

論文 / 著書情報  
Article / Book Information

題目(和文)	
Title(English)	A Study of 360-Degree I/O System by Integrating Spherical Display with Omnidirectional Camera For Panoramic Video Communication
著者(和文)	LiZhengQing
Author(English)	Zhengqing Li
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11621号, 授与年月日:2020年9月25日, 学位の種別:課程博士, 審査員:小池 英樹,徳永 健伸,篠田 浩一,下坂 正倫,伊藤 勇太
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11621号, Conferred date:2020/9/25, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Doctor Thesis

A Study of 360-Degree I/O System by Integrating  
Spherical Display with Omnidirectional Camera  
For Panoramic Video Communication

Tokyo Institute of Technology

School of Computing

Department of Computer Science

Name : Zhengqing Li

Main Supervisor : Professor Hideki Koike

Submission Date : July 2020

# Abstract

This research challenged one of the long-standing issues of video communication systems, that is, the narrow camera field of view. With the advancement of omnidirectional panoramic technology, the omnidirectional camera, which can provide a 360° field of view, has become increasingly popular in the last few years. Previous research indicated that one-way video communication systems with a wider field of view improve task efficiency. Therefore, this research proposes to utilize omnidirectional cameras in a symmetrical video communication system. By proposing novel 360° I/O systems, named OmniEyeball, this research studies how a new camera-display configuration supports symmetrical video communication and remote collaboration.

In the first part of this study, an experiment was conducted based on two conditions: an omnidirectional camera with a hemispherical display (OmniEyeball system) and an omnidirectional camera with a horizontally placed 2D flat display. Under these conditions, the study analyzed how the display types affected video communication and remote collaboration. The results show that participants marginally preferred the spherical display to the 2D flat display. The results also present the advantages and disadvantages of each display. The findings from the study of OmniEyeball contribute to the understanding of how to design a 360° I/O system for remote collaboration that captures and shows a 360° panoramic view of a remote site.

However, the OmniEyeball system still had several issues, e.g., the invisibility of the hemisphere at the opposite side of a user caused the inconvenience of observing the remote environment. To solve these issues, the second part of this thesis introduces OmniEyeball2 system. This novel symmetric full 360° video communication system incorporates an omnidirectional camera, a full spherical display, and several visual or interactive techniques. An experiment indicates that the OmniEyeball2 system effectively reduces the inconvenience of observing the remote environment and increases the remote space awareness and user's gaze awareness to support remote collaboration. The study

also discusses the takeaways, limitations, and application areas in the OmniEyeball2 system, which help improve the system.

In the third part, the thesis proposes a novel collaboration tool named OmniEyeballVR, an asymmetric cooperation application of the OmniEyeball project that supports communication and collaboration between a VR user (occupant) and multiple non-VR users (designers) across the virtual and physical platform. OmniEyeballVR allows designer(s) to explore the VR space from any point of view using two view modes: a 360° first-person mode and a third-person mode. Besides, a shared gaze awareness cue is provided to enhance further communication between the occupant and the designer(s). Finally, the system has a face window feature that allows designer(s) to share their facial expressions and upper body view with the occupant to exchange and express information using nonverbal cues. A user study was conducted to evaluate the OmniEyeballVR, comparing three conditions: (1) first-person mode with the face window, (2) first-person mode with a solid window, and (3) third-person mode with the face window. The results show that the first-person mode with the face window required significantly less mental effort, and provided better spatial presence, usability, and understanding of the partner's focus.

Finally, the thesis summarizes its contributions and discusses design implications and directions for future research of the 360° video communication system.

# Contents

<b>List of Figures</b>	<b>vi</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	1
1.1.1 Limitation of Video Communication Systems . . . . .	1
1.1.2 Omnidirectional Panoramic Technology . . . . .	2
1.2 Research Motivation . . . . .	3
1.2.1 Evolution of I/O Devices for Video Communication . . . . .	4
1.2.2 Spherical Display Systems . . . . .	5
1.2.3 Novel 360-Degree I/O Device for Video Communication . . . . .	7
1.3 Organization . . . . .	7
<b>Chapter 2 Related Work</b>	<b>9</b>
2.1 Effect of Field-of-view in Video Communication . . . . .	9
2.2 Spherical Image Display System . . . . .	14
<b>Chapter 3 Research Proposal - OmniEyeball</b>	<b>19</b>
3.1 Research Approach . . . . .	19
3.1.1 Implementation of the 360° I/O System . . . . .	19
3.1.2 Design OmniEyeball Prototype . . . . .	21
3.1.3 Evaluation of the 360° I/O System for Video Communication . . . . .	25
3.1.4 Exploration for Applications of 360° I/O System . . . . .	26
<b>Chapter 4 OmniEyeball: A Hemispherical 360° I/O System to Study Impact of Display Shapes on Video Communication</b>	<b>27</b>
4.1 Preface . . . . .	27
4.2 System Implementation of OmniEyeball . . . . .	27

.....

4.2.1	Hardware Overview . . . . .	27
4.2.2	Software Overview . . . . .	29
4.2.3	Reversed Space . . . . .	30
4.3	Experiment . . . . .	33
4.3.1	Method . . . . .	33
4.3.2	Conditions . . . . .	35
4.3.3	Apparatus . . . . .	37
4.3.4	Participants . . . . .	37
4.3.5	Task . . . . .	38
4.3.6	Procedure . . . . .	39
4.3.7	Measures . . . . .	40
4.4	Results . . . . .	42
4.4.1	Collaboration Efficiency . . . . .	42
4.4.2	Questionnaire Results . . . . .	44
4.4.3	Findings from Participants' Comments . . . . .	45
4.4.4	Findings from Observational Analysis . . . . .	45
4.5	Discussion . . . . .	50
4.5.1	Visual Perception, Awareness, and Perceived Workload . . . . .	50
4.5.2	Affordances of Displays . . . . .	53
4.5.3	Proper Real-World Scenarios . . . . .	53
4.5.4	Impact of the Reversed Space . . . . .	54
4.6	Limitations and Solutions . . . . .	55
4.7	Summary . . . . .	56

**Chapter 5 OmniEyeball2: A Full Spherical Interactive I/O System For Symmetric 360° Video Communication** **57**

5.1	Preface . . . . .	57
5.2	Improvement Proposal . . . . .	59
5.3	System Implementation of OmniEyeball2 . . . . .	60

.....

- 5.3.1 Full 360° System: Sharing the Lower Half of the Space . . . . . 60
- 5.3.2 360° Video Communication Application . . . . . 62
- 5.3.3 Physical Rotation: Providing Convenient Access to Any Side of  
the Sphere . . . . . 65
- 5.3.4 OmniBar: Supporting A 360° Overview . . . . . 65
- 5.3.5 User Detection . . . . . 69
- 5.3.6 Touch Detection . . . . . 71
- 5.3.7 Quick View: Supporting Focused View & Gaze Awareness . . . . 73
- 5.4 Experiment . . . . . 74
  - 5.4.1 Conditions . . . . . 74
  - 5.4.2 Task Overview . . . . . 76
  - 5.4.3 Participants . . . . . 77
  - 5.4.4 Environment . . . . . 77
  - 5.4.5 Task Procedure . . . . . 77
  - 5.4.6 Measures . . . . . 81
- 5.5 Results . . . . . 82
  - 5.5.1 Objective Results . . . . . 82
  - 5.5.2 Questionnaire Results . . . . . 82
  - 5.5.3 Interview Results . . . . . 84
- 5.6 Discussion . . . . . 84
  - 5.6.1 Fundamental Advantages of OEB2 . . . . . 84
  - 5.6.2 Comparison between Two Conditions . . . . . 85
  - 5.6.3 Perceived Workload . . . . . 85
  - 5.6.4 Affordance . . . . . 86
  - 5.6.5 Effect of OmniBar on Remote Space Awareness . . . . . 87
  - 5.6.6 Supporting Gaze Awareness . . . . . 89
  - 5.6.7 Design Implications . . . . . 90
  - 5.6.8 Possible Application Scenarios . . . . . 91
- 5.7 Limitations . . . . . 91

.....

5.7.1	Limitations of the User Study . . . . .	92
5.7.2	Insufficient Resolution . . . . .	92
5.7.3	Conflict between Physical Rotation and Quick View Use . . . . .	92
5.8	Improvement Suggestions . . . . .	93
5.8.1	OmniEyeball4K . . . . .	93
5.8.2	OmniBar . . . . .	93
5.8.3	Quick View . . . . .	94
5.8.4	360° Audio Support . . . . .	94
5.9	Summary . . . . .	95

**Chapter 6 OmniEyeballVR: A Collaborative 360° Video Communication Application for VR** **96**

6.1	Background and Motivation . . . . .	96
6.1.1	Contribution . . . . .	98
6.2	Related Work . . . . .	100
6.2.1	Collaborative VR/MR Systems . . . . .	100
6.2.2	Communication Systems with Spherical Display . . . . .	101
6.3	System Overview of OmniEyeballVR . . . . .	101
6.3.1	Hardware . . . . .	101
6.3.2	Software . . . . .	103
6.4	Features . . . . .	103
6.4.1	360° First-person Mode . . . . .	105
6.4.2	Third-person Mode . . . . .	106
6.4.3	Face Window in the VR Space . . . . .	108
6.5	Experiment . . . . .	111
6.5.1	Method . . . . .	111
6.5.2	Participants . . . . .	111
6.5.3	Task . . . . .	111
6.5.4	Conditions . . . . .	112

.....

6.5.5	Procedure . . . . .	113
6.6	Results . . . . .	114
6.6.1	User Performance . . . . .	114
6.6.2	User Experience . . . . .	115
6.6.3	Spatial Presence . . . . .	117
6.6.4	Mental Effort . . . . .	117
6.6.5	System Usability . . . . .	117
6.6.6	Post-experiment Results . . . . .	118
6.7	Discussion . . . . .	120
6.7.1	Face Window . . . . .	120
6.7.2	Visual Feedback of Face Window . . . . .	121
6.7.3	First-person and Third-person Mode . . . . .	122
6.8	Limitations and Solutions . . . . .	123
6.9	Summary . . . . .	124
 <b>Chapter 7 Discussion</b>		 <b>126</b>
7.1	Research Goals . . . . .	126
7.2	Research Contributions . . . . .	127
7.2.1	Insights and Findings . . . . .	127
7.2.2	Innovation . . . . .	129
7.3	Design Implication of 360° I/O Systems . . . . .	130
7.3.1	The Synchronized Rotation to Support Gaze Awareness . . . . .	130
7.3.2	Improvement for Reversed Space . . . . .	131
7.4	Limitations . . . . .	133
7.4.1	Vertical Gaze Misalignment . . . . .	133
7.4.2	Privacy Issue . . . . .	134
7.4.3	Limited Use Scenario . . . . .	134
7.4.4	Communications among Three Spaces . . . . .	135
7.5	A Vision for the Future . . . . .	135

**CONTENTS**

.....

7.5.1 Commercial Prospects . . . . . 135

7.5.2 Mobile 360° Video Communication Application . . . . . 137

7.5.3 Computer Vision with 5G Cloud Computing . . . . . 139

7.5.4 VR Communication Scenario . . . . . 140

**Chapter 8 Conclusion 141**

**Bibliography 145**

## List of Figures

1.1	The Theta series 360° cameras produced by RICOH . . . . .	2
1.2	Samples of Omnidirectional cameras produced by Insta360. . . . .	3
1.3	The motivation of this research is to propose a new camera-display configuration. . . . .	3
1.4	Some representative graphics I/O terminal forms. . . . .	4
1.5	Left: Darling VR Smartphone, World's first smartphone with built-in 360° camera. Right: A panoramic video communication setup by using an Insta360Air 360° camera with a laptop. . . . .	5
1.6	Currently how we watch 360-degree panoramic images or videos: VR HMD - HTC Vive (upper left), Dome display (upper right), Cave display (lower left), Twister rotating display (lower right) . . . . .	6
2.1	The VideoWindow teleconferencing system [21] proposed by Fish et al.. . . . .	9
2.2	One is not enough [25]: the MTV set-up by Gaver et al.. . . . .	10
2.3	Virtual Window [26] proposed by Gaver et al.. . . . .	10
2.4	The Tobi system used in Show Objects [47]. . . . .	11
2.5	Can You See Me Now? [38]. . . . .	11
2.6	JackIn Head [41] proposed by Kasahara et al.. . . . .	12
2.7	ScalableBody [53] proposed by Mutsuda et al.. . . . .	12
2.8	Collaboration with 360° Videochat: Challenges and Opportunities. [74] . . . . .	13
2.9	SharedSphere [44] proposed by Lee et al.. . . . .	13
2.10	CamBlend system [58] proposed by Norris et al.. . . . .	14
2.11	t-Room system [86] proposed by Yamashita et al.. . . . .	15
2.12	GEO-COSMOS [51], the world's first full color spherical display, in Japan. . . . .	15
2.13	PufferSphere touch spherical display. . . . .	16
2.14	Sphere [4] proposed by Benko et al.. . . . .	17

.....

2.15	Comparing Flat and Spherical Displays in a Trust Scenario in Avatar-Mediated Interaction. [62] . . . . .	17
2.16	The Qoom system [69] proposed by Miyafuji et al.. . . . .	18
3.1	The blueprint of the proposed 360° I/O device. . . . .	20
3.2	The OmniEyeball (OEB) prototype system proposed in 2016 [46]. . . . .	22
3.3	Diagram of dynamic spherical projection mapping. . . . .	23
3.4	Diagram of 360° panoramic video communication. . . . .	24
4.1	Diagram of the 360° panoramic video communication. The input of the system is from the Kodak camera. The output is either shown on the remote spherical display or the remote 2D flat display (Left: OmniEyeball version; Right: 2D display version). . . . .	28
4.2	The diagram of the panoramic video communication system. Persons B and C are standing across from each other. . . . .	29
4.3	The top view diagram of the reversed space phenomenon. . . . .	31
4.4	Two different mapping methods of the 360° spherical images and the reason behind the reversed space. . . . .	32
4.5	The 360° panoramic video communication systems. Two geographically separated rooms are connected by two different types of 360° panoramic video communication systems, a combination of omnidirectional camera and a spherical (left column) or a horizontally placed 2D flat display (right column). . . . .	34
4.6	The physical layout of both two rooms. . . . .	35
4.7	The real environment of the room when using the spherical display or the 2D flat display. . . . .	36
4.8	The detailed information of the 40 participants. . . . .	37
4.9	Two sets of patterns used in the task. . . . .	39
4.10	Seven questions in the questionnaire. . . . .	41

.....

4.11 Mean scores of questions from Q1 to Q5. “*” and “***” denote significant difference at $p < 0.05$ level and $< 0.01$ level. The error bars depict the standard deviation. For the scores of Q1-Q3, the higher the better. For the scores of Q4-Q5, the lower the better. . . . .	43
4.12 The statistical results of Q6 and Q7. . . . .	45
4.13 The pros and cons of using the spherical display in the panoramic video communication. . . . .	46
4.14 The pros and cons of using the 2D flat display in the panoramic video communication. . . . .	47
4.15 Communication when standing on opposite sides: The local participant directly puts the image on the opposite side to show it to the invisible remote partner (a). By seeing the inverted image, the local participant directly talked to the remote partner who was standing on the opposite side (b). . . . .	48
4.16 The camera occlusion issue was observed in the 2D flat display condition (a)(b)(c). But in the spherical display condition, this issue was not found and participants could naturally show figures (d). . . . .	49
5.1 The OmniEyeball2 system which integrates an omnidirectional camera with a full spherical display (left), and its symmetric 360° panoramic video communication application (right). . . . .	58
5.2 The issue of the invisible hemisphere in 360° video communication. a) The local OmniEyeball2 is surrounded by four people. b) From any viewpoint in the remote site, only one or two people can be seen at once. . . .	59
5.3 The graphic system takes the equirectangular live streaming video(left) as an input, remaps and resizes it onto an azimuthal equidistant projection image(right) for projecting to the spherical display. . . . .	61
5.4 How to convert a point P’s position from equirectangular map to round map. . . . .	62

## LIST OF FIGURES

x

5.5	The diagram of the 360° video communication based on OEB2. . . . .	63
5.6	The theoretical required bandwidths for different configurations of OEB system. . . . .	64
5.7	The Physical Rotation feature of OmniEyeball2 . . . . .	66
5.8	The 360° image captured by OEB2 is segmented into three image areas. .	67
5.9	a) The OmniBar is created by duplicating the main area of image input in the horizontal direction. b) The OmniBar projected on the top of OEB2. c) The Location bar can show the position of the current viewpoint in the OmniBar. . . . .	68
5.10	The overview of touch detection. a) A finger is touching the top area of OEB2. b) Image input of OEB2. The fingertip is captured inside the touch area. c) The RGB image of the touch area. The red point shows the position of a detected fingertip. d) Binarized image of the touch area after extracting the skin color. . . . .	71
5.11	The features of OEB2, the OmniGlobe is a nickname of OEB2. abcd) Quick View. ef) Quick View sharing window. g) The feedback when the Quick View sharing window is activated. h)when the Quick View sharing window is NOT activated. . . . .	75
5.12	The object arrangement map used in the arranging task. . . . .	76
5.13	The detailed information of the Twenty-four participants. . . . .	78
5.14	The real environment of the two rooms in the experiment. . . . .	79
5.15	One set of images showing 10 objects used in the experiment. . . . .	80
5.16	Eight questions in the questionnaire designed on a 6-point scale (1 = Strongly Disagree to 6 = Strongly Agree). . . . .	81
5.17	Mean scores of questions from Q1 to Q8. “*” and “***” denote significant difference at $p < 0.05$ level and $< 0.01$ level. The error bars depict the standard deviation. . . . .	83
5.18	The participant placed the objects casually around OEB2 for checking. . .	85

.....

5.19 L) Participants had to walk around in the plain condition. R) Participants used the proposed feature to replace walking. . . . . 86

5.20 The participants unconsciously tried to use the disabled Physical Rotation in the plain condition. . . . . 86

5.21 a) b) c) An example that the participant observed the OmniBar area effectively. d) The participant lost her remote partner. . . . . 88

5.22 The limitation about the showing direction of the Quick View sharing window. . . . . 89

5.23 The Physical Rotation disturbs the use of the Quick View. . . . . 92

6.1 The DollhouseVR system [36] proposed by Ibayashi et al.. . . . . 97

6.2 The new design workflow [1] proposed by An et al.. . . . . 97

6.3 The CollaVR system [57] proposed by Nguyen et al.. . . . . 99

6.4 The system hardware overview of OEB2 terminal. . . . . 102

6.5 The OmniEyeballVR enables VR users (occupant) to communicate and cooperate with others on the physical world (designer). . . . . 104

6.6 Viewing VR Space in the 360° first-person mode from different viewpoints. 105

6.7 In both first-person and third-person mode, the red translucent frame represents the field of view of the occupant. The occupant is shown as a virtual avatar in the third-person mode. . . . . 106

6.8 In the third-person mode, the designers can rotate or walks around the OmniEyeball in order to observe the VR space at a different direction. . . 107

6.9 The face windows of designers shown in the VR space. . . . . 108

6.10 Face Window optional positions during the design process. . . . . 109

6.11 Left: Based on the position of the face window in the VR space, the occupant can know which direction the designer is looking at. Middle: the face window is also displayed on the OmniEyeball as visual feedback to the designer. Right: Through the face window, designers can use body gestures to express information, such as pointing to a direction by hand. . 110

.....

6.12 (a) The experiment VR scene for the task. The image patch represents the position of face window. (b) The first-person mode with the solid window (1PSW) condition. (c) The first-person mode with the face window (1PFW) condition. (d) The third-person mode with the face window (3PFW) condition. . . . .	112
6.13 Subjective ratings on user experience (*: significant effect). . . . .	115
6.14 Results of Spatial Presence questionnaire (left) and MEQ (right). (*: significant effect) . . . . .	116
6.15 Results of System Usability Scale Questionnaire. . . . .	118
6.16 User preference between 1PSW, 1PFW, 3PFW conditions, based on the three questions against the number of participants (*: statistically significant). . . . .	118
6.17 User preference-based ranking results on the designer side against the number of participants. . . . .	119
6.18 User preference-based ranking results on the occupant side. . . . .	120
6.19 The participant(P12) used the visual feedback to correct his hand gesture shown in the face window. . . . .	122
7.1 Attention from Afar [83] proposed by Xu et al.. . . . .	131
7.2 One solution to the reversed space issue. . . . .	132
7.3 DisplayBowl [54] proposed by Miyafuji et al.. . . . .	132
7.4 The Meeting Owl produced by Owl Labs. . . . .	136
7.5 The 360° video chat application enabled by Insta360 Nano S camera. . .	138
7.6 DL-based computer vision with 5G cloud computing support new OmniEyeball. . . . .	139
7.7 The 360Drops system proposed by Teo et al.. . . . .	140
8.1 A conclusion graph of the whole thesis. . . . .	143

# Chapter 1

## Introduction

### 1.1 Background

#### 1.1.1 Limitation of Video Communication Systems

A long-standing problem with video communication systems was the limitation of the field-of-view. Such an issue caused by narrow field-of-view of a video camera has been discussed since the early studies of video communication systems [21] [24].

Previous researchers pointed out that the narrow field-of-view limits peripheral awareness of activities in a remote scene and often makes remote activities unnoticed. For example, Fish et al. pointed out that reciprocity between participants such as “*if you can see someone else, they can see you*” [21] was not maintained in their VideoWindow system. To challenge this issue, Gaver et al. proposed to use multiple cameras and monitors to cover a wider range in a collaboration space. However, they found that “*the necessity of switching among views increased the difficulty of establishing the relations among them and negotiating a mutual orientation towards the task* [24].”

In their subsequent study, Gaver et al. [25] tested the effect of a remote-controlled camera while it revealed various problems due to unavoidable features of mechanical systems in control. Then, as a future direction, they suggested employing fisheye view image [26]. Recently, Luff et al. [47] have conducted an experiment using multiple telepresence robots. However, they also suffered from the narrow field-of-view problem and suggested employing panoramic technology.

### 1.1.2 Omnidirectional Panoramic Technology

The panorama, as one of the most popular image types, has existed for hundreds of years. A panorama is any wide-angle view or representation of physical space, whether in painting, photography, film, or a three-dimensional model.

The 360° panoramic image is the widest type of panorama. A 360° panoramic image that incorporates a full 180° vertical viewing angle, as well as a 360° horizontal viewing angle, can present all the visual information of one viewpoint in space. All the visual information of four directions, top, and bottom, is included in the 360° panoramic image. Moreover, nowadays, due to the unique panoramic feature of omnidirectional view, the 360° panoramic image is becoming more and more popular in the fields of photography and virtual reality.



Figure 1.1: The Theta series 360° cameras produced by RICOH

With the development of 360° technology, one of the recent trends in digital cameras is omnidirectional cameras. Recently, more than ten different omnidirectional cameras have been released within five years, such as the Ricoh Theta series <sup>1</sup> (Figure 1.1) or Insta360 cameras <sup>2</sup> (Figure 1.2), etc. These kinds of cameras consist of one or two fisheye lenses. The camera with one fisheye lens, which can cover 360° in the horizontal direction, has a field of view from 180° to 240°, while the one with two fisheye lenses can directly capture an image of 180×360 degrees field of view.

Unlike traditional cameras, such new cameras can capture panoramic images in one

---

<sup>1</sup><https://theta360.com>

<sup>2</sup><https://www.insta360.com>



Figure 1.2: Samples of Omnidirectional cameras produced by Insta360.

shot, which provides up to 360° field-of-view. Such technologies seem to open new possibilities to solve the problem of a narrow field-of-view. For example, Johnson et al. [38] compared three kinds of the field-of-view video feed from a telepresence robot, narrow (45°), wide-angle (180°), and panoramic (360°), and their results showed that wider views supported better task efficiency and fewer collisions. However, the panoramic field-of-view increased the visual image's complexity and required a higher cognitive load than a narrower field-of-view.

## 1.2 Research Motivation

Based on the above background, it is interested in knowing whether these issues of increased complexity, workload, and asymmetry of communication could be improved by using a more proper 360° panoramic camera-display configuration (Figure 1.3).

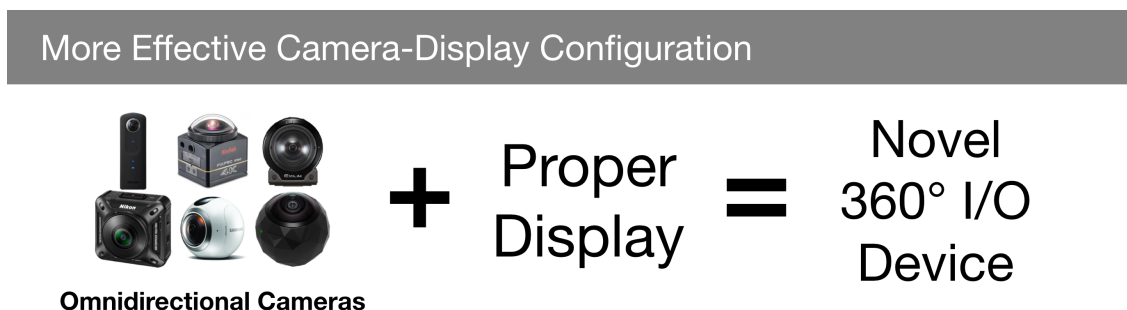


Figure 1.3: The motivation of this research is to propose a new camera-display configuration.

### 1.2.1 Evolution of I/O Devices for Video Communication

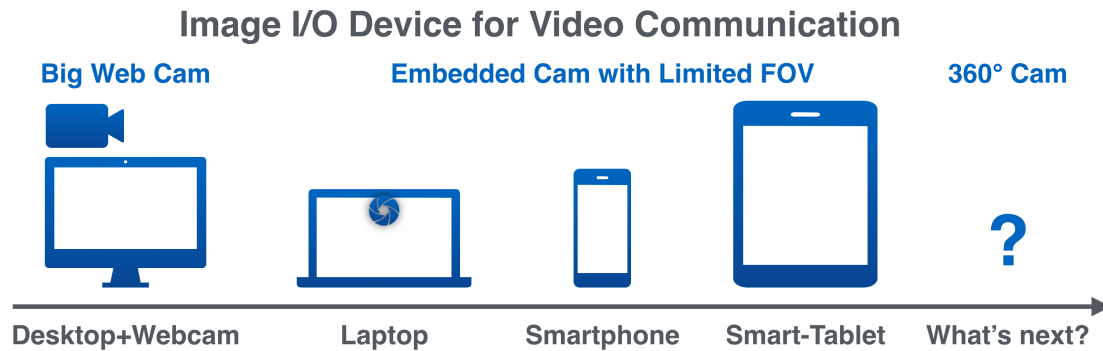


Figure 1.4: Some representative graphics I/O terminal forms.

The combination of image input and output devices, i.e., camera-display configuration, has made significant progress for a long time. Many decades ago, computers and digital cameras were quite huge and heavy that no one tried to combine them. With the development of digital devices, the first low-cost webcam, Trojan Room coffee pot [71], was small enough to be used with personal computers in 1991. Meanwhile, 2D monitors were also getting smaller and lighter. Within three decades, more and more digital I/O devices support video communication. Such as laptops, smartphones, or tablets, combined the 2D display with small embedded cameras as image input and output, see Figure 1.4.

Due to the development of panoramic technology, attempts have been found to combine 360° cameras with existing display devices. For example, Darling VR Smartphone<sup>3</sup> is the world's first smartphone released in 2017, which directly used a build-in 360° panoramic camera (See Figure 1.5 left).

Furthermore, Insta360 tried to implement the panoramic video communication by connecting the Insta360 Air<sup>4</sup> with a laptop, see Figure 1.5 right. The equirectangular format image is directly shown on the 2D display of a laptop, which causes are several issues of this configuration. Although the camera is 360° panoramic, the display is not. The equirectangular format on the 2D display provides a very bad sense of space, which

<sup>3</sup><https://techlog360.com/first-smartphone-built-360-degree-vr-camera/>

<sup>4</sup><https://www.insta360.com/product/insta360-air>



Figure 1.5: Left: Darling VR Smartphone, World's first smartphone with built-in 360° camera. Right: A panoramic video communication setup by using an Insta360Air 360° camera with a laptop.

seems to indicate that the 2D display may not be the perfect match of a 360° camera.

## 1.2.2 Spherical Display Systems

Current popular displays for 360° contents (Figure 1.6) such as the virtual reality head-mounted display (VR HMD) <sup>5</sup>, the cylinder-shaped display, the panoramic dome or cave display, present some issues when used for video communication. For example, the VR HMD covers most of a user's face, hindering facial expressions, which are important in communication [66]. Besides, if users on both sides wear HMDs, it will not be possible for them to view both actual environments. Also, wearing HMD for a long time and walk around may cause physical and eye fatigue. Immersive projection technologies such as a panoramic dome display [3] [23], cave display [16] [17] [28], or twister display [72] [73] confine users inside their space which limit user's interactions with the real environment. In addition, the cylinder-shaped display [87] [88] creates a severe distortion on the top and bottom sides.

An alternative solution may be to use a spherical display. The sphere, as a typical geometric object, is used in lots of designs. Researchers have been proposing new spherical

<sup>5</sup><https://www.vive.com/>

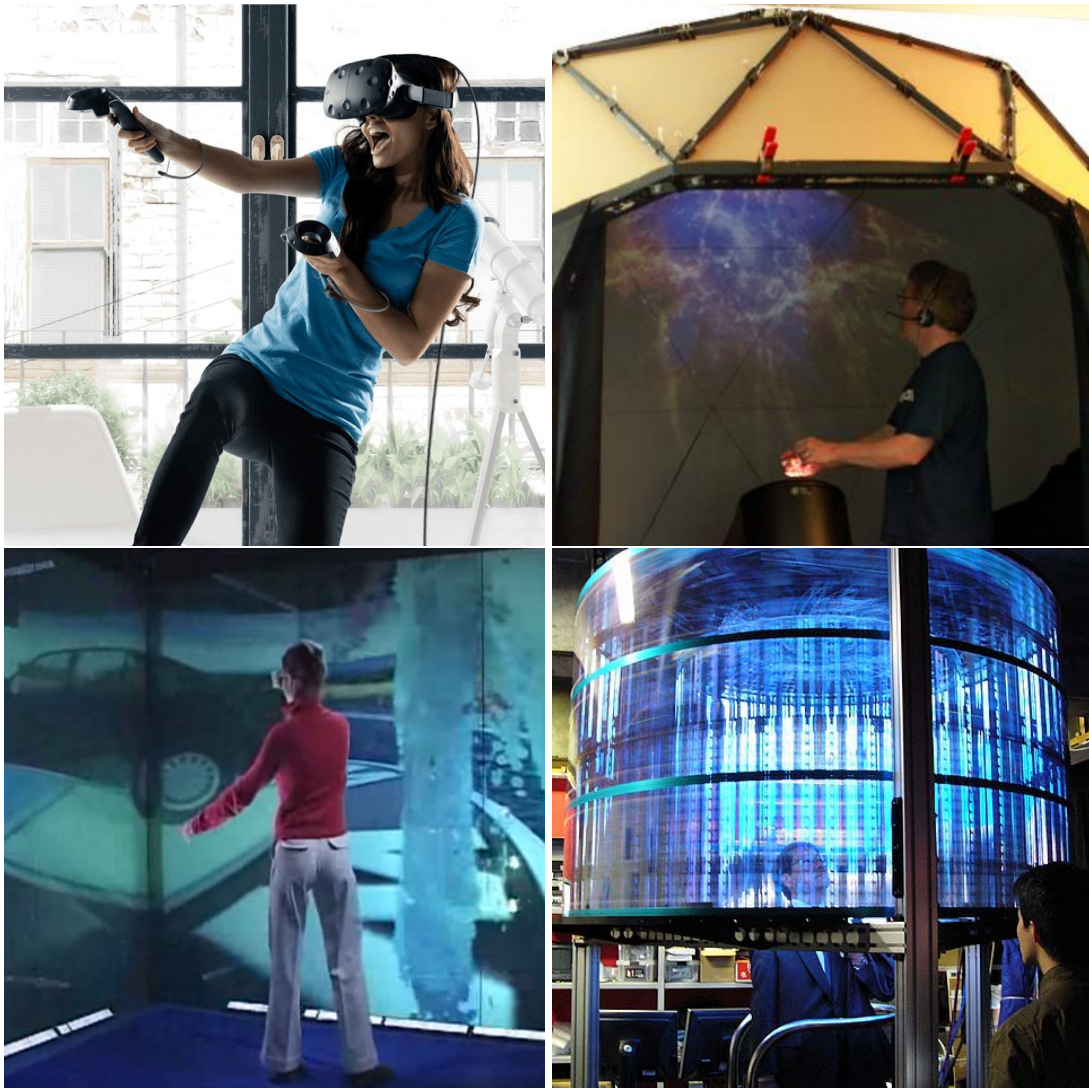


Figure 1.6: Currently how we watch 360-degree panoramic images or videos: VR HMD - HTC Vive (upper left), Dome display (upper right), Cave display (lower left), Twister rotating display (lower right)

display systems for many years. For example, Organic user interfaces [35] created by Holman et al., is the first spherical multitouch display. Moreover, the Sphere [4] proposed by Benko et al. also attracted much attention from the research field. Due to the feature of spherical shape, spherical displays may be a perfect match for 360° panoramic cameras, which could naturally display the 360° images and reduce observers' cognitive load.

