote support for car drivers," Presented at the 4th Augmented Hum. Int. Conf., Stuttgart, Germany, Mar. 2013, doi: 10.1145/2459236.2459249.

[31] A. Kunz, T. Nescher, and M. Küchler, "CollaBoard: A novel interactive electronic whiteboard for remote collaboration with people on content," in *Proc. Int. Conf. Cyberworlds*, Oct. 2010, pp. 430–437, doi: 10.1109/CW.2010.17.

[32] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst, "Mini-me: An adaptive avatar for mixed reality remote collaboration," Presented at the CHI Conf. Hum. Factors Comput. Syst., Montreal QC, Canada, Apr. 2018, doi: 10.1145/3173574.3173620.

[33] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst, "The effects of sharing awareness cues in collaborative mixed reality," *Frontiers Robot. AI*, vol. 6, p. 5, Feb. 2019, doi: 10.3389/frobt.2019.00005.

[34] G. A. Lee, S. Kim, Y. Lee, A. Dey, T. Piumsomboon, M. Norman, and M. Billinghurst, "Improving collaboration in augmented video conference using mutually shared gaze," in *Proc. ICAT-EGVE*, 2017, pp. 197–204.

[35] A. Jing, K. May, B. Matthews, G. Lee, and M. Billinghurst, "The impact of sharing gaze behaviours in collaborative mixed reality," *Proc. ACM Hum.-Comput. Interact.*, vol. 6, Nov. 2022, Art. no. 463, doi: 10.1145/3555564.

[36] W. Huang, S. Kim, M. Billinghurst, and L. Alem, "Sharing hand gesture and sketch cues in remote collaboration," *J. Vis. Commun. Image Represent.*, vol. 58, pp. 428–438, Jan. 2019, doi: 10.1016/j.jvcir.2018.12.010.

[37] H. Tian, G. A. Lee, H. Bai, and M. Billinghurst, "Using virtual replicas to improve mixed reality remote collaboration," *IEEE Trans. Vis. Comput. Graphics*, vol. 29, no. 5, pp. 2785–2795, May 2023, doi: 10.1109/TVCG.2023.3247113.

[38] *Photon Fusion*. [Online]. Available: https://www.photonengine.com/fusion

[39] *Photon Voice*. [Online]. Available: https://www.photonengine.com/voice

[40] J. Brooke, "Sus: A 'quick and dirty' usability," *Usability Eval. Ind.*, vol. 189, no. 3, pp. 189–194, 1996.

[41] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," in *Advances in Psychology*, vol. 52, P. A. Hancock and N. Meshkati Eds. Amsterdam, The Netherlands: North Holland, 1988, pp. 139–183.

[42] C. Harms and F. Biocca, "Internal consistency and reliability of the networked minds measure of social presence," in *Proc. 7th Annu. Int. Workshop, Presence*. Valencia, Spain: Universidad Politecnica de Valencia, 2004, pp. 1–7.

[43] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, "The aligned rank transform for nonparametric factorial analyses using only anova procedures," Presented at the SIGCHI Conf. Hum. Factors Comput. Syst., Vancouver, BC, Canada, May 2011, doi: 10.1145/1978942.1978963.

SUYEON PARK (Student Member, IEEE) received the B.S. degree in computer and information engineering from Chonnam National University, South Korea, in 2022, where she is currently pursuing the M.S. degree in artificial intelligence convergence.

Her research interests include virtual reality and augmented reality, human–computer interaction, and generative artificial intelligence.

GAYUN SUH was born in Hayward, USA. She is currently pursuing the B.S. degree in artificial intelligence with Chonnam National University, Gwangju, South Korea.

She is also an Undergraduate Researcher with the Empathic Computing Laboratory, Chonnam National University, focusing on projects that integrate machine learning into augmented and virtual reality, with a predominant interest in AI applications within these domains.

SOO-HYUNG KIM received the B.S. degree in computer engineering from Seoul National University, in 1986, and the M.S. and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology, in 1988 and 1993, respectively.

Since 1997, he has been a Professor with the Department of Artificial Intelligence Convergence, Chonnam National University, South Korea. His research interests include pattern recognition, document image processing, medical image processing, and ubiquitous computing.

HYUNG-JEONG YANG received the B.S., M.S., and Ph.D. degrees from Chonbuk National University, South Korea.

She is currently a Professor with the Department of Artificial Intelligence Convergence, Chonnam National University, Gwangju, South Korea. Her main research interests include multimedia data mining, medical data analysis, social network service data mining, and video data understanding.

GUN LEE received the Ph.D. degree in computer science and engineering from POSTECH, in 2009, researching immersive authoring methods for creating virtual reality (VR) and augmented reality (AR) contents, developing the first system in the world that allowed people to create AR applications within the AR environment.

He is currently a Senior Research Scientist with the Empathic Computing Laboratory, University of South Australia, researching interaction and visualization methods for mobile and wearable AR and VR systems. He has produced over 60 publications in the areas, such as augmented reality and virtual reality. He has been with ETRI (2005–2011) as a Senior Researcher, where he developed VR and AR systems for industrial applications.

SEUNGWON KIM received the B.S. and M.S. degrees from the University of Tasmania, in 2008 and 2010, respectively, and the Ph.D. degree from the HIT Laboratory NZ, New Zealand, in 2016, under the supervision of Prof. M. Billinghurst.

During the Ph.D. degree, he developed a remote collaboration system supporting stabilized sketch and pointer cues. He is currently a Professor with the Department of Artificial Intelligence Convergence, Chonnam National University, South Korea. His research interests include remote collaboration using augmented virtual communication cues and sharing experiences between distance users.

● ● ●