### 1.2.3 Novel 360-Degree I/O Device for Video Communication

Therefore, the motivation of this research is to develop a new 360° I/O device that captures 360° panoramic images as well as displays it intuitively. The proposed device combines omnidirectional cameras with a spherical display system. It can capture a 360° live video stream, send the data via the wireless connection, and display the 360° live video feed on its spherical surface. Since the video communication system utilizes this 360° I/O Device, it can exchange the image of the whole environment around the devices. Therefore users are not constrained by the camera position or display orientation, i.e., users can position themselves anywhere around the device. This new I/O device for 360° images may bring users more interactive possibilities in daily life.

## 1.3 Organization

The organization of this thesis is as follows:

Chapter 1: In this introductory section, the background, as well as the motivation of the research, have been presented.

Chapter 2: The related works section includes two parts: The video communication systems and the 360° graphic displays.

Chapter 3: The research proposal, methods, prototype (OEBP) of the OmniEyeball system is introduced.

Chapter 4: The OmniEyeball system (OEB) is introduced. Then, a user study about the impact of display shapes on 360° video communication is presented.

Chapter 5: Based on the previous user study's findings, an improved system, OmniEyeball2 (OEB2), is introduced. An experiment is then presented to study how its interactive features support symmetric 360° video communication.

Chapter 6: An application tool of OmniEyeball project, named as OmniEyeballVR, is introduced. It works as a collaborative, 360° video communication tool for VR.

Chapter 7: This section discusses the contribution of the research as well as research findings, insights, design implication, and future works.

Chapter 8: This section summarized the whole research.

---

## Chapter 2

### Related Work

#### 2.1 Effect of Field-of-view in Video Communication

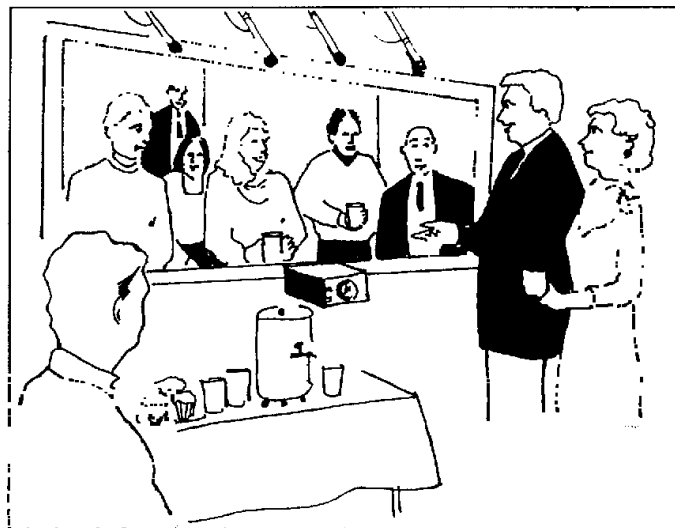


Figure 2.1: The VideoWindow teleconferencing system [21] proposed by Fish et al..

Problems caused by the narrow field of view of a video camera have been discussed since early studies on video communication systems [21] [22] [24] [39]. Fish et al. pointed out that reciprocity between participants such as “if you can see someone else, they can see you” [21] was not maintained in their VideoWindow system (Figure 2.1). For example, when a participant stands outside the field of view of a camera used in a video conference, he or she can see the remote participants, but they cannot see him or her. These researchers further suggested that they needed new ideas for design as well as the placement of cameras and monitors. To challenge this issue, Gaver et al.

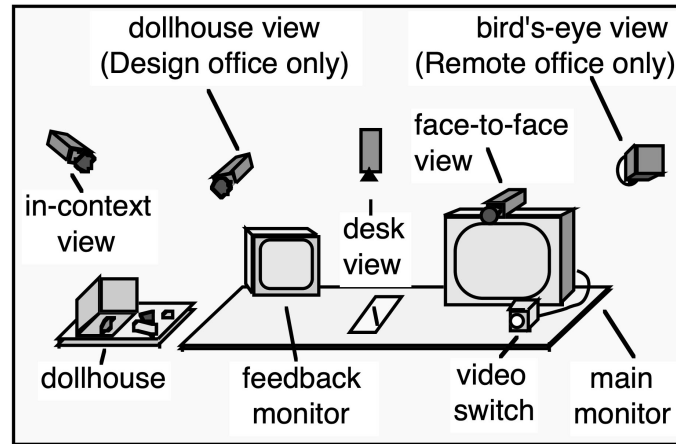


Figure 2.2: One is not enough [25]: the MTV set-up by Gaver et al..

proposed using multiple cameras and monitors to cover a wider range in a collaboration space. However, they found that “the necessity of switching among views increased the difficulty of establishing the relations among them and negotiating a mutual orientation towards the task [25] (Figure 2.2).”

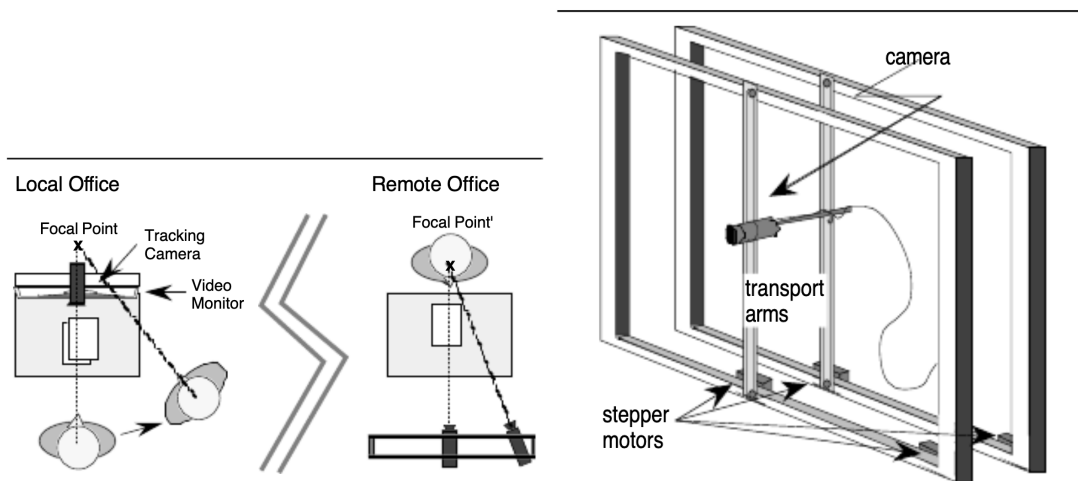


Figure 2.3: Virtual Window [26] proposed by Gaver et al..

In their subsequent study, Gaver et al. [26] tested the effect of a remote-controlled camera that was expected to overcome the limitation of a narrow field of view (Figure 2.3). However, the system revealed various problems due to the unavoidable features of mechanical systems concerning control delay, accuracy, and noise. Then, as a future di-



Figure 2.4: The Tobi system used in Show Objects [47].

rection, they suggested employing fisheye view imaging. Recently, Luff et al. [47] have conducted an experiment using multiple telepresence robots, Tobi<sup>1</sup> (Figure 2.4). However, they also suffered from the narrow field-of-view problem and suggested employing spherical images.

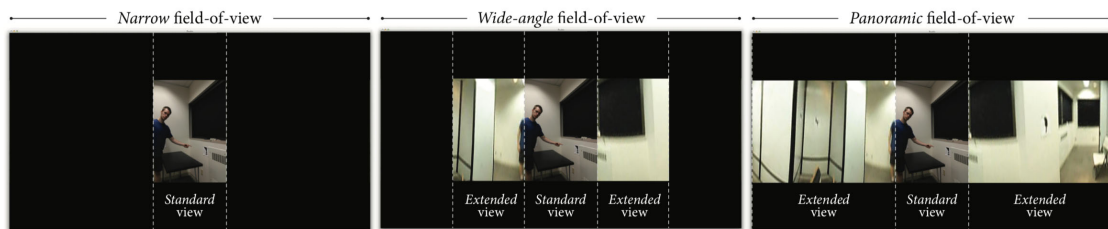


Figure 2.5: Can You See Me Now? [38].

To investigate how the different fields of view affect remote collaboration, Johnson et al. compared the performance of three field-of-view angles (narrow (45°), wide-angle (180°), and panoramic (360°)) of a camera on a telepresence robot [38], see Figure 2.5. The results of these experiments indicated that a wider field of view supported better task efficiency. However, these researchers also pointed out that users felt more difficulty in perceiving a panoramic view, suggesting the need for further exploration of user interfaces for a panoramic video communication system. Even though there are some other studies that have explored the use of a 360° panoramic video system. JackIn Head

<sup>1</sup><https://www.revolverobotics.com/>

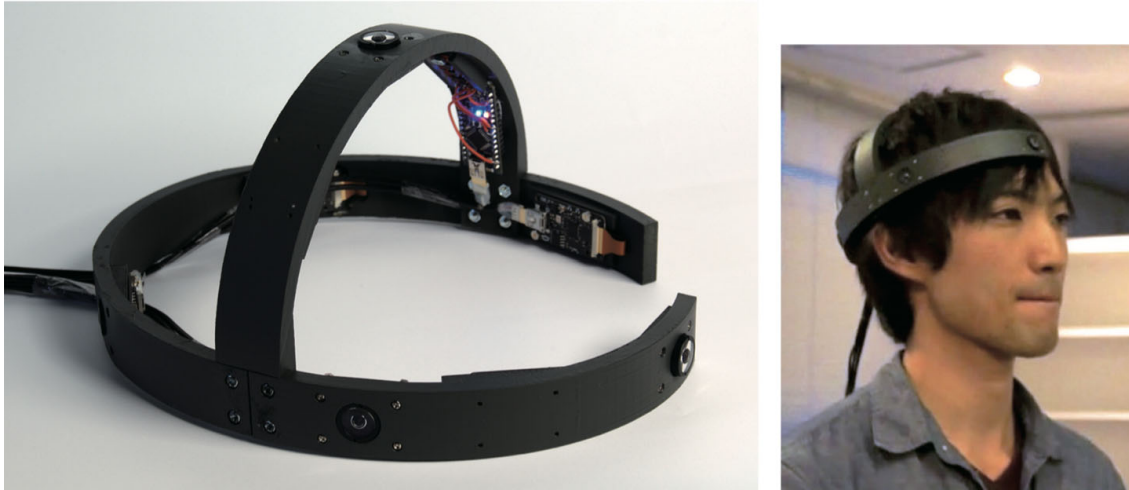


Figure 2.6: JackIn Head [41] proposed by Kasahara et al..



Figure 2.7: ScalableBody [53] proposed by Mutsuda et al..

(Figure 2.6) proposed by Kasahara et al. [41] is a wearable camera that can shoot and transmit 360° surroundings, and provide remote collaboration support by transmitting a stabilized video to a remote user wearing an HMD. ScalableBody [53] (Figure 2.7) is a telepresence robot supporting socially acceptable interactions and human augmentation through vertical actuation. Tang et al.[74] created a system for observing the image of a remote 360° camera with a tablet terminal, see Figure 2.8. Lee et al. [44] combined the 360° camera with VR-HMD to improve the remote collaboration (Figure 2.9).

However, the issue for these systems is that their system configurations were asymmetrical. In other words, users at a local site could see the 360° panoramic video that captured the remote site but not vice versa—the video configuration did not allow equal



Figure 2.8: Collaboration with 360° Videochat: Challenges and Opportunities. [74]

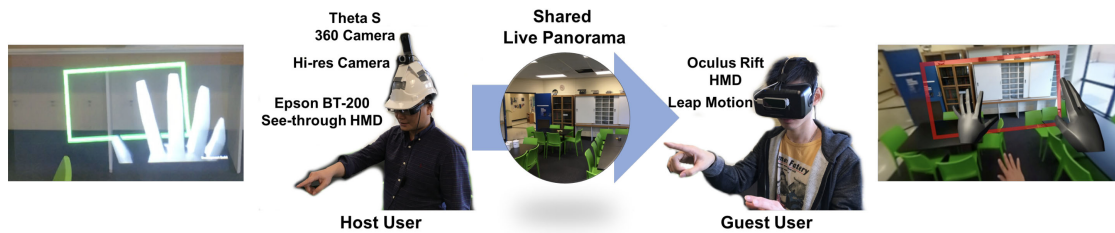


Figure 2.9: SharedSphere [44] proposed by Lee et al..

contribution from both local and remote sites.

Meanwhile, some studies have already proposed video communication systems that enable the exchange of wide field-of-view images symmetrically. CamBlend [58] [59] was designed as a symmetrical video conferencing system that integrates wide field of view(180°) images with high-precision in-context views (Figure 2.10). The system requires a camera and large display placed on one side of the wall in the room, which almost covers the whole room. However, the wall that the camera and the display are mounted becomes the blind zone of the system. t-Room [85] [86] (Figure 2.11) is a room-size video communication system that consists of multiple screens, cameras, and tables and enables the exchange of panoramic images of the two rooms on either side over a distance. The system creates an illusion of merging two geographically separate rooms. However, the shared area in the t-Room is constrained to the area close to the screen



Figure 2.10: CamBlend system [58] proposed by Norris et al..

surfaces. Therefore, it is still challenging to realize a symmetrical video communication system that enables the exchange of spherical images over a distance.

Furthermore, compared to t-Room and CamBlend, the proposed 360° system in this thesis is much more flexible in that it can be placed anywhere in the existing work environment and enables sharing the whole space and objects around the terminal.

## 2.2 Spherical Image Display System

As suggested by Gaver et al. [26] and Liccope et al. [47], it is interested in using spherical images (360° panoramic images) for video communication. This section briefly reviews the technologies related to spherical images. One of the recent trends in digital (video) cameras is the omnidirectional camera. Although this technology has been explored for a relatively long time, recent advancement in digital technology has resulted in high-definition omnidirectional panoramic cameras affordable to general consumers.

Various display methods have been proposed to view such spherical images. One approach is to use an ordinary 2D flat display. For example, the whole spherical image

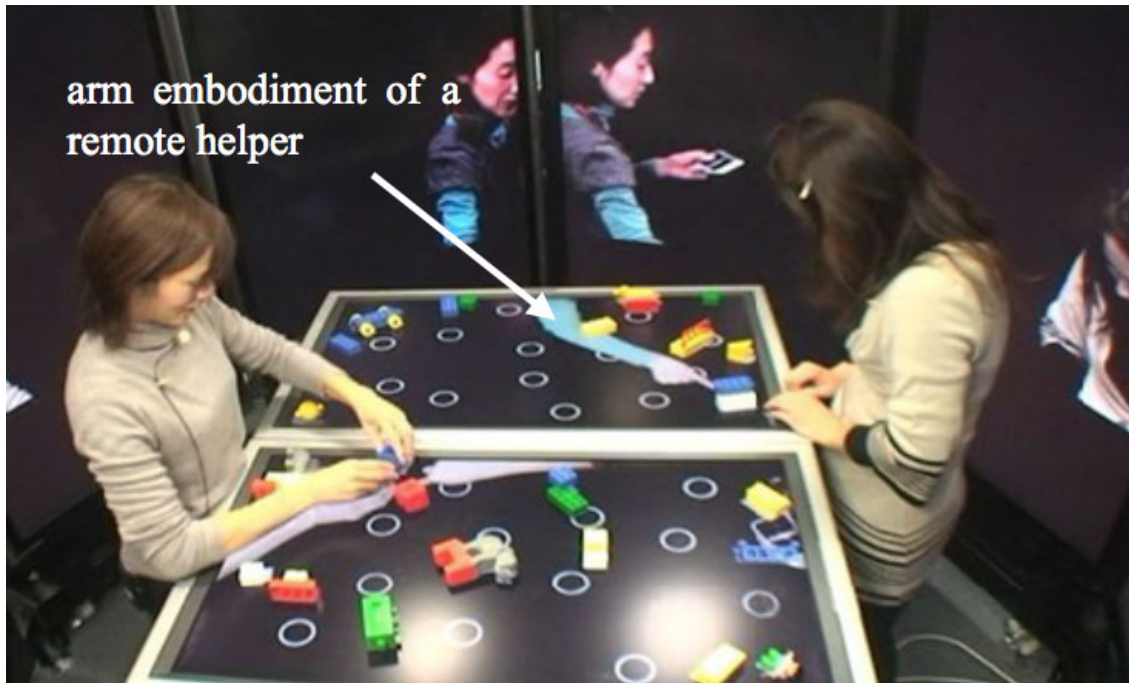


Figure 2.11: t-Room system [86] proposed by Yamashita et al..

can be shown by distorting the image into a circular shape or, to reduce the distortion, by showing part of the image and allowing users to control the viewing orientation using cursor keys or a pointing device (QuickTime VR or Application of Ricoh Theta <sup>2</sup>).



Figure 2.12: GEO-COSMOS [51], the world's first full color spherical display, in Japan.

<sup>2</sup><https://theta360.com/en/about/application/>

.....

Instead of using an ordinary 2D flat display, there are some alternative methods, such as using an HMD or a mobile terminal. Another advanced method is to use a spherical display. GEO-COSMOS [51] (Figure 2.12) is the world's first full color spherical display. However, because its diameter is 6m and it weighs 13-tons, it is not suitable for ordinary users. With the advance of projection technologies, however, several smaller spherical displays have been prototyped [7] [9] [14] [15] [42] [70] [79] [80] [81] [82] [84] and are even commercially available now <sup>3</sup>, see Figure 2.13.

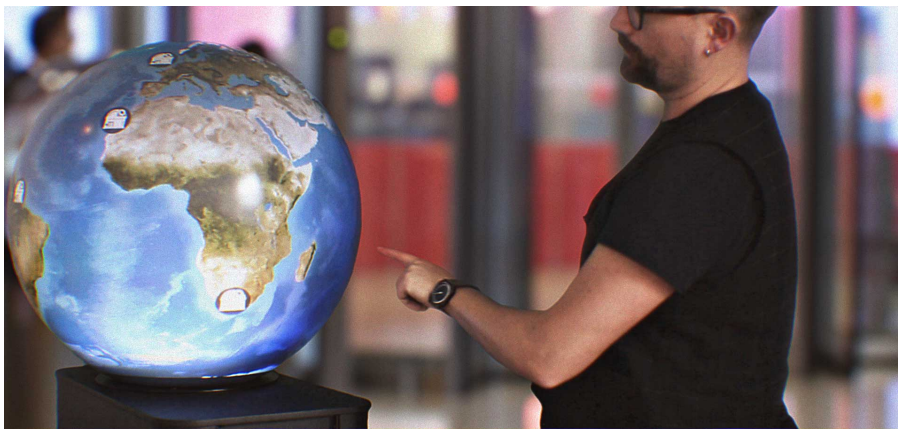


Figure 2.13: PufferSphere touch spherical display.

Sphere [4] (Figure 2.14), Spheree [20] are 360° spherical displays. They enable users to interact with the spherical display and allow viewers to explore different perspectives of the display. To project through 360 degrees, these systems have a projector under or beside the spherical display.

Vortex [40], Magic Crystal Ball [11], and The Object Inside[5] are spherical display systems that allow the users to see virtual objects/scenes appearing inside a glass sphere. Users can control the display by touching it. They did not consider to show 360° image on the spherical display.

Pan et al. [61] [62] and Oyekoya et al. [60] researched the usage of avatars for performing remote tasks on spherical displays. They researched how accurate local people can recognize the gaze direction of a remote avatar appearing in the spherical display.

---

<sup>3</sup><https://pufferfishdisplays.com>

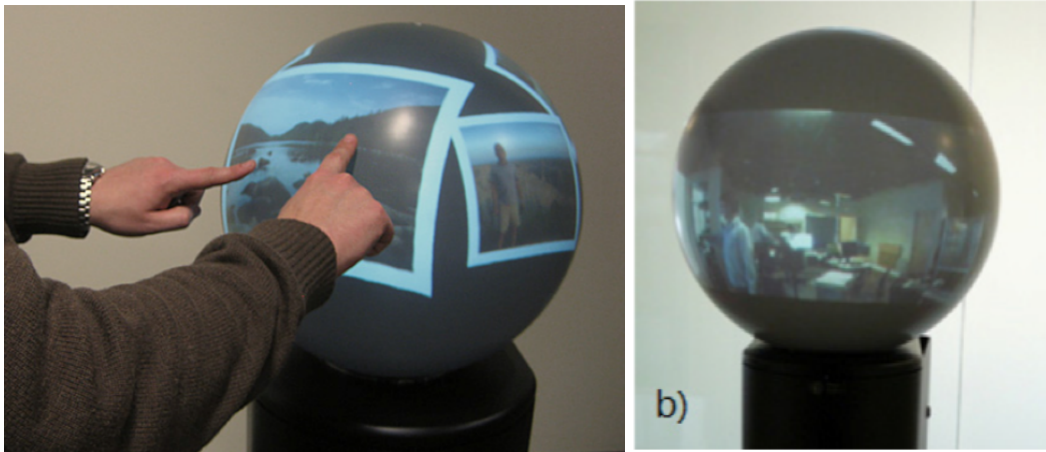


Figure 2.14: Sphere [4] proposed by Benko et al..

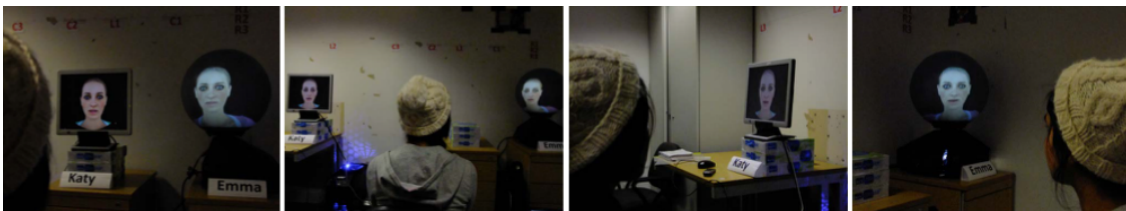


Figure 2.15: Comparing Flat and Spherical Displays in a Trust Scenario in Avatar-Mediated Interaction. [62]

They also compared when avatars appear inside a spherical display and a 2D flat display, people are willing to trust which avatar, as shown in Figure 2.15.

Miyafuji et al. [55] [69] presented a sphere-shaped interactive display system, named Qoom shown in Figure 2.16 , as a new input and output device of the future. Unlike existing sphere-shaped displays, Qoom has perfect sphere shape, like a ball, which can be rotated, thrown, or even kicked. It is the first spherical display that is movable, touchable, and also allowed to be thrown or even kicked. The application called Pano Ball in Qoom can enable users to see the 360° live stream of another space in Qoom. However, it is also one-way communication, just like the application in Sphere[4]. The difference between Qoom and this work is that the OmniEyeball concentrates on the integration of spherical display and omnidirectional camera and its application on panoramic videoconferencing.

To overcome the problem of a narrow camera field of view, it is interested in using

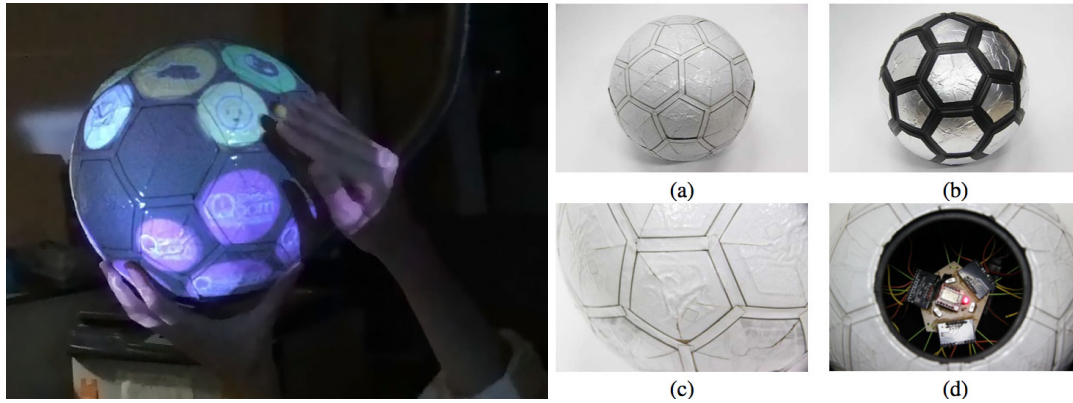


Figure 2.16: The Qoom system [69] proposed by Miyafuji et al..

spherical images for video communication. To the best of my knowledge, despite the availability of spherical cameras and displays, no studies have been conducted to explore how the spherical display technologies can support symmetrical video communication. Especially, it is important to explore its usability issue of such a system.

---

## Chapter 3

# Research Proposal - OmniEyeball

Based on the above background and related work, it is interested in knowing whether these issues of increased complexity, workload, and asymmetry of communication could be improved by using a more proper 360° panoramic camera-display configuration.

### 3.1 Research Approach

This research aims to build the new 360° I/O device or system integrating spherical display with omnidirectional cameras, as shown in the Figure 3.1, and discover how it supports the panoramic video communication.

Therefore, the research will proceed as follows:

#### 3.1.1 Implementation of the 360° I/O System

##### Basic functions of the 360° I/O device

This new 360° I/O device can capture a 360° panoramic live stream as well as display it. To achieve that, omnidirectional panoramic cameras or lens are embedded into the spherical display. Each camera or lens can cover a 180° field of view or even more than that. Then, two cameras or lens can capture the whole space. The purpose of this device is to enable a 360° panoramic image as input and output. To enable remote panoramic video communication, it should be possible to do a panoramic live stream through a wire or wireless connection. The panoramic live stream can be seen directly projected onto the spherical display itself.

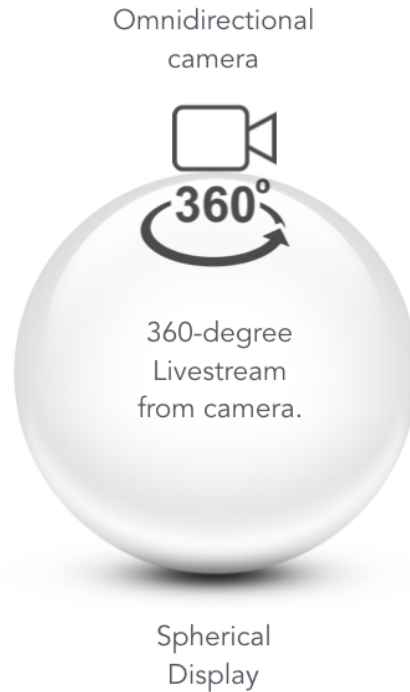


Figure 3.1: The blueprint of the proposed 360° I/O device.

### Implementation of the Spherical Display

Based on the current cutting-edge implementation of spherical displays, high-quality spherical displays are mainly divided into two types: External projector type and Built-in projector type.

1. External projector type: Taking Qoom [69] as an example, this type of spherical display mainly relies on one or several external projectors placed in space to project a 360° image on a spherical screen through projection mapping.
2. Built-in projector type: Taking Sphere [4] and FTVR [18] as examples, this type of spherical display mainly relies on the built-in one or several projectors under a spherical screen, and through the image correction of the optical lens, the image projected by the projector is displayed.

### **Interactive features of the 360° I/O System**

Besides the hardware implementation, it is also essential to implement the software of the 360° I/O system, such as image processing and image converting approaches for 360° live streaming videos.

Furthermore, it is interesting to explore the interactive features (such as gestures input or touch manipulation) suitable for the proposed system to make it perform better in supporting panoramic video communication.

#### **3.1.2 Design OmniEyeball Prototype**

The first prototype system is implemented based on the spherical display of the external projector type. To create the new 360° I/O spherical device which can capture 360° panoramic image and display it, a 360° panoramic camera is embedded into an External projector type spherical display system, as shown in Figure 3.2. The device is named as OmniEyeball (OEB) because the embedded camera looks like an eye on the device, which has a sphere shape. It can capture 360° panoramic images and also display a spherical image. It can do a panoramic live stream through a wireless connection. The real-time panoramic image can be seen directly projected onto the spherical display itself with dynamic spherical projection mapping.

#### **Prototype Overview**

The spherical device consists of one rubber ball embedded with one PIXPRO SP360 4K digital camera on the pole. The SP360 camera is equipped with a fisheye lens that can cover 360° in the horizontal direction has a 235° field of view. Therefore, most of the surrounding environment of the spherical device can be covered by this SP360 camera with a limited blind zone. The proposed system streams live video to the PC, then integrates a dynamic spherical projection mapping technology and motion tracking system to project the live stream image onto its spherical display. Therefore, the device seems like a spherical mirror.

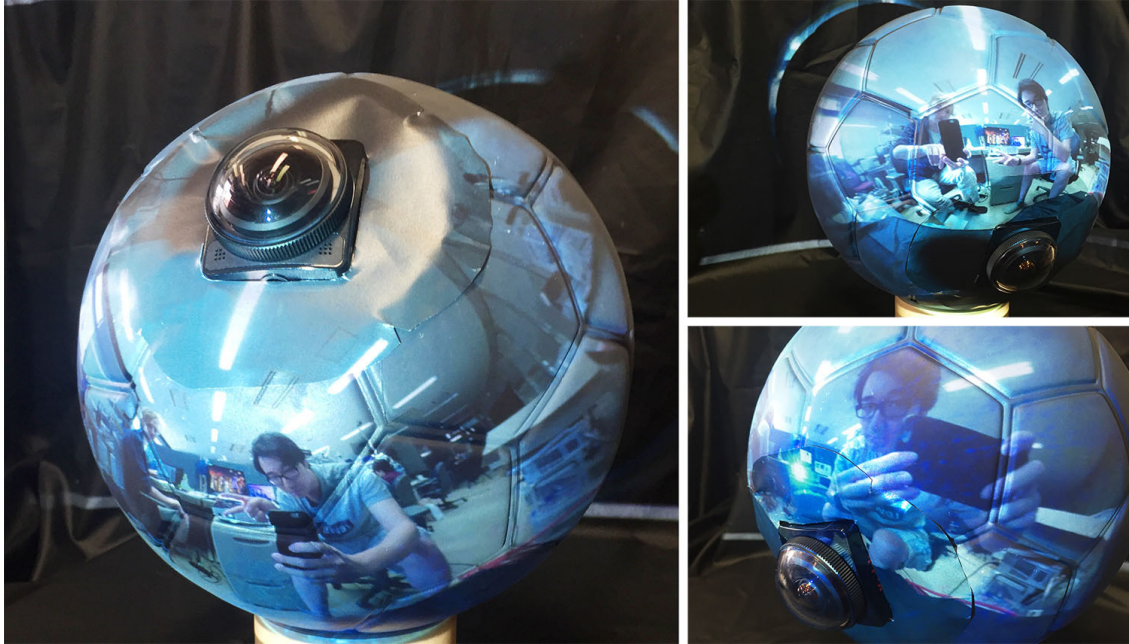


Figure 3.2: The OmniEyeball (OEB) prototype system proposed in 2016 [46].

The live stream received from the SP360 4K camera has a  $1280 \times 720$  resolution with 15fps when connected to the PC by WiFi. The format of the image is fisheye motion-JPEG. In order to integrate the SP360 4K camera's live stream into the system, the raw image is transferred into an equirectangular format by using a fragment shader in the graphic system. This shader does camera calibration to convert the fisheye image to a square image with high quality and low cost.

### **Dynamic Spherical mapping**

The dynamic spherical mapping of the OmniEyeball prototype is based on technology from one previous work [55]. It can make a  $360^\circ$  projection to moving balls in real-time. The positional relationship between the ball and projectors can be calculated by transforming projector positions in the real world to the camera coordinate system. A motion tracking system is built by consisting of 6 motion tracking cameras to track the position of the moving ball. By combining three-dimensional ball positions from the motion tracking system, the system generates the projection images corresponding to each

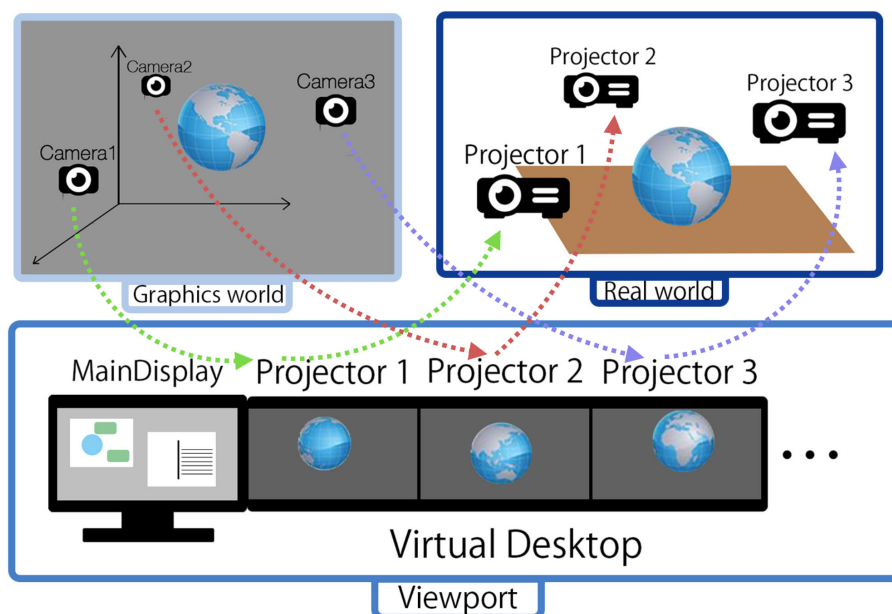


Figure 3.3: Diagram of dynamic spherical projection mapping.

projector's position.

Figure 3.3 shows the diagram of spherical mapping. The PC computes the different virtual desktop images based on the position of the ball and projectors from the motion tracking system. Then, these images are projected by corresponding projectors from different directions to cover the whole area of the ball according to the ball's position and orientation. To put real-time projection into practice, this system overlays images without optimizing the overlapping parts.

### Panoramic Video Communication Application

Figure 3.4 shows a practical video communication application. This application allows users in different places to do video communication with a 360° panoramic image by arranging two spherical devices in different places, such as conference rooms. Each spherical device will use its embedded cameras to capture the surrounding environment, and the image data will be projected onto the other spherical device by the system.

The OmniEyeball prototype is designed to improve remote conference meetings based

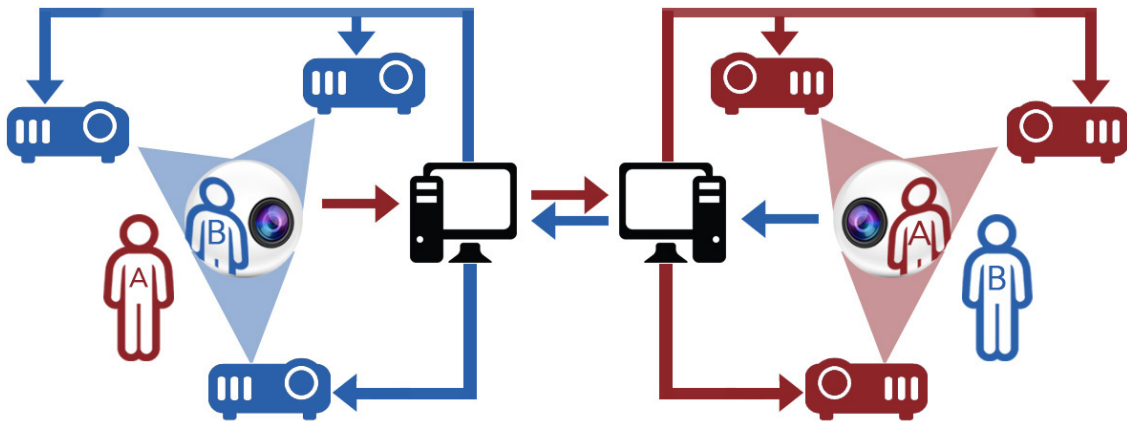


Figure 3.4: Diagram of 360° panoramic video communication.

on its panoramic and movable features. It will be efficient for users to just carry our spherical device to view every detail in the other place and simultaneously show relevant information in their surrounding environment.

### Limitations and Discussion

Due to the projection mapping approach, the proposed OmniEyeball prototype (OEBP) encountered several limitations that restrict its possibilities for further user study.

1. Glare Problem: There is strong glare appearing when the projectors are emitting light toward the omnidirectional cameras. As for a 360° device, it is unavoidable that the omnidirectional cameras capture the projectors. The glare will affect the brightness, color, and sharpness of the area near the projectors
2. Occlusion: Since the projectors are located outside of the OEBP, there is no doubt that when people walk around the OEBP. People will block the projection. Even though this problem can be reduced by increasing the number of projectors or put the projectors between the OEBP and people, it still happens when people touch the display.

3. Overlapping Area: Using multiple projectors to do projection mapping in the same area will cause an overlapping problem. The image on the overlapping area is a composition of multiple projections. Therefore, the image inside this area is always in a mess. Usually, a double ghost image will be seen. Using only two projectors can reduce this problem while lowering the quality of the image in the boundary.
4. Darkness Environment: Projection mapping performs best when the environment is dark. However, the goal of the OEBP is to improve the video communication experience, so keeping the room dark is quite impossible.

Because of the limitations mentioned above, the formal devices introduced in the following chapters are all designed based on the Built-in projector type spherical display.

### **3.1.3 Evaluation of the 360° I/O System for Video Communication**

The user study in this research is mainly divided into two parts:

#### **Comparison between Spherical Displays with Other Types of Displays**

The first user study is to compare 360° I/O devices between equipped with spherical displays or other possible types of displays. Using two types of symmetrical panoramic video communication systems, an experiment is conducted to explore the possibility of using omnidirectional cameras for video communication and remote collaboration. Specifically, it is interested in the impact of display types on 360° panoramic video communication. Therefore, the key questions of this study are shown as follows:

1. How do users perceive the video images differently between the 2D flat display and the spherical display?
2. Is the spherical display a suitable choice for a 360° I/O system in video communication?

---

**Evaluation of Interactive Features of the 360° I/O System**

The second user study explores how the proposed 360° I/O device and its interactive features support remote collaboration by improving the awareness of remote sites. Through the user study, it is planned to evaluate the system usability, user experience, affordance, the mental effort of the proposed system.

**3.1.4 Exploration for Applications of 360° I/O System**

Besides the video communication application, it is interesting to explore more possible application scenarios for the proposed 360° I/O system. For example, VR, AR, MR fields have received significant attention from the academic community. Due to the high correlation between 360° panoramic images and VR, it becomes quite valuable to explore whether the proposed system has a landing possibility in related scenes in the VR field.

---

## **Chapter 4**

# **OmniEyeball: A Hemispherical 360° I/O System to Study Impact of Display Shapes on Video Communication**

### **4.1 Preface**

One goal of this study is to explore the possibility of using a symmetric 360° video communication system in remote collaboration. Meanwhile, it is important to investigate how the shape of the displays affects remote collaboration involving a real physical environment. Especially, it is essential to explore the usability issue. Therefore, the study introduced in this chapter conducts user study about remote collaboration to compare two display methods: 2D flat displays and spherical displays.

The findings contribute to the understanding of how to design an environment for remote collaboration and to capture and display a 360° image of a remote site.

### **4.2 System Implementation of OmniEyeball**

#### **4.2.1 Hardware Overview**

The panoramic video device used for the experiment is created by integrating the omnidirectional camera with a physical display, which could be a spherical display or a 2D flat display. For example, the re-designed OmniEyeball (OEB) integrated the World-

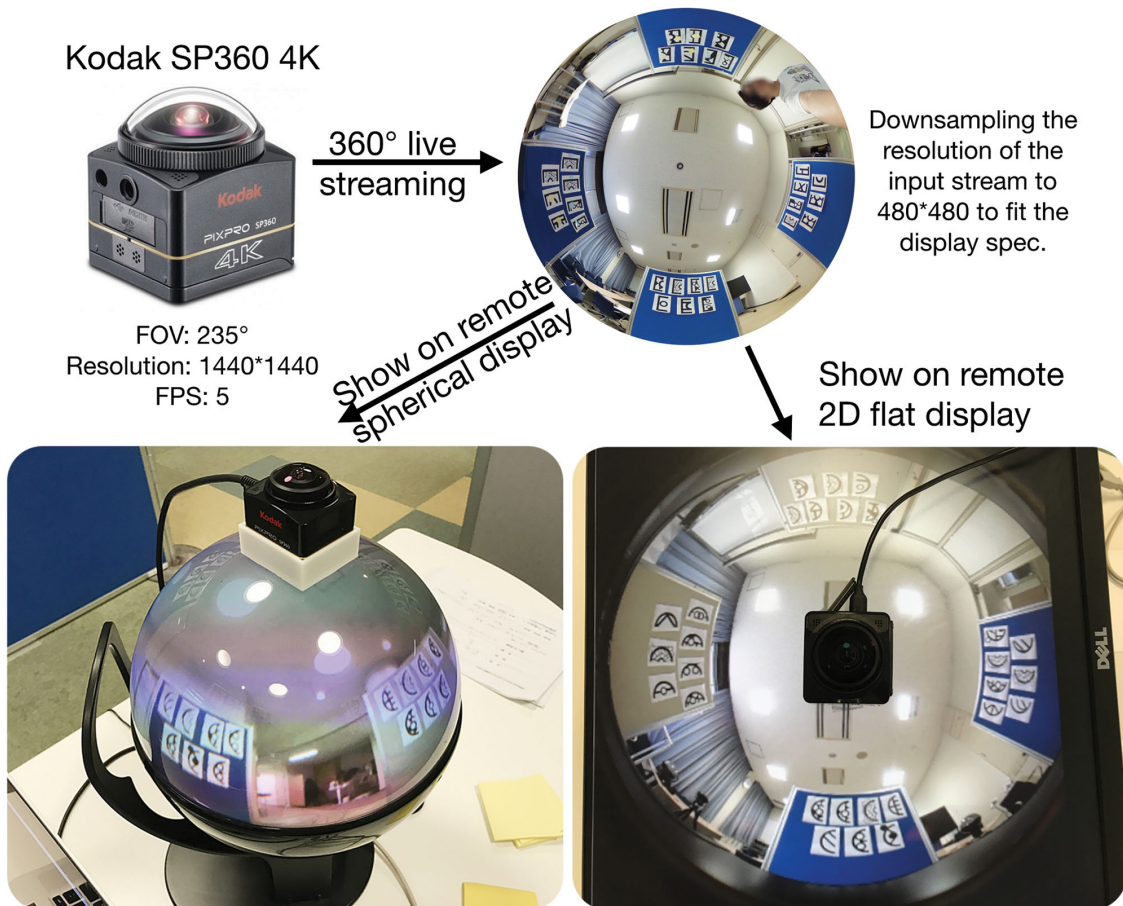


Figure 4.1: Diagram of the 360° panoramic video communication. The input of the system is from the Kodak camera. The output is either shown on the remote spherical display or the remote 2D flat display (Left: OmniEyeball version; Right: 2D display version).

Eye spherical display with the Kodak PIXPRO SP360 4K camera to create the spherical panoramic video system, shown in Figure 4.1. The Kodak omnidirectional camera is put on the top of the WorldEye display. The WorldEye display is a hemispherical display sold by Gakken Sta:Ful Co with a display resolution of  $480 \times 480$ . By showing a round image in the WorldEye display, it can directly map the round image onto its hemispherical screen. The Kodak camera is equipped with a fisheye lens which covers a 235° field-of-view. It compresses the 235° space into a fisheye image by equidistant projection.

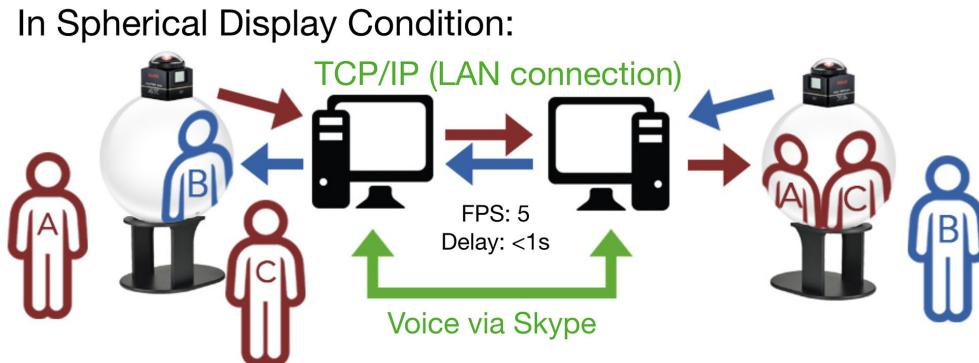


Figure 4.2: The diagram of the panoramic video communication system. Persons B and C are standing across from each other.

The Kodak camera can output a live video stream with  $1440 \times 1440$  resolution with 5fps over a USB connection. Then the system receives the live video stream from the Kodak camera and converts the resolution to  $480 \times 480$  for data transmission.

For the 2D flat display type, a DELL 24 inch display was employed. To minimize the difference between the two types of displays, the PPI (pixels per inch) were set to be the same to achieve the same display effect under both conditions. Otherwise, audio/video transmission was employed with the same technologies.

#### 4.2.2 Software Overview

For panoramic video communication, two identical panoramic video terminals were interconnected in different rooms via LAN (local area network). To be more specific, a live video stream captured by the local omnidirectional camera is transformed to  $480 \times 480$  pixels by downsampling. The live stream is then transferred to the remote counterpart terminal using TCP/IP protocol and displayed on the remote hemispherical or 2D flat display. The audio transmission was based on Skype. The FPS of the whole system can be maintained as 5, and the communication delay is controlled under 1 second.

This omnidirectional communication system requires careful directional calibration to enable one-to-one communication between two environments. As shown in the bottom

of Figure 4.2, the orientations of cameras and spherical displays should be appropriately configured in a way that if a local person stands across from a remote person (in the display) at one end, the remote person should stand across from the local person at the other end. In the Figure, C stands across from B (in the display) at one end, then B should stand across from C at the other end.

### 4.2.3 Reversed Space

The system design proposed above invited one phenomenon, which is named as “reversed space.” Assume that there are two rooms with the same arrangement, and the four walls are marked as wall north, east, south, and west, corresponding to the actual directions (Figure 4.3). These two rooms are connected by 360° video communication systems located in the center of the rooms. Therefore, the 360° image of room 1 is shown in the display in room 2 (which is the circular image in the center of room 2). When the actual wall N and wall S of room 2 face toward the images of walls N and S of room 1 in the display, the image position of walls E and W are reversed. Therefore, as shown by the red arrow, the wall E in room 2 faces toward the image of the wall W in room 1.

The reversed space is an inherent phenomenon in the symmetric 360° video communication, no matter using the spherical display or 2D flat display in the proposed system. The reason behind this phenomenon is due to the mapping method of 360° spherical images. Let us assume a 360° camera captures a 360° spherical image of a space where the eight directions (North, East, South, West, Northeast, Northwest, Southeast, Southwest) are marked (Figure 4.4 A). The captured 360° image in equirectangular format is shown in Figure 4.4 B. There are two ways to map the 360° spherical image to a sphere, which is either on the outside surface or the inside surface of the sphere.

Mapping the 360° image on the outside surface of the sphere is usually used for the spherical display or the cylindrical display. Users naturally see 360° spherical images from the outside of the sphere (Figure 4.4 C). In such a method, the left-right direction is not flipped, which is similar to the way that users observe the remote space by using

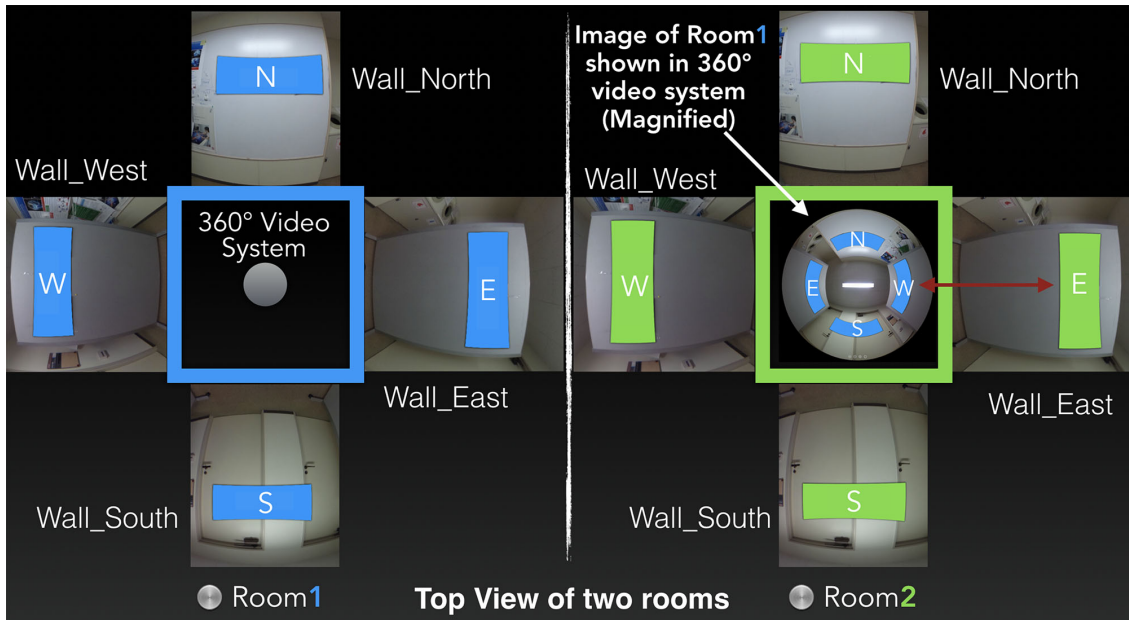
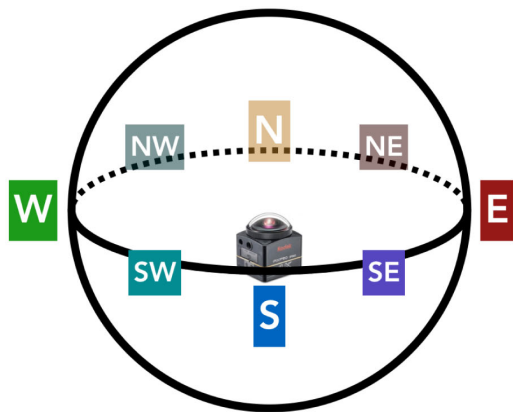


Figure 4.3: The top view diagram of the reversed space phenomenon.

a conventional teleconference system, e.g., Skype. Therefore, this method is named as “teleconference metaphor”. However, mapping the 360° image on the outside surface of the sphere makes the spatial configuration in the spherical images different from the spatial configuration in the original space (i.e., the difference between Figure 4.4 C and Figure 4.4 A), which triggers the reversed space.

The other mapping method is to map the 360° image on the inside surface of the sphere, which is usually used for the VR HMD or the dome display so that users naturally look around 360° spherical images from the inside of the sphere (Figure 4.4 D). Although this mapping method does not cause the reversed space (i.e., no spatial difference between Figure 4.4 D and Figure 4.4 A), however, since the proposed system (e.g., spherical display and 2D display) requires users to observe from the outside of the sphere, utilizing this mapping method for the spherical display leads to a flipped left-right direction effect. That is, all the objects, including texts and pictures, are horizontally flipped in the images, which seems like a mirrored effect. Therefore, this mapping method is called as “mirror metaphor”.

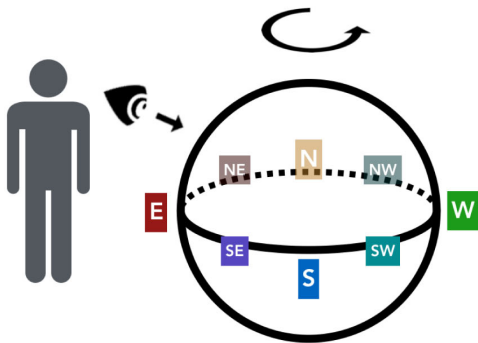


A) A 360° camera captures an image of a space where 8 directions are marked.



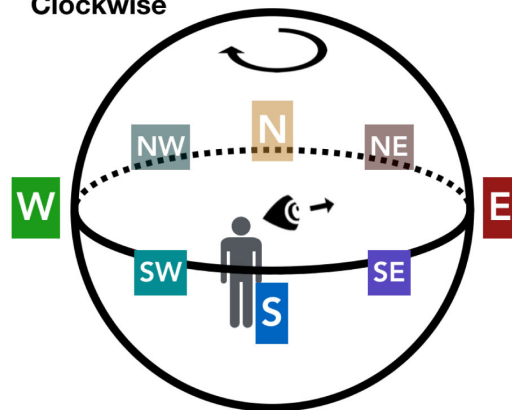
B) The captured 360° equirectangular image with the N, E, S, W, NE, SE, SW, NW 8 directions.

Mapping Direction:  
Counterclockwise



C) Mapping the 360° image on the **outside** of a sphere is usually for users to see from the **outside** of the spherical image. (Spherical Display, Cylinder Display, etc.)

Mapping Direction:  
Clockwise



D) Mapping the 360° image on the **inside** of a sphere is usually for users to see from the **inside** of the spherical image. (VR, Dome Display, etc.)

Figure 4.4: Two different mapping methods of the 360° spherical images and the reason behind the reversed space.

.....

Due to two reasons, the decision was made to choose the teleconference metaphor instead of the mirror metaphor for the proposed system. First, the mirror metaphor results in flipped texts and pictures, which is unacceptable in the experiment that required the participants to match geographically distributed patterns (pictures) through video communication. Second, the directional expression related to left and right in the video communication might be complicated to understand in the mirror metaphor, because users do not see mirrored images in regular teleconference configuration.

## 4.3 Experiment

Using two types of symmetrical panoramic video communication systems, an experiment was conducted to explore the possibility of using omnidirectional cameras for remote collaboration. Specifically, it is interested in the impact of display types on 360° panoramic video communication. Therefore, the key questions of this study are shown as follows:

1. How do users perceive the video images differently between the 2D flat display and the spherical display?
2. How does the display type affect participants' actions during remote collaboration in a real physical environment?

### 4.3.1 Method

The research is based on a four-people remote collaboration because panoramic video communication can support group-to-group remote communication. To study the features of a 360° field of view, a pattern-matching task was designed that required the participants to walk around the video communication system.

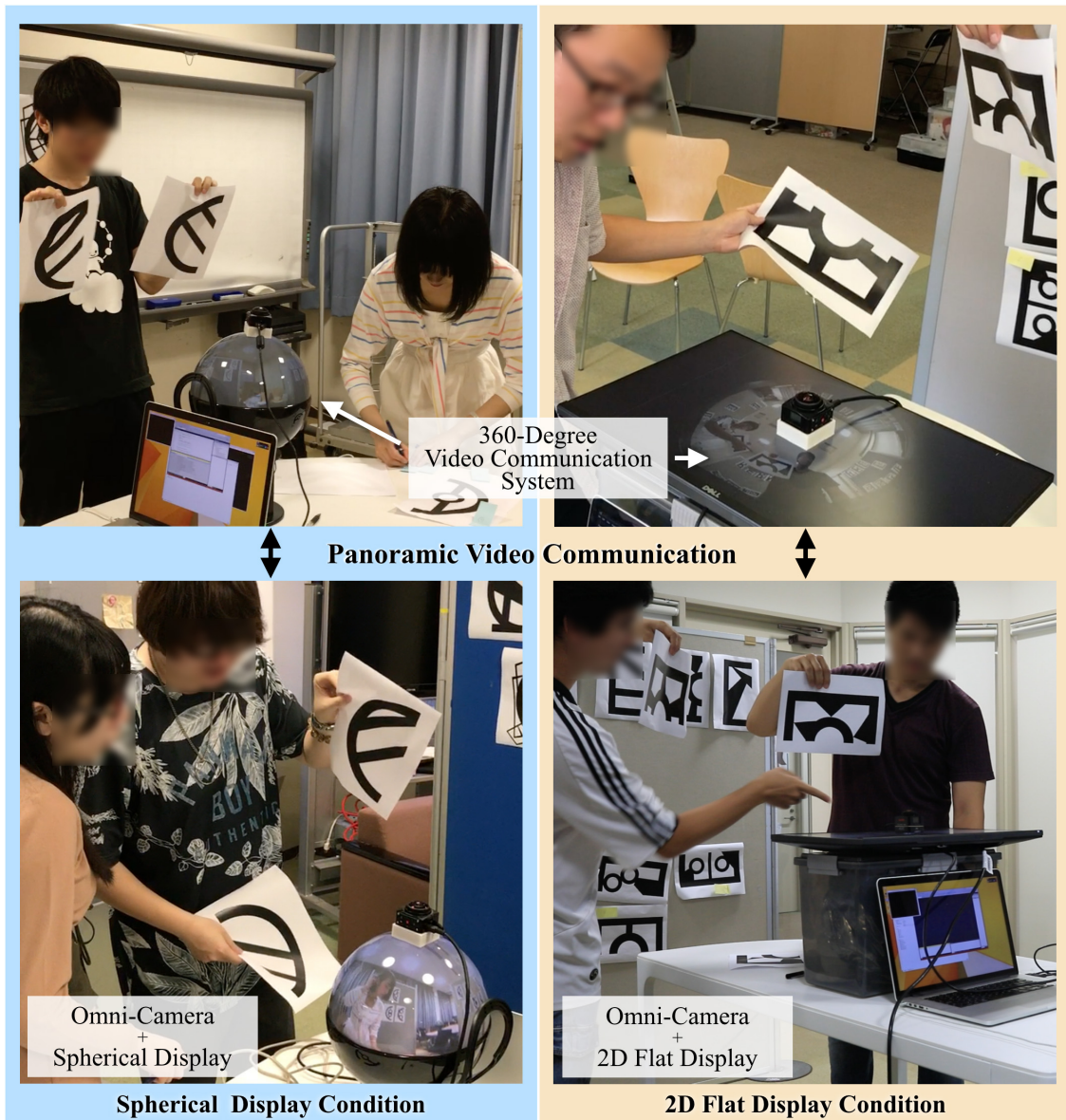


Figure 4.5: The 360° panoramic video communication systems. Two geographically separated rooms are connected by two different types of 360° panoramic video communication systems, a combination of omnidirectional camera and a spherical (left column) or a horizontally placed 2D flat display (right column).

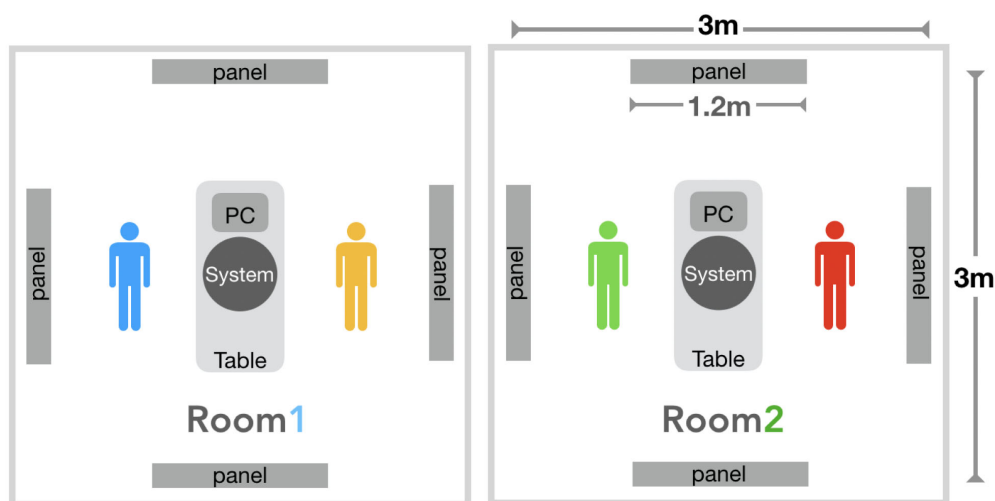


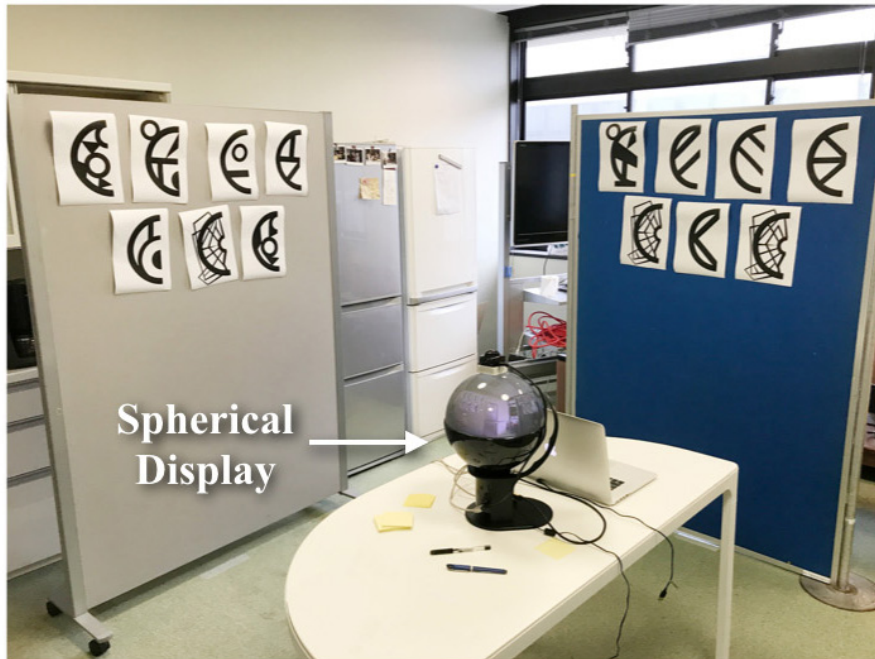
Figure 4.6: The physical layout of both two rooms.

### 4.3.2 Conditions

The experiment compared the 360° panoramic video communication efficacy using two conditions: the first with a hemispherical camera and a hemispherical display (spherical display condition) and the second with a hemispherical camera and a horizontally placed 2D flat display (2D flat display condition)(see Figure 4.5). For this experiment, the study employed OEB technology (spherical display condition) because it is the only existing system that enables synchronous communication in a real physical environment. However, its apparent drawback is that one should walk around the terminal to observe the whole hemispherical image. Therefore, a system that employs a 2D flat display is designed. Normally, a 2D flat display is configured vertically. In the experiment, however, such configuration disabled the participants who happened to be behind the display to see the image.

Therefore, for equivalent visibility from all the participants standing anywhere in the room, the display is placed horizontally. The intention is not to prove one is better than the other. Instead, the study aims to investigate how symmetrical video communication in a real physical environment can be supported by two possible technologies and arrive at the pros and cons of each technology.

When using the spherical display



When using the 2D flat display



Figure 4.7: The real environment of the room when using the spherical display or the 2D flat display.

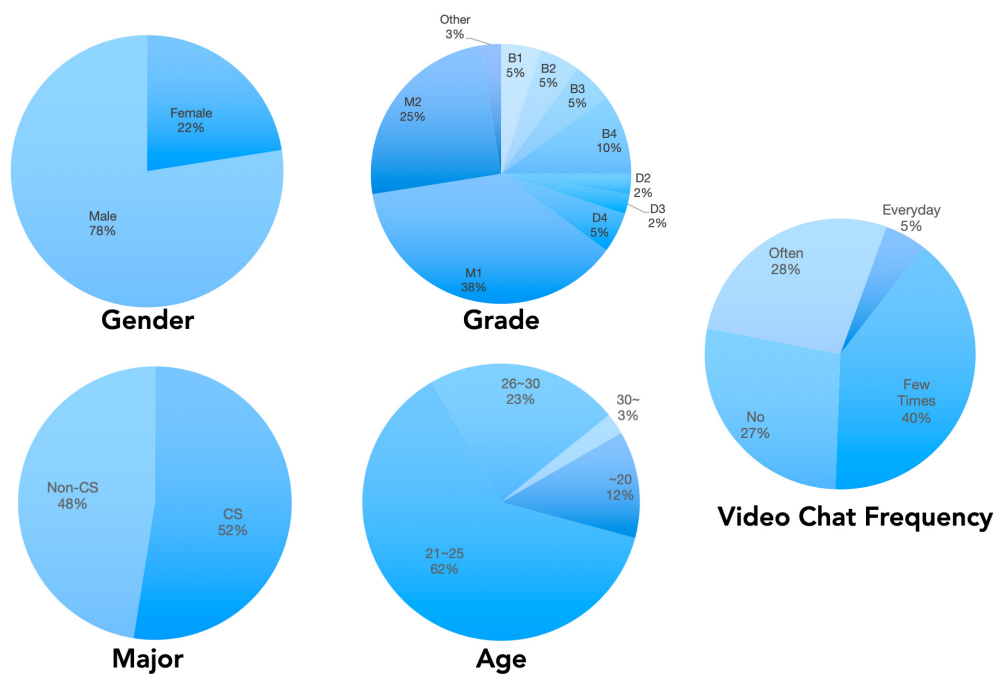


Figure 4.8: The detailed information of the 40 participants.

### 4.3.3 Apparatus

The experiment used two mutually isolated rooms equipped with the identical display systems (either two spherical displays or two 2D flat displays) (Figure 4.6 and 4.7). There were four panels (representing four walls) and one table in each room. The table was set up in the middle, and the distances between the table and four panels were the same. The two rooms were connected by the panoramic video communication system using a gigabit network. The displays were located at the same height for both conditions.

### 4.3.4 Participants

Ten groups of four participants (40 people in total) were recruited for this experiment. The average age of the participants was 23.5 years old. Of these participants, 21 were students in the department of computer science, and 29 had used video communication applications in their daily life. No participant had prior experience with panoramic video

.....

communication systems. The detailed information of the participants can be checked in Figure 4.8.

### 4.3.5 Task

Each task consists of four participants, with two participants located in one room. One room was provided with a set of 28 geometric pattern figures; each figure was printed on a sheet of paper, as shown in Figure 4.9. These figures were put up on four panels (seven figures on each panel) in the room. In the other room, an identical set of figures were put up at different positions on the four panels. The participants were asked to match all the figures collaboratively across the rooms. Time was measured during the task, and all participants were asked to finish the task as quickly and correctly as possible. If needed, they could take off the figures from the panels freely to show them to the remote partners or to just rearrange the figures. Scotch tapes were used on the back of the figures whereby the figures could easily be placed back on the panels.

The design of the collaborative task is affected by [38] [47] and [59], in that it requires participants to move around in a shared space and show objects to remote participants, identify remote objects, refer to remote objects, and ask the remote participants to manipulate remote objects. Although it may not be naturalistic, the task was designed to incorporate as many of these actions as possible.

The experiment was based on a within-participants design. Participants conducted the task in two rounds, once in the spherical display condition and once in the 2D flat display condition. Therefore, two different sets of figures (round set and rectangular set) were prepared whereby the same figures would not be used for two rounds of tasks. In addition, the two different sets of figures were used for the two conditions in alternation; thus, the difference of difficulty between the two sets did not affect the results of the experiment.

Each group had to complete two rounds of tasks. To minimize the learning effect, half of the groups used the spherical display in the first round and used the 2D flat display in the second round, whereas the other half did it in the opposite order.

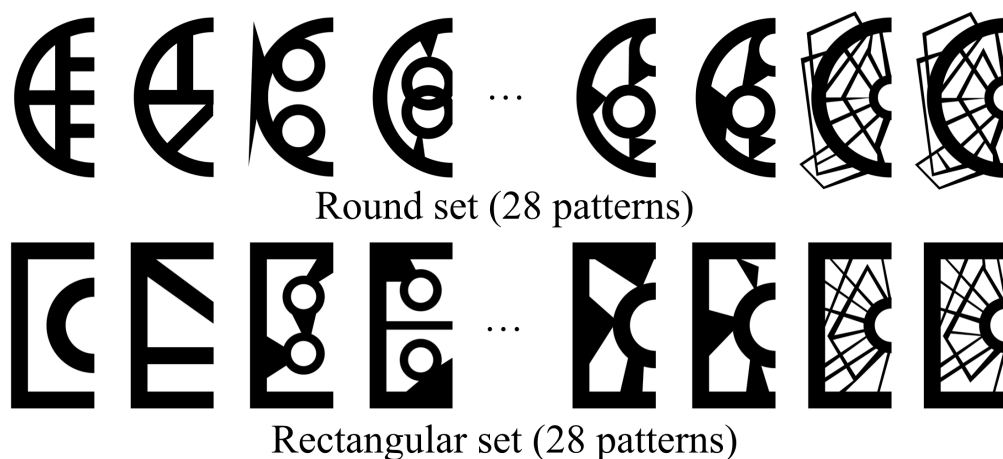


Figure 4.9: Two sets of patterns used in the task.

#### 4.3.6 Procedure

1. Preparation: To begin with, participants were given a brief introduction to the experiment and panoramic video communication system; they were then introduced to each other and split into two groups; after which, they did a practice round with six figures. After the practice, the correctness of their answers was checked, and they were given 2 min to discuss how to complete the task more efficiently in the main experiment.
2. Main experiment: Five of the ten groups started with panoramic video communication using the spherical display, and the other five groups began using the 2D flat display. They were asked to match all the 28 different figures with the corresponding figures in the other room. Stickers and pens were prepared on the table whereby the participants could number the figures by writing the number on the sticker and sticking it on the figure. When they finished all the figures for the first round, participants were asked to complete a short questionnaire concerning their thoughts about the condition in the first round. Then the display was replaced by the other type, and the figures were also changed to the other set. Next, participants were asked to carry out the second round of the task with the same rules. When

.....

they finished the task, participants completed another short questionnaire about the current condition.

3. Debriefing: After finishing two rounds of the task, participants filled in a final questionnaire about their opinions and comments on the two conditions and their overall impressions of panoramic video communication.

### 4.3.7 Measures

The study employed objective and subjective measures to examine task efficiency and users' perception of video images during the remote task. Each session was videotaped for post-analysis.

#### Objective Measures

The study evaluated the collaborative efficiency based on completion time (the time it took the participants to complete the task) and completion accuracy (the number of correct pairs/total pairs).

#### Subjective Measures

The study conducted a post-task questionnaire consisting of five questions per condition (10 in total), designed on a 5-point scale (1 = Strongly Disagree to 5 = Strongly Agree) to measure participants' feelings after using the panoramic video communication systems with two different displays. It was expected that most of the variance in the answers of the questionnaire could be explained by four sets of correlated questions, which are "Visual Cognition", "Presence", "Perceived Workload" and "Communication Difficulty", respectively. The detailed questions are listed in Figure 4.10.

In addition, "Compatibility" – a measure of the participants' feeling concerning which display is more suitable for the 360° video communication system, was evaluated by two questions shown in Figure 4.10. The participants could answer these questions by choosing from "Spherical display", "2D flat display" or "Same".

Category	Question
Visual Cognition	Q1. It was easy to identify the objects/figures shown in the display.
Presence	Q2. I felt as if the remote participants were in the same room.
	Q3. I had a feeling of closeness to the remote participants.
Perceived Workload	Q4. I felt tired to work with the remote participants.
Communication Difficulty	Q5. I had difficulty in talking to only one of the two remote participants.
Compatibility	Q6. For displaying 360 panoramic image, which type of display do you think is better?
	Q7. Which display is easier and faster for searching?

Figure 4.10: Seven questions in the questionnaire.

.....

An additional free-response question investigated the features, including the pros and cons of both displays. Question contributing to this measure was: *Please write down the advantages and disadvantages of each display, which are found in this experiment.* The participants' qualitative comments for the free-response question were used to sum up the pros and cons of the two displays, which might also help to support the results of the above questions.

## 4.4 Results

To find answers to the research questions given above, quantitative and qualitative analyses were conducted on the experimental data. Firstly, the study analyzed collaboration efficiency and then analyzed the questionnaire results to see how the display types affect user perception. Finally, participant comments and video recordings were analyzed to see how the two display types affected remote collaboration.

### 4.4.1 Collaboration Efficiency

The average task completion time was 485.6s ( $SD = 71.65$ ) with the spherical display and 513.6s ( $SD = 124.29$ ) with the 2D flat display. The Wilcoxon signed-rank test showed no significant difference between the two conditions ( $Z = -.7139$ ,  $p = .4752$ ). As for the completion accuracy, the average task completion accuracies were .9892 ( $SD = .0339$ ) and .9928 ( $SD = .0226$ ), respectively. The Wilcoxon signed-rank test showed no significant difference between the two conditions ( $Z = -.0743$ ,  $p = .9407$ ).

The results of the questionnaires, because the Kolmogorov-Smirnov test showed a significant difference between the distributions of the two sample sets, the analysis was by a Wilcoxon matched-pairs test.

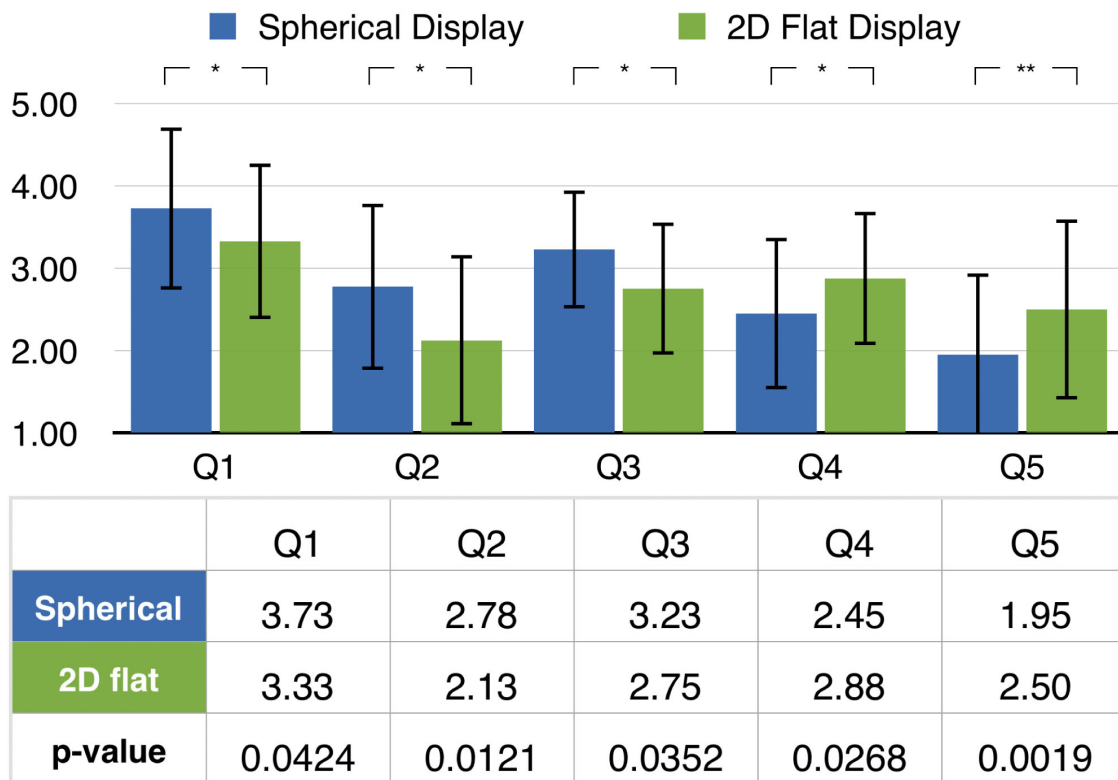


Figure 4.11: Mean scores of questions from Q1 to Q5. “\*” and “\*\*” denote significant difference at  $p < 0.05$  level and  $< 0.01$  level. The error bars depict the standard deviation. For the scores of Q1-Q3, the higher the better. For the scores of Q4-Q5, the lower the better.

#### 4.4.2 Questionnaire Results

As shown in Figure 4.11, the results of the questionnaires were analyzed using Wilcoxon signed-rank test with the display condition as the factor.

1. Visual Cognition – The average scores of Q1 were 3.73 for the spherical display and 3.33 for the 2D flat display; also, a significant difference ( $Z = 2.0296$ ,  $p = .0424$ ) was found between the two conditions. According to the results, the participants seemed to feel that the spherical display images were easier to recognize than the images shown in the 2D flat display.
2. Presence – The average scores of Q2 were 2.78 for the spherical display and 2.13 for the 2D flat display, and there was a significant difference ( $Z = 2.5098$ ,  $p = .0121$ ). The average scores of Q3 were 3.23 for the spherical display, 2.75 for the 2D flat display and there was a significant difference ( $Z = 2.1059$ ,  $p = .0352$ ). Overall, it seems that the spherical display provided a stronger feeling of telepresence than the 2D flat display.
3. Perceived Workload – The average scores of Q4 was 2.45 for the spherical display and 2.88 for the 2D flat display, and a significant difference ( $Z = -2.2141$ ,  $p = .0268$ ) was found between the two conditions. The results indicate that the participants felt more fatigued when using the 2D flat display.
4. Communication Difficulty – The average scores of Q5 were 1.95 for the spherical display and 2.50 for the 2D flat display. A significant difference was found between the two conditions ( $Z = -3.0918$ ,  $p = .0019$ ). The results indicate that the participants found it more difficult to have one-to-one conversations when using the 2D flat display than when the spherical display was used.
5. Compatibility – According to the results of Q6, 60% of the participants thought that the spherical display is better for displaying the 360° panoramic image. The results of Q7 indicate that the 2D flat display is equally fast and easy to use as the spherical display is to search within the display, as shown in Figure 4.12.

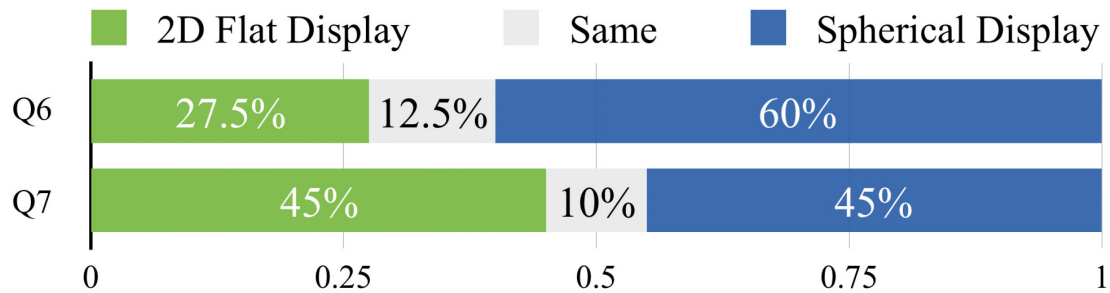


Figure 4.12: The statistical results of Q6 and Q7.

#### 4.4.3 Findings from Participants' Comments

Participants' comments were obtained from the free-response question in the questionnaire. In Figure 4.13, the spherical display's advantages and disadvantages are summarized with the corresponding number of participants.

Opinions about the 2D flat display were also gained from the results of the free-response question. The advantages and disadvantages of the 2D flat display are summarized in Figure 4.14 with the corresponding number of participants.

#### 4.4.4 Findings from Observational Analysis

In addition, distinctive scenes were found in the video recordings of the experiment. For example, sometimes, participants directly communicated when standing on different sides, which was not common in the experiment. Furthermore, participants seemed to have difficulties showing an object (i.e., a sheet of paper with printed images) to remote partners in the 2D flat display condition because of camera occlusion.

##### Communication when Standing on Different or Opposite Sides

In the experiments, four participants were always split into two groups for parallel working to increase efficiency. Most of the time, participants stood across from their remote partner or at least on the same side, which gives them a more natural feeling of communication. By standing across from each other, participants were not only visible

Advantages of the spherical display	Number of people
Strong three-dimensional sense	20
The spherical display is easier to be seen because of less distortion	20
The spherical display is more natural to be seen. Users just need to lower their head a little bit to see the display	10
Since the spherical display only shows half of the 360-degree images, it brings a better feeling of private conversation with remote partner and they would not be disturbed by the information from the opposite side	8
The spherical display provides better sense of position and directional relation	6
The sense of telepresence is better, which leads to a good feeling of conversation	6
The image on the spherical display is visible from anywhere in the room	5
Users can better be aware of the space range where the remote partner can see, which helps to understand what they are saying	3
Disadvantages of the spherical display	Number of people
Users often have to walk around to check the other sides of images, which is not convenient	30
Sometimes, users get lost with their remote partners or just hear the voice of remote partners without seeing them	5

Figure 4.13: The pros and cons of using the spherical display in the panoramic video communication.

Advantages of the 2D flat display	Number of people
Users can directly see the whole 360-degree image	32
They could put lots of objects above the omnidirectional camera to show them and they knew that these objects could be seen by the remote partners	3
Disadvantages of the 2D flat display	Number of people
The distortion is too serious to make it easy to be seen	23
It is hard to see the display clearly from a slight distance	9
Always have to lower head which is tiring	8
The sense of distance and space is bad	5
Sometime, it is hard to tell which remote partner is talking to you	4
Figures shown on the other side are inverted. If users want to see it clearly, walking is still needed	2

Figure 4.14: The pros and cons of using the 2D flat display in the panoramic video communication.

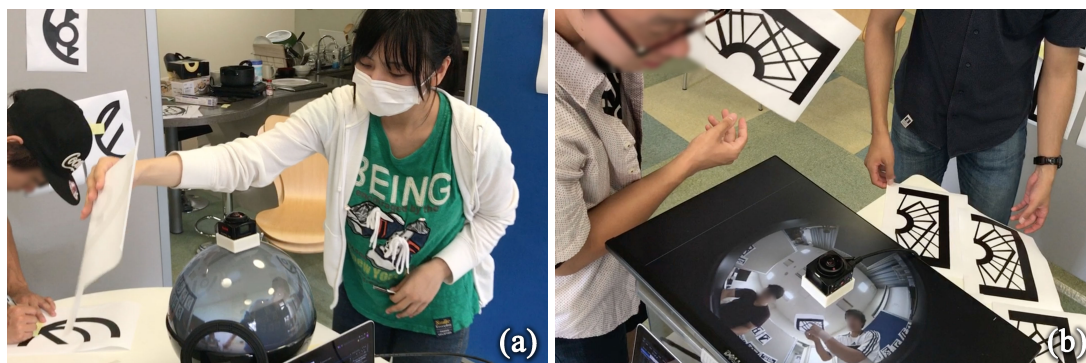


Figure 4.15: Communication when standing on opposite sides: The local participant directly puts the image on the opposite side to show it to the invisible remote partner (a). By seeing the inverted image, the local participant directly talked to the remote partner who was standing on the opposite side (b).

to each other during the conversation in the spherical display or 2D flat display condition, but they could also avoid seeing an inverted figure or object. However, some distinctive scenes that observed were that participants communicated when standing on opposite sides, which means that person A in the local room was standing on one side of the system, while person A's remote partner in the remote room was standing on the opposite side of the remote system. In the 2D flat display condition, this positioning represents the fact that the remote partner's image was inverted from the local A's point of view. However, in the spherical display condition, this means that the image of the remote partner was on the opposite side and was thus invisible to the local A. Three out of 10 groups were found to be engaged in this type of communications (four times in total) in the spherical display condition. Because the remote partner was invisible, the conversation seemed to mainly rely on voice communication. As an extreme case, there was one occurrence of a participant directly showing an image to an invisible remote partner, as shown in Figure 4.15(a).

On the other hand, such conversations were observed more frequently in the 2D flat display condition (19 times in seven groups). As shown in Figure 4.15(b), the person holding a figure was talking to the remote partner who was also holding a figure. Al-

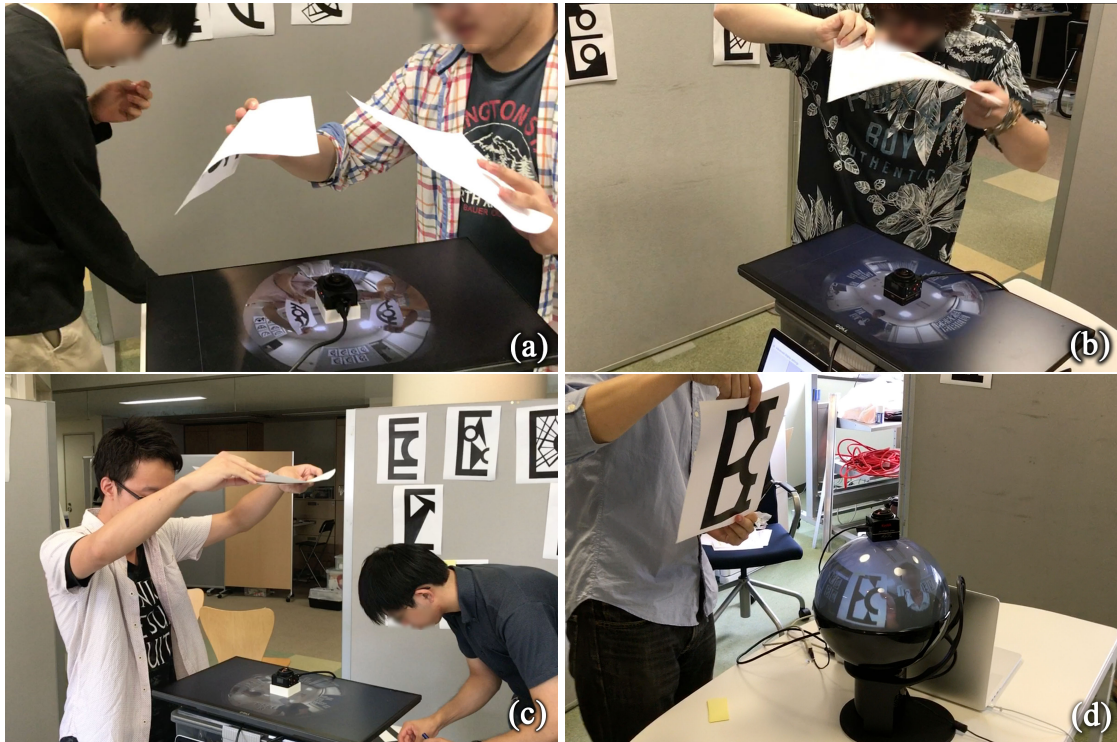


Figure 4.16: The camera occlusion issue was observed in the 2D flat display condition (a)(b)(c). But in the spherical display condition, this issue was not found and participants could naturally show figures (d).

though the remote partner's image was inverted from the local person's point of view, the 2D flat display seemed to help facilitate this type of conversation because the remote partner was visible. In addition, it was observed that when such scenes occurred in the 2D condition, the four participants usually paused the parallel one-to-one communications and communicated together.

#### **Difficulties of Showing Objects with 2D Flat Display**

Other distinctive behaviors were observed in the 2D flat display condition. The participants encountered the camera occlusion issue when showing the figures to their remote partners. As shown in Figure 4.16 (a)(b)(c), participants tried to place the figures exactly on top of the 2D flat display to show it. However, the image was blocked by the camera on the remote side because the object located on the top of the Kodak camera will be

.....

captured and shown at the center of the image. Four out of 10 groups were observed to suffer from this issue when only using the 2D flat display. Normally showing the image at the side of the 2D flat display could solve the occlusion issue. However, in the free-response question, two participants indicated that *“They had to look down to the 2D flat display, while showing the objects at the side of the display instead of above the display made them feel inconsistent.”* In the spherical display condition, this issue was not found and participants could naturally show figures, as shown in Figure 4.16(d). The occlusion issue of the 2D flat display was caused by the unusual configuration of the 2D flat display’s video system. If the omnidirectional camera becomes much smaller or users get more used to this system, it might not be a serious issue. However, as mentioned by a participant, *“it is not intuitive to look down on the 2D flat display, while showing the objects at the side of the display.”*

## 4.5 Discussion

The results of the objective measures, including completion time and accuracy, showed there was no significant difference in collaboration efficiency between the two conditions. Meanwhile, based on the questionnaire results, the results of six out of seven questions showed that the spherical display was significantly better than the 2D flat display, indicating that the spherical display was slightly preferred over the 2D flat display as a whole. Below, the potential reasons for this result are discussed by focusing on other factors, for example, visual perception, awareness, perceived workload, and affordance.

### 4.5.1 Visual Perception, Awareness, and Perceived Workload

It was found that the 2D flat display enabled the participants to see everything in one view but increased their cognitive load because of image distortion and the necessity for mental rotation. The spherical display could only show half of the environment at once but provided a better visual cognition, sense of space, and consistency and coherency of the two distant spaces.

.....

Such findings are supported by the questionnaire results and comments from the participants. The results of Q1 showed that the spherical display provided users a better visual cognition, which is supported by the participants' comments such as "*The spherical display has less distortion and stronger three-dimensional sense*" or "*The 2D flat display has severe distortion*". Besides the easiness to recognize the images displayed on the screen, participants also mentioned that "*The spherical display provides a better sense of space, distance, and directional relation*" and "*The sense of distance and space is bad in 2D flat display condition*", which might explain the results of the questions related to the presence (Q2 and Q3). Overall, these results indicate that the images displayed on the spherical display enabled the recognition of spatial relationships more intuitively and brought a better feeling of telepresence. It suspects that such a good sense of space, direction, and telepresence helped generate the consistency and coherency of the two distant spaces.

On the other hand, the advantage of the 2D flat display was mainly reflected by the fact that users could see the whole 360° images at a glance, which presents several merits. For instance, some participants mentioned that "*They could quickly locate remote partners or obvious objects by seeing the whole 360° image.*"

However, seeing 360° images in one view might also give rise to discomfort when using the 2D flat display. The results of Q4 indicated that the use of the 2D flat display resulted in the users feeling more fatigued. One explanation might lie in what nine participants indicated that "*They always have to lower head, which is tiring when using the 2D flat display*". Another reason might be due to overloaded visual information. Johnson et al. [38] indicated that 360° panoramic images were found to be more challenging to use than a 180° interface, which they believed supported the theory that a wider field of view requires higher amounts of cognitive processing to synthesize the larger quantity of visual information provided. In such a case, the participants thought that using the 2D flat display made them feel more fatigued because the 360° panoramic images include more visual information that adds to the cognitive processing burden. "*Sometime, it is hard to tell which remote partner is talking to you*" was conveyed by four participants. This

.....

was an issue caused by overloaded visual information. Furthermore, two participants also mentioned that *“Figures shown on the other side are inverted. If users want to see it clearly, walking is still needed”*. In fact, there were some instances where participants chose not to move around but observed the inverted figure, which might cause the mental rotation process. On the other hand, in the spherical display condition, participants did think likewise because the opposite side of the spherical display was invisible and they could only see half of the 360° image. But some participants expressed that *“it provided them a more private conversation environment because they would not be distracted by the opposite half of the 360° image”*; thus, they could concentrate on the communication without distraction. In addition, 10 participants commented that *“The spherical display is natural to be seen. They just have to lower head a little bit to see the screen”*. These factors might help the participants to lower the perceived workload when using the spherical display.

Moreover, it seems that restricting the visual space (i.e., showing only half of the space) gives the spherical display more advantages concerning awareness of remote communication. Luff et al. [49] proposed an issue about the ecology of video communication, and they noted that the local person might not be able to make out where the remote person is looking during the video communication, particularly when the field of view is wide. It was found that the same issue occurred in the symmetrical 360° video communication as the participants could see the whole remote space in the 2D flat display condition. However, the issue may be resolved by using the spherical display. Indeed, three participants commented that *“When using the spherical display, they can better be aware of the space range where the remote partner is able to see, which helps to understand what they are saying.”* As an example, the girl in Figure 4.15(a) did not see her remote partner who was standing on the opposite side, looking for her. However, she could directly show the image to the opposite side because she was aware of the space range in which her remote partner was able to see. As for the 2D display condition, one participant wrote that *“he had no idea where the remote partner was looking at in the 2D flat display condition.”*

### 4.5.2 Affordances of Displays

The scenes arising from the camera occlusion might indicate that the upward orientation of the 2D flat display affords participants to hold objects right above the display. The spherical display affords participants to hold objects in front of their chest naturally. The 2D flat display requires participants to look down to see the display, which is quite unnatural for the participants when they are showing objects. A similar issue was also reported by Licoppe et al. [47]. The spherical display does not have such issues.

In addition, it was observed that participants in the 2D condition communicated a little more as a group (with four people) than parallel one-to-one communication, compared with the communication pattern in the spherical display condition. It is suspected that such findings might indicate that the spherical display better affords one-to-one conversation, whereas the 2D flat display better affords many-to-many communication. This point is also supported by eight participants' comments, "*The spherical display brings a better feeling of private one-to-one conversation*". Moreover, the results of Q5 showed that it was easier to talk to a specific remote partner when using the spherical display.

### 4.5.3 Proper Real-World Scenarios

Comments about the 360° panoramic video communication were obtained from participants. They thought that "*The 360° video communication is not able to protect their privacy during video communication with families or friends in daily life since the whole room is shown to the remote side*". But they did agree that "*The 360° field-of-view shows much more information than regular video communication and they felt a better sense of the space, directional relation and telepresence of the remote side.*" Participants indicated that "*The 360° video communication may be helpful in the video conference since there are multiple users on both sides to use the system*". I agree with this opinion and believe that the 360° video communication may also help to improve the collaborative efficiency of the remote collaboration tasks when users need to have comprehensive information about the remote side.

Based on the objective results, the type of displays did not significantly change the collaborative efficiency. However, the questionnaire results showed that the spherical display was given a slight edge. Also, the participants felt that the spherical display's images had less distortion and provided a better sense of space and direction. As a result, they appeared to perceive less workload when using the spherical display. Therefore, the spherical display is proposed as a more proper display for 360° video communication, particularly for group-to-group collaboration that requires showing and seeing real-world objects.

As a realistic scenario, recent discussion rooms tend to use a whole wall or multiple walls in a room as whiteboards. A company like Smart Wall Paint is proposing to transform any smooth surface in a room into a write-on wipe-off surface. With the advancement of video technologies, researchers have been envisioning the use of most of the surfaces in a room as video displays. In such a room, participants inevitably walk around in the room to use all the available surfaces, and refer to, manipulate, and discuss the objects and drawings in the room. These trends strengthen the necessity for video communication systems to support collaboration between such rooms by solving the narrow field-of-view problem. I believe that remotely connecting such two discussion rooms through symmetrical 360° video communication may improve the remote communication efficiency.

#### 4.5.4 Impact of the Reversed Space

Although the system design choice for the experiment caused the reversed space phenomenon, according to the observation and user interviews, it seems that the participants were unaware of this phenomenon. Instead, since the teleconference metaphor was chosen for the proposed system, the participants could see the images of the remote site in a familiar way, and also they could use the directional expressions like “*on your left/right*” or “*on your left/right hand side*” naturally. In fact, seven groups out of ten used such directional expressions. If the mirror metaphor was chosen, the participants might have

.....

difficulty matching the patterns and explaining directions.

In summary, although the 360° image mapping method that used caused the reversed space phenomenon, it enabled to show the texts and images(patterns) properly. The method also allowed the users to use familiar directional expressions in the 360° video communication.

However, the reversed space phenomenon may become an issue in some cases. The reason why the participants were not affected by the reversed space in the experiment might be that the remote collaboration task did not require participants to match spatial arrangement between two spaces. For example, if an application is arranged to arrange the furniture of two rooms in the same layout, the reversed space may become an issue and cause trouble. Therefore, it supposes that the metaphor should be chosen based on the specific use scenario. The teleconference metaphor is more proper to be used when showing objects with texts or figures are needed, or the cues of the left-right directions are often used. On the other hand, the mirror metaphor without the reversed space issue may be useful in cases in which users need to have a sense of spacial directions about the remote site.

## 4.6 Limitations and Solutions

One of the limitations of the proposed system was that the OEB only supported hemispherical live video streams. It was occasionally observed instances where a local participant disappeared from the field of view of the camera when he or she bent down to put the image on the lower part of the wall. In such cases, a remote partner had to pause his or her talk and wait for the local participant to come back into the field of view. It was interesting to see that even a hemispherical field of view was not wide enough to cover all the possibilities of the participants' bodies' spatial positions during the interaction. To solve this problem, a full-spherical version of the panoramic video communication system was planned.

Another limitation was that the frame rate of the current live stream is limited to 5fps

.....

due to the transmission speed of the omnidirectional camera that used. Although no one complained about the slow frame rate in the pilot test interview, some tasks may require a higher frame rate. Therefore, it is future work to improve the frame rate of the system.

Furthermore, a possible research topic is to use a shape-changeable display [77] to utilize the advantages of both display types.

## 4.7 Summary

In this chapter, a 360° hemispherical video communication system, named as OEB, was proposed to explore the possibilities in assisting symmetrical remote collaboration. One of the research interests was how the display shape, the 2D flat display, or the spherical display affects the remote collaboration.

The experimental results showed that both systems (OEB and the 2D display version device) enabled two pairs of people to communicate over the distance simultaneously from anywhere around the devices. However, subjective results indicated that participants preferred the OEB over the 2D display version. This could possibly be due to better image quality (less distortion) and a better sense of three-dimensional space, directional relation, and telepresence.

The study also exposed the shortages of the OEB system. The participants commented that they had to walk around to check the other side of the display in the spherical display condition. It may be better to add an alternative way to allow users to see the other side of the spherical display without walking around. The solution to such issues is expected in the continued study.

---

## Chapter 5

# OmniEyeball2: A Full Spherical Interactive I/O System For Symmetric 360° Video Communication

### 5.1 Preface

One of the systems that have the potential to support such collaboration is OmniEyeball (OEB), which was presented in the previous chapter. OEB is a hemispherical display equipped with an omnidirectional camera. By placing an OEB terminal at each distributed site and mutually sending the whole 360° video images of each environment around the terminal, the system enables bidirectional video communication. Although the study showed that OEB is effective in supporting object-centered collaboration, they also reported the drawbacks of the system. For instance, the display only covered the upper hemispherical view, leaving the lower hemispherical view invisible.

Meanwhile, it is reported that spherical display suffered from the issue of the invisible hemisphere, as shown in Figure 5.2. According to their results, the invisibility problem sometimes led to inefficient collaboration, for example:

1. Users often had to walk around the spherical display to get the overall picture of the remote space.
2. Sometimes, users lost sight of their remote partners.



Figure 5.1: The OmniEyeball2 system which integrates an omnidirectional camera with a full spherical display (left), and its symmetric 360° panoramic video communication application (right).



Figure 5.2: The issue of the invisible hemisphere in 360° video communication. a) The local OmniEyeball2 is surrounded by four people. b) From any viewpoint in the remote site, only one or two people can be seen at once.

According to Luff et al., these are essential features that must be equipped in video communication systems for object-centered collaboration [50]. And the problem of invisibility, especially not being able to see the opposite side of the sphere, is inherent in the spherical display. Solving this issue is essential to support symmetric 360° video collaboration.

Recently, researchers used a novel interface design to address the problems of 360° video [13] [56]. This chapter investigates how the novel interactive features support users' interaction with 360° video image and address the issues of invisibility in video communication. Although there have been some existing studies that proposed various interaction techniques for spherical displays or 360° video image [4] [30], this is the first to explore the design of user interface when a spherical display is employed to show 360° video image in symmetric video collaboration.

## 5.2 Improvement Proposal

To address these issues, the study proposes a novel video communication system, OmniEyeball2 (Figure 4.5), an improvement of OEB. OmniEyeball2 (OEB2) mainly obtains four new features:

1. **Sharing the lower half of the environment.** While the spherical display of OEB only covered the upper hemispherical view, OEB2 covers a full spherical panorama. This allows the users to share an awareness of the lower half of the environment.
2. **Providing convenient access to any side of the sphere.** A mechanism that allows users to rotate the spherical display physically was added to the OEB2. This mechanism allows the users to see the focused view of any side of the display by simply rotating the sphere.
3. **Supporting a 360° overview.** A bird's eye view segment is added at the top of the spherical display, which supports the awareness of the opposite side of the spherical display. This design allows the users to simultaneously monitor the overview and the regular view of the remote site.
4. **Supporting focused view and gaze awareness.** When a user touches an image in this bird's eye view segment, a full-size image of the touched region ("focused view") appears so that the user can quickly view any region of the remote space [10]. Additionally, a user can share his/her focused view with the remote partner by a long press of the region.

## 5.3 System Implementation of OmniEyeball2

This section presents how the system of OEB2 is designed to resolve the issues introduced by the use of the spherical display.

### 5.3.1 Full 360° System: Sharing the Lower Half of the Space

To support a 360° full-spherical view, the OEB2 system is created by integrating the Ricoh Theta V <sup>1</sup> with the Glomal350 spherical display <sup>2</sup>. As shown in Figure 5.1, the

---

<sup>1</sup><https://theta360.com/>

<sup>2</sup><http://glomal.jp>

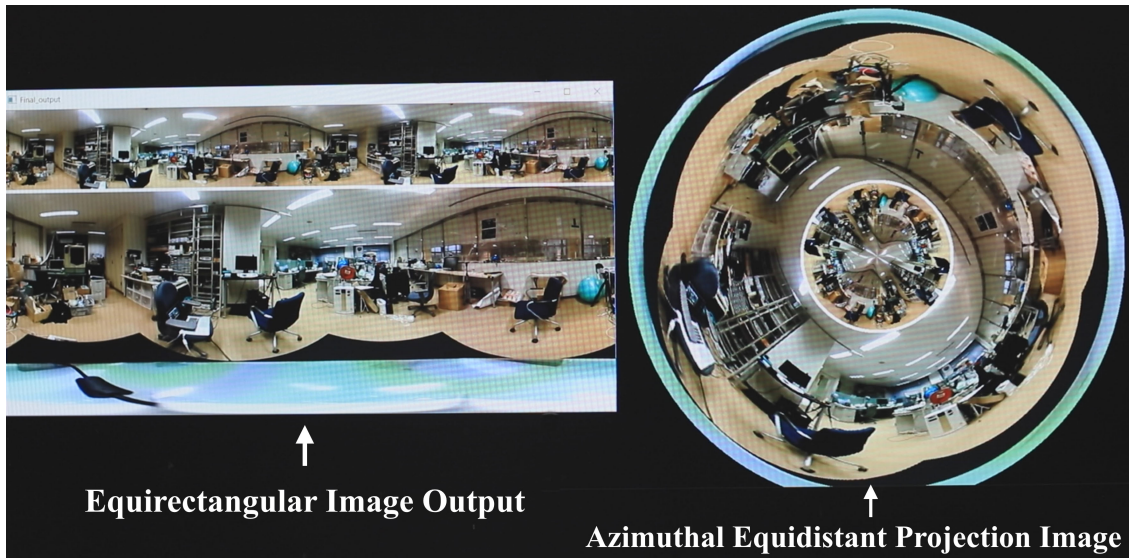


Figure 5.3: The graphic system takes the equirectangular live streaming video(left) as an input, remaps and resizes it onto an azimuthal equidistant projection image(right) for projecting to the spherical display.

Ricoh Theta V is attached at the top of the spherical display. The Ricoh Theta V is an omnidirectional camera equipped with two fisheye lenses, which can provide an omnidirectional panoramic live streaming video with  $1920 \times 960$  pixels over a USB connection.

The display part is based on the Glomal350 spherical display from Shibuya Optical CO., LTD. The Glomal350 spherical display uses a wide-angle lens to project a rounded image from its Casio projector onto a spherical surface ( $\varnothing 350\text{mm}$ ). The bottom of the spherical surface is occupied for the connecting parts, making the displayable portion of the sphere at around  $300^\circ$  vertically and  $360^\circ$  horizontally.

The image format obtained from the omnidirectional camera is an equirectangular format. For projecting to the spherical display, the image needs to be converted from equirectangular format to azimuthal round map format, shown in Figure 5.3. The conversion is shown in Figure 5.4. For an equirectangular image, if the height is treated to be the radius of the round map and the width is treated to be from 0 to  $2\pi$ . Then, given a pixel  $P'(x',y')$  in a round map, it is easy to compute its position  $P(x,y)$  in the equirectangular

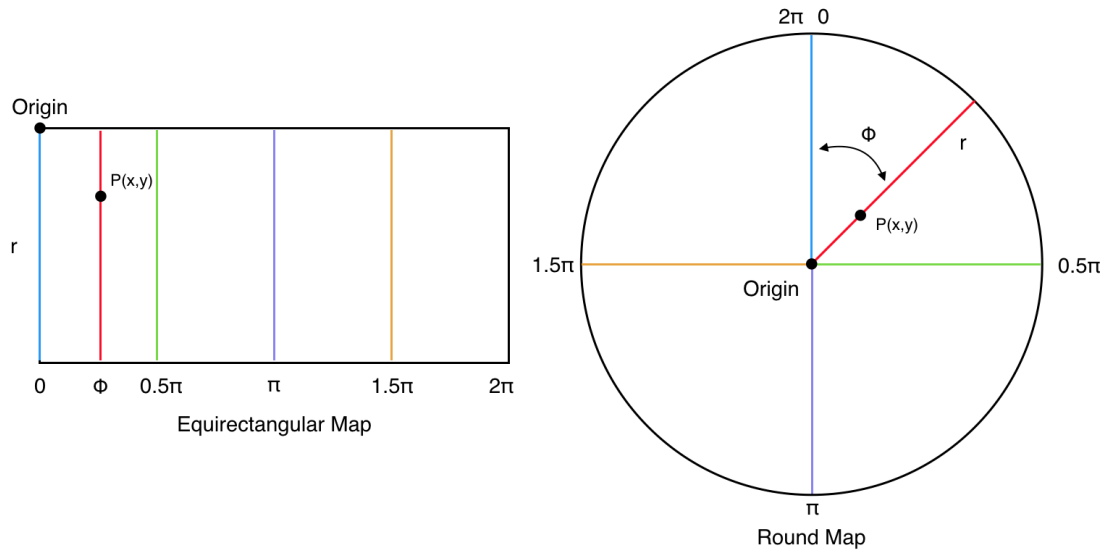


Figure 5.4: How to convert a point  $P$ 's position from equirectangular map to round map.

image. The color of  $P$  in the equirectangular image will be assigned to the pixel  $P'$  in the round map.

$$\theta = x/y$$

$$x' = r + y \times \sin \theta$$

$$y' = r + y \times \cos \theta$$

The projector can output images with  $1024 \times 768$  pixels. However, the visible part of the projected image is limited to a circle with a diameter of 768 pixels. The final display quality is also affected by the setting of the projector as well as the ambient light.

### 5.3.2 360° Video Communication Application

Two OEB2 terminals and two PCs are arranged in two different places to set up the 360° panoramic video communication. As shown in Figure 5.5, each OEB2's camera captures its surrounding environment and streams the live video to the local PC. Then, the two live streaming videos are transferred to a PC server, which equips with an NVIDIA

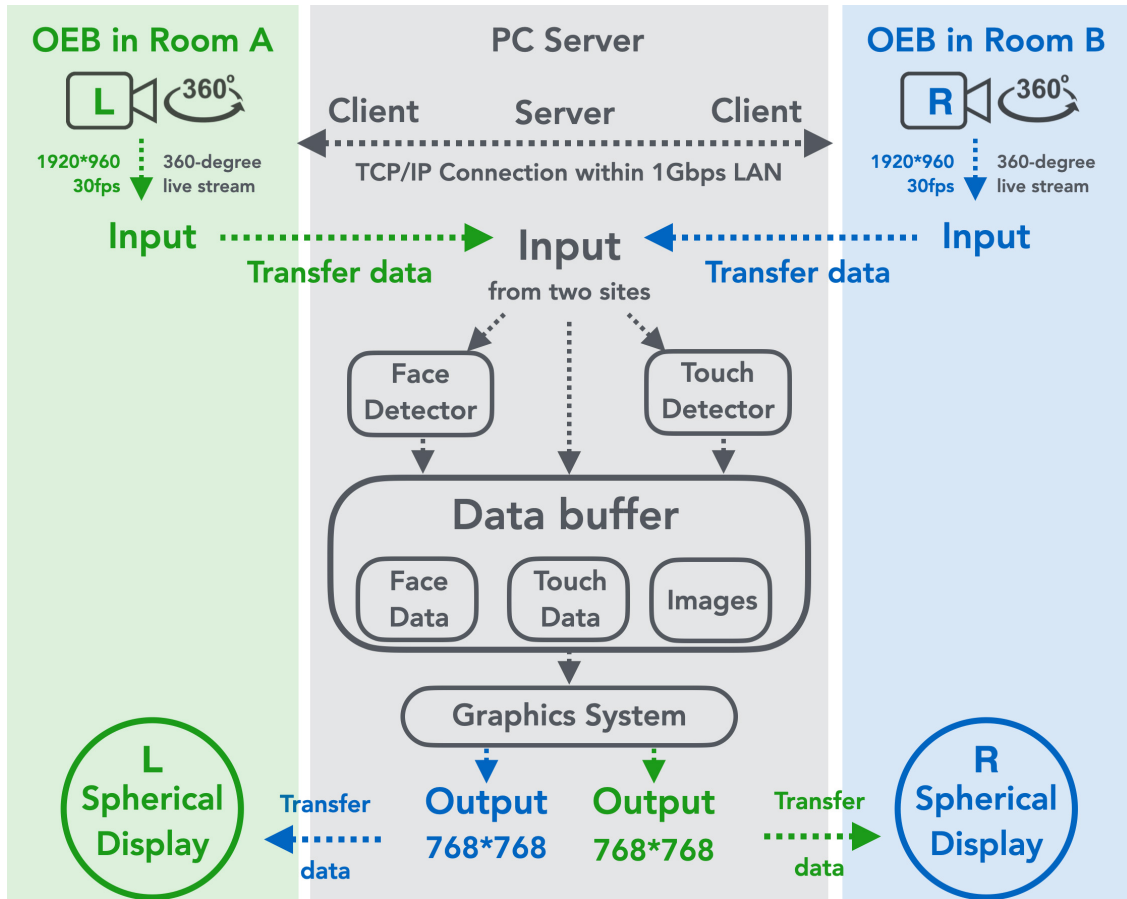


Figure 5.5: The diagram of the 360° video communication based on OEB2.

Geforce 1080 card (12GB memory). The PCs are connected by 1Gbps LAN (local area network) through TCP/IP protocol. Once the server receives the input, the face and touch detectors process the live streaming videos and provide the detected data to the graphics system. Then, the graphics system generates the live output videos, remap and resize it onto an azimuthal equidistant projection image for projecting to the OEB2's display. Finally, the server transfers the live output videos back to two OEB2 terminals for display. The frame rate of the whole system is 10 fps on average.

The frame rate of the whole OEB2 system (10 fps on average) is mainly affected by the vision-based sensing techniques used in the OEB2 system, which will be introduced in the following sections. Therefore, regardless the influence of the sensing technique to

Condition	Resolution for Stream Transmission	FPS	Theoretical Required Bandwidths
OEB2 plain mode	768 × 768 (OEB3 Display)	30	>1Mbps (up/down)
OEB2 interactive mode	1920 × 1080 (1080p)	30	>2Mbps (up/down)
OEB4K	3840 × 2160 (4K)	30	>20Mbps (up/down)

Figure 5.6: The theoretical required bandwidths for different configurations of OEB system.

the system's frame rate, the 360° panoramic video transmission can easily achieve 30fps. Since the data transmission is not the goal of this research, the current data transmission method used in the OEB2 system is a simple approach based on TCP/IP protocol implemented by OpenCV. Due to the 1Gbps LAN connection, the connection bandwidths did not become the bottleneck of the system's frame rate. Furthermore, by referencing the configurations of the current popular video communication platform, such as Zoom or Skype, the upload and download required bandwidths for the OEB system to achieve 30fps video communication can be predicted and shown in Figure 5.6. With the rapid development of technology in the field of video transmission and compression, it is expected that a 4K version of OmniEyeball implemented in the future.

### Calibration of Orientations

After the video communication connection is established, a careful orientational calibration is needed to enable face-to-face communication as well as to calculate azimuth of faces and touches between two sites correctly. For example, the orientations of cameras and spherical displays should be configured appropriately, either in physical or digital way, to achieve the face-to-face status, which means that if person A stands across from person B (shown in the display) at one site, A's image should stand across from B at the

other site.

### **5.3.3 Physical Rotation: Providing Convenient Access to Any Side of the Sphere**

Although OEB2 can show the whole 360° image on its spherical display at once, users can only view half of the image due to the issue of the invisible hemisphere. Therefore, It is expected that an image rotation function might be necessary for OEB2. One straightforward but intuitive solution is to enable physical rotation. For this purpose, a rotating platform is built which is similar to a Lazy Susan and mounted the spherical display on it. The rotating platform allows users to see the other sides of OEB2 by rotating the spherical surface intuitively (Figure 5.7).

Through this function, not only can users view other sides of the OEB2's display, but also the camera on the top of OEB2 will synchronously rotate with the display. This function can keep the video communication system in face-to-face status, that is, if a user physically rotates OEB2 to check one direction of the remote environment, the image of this user on the remote OEB2 will synchronously be rotated to face that direction in the remote site. Thanks to the synchronization, the standpoint locations and orientations of a user can be continuously shared over distance.

### **5.3.4 OmniBar: Supporting A 360° Overview**

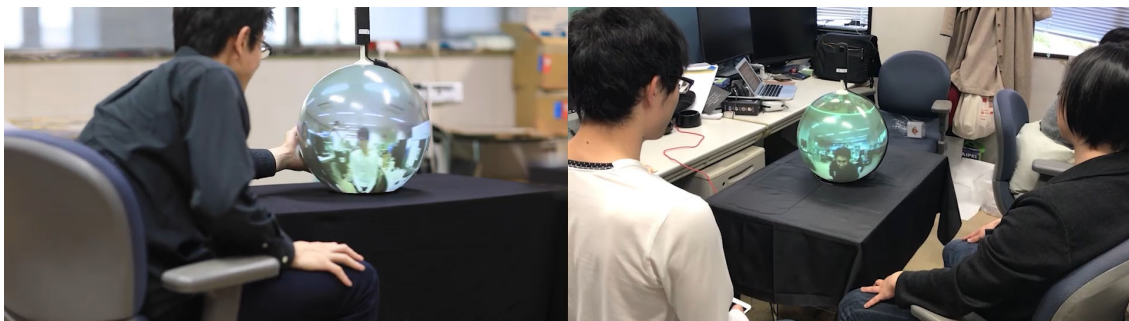
The OmniBar is a 360° ring-shaped image bar displayed on the top of OEB2's display (Figure 5.9b) to provide a bird's eye view, which supports the awareness of the opposite side of the spherical display. It also works as a natural touch interface to enable intuitive observing the remote environment (Quick View). To implement the OmniBar, a segmentation design is made of the system and also enabled the face and touch detection.



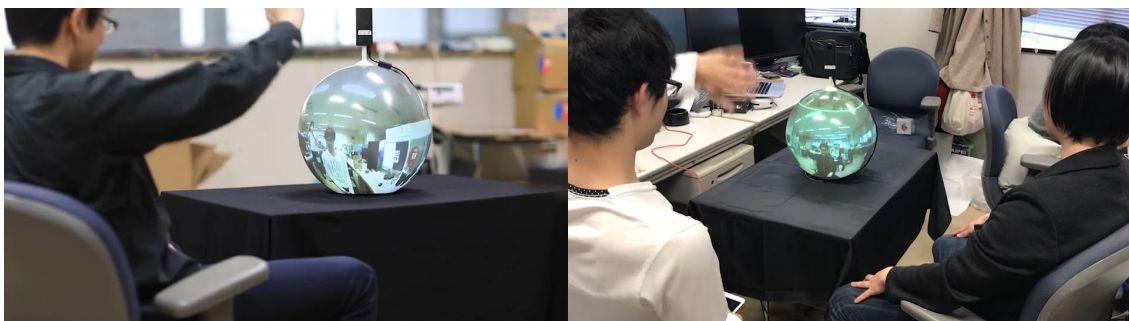
(a)



(b)



(c)



(d)

Figure 5.7: The Physical Rotation feature of OmniEyeball2

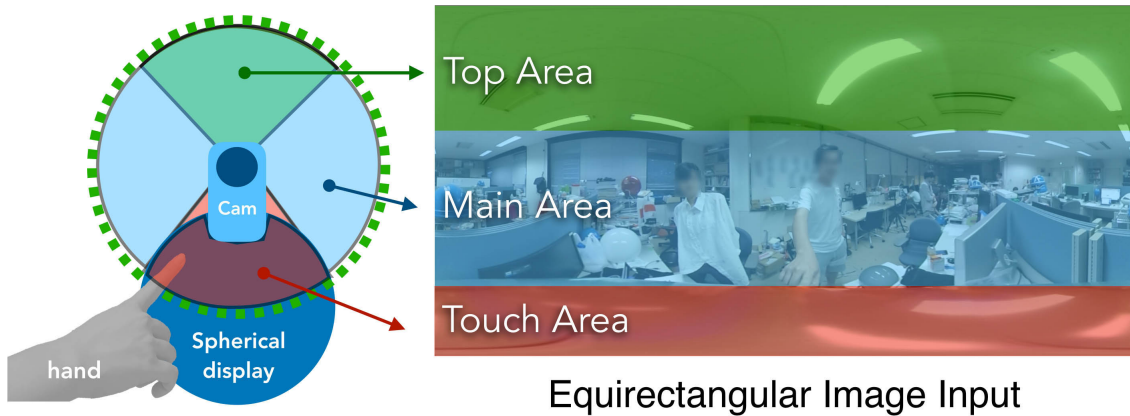


Figure 5.8: The 360° image captured by OEB2 is segmented into three image areas.

### Segmentation Design

The image fed to the software system is a live streaming video captured by Ricoh Theta V in an equirectangular format, with a resolution of  $1920 \times 960$  pixels, 30fps. The input image is divided into three image areas, “top area”, “main area” and “touch area” (Figure 5.8). When OEB2 is placed in a room, the image of the ceiling always takes up around the top third area of the input image, which is named the top area. Moreover, the most bottom area of the input image constantly shows the distorted image of OEB2’s display itself. Since it would be used for detecting a user touches OEB2 surface, named as the touch area. The remaining middle area is named the main area, containing the most useful visual information.

### Design of OmniBar

Since the original image at the top area was always showing the ceiling of the room, which was useless visual information at most times, so an idea was out to use this area as a user interface for a user. Therefore, it replaced the image of the ceiling with the OmniBar to provide a bird’s eye view for the user. As shown in Figure 5.9a, the OmniBar is made by cropping the main area of the input image and duplicating it in the horizontal direction to create an image of 720° field of view. Therefore, no matter where a user stands around

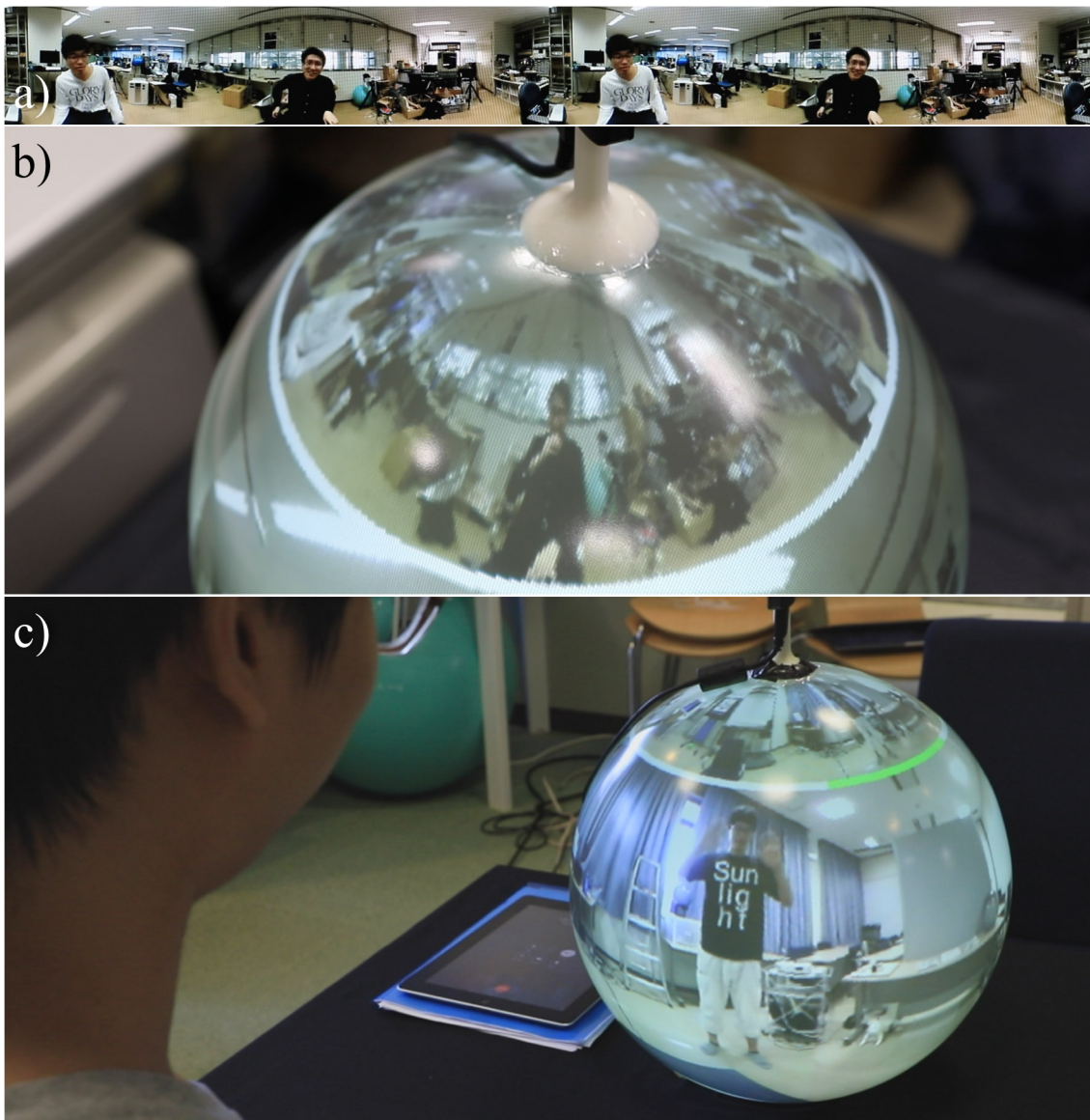


Figure 5.9: a) The OmniBar is created by duplicating the main area of image input in the horizontal direction. b) The OmniBar projected on the top of OEB2. c) The Location bar can show the position of the current viewpoint in the OmniBar.

OEB2, the user can see the 360° field of view overview of the remote site in half of the OmniBar area. Similar to showing an equirectangular 360° image on the 2D display, the OmniBar is not suitable for expressing the directional and spatial sense. However, it is expected that the OmniBar is a good visual support for users to perceive the remote site's overview and be aware of the activities there.

The feature also proposes the Location bar, which is a green bar shown in Figure 5.9c. It indicates where the user's current point of view is located in the OmniBar.

### 5.3.5 User Detection

Sensing technology is an essential part of the OEB2 system. Unlike traditional interactive terminal devices, OEB2 systems that can support 360° panoramic display need to rely on user detection technology to obtain the user's relative orientation to the device in order to provide a proper interactive interface. The 360° camera equipped on the top of OEB2 becomes a natural 360° sensing device. By referencing to the latest user detection technology and considering the trade-off between detection performance and real-time, the OEB2 system can mainly adopt the following two RGB vision-based solutions:

1. AR marker detection.
2. Vision-based face detection.

AR marker solutions, such as QR code detection, can provide high accuracy and real-time performance for detecting the user position. However, this solution requires the user to wear the AR marker, which may be inconvenient in some scenes. Therefore, the OEB2 applied the vision-based user face detection.

The face detection is kept activated in the OEB2's software system. It utilized deep learning (DL) based face detecting function in the Dlib C++ library [43] to detect faces that are appearing in the live streaming video. While using OEB2, users have to keep away from OEB2 for a distance (usually longer than 0.5m) to naturally view the display, which leads to a condition that the faces of users mostly only appear in the "main area" of

the input image. Therefore, to save the computation cost of the system, the face detector is configured to only process the “main area” of the input image.

However, running DL-based face detection algorithms in real-time can consume a tremendous amount of computational power. Even with an Nvidia GeForce GTX 1080 graphics card, It is nearly impossible to do face detection in a 360-degree panoramic input in real-time. Meanwhile, the pixel size of a 360° 1080p omnidirectional panoramic image input is huge. Decreasing the resolution of the image input can speed up face detection. However, it may bring problems with detecting a user’s face at a distance.

To solve this issue, a quick object tracking method is applied to support the DL-based face detector. The final face detection method used in the OEB2 is a combination of DL-based face detection and quick object tracking. The specific process is:

1. For the 1st frame, the DL-based face detection detects the positions of all faces appearing in the main area of the 1st frame from 360° image input.
2. For the 2nd frame, the object tracker takes the positions and bounding boxes of faces from the 1st frame as input and then detects the positions of faces in the 2nd frame.
3. For the 3rd-9th frame, the object tracker updates the positions of faces based on a quick object tracking method.
4. For the 10th frame, the DL-based face detection is activated again to detect the positions of all faces appearing in the main area of the 10th frame from 360° image input.

Following the strategy described above, the customized face detection method can detect faces from the 360° 1080p omnidirectional panoramic image input in real-time. The positions of the detected faces are used to implement the proposed features in OmniBar, which require the relative position between users and OEB2.

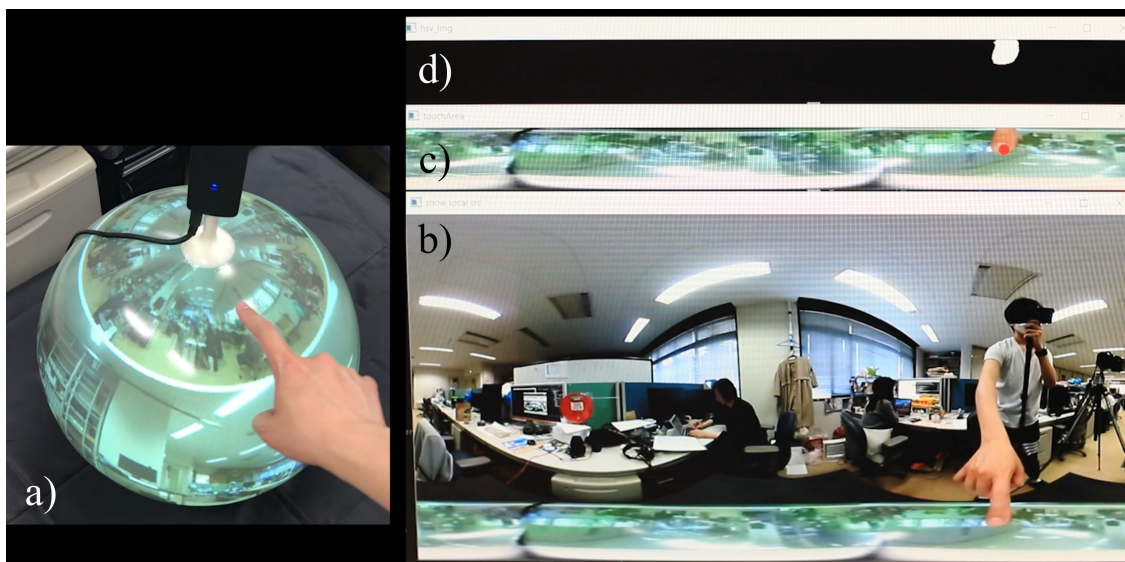


Figure 5.10: The overview of touch detection. a) A finger is touching the top area of OEB2. b) Image input of OEB2. The fingertip is captured inside the touch area. c) The RGB image of the touch area. The red point shows the position of a detected fingertip. d) Binarized image of the touch area after extracting the skin color.

### 5.3.6 Touch Detection

In addition to user detection, touch detection is also an important sensing function of the OEB2 system. By referencing to the latest touch detection technology on spherical display, the OEB2 system can mainly adopt the following three solutions:

1. Infrared light multi-touch sensing solution.
2. DL-based hand detection.
3. Skin-color-based touch detection.

The infrared light multi-touch sensing method is a popular touch detection solution for the spherical display. Several research [4] or commercial product <sup>3</sup> of spherical display

<sup>3</sup><https://pufferfishdisplays.com>

utilized this approach. A simplified version of infrared light touch detection was implemented in the OEB system and used for the pilot test. Unfortunately, the same approach can not be applied to the OEB2 system due to the hardware structure of the Glomal350 spherical display, which does not have enough space for the beam splitter, that is an essential part for infrared light touch sensing.

Therefore, the solutions left are all vision-based touch detection methods. Although the 360° camera equipped on the top of OEB2 is a natural 360° sensing device, the high resolution and wide field-of-view become challenges for the DL-based touch detection. Running a DL-based hand detection on input with 1080P resolution can generate high computation cost to the OEB2 system. Considering the DL-based face detection which already takes a certain degree of the computation power from the OEB2 system, the touch detection approach must be light-weight enough to reach the real-time running of the whole system. Therefore, the skin-color-based touch detection is applied in the OEB2 system, and it is found that the OEB2 hardware configuration can naturally support skin-color-based touch detection well.

As shown in Figure 5.10b, the touch area located at the bottom of the input image is always displaying the spherical display itself. It is because the camera has a 360° field of view, which can capture the whole space, including the display surface under it. Since the relative positions of the camera and display are fixed, the touch area constantly takes up ~17.3% area of the input image. The image of the display surface captured by the camera is found to turn greenish and bluish from its real color, which may be caused by optic reason or false color. Therefore the position of finger color can be stably detected from the touch area by using general skin color detection (Figure 5.10c and 5.10d). Based on this special condition of OEB2, touch detection is proposed by using image processing to detect the fingertip appearing in the touch area (Figure 5.10a). If the detected fingertip appears in the touch area for more than ~1s (determined by testing experience), it is detected as a touch. The location of the fingertip, which is calculated by the pixels near the lowest point of the finger contour, is recorded.

However, during the pilot test, it was found that the touch detection method was quite

.....

sensitive to the red color appearing in the OmniBar. Some pure red objects appearing in the OmniBar can not be turned to a bluish or greenish color, which may be wrongly detected as skin color. To solve this issue, I applied a simple trick to improve the touch detection more robust. Since the finger area (Figure 5.10c and 5.10d) always connected with the top boundary of the touch area, I removed the detected areas which are not connected with the top edge, since they can not be a finger. By using this simple solution, the detection is found to be more robust.

### 5.3.7 Quick View: Supporting Focused View & Gaze Awareness

Although the OEB2 already has the Physical Rotation function, while being physically rotated, the image shown in the remote display synchronously rotates with the local OEB2, which may disturb the remote user. It is believed it is essential to provide an additional rotation function that lets local users quickly check other sides of the image without distracting other local or remote users. Therefore, a digital way is designed to allow users see the different sides of the OEB2 by utilizing the OmniBar as an interactive interface, shown in Figure 5.11abcd.

Based on this feature, users can glance at any direction's image by touching on the OmniBar. The feature can display around 120° image of the touched direction as an independent focused window toward the user. Furthermore, users can swipe their fingers on the OmniBar as if they were digitally rotating the display to check any side of OEB2. This feature will not distract remote users or other local users standing on the opposite side, so they can still view the original 360° image.

Also, there is a subsequent gaze sharing function of the Quick View. By swiping on the OmniBar, a user can use the Quick View to observe the remote environment. Once he/she stops the finger for a long press, then, a Quick View sharing window will be activated in the remote site. As shown in Figure 5.11ef, by long press on a region displayed in the OmniBar, the image of the touched area will be cropped and shown as a small PIP window above the touching user's head shown in the remote OEB2 (Figure 5.11ef). This

.....

feature allows a user to share his/her focused view with the remote partner. It is designed that the long press as the key to activation. Otherwise, the vertiginous sharing window may become an excessive distraction to the remote partner.

The feature also has feedback designed for the sharing window by turning the white border of the Quick View window to red (Figure 5.11g).

## 5.4 Experiment

An experiment was conducted to study how the proposed features address the invisibility issues, improve the system's usability, and affect distant collaboration by enhancing the awareness of remote sites. It was also interested in studying how each of these features impacted the users' understanding and perception of the remote environment. Through the user study, it is planned to answer the following questions:

1. How the proposed features address the issue of the invisible hemisphere or improve the usability of the system in a collaborative task?
2. How the proposed features affect the users' understanding of the remote environment? How the users use each feature differently?

### 5.4.1 Conditions

The study compared two conditions: 1) the plain OEB2 condition, and 2) the interactive OEB2 condition. In both conditions, there is 360° video communication using two OEB2s set up on each site. The former condition used plain OEB2 without the proposed features, i.e., the OmniBar, Physical Rotation, and Quick View. Note that the capability of plain OEB2 is a full-sphere version of the OEB proposed in chapter 4. On the other hand, all the proposed features were enabled in the interactive OEB2 condition.

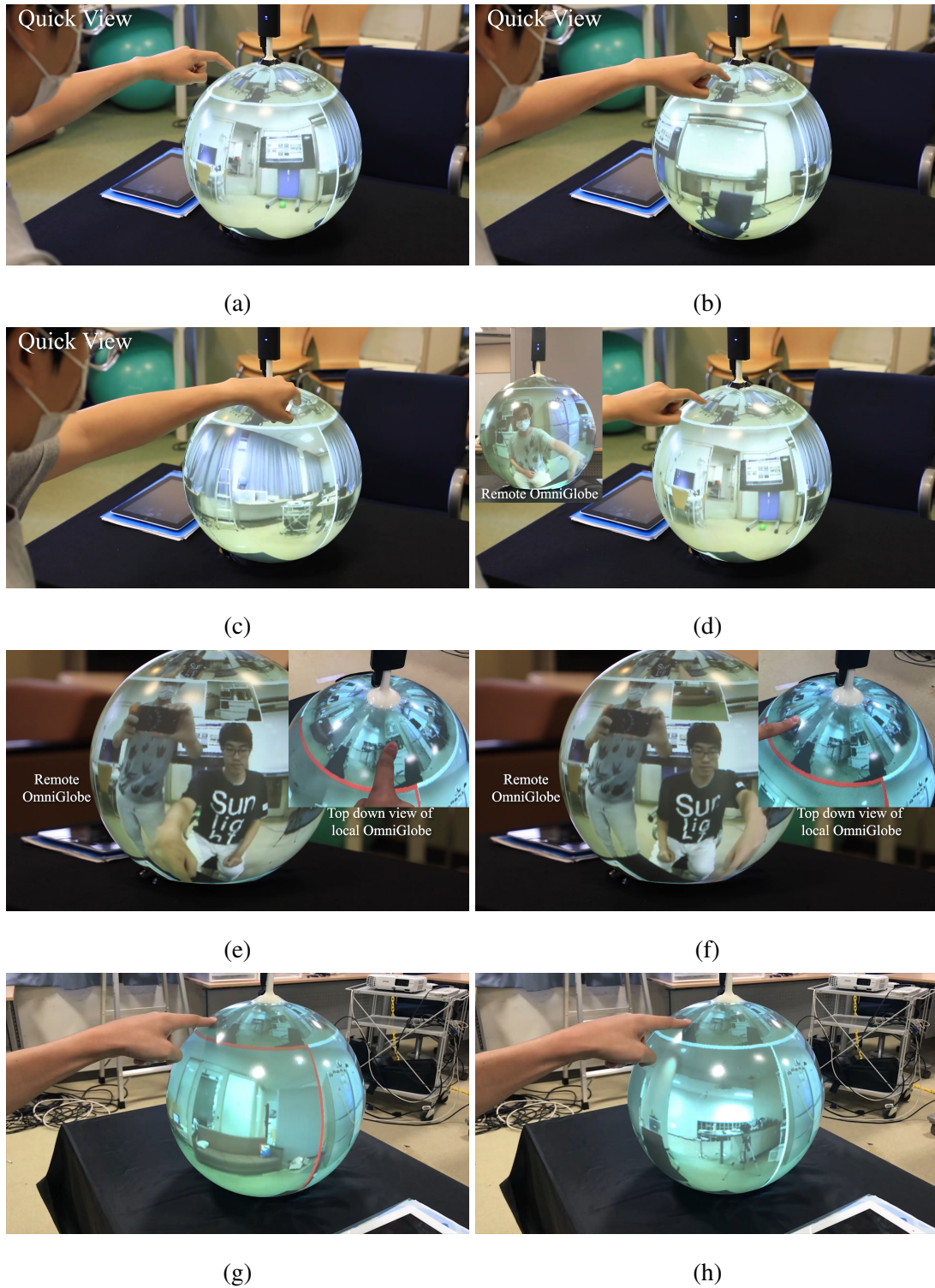


Figure 5.11: The features of OEB2, the OmniGlobe is a nickname of OEB2. abcd) Quick View. ef) Quick View sharing window. g) The feedback when the Quick View sharing window is activated. h) when the Quick View sharing window is NOT activated.

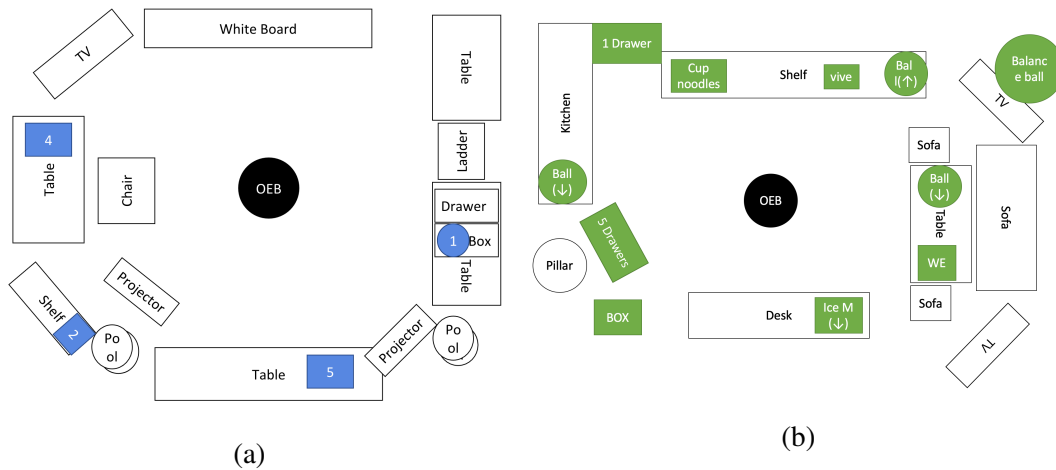


Figure 5.12: The object arrangement map used in the arranging task.

### 5.4.2 Task Overview

Inspired by the tasks employed by Norris et al. [59], the study designed a similar remote collaborative finding task between two rooms. In addition, a collaborative arranging task is designed. Both of these tasks are essential elements of distant collaboration using real objects, such as distant war room collaboration. Each task consisted of two participants, with one participant located in one room. The finding task provided each participant with several object images and asked them to find the actual objects placed in the two rooms collaboratively. The arranging task provided each participant with an object arrangement map (Figure 5.12), which showed how they had to arrange the objects found during the finding task. Because all the objects were in the remote room, each participant had to ask the remote partner to carry and place each object. The task is designed as a two-way task, which means that each participant had to ask his/her partner to find objects and arrange them in the partner's room.

The experiment took a within-subjects design, that is, participants engaged in the set of tasks in both two conditions, once in the plain OEB2 condition and once in the interactive OEB2 condition. Two different sets of objects were prepared and placed in different places of the rooms so that the participants would not use the same objects in the two

tasks. To minimize the learning effect, two task sets, two object sets, and the order of conditions were counter-balanced.

### 5.4.3 Participants

Twenty-four university students (21 to 29 years old, 17 males and seven females) were recruited from the Tokyo Institute of Technology, which constituted 12 pairs. The participants were recruited no matter they came as individuals or pairs. As a result, six pairs of participants knew their partners before the experiment. All participants had not experienced the proposed system before. Also, all participants received 1,000 Yen compensation for participating in the study. The detailed information of the participants can be checked in Figure 5.13

### 5.4.4 Environment

As shown in Figure 5.14, it used two mutually isolated rooms connected by the identical OEB2 video communication systems. The two rooms were arranged with different pieces of furniture, and they were decorated differently so that one emulated a research laboratory and the other emulated a living room. In this way, the participants had to figure out the remote environment by communicating with their partners through OEB2. It enabled audio communication between two rooms using two Apple iPads.

### 5.4.5 Task Procedure

The study first spent about five minutes to introduce the system to the participants. Then, it formed pairs and provided each participant with three images, each printed with a different object (six objects in total). The participants were asked to engage in a practice session where each pair worked on collaborative finding and arranging tasks using those images. The whole experiment took more than 1 hour, including the interviews. Since it was afraid that the participants would be too tired before the 2nd round of tasks, the study is decided to have the practice session only once. The interactive OEB2 was used

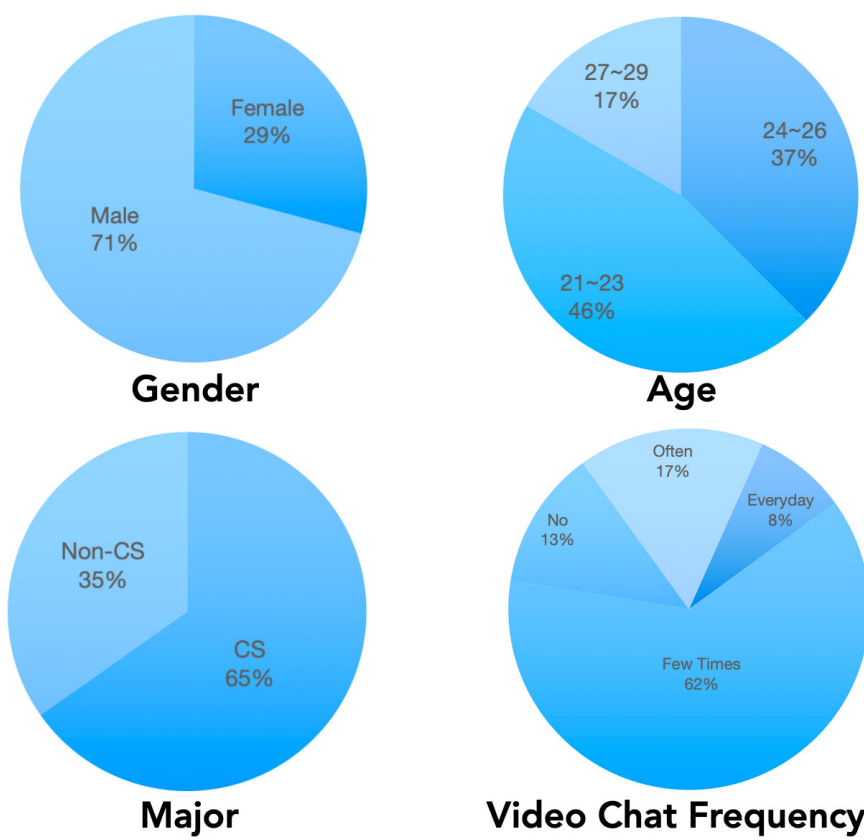


Figure 5.13: The detailed information of the Twenty-four participants.



(a)



(b)

Figure 5.14: The real environment of the two rooms in the experiment.

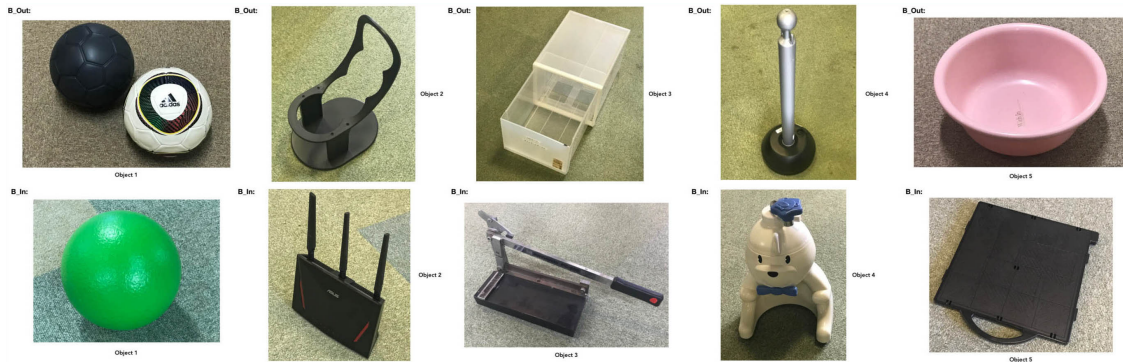


Figure 5.15: One set of images showing 10 objects used in the experiment.

in the practice session so that participants can understand the basic mechanism of OEB2 and get familiar with the new features. However, it explained the participants about the availability of features of OEB2 well before each condition. Participants engaged in this practice session for ten minutes. If the participants finished the task in less than ten minutes, they were asked to practice using the features for the rest of the time.

In the main session, the participants were asked to do the same tasks with different images/objects. However, to make the task slightly difficult, each participant was provided with five object images. Four actual objects were located in the remote room, while one was non-existent. Therefore, two participants had to collaboratively find a total of eight objects out of ten printed objects (Figure 5.15). During the following collaborative arrangement task, the participants were asked to arrange all eight objects in the room as instructed in the object arrangement map.

After the participants finished two tasks in one of the conditions, participants were asked to complete a short post-task questionnaire, which asked about their impressions of the system and the collaboration they have just had. The participants were then asked to proceed to the second round with the other condition and answer to the same post-task questionnaire. Finally, the participants were interviewed about their different experiences of collaboration between the two conditions.

Category	Question
Awareness	Q1. It was easy for me to be aware of the remote partner's movement.
Collaboration	Q2. I could communicate with the remote partner easily.
Guiding	Q3. I could easily indicate areas to the remote partner.
Gaze Awareness	Q4. I could recognize which area the remote partner is observing.
Guiding	Q5. I could easily guide the remote partner's attention to a specific object which is in my own room.
Locating	Q6. I could quickly locate the remote partner's position in the display
Locating	Q7. I could easily locate the area I want to observe in the display.
Perceived Workload	Q8. I needed energy to observe the remote space in the display.

Figure 5.16: Eight questions in the questionnaire designed on a 6-point scale (1 = Strongly Disagree to 6 = Strongly Agree).

### 5.4.6 Measures

#### Objective Measures

It evaluated the efficiency based on task completion time (the time it took for the participants to complete the task). Furthermore, since it was interested in whether the proposed features changed the ways participants accessed the opposite side of the sphere, it analyzed the video recordings of the experiment and measured the time that the participants walked around during the tasks. The study also measured the frequency of interactive features (Physical Rotation or Quick View) used in the experiment.

#### Subjective Measures

After each round, the participants answered a post-task questionnaire, which consisted of eight questions (Figure 5.16). The participants answered the questions in a 6-point

.....

scale (1 = strongly disagree, 3 = possibly disagree, 4 = possibly agree to 6 = strongly agree).

## 5.5 Results

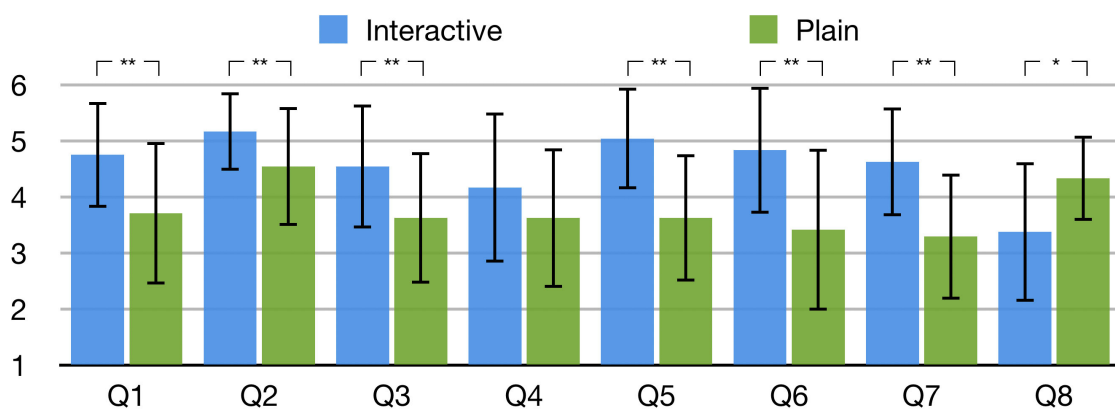
### 5.5.1 Objective Results

In the experiment, the mean completion time was 636s (SD = 121.1) with the plain condition and 638s (SD = 153.1) with the interactive condition. The Wilcoxon signed-rank test showed no significant difference between the two conditions ( $Z = -.08$ ,  $p = .9375$ ). By analyzing the video data, it also calculated the average time that the participants walked around during the task. The mean walking time was 269s (SD = 111.1) in the plain condition and 205s (SD = 84.6) in the interactive condition. The Wilcoxon signed-rank test showed there is a significant difference between the two conditions ( $Z = 3.24$ ,  $p = .0012$ ).

The interactive feature usage analysis showed that out of 24 participants, 18 used both Physical Rotation and Quick View, 5 used only Physical Rotation, and 1 used only Quick View. On average, they used Physical Rotation 7.8 times and Quick View 3.6 times. The Wilcoxon signed-rank test showed a significant difference between the two features ( $Z = 2.30$ ,  $p = .0213$ ), indicating that the participants preferred Physical Rotation more than Quick View.

### 5.5.2 Questionnaire Results

As shown in Figure 5.17, questionnaire responses were analyzed using the Wilcoxon signed-rank test. The mean scores, as well as p-value, are presented in the chart. Results show that the interactive OEB2 was more preferred in seven of eight questions (significant differences were found).



(a)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
<b>Inter</b>	4.8	5.2	4.5	4.2	5.0	4.8	4.6	3.4
<b>Plain</b>	3.7	4.5	3.6	3.6	3.6	3.4	3.3	4.3
<b>p-value</b>	0.0027	0.0037	0.0089	0.0800	0.0007	0.0001	0.0021	0.0115

(b)

Figure 5.17: Mean scores of questions from Q1 to Q8. “\*” and “\*\*” denote significant difference at  $p < 0.05$  level and  $< 0.01$  level. The error bars depict the standard deviation.

### 5.5.3 Interview Results

While the questionnaire results indicated that the participants generally found the added features useful, four participants commented that they preferred the plain OEB2 because the display was not rotatable, which was simpler to understand the fixed directional relationship between the two sites. On the other hand, 20 participants indicated that they preferred using the interactive OEB2 in the task because it is easier and more convenient to use when checking for the remote objects and space.

The study also gathered different opinions for two rotating features. 12 participants indicated that they preferred Physical Rotation since it is intuitive. On the contrary, six participants expressed that they preferred the Quick View more.

## 5.6 Discussion

### 5.6.1 Fundamental Advantages of OEB2

During the finding task, searching for objects mainly relied on local participants. In other words, the participants mostly searched for objects in their own environment and did not search for objects in the remote site very often. However, since OEB2 allowed the participants to view the whole remote space, some instances were observed in which a remote participant found a local object faster than the local participant and tried to explain the location through the system.

One interesting strategy observed was that one participant placed all the objects he found around OEB2 so that the remote participant could check if he got the right objects (Figure 5.18). With conventional video communication systems, users always have to mind the field of camera view and how it is seen by a remote participant [47]. It infers that 360° video and the spherical display shape provided the participants the feeling of sharing everything around the table, which supports the local participant's casual placement of the objects. Furthermore, thanks to the omnidirectional camera, all the objects were successfully captured and shown on the other OEB2 at the remote site. It seems that this

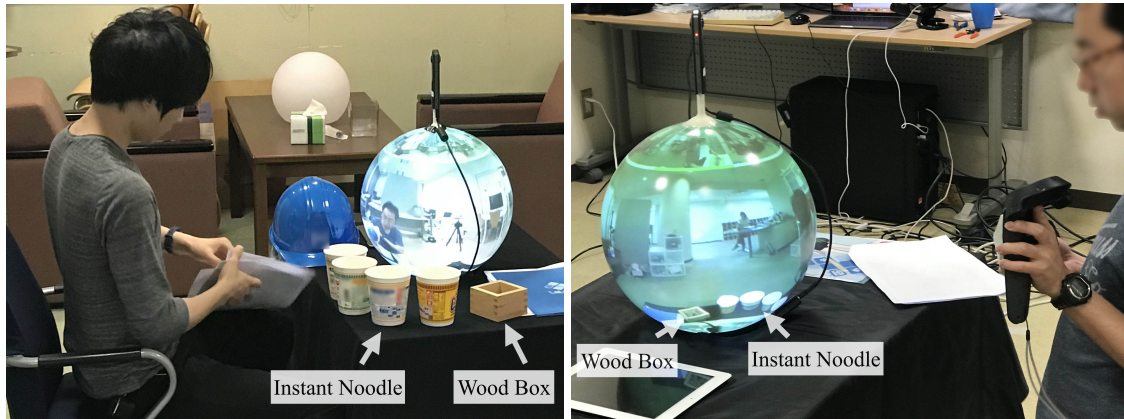


Figure 5.18: The participant placed the objects casually around OEB2 for checking.

instance indicates that OEB2, which supports a 360° full-spherical view, affords casual visual interaction that utilizes the whole 360° environment.

### 5.6.2 Comparison between Two Conditions

The results of the task completion time did not show a significant difference between the two conditions. However, subjective measures and other qualitative data showed that the participants had different feelings about the two conditions. The following sections present several interesting scenes extracted from the video data and discuss the potential reasons for these results.

### 5.6.3 Perceived Workload

Based on the objective results, it is found that most of the participants used the new features, especially Physical Rotation, and Quick View, when searching for the objects in the remote space or when checking if the objects were arranged in the correct places (Figure 5.19 (right)). However, in the plain condition, they had to walk around the OEB2 (Figure 5.19 (left)). These scenes support the result of Q8 in the questionnaire as well as the significant difference in the average walking time between the two conditions. 17 participants clearly expressed that the interactive features (Physical Rotation or Quick



Figure 5.19: L) Participants had to walk around in the plain condition. R) Participants used the proposed feature to replace walking.

View) reduced their physical workload, e.g., “*The interactive OEB2 was easy to use when checking the found objects or the correctness of the arranging places because I do not have to always walk around.*” It seems that the proposed features (Physical Rotation or Quick View) are easy and convenient to use when checking for the remote objects and space.

#### 5.6.4 Affordance

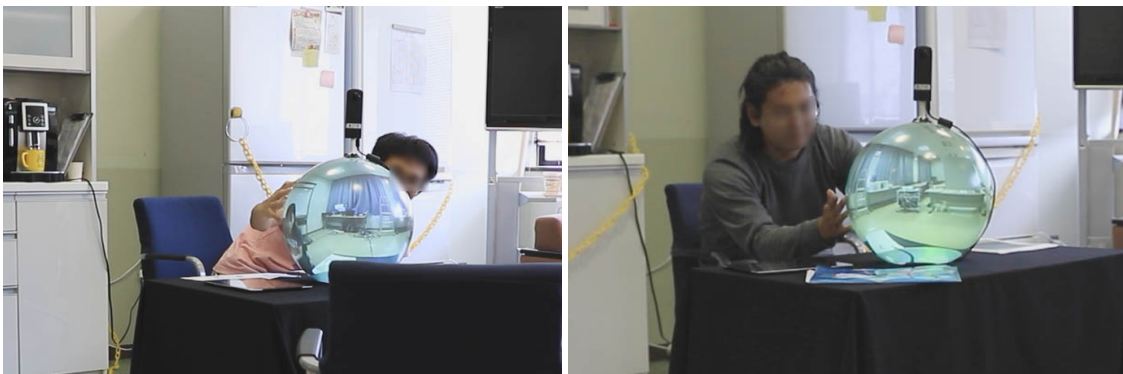


Figure 5.20: The participants unconsciously tried to use the disabled Physical Rotation in the plain condition.

.....

The interview data showed that most of the participants thought the Physical Rotation as an intuitive design. As shown in Figure 5.20, although the scenes happened in the plain condition, which means that the participants were not allowed to use the Physical Rotation. It is observed that two participants unconsciously tried to use the disabled Physical Rotation to check the correctness of the arranged object. Although they were reminded not to use the feature immediately, these findings seem to prove that the Physical Rotation feature is quite intuitive and it may afford the user to use it to check the other sides of OEB2. According to the observation results, most of the participants used Physical Rotation and were well adapted to it in the interactive condition.

### 5.6.5 Effect of OmniBar on Remote Space Awareness

Some instances were observed that the participants used the OmniBar effectively. Usually, the participants observed the main area of OEB2 to observe the remote participants, remote objects, and their positions because the image quality is much better there compare to the OmniBar area. For example, in Figure 5.21a, the participant was watching the main area to observe that the remote partner took the object as he asked. Then, as he told the partner to place the object on the right side of a sofa, he looked down at the arranging guide sheet (Figure 5.21b). While he was reading the guide, the partner moved to the sofa and disappeared from his sight. When he looked up the main area again, he found that the partner had disappeared, so he started to rotate the OEB2 physically. However, he soon was aware that he could see the partner in the OmniBar area, stopped turning the OEB2, and made sure that the object was placed correctly (Figure 5.21c). After this scene, the participant looked at the main area again and started to ask the partner to find the next object. As this example shows, OmniBar was effective in quickly figure out the latest status of the remote site, especially when the local participant could not find his/her partner or objects in the main area.

The findings above are also consistent with the results of Q1. Furthermore, nine participants indicated that it helped them to observe and understand the remote space. As

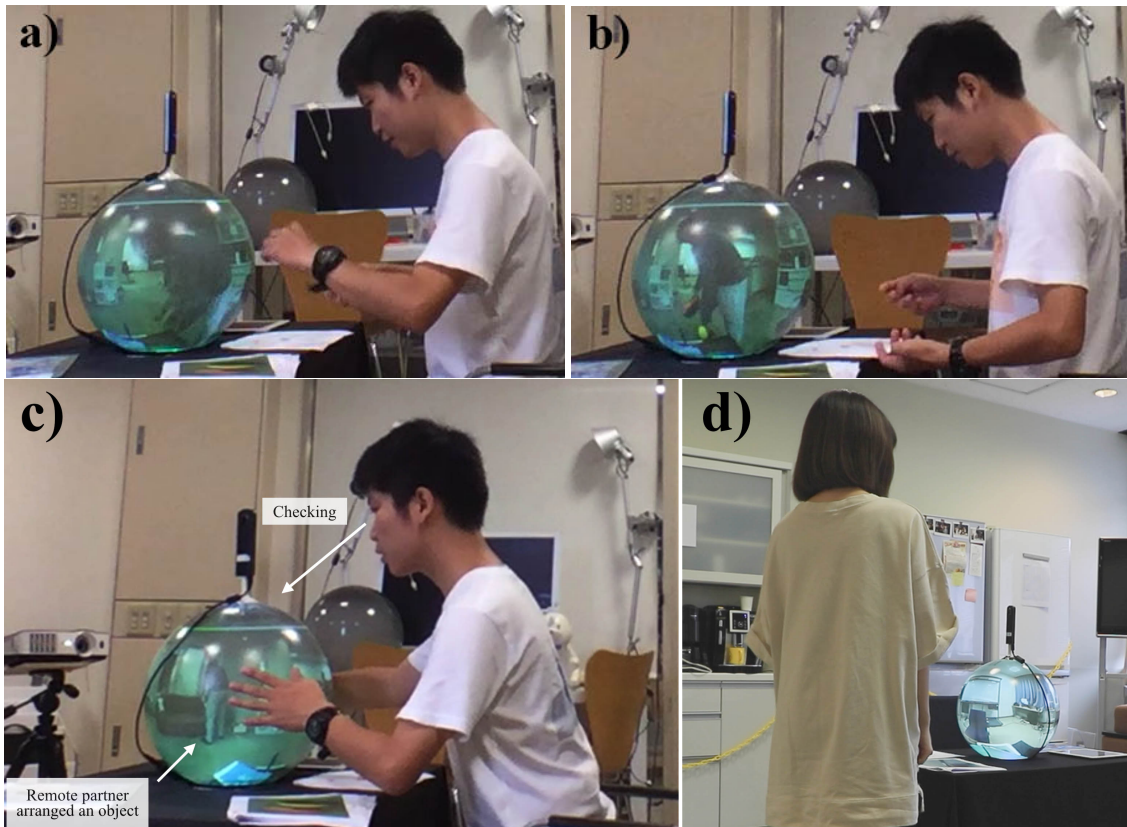


Figure 5.21: a) b) c) An example that the participant observed the OmniBar area effectively. d) The participant lost her remote partner.

an example, one participant commented that “*he usually checked the OmniBar before movement. Through the OmniBar, he can quickly know which region the remote partner was searching, and then he could help to search on the other regions on the OEB2.*”

In the case of the plain OEB2 condition, a pair of participants tend to converge to a certain position to stand when talking with each other. For instance, although they need to move around the room to find objects and arrange them in the room, they always came back to a certain position when they needed to have conversations. However, when the participant walked away from the OEB2 and came back to talk with his/her partner, there were some cases that a participant could not find his/her partner in the OEB2 (Figure 5.21d). This kind of experience might be the reason for the result of Q6.



Figure 5.22: The limitation about the showing direction of the Quick View sharing window.

## 5.6.6 Supporting Gaze Awareness

### Visibility of Quick View Sharing Window

The Quick View sharing window can be used to share the user's focused view with the remote partner. However, it is found that this feature did not always work as expected. For example, as shown in Figure 5.22, the remote participant in OEB2 was trying to use the Quick View to indicate the location to arrange an object. However, since the Quick View sharing window appeared above the remote participant's head, which was shown on the opposite side of the local participant, the local participant could not see it. This incident happened because the remote participant did not notice that his image's facing direction was not facing towards the local participant.

### Pseudo Gaze Leading Effect

Even though the Quick View sharing window was not used very often, the results of the Q3 and Q5 indicate that, for leading the remote participant's attention to a region, the interactive OEB2 was better than the plain OEB2. From the video observation, it seemed

.....

that the ability to share the gaze awareness over distance relied more on the Physical Rotation. In fact, Ten participants expressed that they thought the Physical Rotation contributes to indicating a region, e.g., “*Not only could I indicate a remote region through my facing direction, but also I could guide my remote partner’s attention to a local region by rotating the OEB2.*” This effect is named as “pseudo gaze leading effect” and found that there are two kinds, remote gaze leading effect, and local gaze leading effect.

The remote gaze leading effect was observed when a local participant rotated OEB2 to find specific objects or regions in the remote site. In this instance, since the local participant’s face displayed in the remote site also oriented toward the direction that the local participant was seeing, the remote participant could figure out the gaze direction of the local participant, and then the remote participant naturally followed the direction.

Meanwhile, around half of the participants said that they could guide the remote partner’s attention to a specific local region simply by physically rotating the OEB2 to make the remote partner’s face orient toward that region. Meanwhile, this action brought the target region right in front of the remote partner, so the remote partner could be easily led to see that region. This effect is named as local gaze leading effect. This effect was surprising because, although the researcher did not anticipate this strategy, participants could naturally use it and soon get used to it.

### 5.6.7 Design Implications

Based on the results, the takeaways are summarized in twofold:

1. The analysis of feature frequency indicates that the participants preferred Physical Rotation more than Quick View. In fact, 12 participants said that they preferred Physical Rotation, e.g., “*I think Physical Rotation is quite intuitive. It made me feel like I was moving around in the remote room.*” These points may indicate that the spherical display affords Physical Rotation rather than touch interaction. This observation suggests that designers of spherical display interface should consider tangible interaction methods such as Physical Rotation.

2. Many studies (e.g.[83]) have been trying to utilize remote gaze awareness functions. Pseudo gaze leading effect indicates that a video communication system should provide users an intuitive way to lead a remote partner's gaze to the intended orientation and present the user's gaze orientation to the remote partner. Utilizing symmetric spherical displays with physical rotation is one of the effective solutions. Even without using the spherical display, it suggests designers of 360° video communication system implement both remote and local gaze leading effects in their systems.

### 5.6.8 Possible Application Scenarios

The task was designed to reflect the daily activities, such as finding objects or giving suggestions for arranging objects. Seven participants commented that *“they would like to use the 360° video communication system with their families or lover when they are living in different places.”* Among them, three participants mentioned that *“it is good to use this system to help the child or aged person who is living alone in the house.”* Also, five participants commented that *“the system supports the remote collaboration of spatial decoration. Through the system, they could easily understand the remote space and give suggestions for arranging the room.”* Furthermore, previous studies [50] [59] demonstrated a new remote meeting environment that merges distributed rooms and allows participants to move around. It is believed OEB2 can be used for such a scenario. Typically, I envision its contribution to a more complex remote collaboration scenario, i.e., distributed war room collaboration [52] where various physical objects are scattered in the environment. Since OEB2 is quite flexible, it can fit in a varied environment of the collaborative war room.

## 5.7 Limitations

The proposed system and user study also have some limitations.

### 5.7.1 Limitations of the User Study

The user study has two limitations. First, it recruited only 12 pairs of participants. It needs more participants to get more solid results. Second, it used only interactive OEB2 for the practice session, which may slightly affect the participants' perceptions of the two conditions in the experiment.

### 5.7.2 Insufficient Resolution

Ten participants reported that the resolution of the OmniBar is low. As an interactive interface and also supporting a 360° overview, a higher resolution may enhance its effect.

### 5.7.3 Conflict between Physical Rotation and Quick View Use

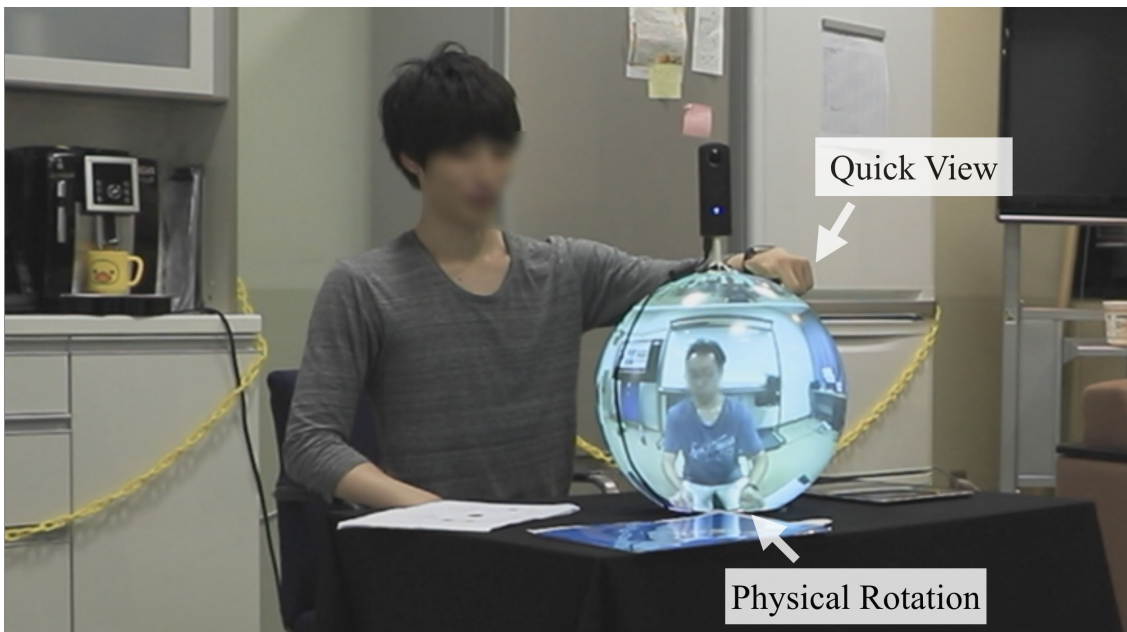


Figure 5.23: The Physical Rotation disturbs the use of the Quick View.

During the experiment, a few cases were observed that a remote participant physically rotate OEB2 while a local participant was touching OmniBar. For example, in Figure 5.23, the local participant was using the Quick View by touching the OmniBar. However,

.....

the remote participant was also physically rotating the remote OEB2 at the same time. Because Physical Rotation on the remote site rotated the images of both the main area and the OmniBar area of the local OEB2, the local participant had difficulty in using the Quick View feature since the OmniBar was rotating. This indicates that it is needed to find a way to avoid the conflict between Physical Rotation and Quick View use.

## 5.8 Improvement Suggestions

This section proposes some future works to improve the limitations of OEB2.

### 5.8.1 OmniEyeball4K

It seemed that the low image resolution diminished the effect of OmniBar as well as the whole system. The visual quality of the current system is mainly limited by the resolution of the Glomal350 spherical display. Because the camera already can provide HD image, my colleagues and I are considering to replace the spherical display with those that have higher 4K resolution to create the next generation of the OmniEyeball.

### 5.8.2 OmniBar

As described as one of the system's limitations, there is a possible conflict when a user uses Physical Rotation on one side, and another user uses Quick View on the other side simultaneously. Although it is seen only a few cases during the formal experiment, it is anticipated that there will be more chances of this conflict if multiple users in each site conduct actual tasks in a more casual manner. Therefore, to avoid this kind of conflict, I plan to make the OmniBar area not rotate no matter the local or the remote user physically rotates the OEB2. For this purpose, it can incorporate a rotary encoder to measure the physical rotating angle of OEB2.

### 5.8.3 Quick View

There are two possible areas regarding the improvement of Quick View. The first area is for a local site (a participant who points at an object or a spatial region). Although the remote referencing feature was used, several participants commented that it is hard to find the object they wanted to reference in the OmniBar because the OmniBar area was small, and its image resolution was low. Meanwhile, the study recorded a few instances that the participants directly pointed at the object in the main area of the OEB2 when he/she wanted to refer to the object. As Jones et al. revealed [39], participants in video communication cannot help such gestures even though the gestures cannot be seen by the remote participants. Therefore, it seems that it is better to allow users to use the Remote Referencing feature by touching not only the OmniBar area but also the main area. For this purpose, it is better to implement the full-sphere touch detection capability using the technologies developed by Benko et al. [4].

The second area for improvement is for the remote site (a participant who should see/manipulate the object specified by the local participant). As described in the “Visibility of Quick View Sharing Window” section, the visibility of the Quick View sharing window should be improved. In the current implementation, the referencing window always popped up above the user’s head. Instead, by detecting the user’s remote partner’s face position, I am planning to make the window pop up toward the direction of the remote partner.

### 5.8.4 360° Audio Support

The sound is important for video communication. Considering the shape of OEB, it appears that circular microphone and speaker arrays [37] may be suitable. For intuitive audio interaction, the directivity of 360° sound has to match the 360° video. For instances, voice from the opposite side of a user’s view point should be attenuated. We are also interested in deeply combining the 360° sound with video for new interactions.

---

## 5.9 Summary

This chapter introduced a symmetric full 360° video communication system, OEB2, which has four novel features to support remote observing and increased shared awareness. Based on the evaluation of the comparative experiments, the four features were effective in reducing the inconvenience mainly caused by the partial invisibility of the spherical display. The results also indicate how the proposed features supported the remote space awareness and gazed leading intuitively and effectively.

One shortage of the study is that the experiment is not designed for multiple users (more than two people on each side). As a future plan, it is meaningful to extend the current user interface to allow multiple users to participate at each site and conduct a qualitative user study by asking two groups of participants to accomplish quasi-naturalistic tasks [34] using the extended OEB2 system. Such studies are expected to give the 360° research community a more in-depth understanding of how the 360° video communication can support remote multi-party collaboration.

---

## Chapter 6

# OmniEyeballVR: A Collaborative 360° Video Communication Application for VR

### 6.1 Background and Motivation

In recent years, VR technology has continued to develop rapidly. In addition to popular immersive gaming experiences, VR also has possibilities in many other application areas such as design, education, and content creation, where sharing VR experiences with others is desirable. However, a highly immersive VR experience usually requires wearing a Head-Mounted Display (HMD) which blocks the view of the surrounding real world, making it difficult for the VR user to communicate with the people around them outside the VR environment.

Researchers have worked on systems that allow VR users (occupants) to communicate and collaborate with others in the physical world (designers) [6] [33]. For example, with CollaVR [57] and On the Shoulder of the Giant [65], the collaborators wear an HMD in order to share the scene and collaborate in the VR space. However, those systems had difficulties in allowing the designers to access physical objects or content in their real environment. For example, designers might need to view plans or make notes in the physical space while cooperating with the VR occupant.

Ibayashi et al. [36] and An et al. [1] demonstrated the usefulness of using an asym-



Figure 6.1: The DollhouseVR system [36] proposed by Ibayashi et al..



Figure 6.2: The new design workflow [1] proposed by An et al..

metric collaboration system for VR and non-VR users working together on architectural or car interior design tasks. In these examples, the occupant in the VR space can share the virtual world and cooperate with other designers in the physical world through a 2D display which shows the VR scene. In FrontFace [12] and FaceDisplay [31], touch screens were added to the VR occupant's HMD to facilitate casual collaboration between designers and the occupant through touch interaction. However, none of those systems could provide equivalent visual information to the designers in the physical world. For example, the designer could not observe the VR space from a different perspective, as the shared view is restricted to the field of view of the occupant or a static viewing perspective.

Compared to this prior work, this work propose to use an interactive spherical display that can introduce new possibilities for collaborative design. Spherical displays have previously been used by researchers as a cooperation platform [2] [18] [19] [20] [32] [89], or communication terminal for sharing the VR or 360° content between collaborators in the same physical environment. Compared to some of the other similar platforms (e.g., Dome Display [3]), these systems can be better integrated into a physical environment, such as an office or design lab. This chapter presents OmniEyeballVR, which is a novel tool for collaboration between VR occupants and designers in the physical environment that uses the OmniEyeball2 (OEB2 presented in last chapter), as a terminal in the physical world (see Figure 1), therefore the OmniEyeball in this chapter stands for OEB2 (chapter5) instead of OEB (chapter4). A user study was conducted to evaluate the system and report significant results from comparing three design alternatives of the OmniEyeballVR. The main contributions and novelty of the presented work include:

### 6.1.1 Contribution

The main contributions of the OmniEyeballVR include:

1. A novel collaboration system that allows a VR occupant to cooperate with designers in the physical environment while also allowing the designers to share their

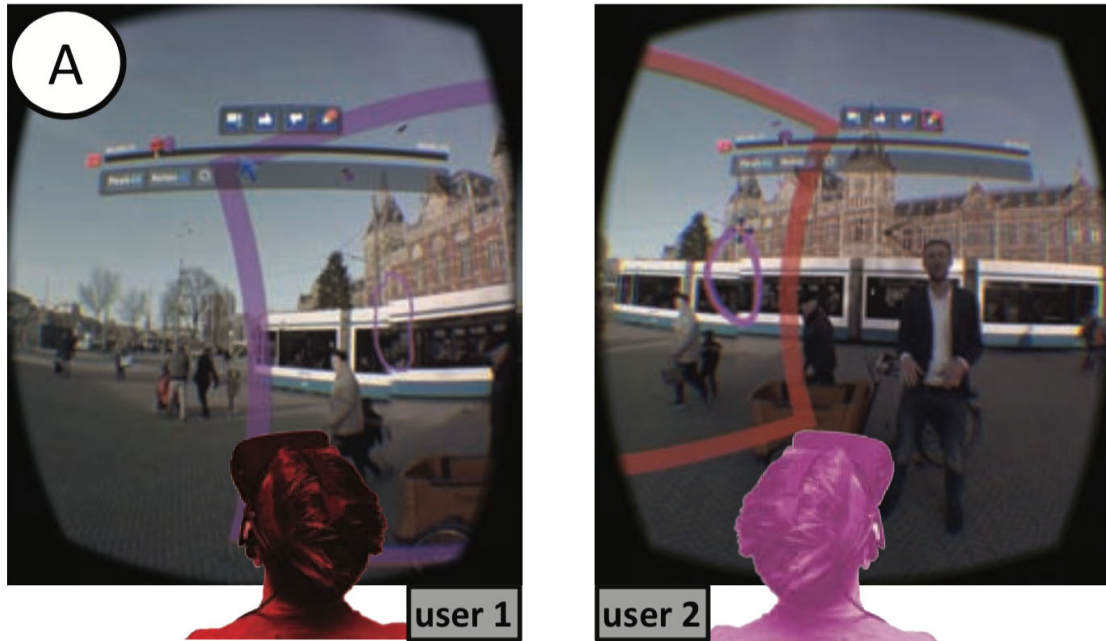


Figure 6.3: The CollaVR system [57] proposed by Nguyen et al..

facial expressions and gestures with the occupant. The occupant and designers can share their gaze and visual cues with each other to enhance the communication between the cooperators.

2. It allows designers to easily access the content in the VR space by actively controlling the view perspective to watch it from any direction without wearing a VR HMD. Designers can also access both the VR space and the physical world simultaneously.
3. The first user study that explores the benefits and implications of integrating a 360° spherical I/O device into a Mixed Reality (MR) collaboration system.

---

## 6.2 Related Work

### 6.2.1 Collaborative VR/MR Systems

VR/MR collaboration systems that allow VR users (occupants) to communicate and cooperate with others in the physical world (designers) are getting more and more popular. A local non-VR user can share his/her surroundings with a VR user to collaborate together. This allows the VR user to provide non-verbal communication cues such as view awareness [45], 3D virtual avatar [64], or ray pointer [76] to a local non-VR user. For instance, CollaVR [57] and On the Shoulder of the Giant [65] enable users to wear HMDs to share the scene and collaborate in the MR space. As a result, the systems can enhance the collaborative experience and task performance between VR users and non-VR users [63]. However, these systems could only be useful in particular design scenarios, such as a co-review of 360° videos, as they require the VR user to wear an HMD in order to see the shared information.

Another kind of system integrated a 2D display on the surfaces of the HMD. For example, FrontFace [12] and FaceDisplay [31] allowed casual collaboration between designers and the occupant by adding touch screens to the VR occupant's HMD.

Systems have also been designed to support asymmetric collaboration between VR and non-VR users. For instance, DollhouseVR [36] provided the designers with a top-down static viewing perspective of the VR environment through a horizontally placed 2D display. The work proposed by An et al. [1] focused more on the haptic feedback during the VR cooperation design. All non-VR users could view the VR world through the field of view of the VR occupant. However, none of the display technologies in those systems can provide symmetric visual information to the designers, so they will not be able to observe the VR space from a different perspective.

## 6.2.2 Communication Systems with Spherical Display

With advances in technology, sharing a 360° scene has become popular due to the amount of content that can be shared at once. This allows a user to be immersed in a remote environment by connecting with a 360 camera [44] [53] [74]. Besides HMDs or 2D monitors, several other display methods have been explored for enabling users to observe VR or 360° content, e.g., panoramic dome [3] or CAVE [29] displays. However, dome and CAVE displays confine users inside their spaces, which support low flexibility and prevent users from interacting with the real environment.

In summary, there has been little previous work on using spherical displays for remote collaboration. Therefore, one of the innovations of our researches is that it is the first that connects a VR user with observers gathered around a spherical display. Furthermore, by utilizing the spherical display technology, it is interesting in if the proposed system can support multiple designer users and provide symmetric information (VR scene and physical world) while allowing them to stay in the physical environment. In the next section, more detail will be introduced about the prototype system that have been developed.

## 6.3 System Overview of OmniEyeballVR

### 6.3.1 Hardware

The OmniEyeballVR (OEBVR) system consists of two parts: the occupant side that uses a VR HMD (HTC Vive) and the designer side that has an OmniEyeball2 terminal (OEB2, chapter 5).

Therefore, the term “OmniEyeball” in this chapter represents the OEB2 terminal, which is made from a spherical display (Glomal350 [27]) with a 360° camera (Ricoh Theta V [67]) mounted on the top (see Figure 6.4). The Ricoh Theta V is an omnidirectional camera equipped with two fisheye lenses, which provides an equirectangular panoramic live streaming video with  $1920 \times 960$  pixels over a USB connection. The Glomal350 spherical display uses a wide-angle lens to project a round image from its

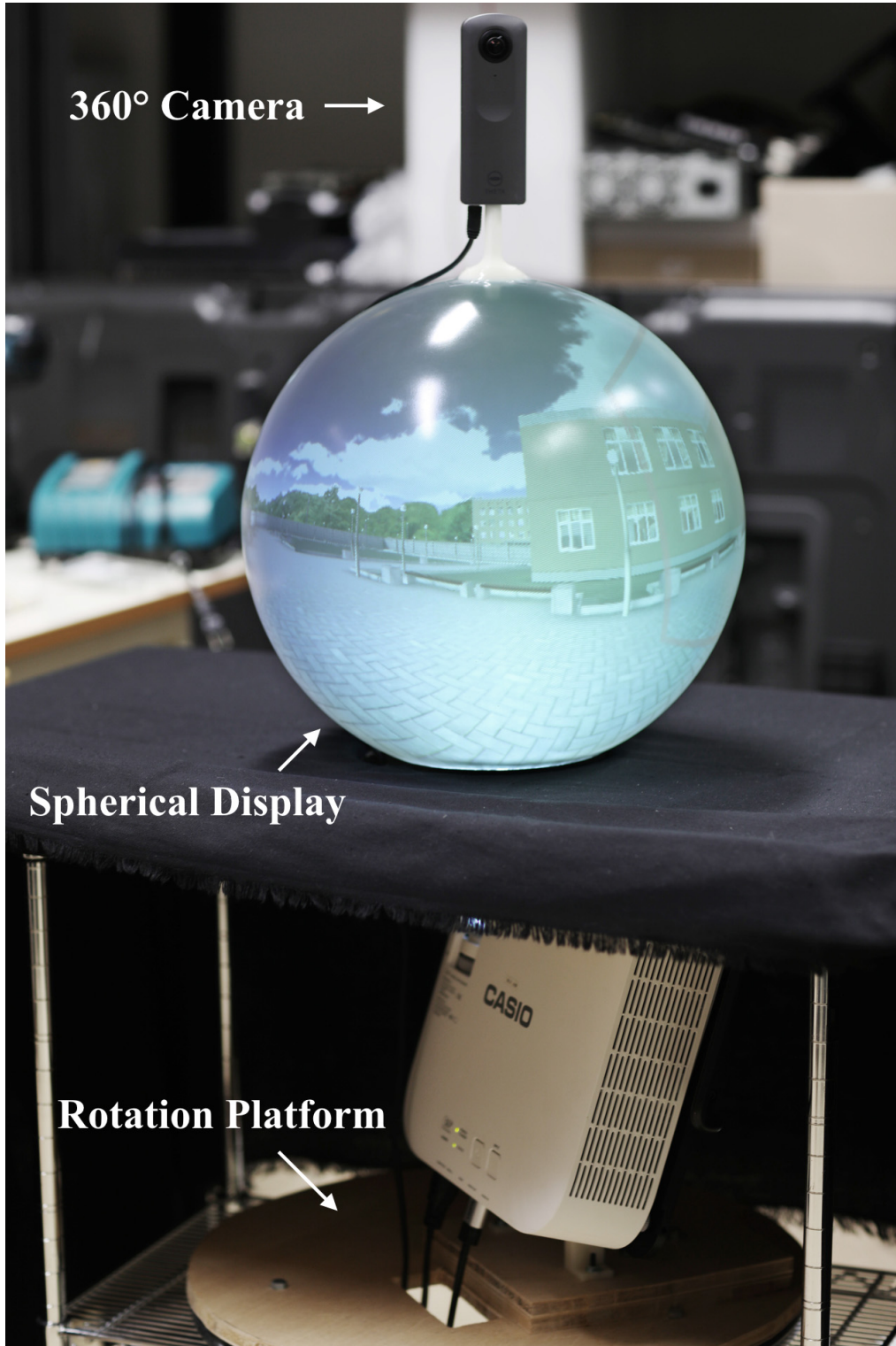


Figure 6.4: The system hardware overview of OEB2 terminal.

.....

Casio projector onto a spherical surface ( $\varnothing 350\text{mm}$ ). The bottom of the spherical surface is reserved for the connecting parts, making the displayable portion of the sphere at around  $300^\circ$  vertically and  $360^\circ$  horizontally. Although the projector can output images with  $1024 \times 768$  pixels, the actual visible part of the projected image is limited to a circle with a diameter of 768 pixels. In addition, a rotating platform is used as the base for the spherical display to allow designers to check other sides of the OmniEyeball by spinning the spherical surface when the space is difficult for the designers to move around physically.

### 6.3.2 Software

The VR HMD and the OmniEyeball are connected through a software developed in the Unity 3D (2018.4.14f1) game engine. To naturally show the  $360^\circ$  image on the spherical display, a graphic transforming shader is implemented to map the pixels from a skybox format to an azimuthal equidistant projection format. Since the visual information of the input image is a  $360^\circ$  content, the remapped image can be naturally projected onto the spherical display. Due to the display specification, the OmniEyeballVR system streams a live video with  $1024 \times 768$  pixel resolution to the OmniEyeball side and with  $1080 \times 1200$  pixel resolution (standard resolution of HTC Vive) to the VR HMD side. The OmniEyeballVR requires the position of each designer user relative to the OmniEyeball. To achieve that, a QR marker is attached to the chest of each designer and tracked by the system to determine the position of the designer. The frame rate of the whole system can be maintained at 20 fps on average. Also, the rotation interaction has a latency of 0.2s.

## 6.4 Features

The OmniEyeballVR system has two view modes: a  $360^\circ$  first-person mode and a third-person mode. In general, the  $360^\circ$  first-person mode allows the designer to see the VR space of the occupant from a panoramic view at a standard eye-level, whereas the third-person mode allows the designer to see it from a higher point of view. Two view

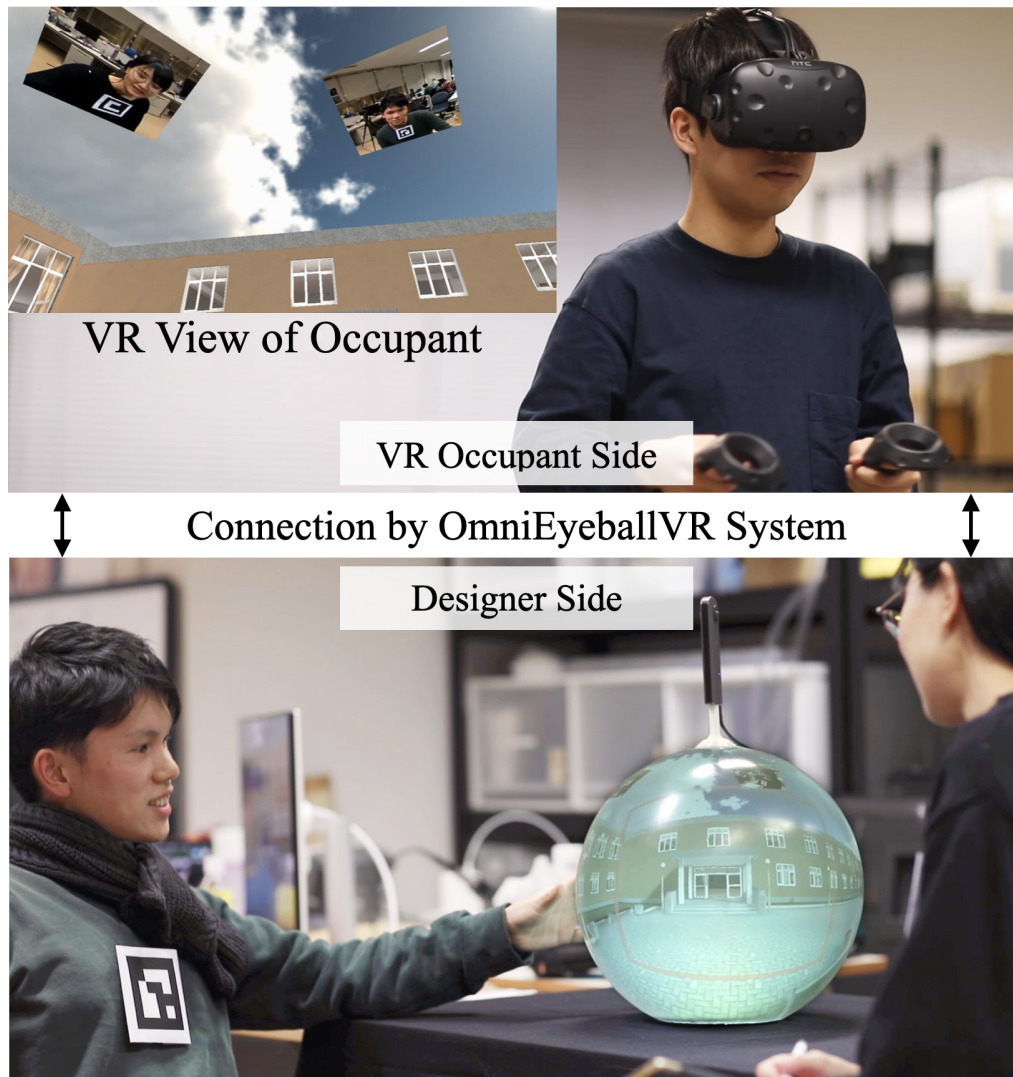


Figure 6.5: The OmniEyeballVR enables VR users (occupant) to communicate and cooperate with others on the physical world (designer).



Figure 6.6: Viewing VR Space in the 360° first-person mode from different viewpoints.

modes allow the designers to observe the VR space from free viewpoints and directions. On top of viewing the VR space, both view modes allow designers to share their gaze and face with the occupant and also be aware of the occupant's visual cue. The remaining sections will describe each of these features in detail.

### 6.4.1 360° First-person Mode

#### Viewing VR Space on the OmniEyeball

In this view mode, the 360° surroundings of the occupant are shared with the designers. This is achieved by configuring a virtual 360° camera at the head position of the occupant and streaming the live 360° video onto the OmniEyeball's spherical display (Figure 6.6). This allows the designers to understand the close surroundings of the occupant by seeing the VR space through the 360° first-person view of the occupant. Furthermore, the designers can also change the viewpoint of the VR space by either physically rotating the OmniEyeball in any direction they prefer, or walking around it.

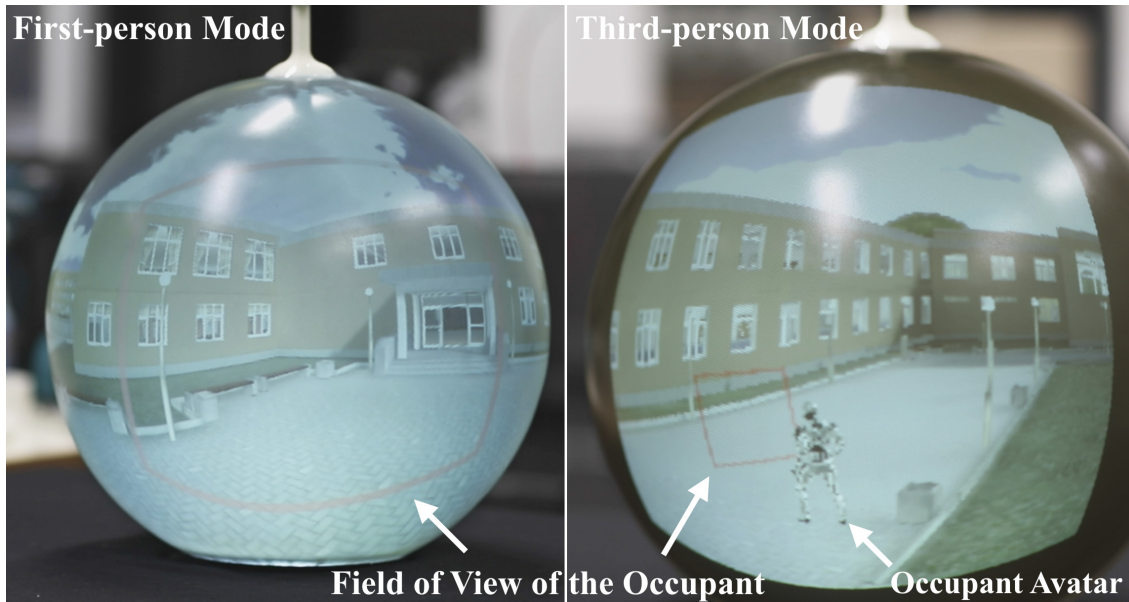


Figure 6.7: In both first-person and third-person mode, the red translucent frame represents the field of view of the occupant. The occupant is shown as a virtual avatar in the third-person mode.

### Shared Gaze on the OmniEyeball

Inspired by CollaVR [57], the system renders a red translucent frame that represents the field of view of the occupant on the OmniEyeball (see Figure 6.7). This allows the designers to quickly understand the current point-of-interest of the occupant, thus improving the collaborative performance between the users [45]. Furthermore, the red translucent frame updates synchronously as the occupant changes his/her view direction.

## 6.4.2 Third-person Mode

### Viewing VR Space on the OmniEyeball

In this view mode, the content in the VR space is shared from a third-person perspective of the designer, similar to a fish-tank VR display [18]. This allows the designers to better understand the virtual position of the occupant for tasks requiring spatial un-



Figure 6.8: In the third-person mode, the designers can rotate or walks around the OmniEyeball in order to observe the VR space at a different direction.

derstanding. To achieve this, a 3D avatar is used to represent the virtual position of the occupant. Then, a virtual camera with a  $90^\circ$  field-of-view (FOV) is attached away from the 3D avatar in the VR space. This captures a live video of the surroundings, including the occupant from a third-person perspective and streams as an independent sector window on the OmniEyeball facing towards each individual designer (Figure 6.7 right). To promote the spatial relationship of the independent sector window, the OmniEyeballVR measures the direction of the window and the virtual camera according to the direction of the designer relative to the OmniEyeball, so that the third-person view image can be shown in the correct direction. Meanwhile, a graphic shader is implemented to transform the rectangular format image captured by the virtual camera to a fan-shaped format, in order to show on the OmniEyeball without significant distortion. Similar to the  $360^\circ$  first-person mode, the designers can also rotate or walk around the OmniEyeball in order to observe the VR space in a different direction (Figure 6.8). Through the spherical display, multiple designers can observe the VR space from their preferred perspective without interrupting each other.

### Shared Gaze on the OmniEyeball

The occupant is displayed at the center of the OmniEyeball as a virtual avatar (low poly character) in the third-person mode. To maintain consistency, a similar red translucent

frame that refers to the view direction of the occupant in the VR space is used (Figure 6.7). The red translucent frame will rotate accordingly on the OmniEyeball to show the designers where the occupant is looking at when the occupant changes the view direction.

### 6.4.3 Face Window in the VR Space

The red translucent frame indicates the view direction of the occupant to the designer, but it does not support the opposite way, so the occupant will not be able to see or understand the view direction of the designer. To overcome this, the system integrates face windows in the VR space, which are image patches that show the occupant the viewpoints of the designers around the OmniEyeball (Figure 6.9).



Figure 6.9: The face windows of designers shown in the VR space.

A QR marker is attached to the chest of each designer to allow the system to detect the position of designers. Then, the 360° camera on the OmniEyeball tracks each individual designer, crops the 360 camera image to their faces with upper bodies and renders this as an image patch in the VR space. The position of the image patch is determined by the position of the designer relative to the OmniEyeball (Figure 6.11 left). For example, when a designer is looking at the OmniEyeball, the face window of the designer will be displayed in the corresponding direction in the VR space.

Although the direction is decided, the precise position of the face window is not unique. During the design process, there were three optional positions for the face window. They are represented in the Figure 6.10 respectively. For example, suppose that when a de-

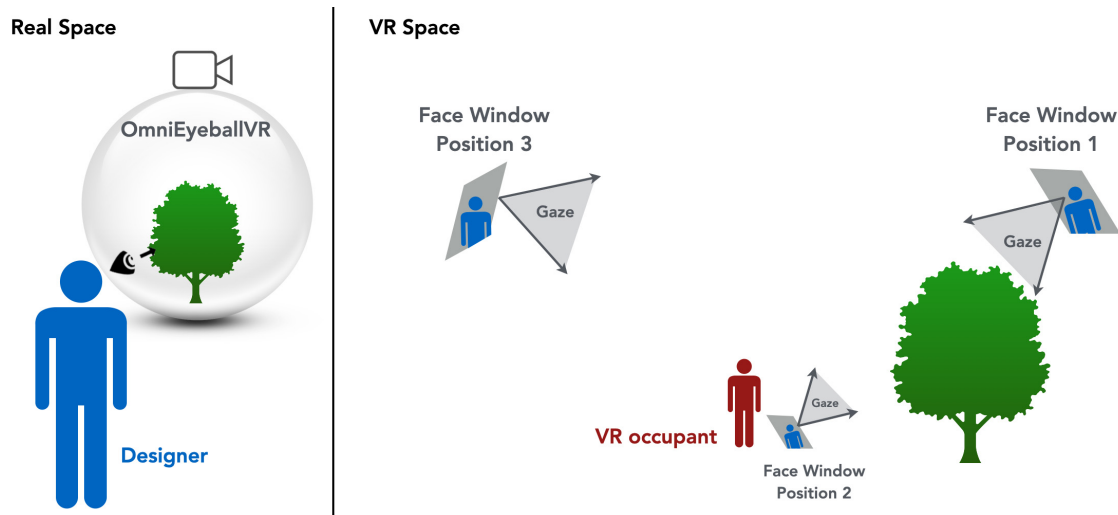


Figure 6.10: Face Window optional positions during the design process.

signer looks at a tree in the VR world displayed on OmniEyeball (Figure 6.10 left), FaceWindow's relative position in the VR world can have the following three designs (Figure 6.10 right).

1. Position One of the face window is designed to be positioned backward from the top of the observed object. The altitude and rotation of the face window is optimized so that the designer's gaze in the window appears to be looking down at the object.
2. Position Two is designed close to the VR occupant. In order not to block the gaze of the occupant, the face window is positioned downward in front of the occupant.
3. Position Three is designed at the position where is the opposite side of the observed object to the VR occupant.

A pilot test was conducted to compare the three optional positions and decide the final configuration of the face window. The pilot test results showed that Position Three is not convenient since the VR occupant always has to look back to check which direction the designer is looking at. As for Position Two, some reviews said that it's not very



Figure 6.11: Left: Based on the position of the face window in the VR space, the occupant can know which direction the designer is looking at. Middle: the face window is also displayed on the OmniEyeball as visual feedback to the designer. Right: Through the face window, designers can use body gestures to express information, such as pointing to a direction by hand.

convenient to look down often to see the face window. Meanwhile, the image of the designer in the face window is hard to be seen.

In the end, Position Three was chosen as the location configuration for the face window. The altitude and rotation of the face window are carefully set as if the face inside is looking downwards in the given direction of the VR space from above. In addition to its use as a view indicator, the face window also enables face-to-face communication between the occupant and the designers. One benefit is that designers can use body gestures to express information, such as pointing to a direction by hand (Figure 6.11 right). Meanwhile, the face window is also displayed on the OmniEyeball as visual feedback so the designers can know how they appear in the VR space (Figure 6.11 middle).

## 6.5 Experiment

Based on the OmniEyeballVR system, an experiment was conducted to explore how people are likely to use the OmniEyeball terminal and its different features in a collaboration task. The study planned to answer the following questions:

1. What is the difference between using OmniEyeball in the first-person mode, and the third-person mode in terms of user interaction and performance?
2. How could the face window feature be used to support collaboration between the two sides?

### 6.5.1 Method

Based on the characteristics of OmniEyeballVR, this study designed a collaborative VR task for two people. In this task, the designers receive a document with additional information, and they need to guide the occupants to complete the task according to the instructions on the document.

### 6.5.2 Participants

Twelve participants (in six pairs) were recruited for this experiment, including ten males and two females. The average age was 25 years old ( $SD=2.6227$ ). Among them, two participants had never used a VR device before, seven participants had used VR devices a few times per year, while the rest used VR more frequently.

### 6.5.3 Task

The task required the two participants to collaborate, one being on the occupant side and the other on the designer side. In the VR scene, 24 objects were placed around the occupant, divided into two rows, surrounding the occupant in a circle with a 30-degree interval between each column of objects (see Figure 6.12a). Among the 24 objects, there

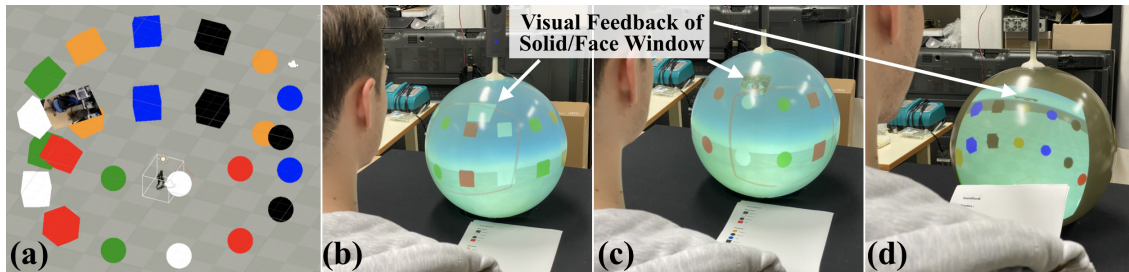


Figure 6.12: (a) The experiment VR scene for the task. The image patch represents the position of face window. (b) The first-person mode with the solid window (1PSW) condition. (c) The first-person mode with the face window (1PFW) condition. (d) The third-person mode with the face window (3PFW) condition.

were 12 cubes and 12 spheres in 6 different colors. The shape and color of the objects were the same in each column. The designer was given an instruction note with description information (such as “*red sphere top*”, or “*blue cube bottom*”) that described a target object in the VR space. The designer needed to indicate the listed objects to the VR occupant one by one. In the actual use case, many design objects are abstract and cannot be described by verbal cues, so the designer was restricted not to use verbal cues to prompt color, shape (sphere or cube), or elevation (top and bottom) to convey information. Other than that, no any other restrictions was placed on this task. After the occupant has identified the target object, they call out the target object to confirm with the designer (by describing the color, shape, and elevation). At this time, the designer could confirm whether it is correct. If it is correct, the designer moved to the next target object instruction. Otherwise, the instruction was given again on the current target object.

#### 6.5.4 Conditions

The study compared performance with the OmniEyeballVR system between the three configurations. In the first condition, the system is configured in the 360° first-person mode with the face window (Figure 6.12c). In the second condition, the system is also configured in the 360° first-person mode, but with a white solid window. The solid win-

.....

dow, like the face window, adjusts its position in the VR space according to the relative position between the designer and the OmniEyeball. The only difference is that there is no live video of the designer side shown in the solid window (Figure 6.12a).

In the 3rd condition, the OmniEyeballVR system is configured in the third-person mode with the face window (Figure 6.12d). The face window is similar to the one in the first condition, which means that it looks almost the same in the VR occupant's view. However, the visual feedback of face windows is shown differently due to the viewpoint of the third-person mode.

### 6.5.5 Procedure

The experiment first began with a pair of participants who went through a brief introduction to the OmniEyeballVR system and the details of the experiment task. Then, the participants were asked to fill out a pre-experiment questionnaire for collecting demographic information.

In the task, one of the two participants was randomly assigned to become a designer and the other as a VR occupant. The designer participant was asked to sit near the OmniEyeball and was given an instruction note which showed the task targets. The VR occupant was asked to wear the VR HMD and sit at a distance of 1m away from the designer so that they could verbally communicate with each other.

The study was a within-group design with each pair of participants trying all three conditions. The order of conditions, as well as task targets, was counterbalanced between the six groups of participants. The first condition began when the two participants were ready. In each condition, two participants were required to find 12 target objects. The first four were used in a practice session, and the last eight were the formal tasks. In the practice session, participants were allowed to talk to each other freely, as well as developing strategies without being restricted by verbal cues. The participants were video-recorded throughout the task process for later analysis. After completing the task, participants were asked to fill the per-condition questionnaire, which included the Single Ease Ques-

.....

tion (SEQ) [68], the Spatial Presence Questionnaire (SPQ) [78], the Subjective Mental Effort Questionnaire (MEQ) [90] and the System Usability Scale (SUS) [8]. Then, the experiment was continued with the next condition. Once all three conditions were completed, participants were asked to complete an additional post-experiment questionnaire exploring their preference for each of the conditions.

After that, the participants' roles were switched, and similar task procedures and questionnaires were repeated. Overall, the entire experiment took about one hour for each pair.

## 6.6 Results

This section reports the study results with statistical analyses and summarise qualitative feedback collected from the participants ( $n = 12$ ) among the three conditions: 1st-person solid window (1PSW) condition, 1st-person face window (1PFW) condition and 3rd-person face window (3PFW) condition. Considering the relationship between the three conditions, the study utilizes a pairwise comparison of the three conditions. Since both were in 1st-person mode, the analysis use the Wilcoxon Signed Rank tests ( $\alpha = 0.05$ ) to compare the two conditions of 1PSW condition and 1PFW condition in order to observe the effect of the face window on experiments and participants. Meanwhile, since both conditions had the face window feature, the analysis applied the Wilcoxon Signed Rank tests to compare the two conditions of 1PFW and 3PFW to discuss how the view modes on the OmniEyeball affected the evaluation of the system. Furthermore, the study compares the user's perception between the two roles (designer and occupant) toward the OmniEyeballVR system.

### 6.6.1 User Performance

To evaluate the performance, the study recorded the correct rate of occupant's first answer. The average correct answer rates were 0.812 (SD = 1.931, 1PSW condition), 0.989 (SD = 0.289, 1PFW condition) and 1.000 (SD = 0, 3PFW condition). A significant differ-

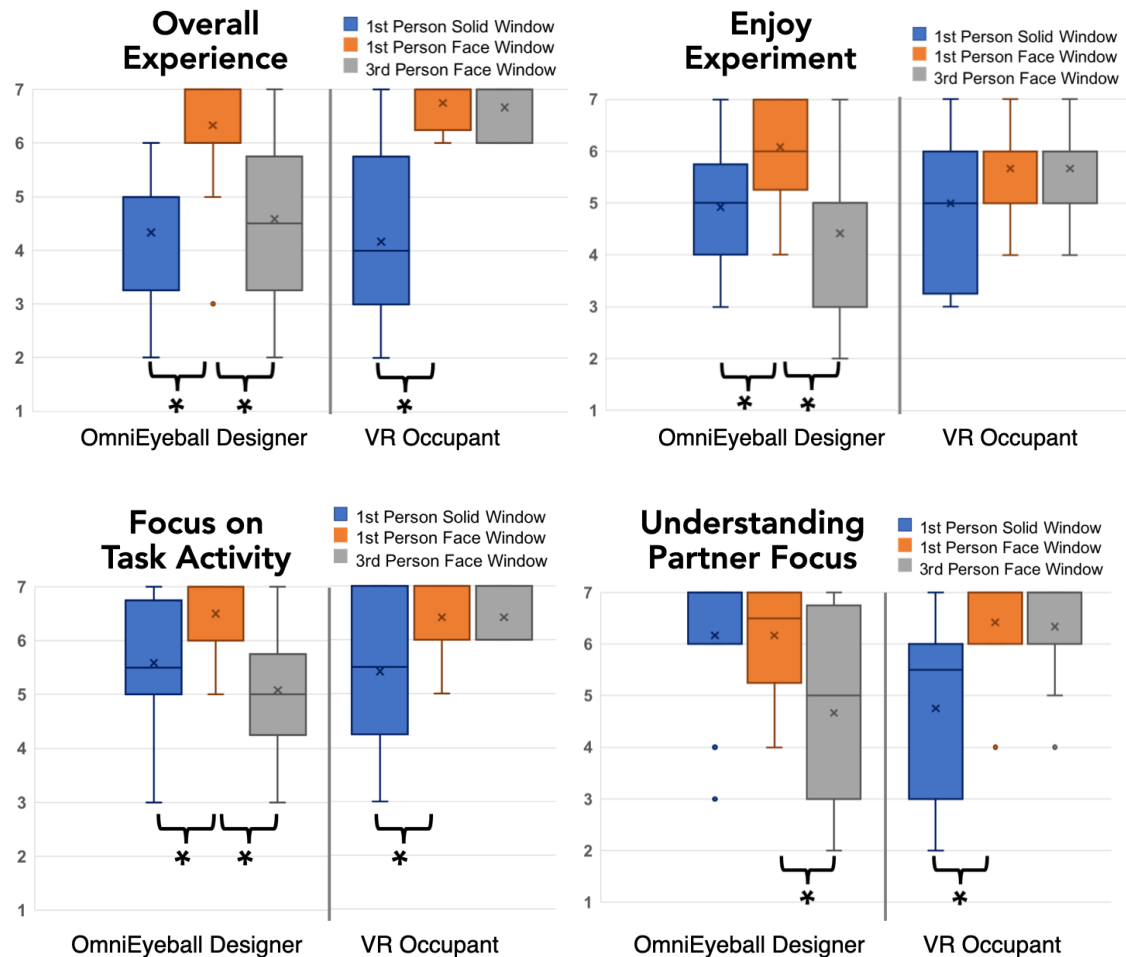


Figure 6.13: Subjective ratings on user experience (\*: significant effect).

ence was found between 1PSW and 1PFW ( $Z = -2.219, p = .026$ ), while no significant difference between 3PFW and 1PFW ( $Z = 1.000, p = .317$ ).

## 6.6.2 User Experience

The results from SEQ [68] and custom rating items are shown in Figure 6.13. All of the rating items were answered on a 7-point Likert scale (SEQ 1: very difficult – 7: very easy; All other items - 1: strongly disagree – 7: strongly agree).

1PFW vs 1PSW: On the designer side, the results showed that using the 1PFW mode to

finish the task was significantly easier than the 1PSW mode ( $Z = 2.963$ ,  $p = .003$ ). It also showed that the 1PFW mode provides a significantly better enjoyment of the experience ( $Z = 2.392$ ,  $p = .017$ ) as well as better focus on the task activities ( $Z = 2.414$ ,  $p = .016$ ). On the occupant side, the results showed that using the 1PFW mode was also significantly easier than the 1PSW mode ( $Z = 2.949$ ,  $p = .003$ ). It also showed that the 1PFW mode provides a significantly better understanding of partner's focus ( $Z = 2.687$ ,  $p = .007$ ) as well as a better focus on the task activities ( $Z = 2.46$ ,  $p = .014$ ).

1PFW vs 3PFW: On the designer side, it was found that using the 1PFW mode was significantly easier than the 3PFW mode ( $Z = 2.687$ ,  $p = .007$ ). It was also found that the 1PFW mode provided a significantly better understanding of the partner's focus ( $Z = 2.265$ ,  $p = .024$ ), better enjoyment of the experience ( $Z = 2.831$ ,  $p = .005$ ) as well as better focus on the task activities ( $Z = 2.46$ ,  $p = .014$ ). On the other hand, no significant differences were found on the occupant side.

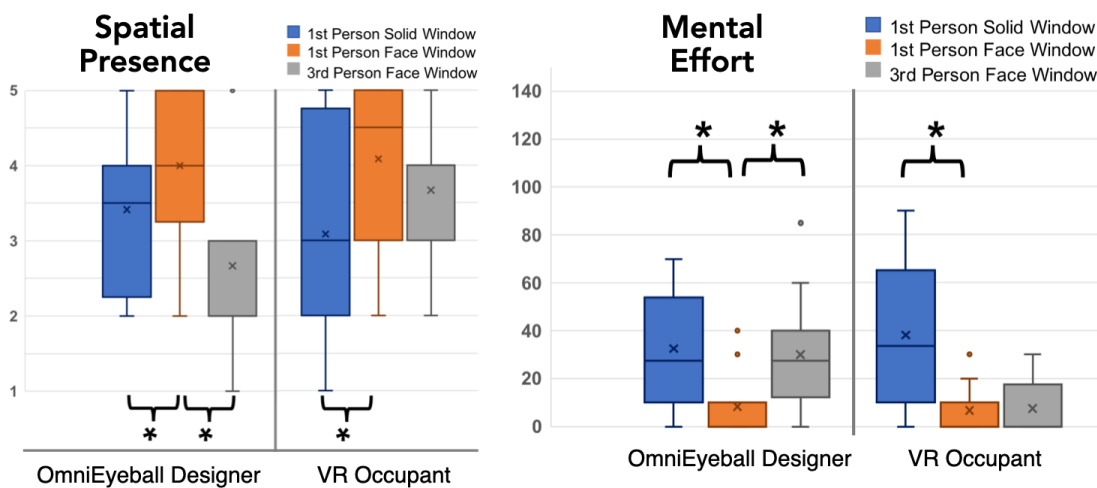


Figure 6.14: Results of Spatial Presence questionnaire (left) and MEQ (right). (\*: significant effect)

### 6.6.3 Spatial Presence

1PFW vs 1PSW: The spatial presence was evaluated by answering the question “*I felt like the partner and I was in the same space.*” with a 5-point Likert scale (1: strongly disagree – 5: strongly agree), used in a previous work. As shown in Figure 6.14 (left), it was found that the participants felt a significantly stronger spatial presence in the 1PFW condition on both sides. (designer side:  $Z = 2.333$ ,  $p = .02$ ; occupant side:  $Z = 2.232$ ,  $p = .026$ )

1PFW vs 3PFW: On the designer side, it was found that using the 1PFW mode provided a significantly stronger spatial presence ( $Z = 2.599$ ,  $p = .009$ ). While on the occupant side, the results reported no significant difference.

### 6.6.4 Mental Effort

1PFW vs 1PSW: The results from MEQ [90] are shown in Figure 6.14 (right). Significant differences were found on both sides. (designer side:  $Z = 2.965$ ,  $p = .003$ ; occupant side:  $Z = 2.943$ ,  $p = .003$ ), which shows that the participants felt that it took less mental effort to use the 1PFW mode.

1PFW vs 3PFW: On the designer side, it was found that using the 1PFW resulted in less mental effort ( $Z = 2.814$ ,  $p = .005$ ). While on the occupant side, no significant difference was found in the results.

### 6.6.5 System Usability

1PFW vs 1PSW: In terms of SUS [8], it was found significant differences on both sides (see Figure 6.15). (designer side:  $Z = 3.062$ ,  $p = .002$ ; occupant side:  $Z = 2.909$ ,  $p = .004$ ), which shows that the participants thought the system usability was significantly better in the 1PFW mode.

1PFW vs 3PFW: On the designer side, it was found that using the 1PFW resulted in better system usability ( $Z = 3.066$ ,  $p = .002$ ). While on the occupant side, the results reported no significant difference.

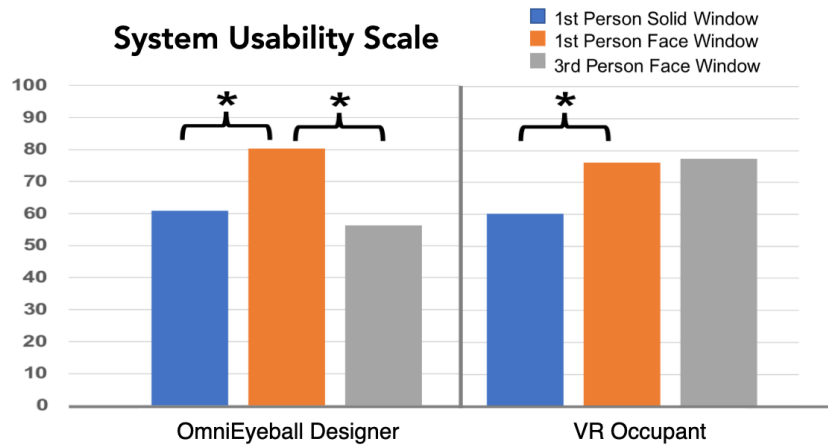


Figure 6.15: Results of System Usability Scale Questionnaire.

### 6.6.6 Post-experiment Results

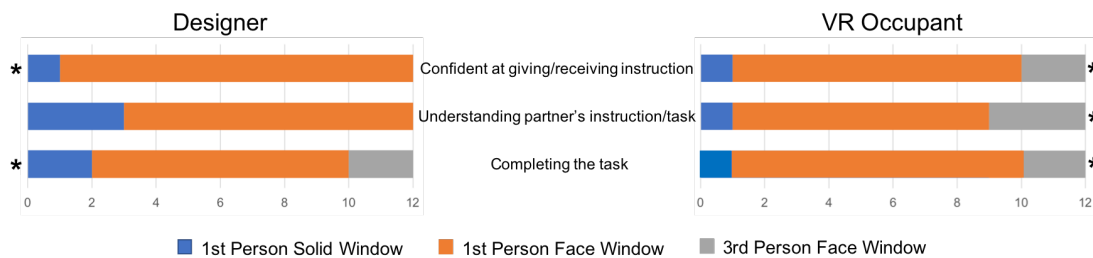


Figure 6.16: User preference between 1PSW, 1PFW, 3PFW conditions, based on the three questions against the number of participants (\*: statistically significant).

Four custom ranking questions were included in the post-experiment questionnaire in order to access the participant's choice of preferences and subjective ratings on both sides. Figure 6.16 shows their preferences based on three aspects (1) Confident at giving/receiving instruction, (2) Understanding a partner's instruction/task, and (3) Completing the task. Participants were asked to select for a condition as their best preferred condition in each of the question. Overall, participants on both sides had a huge favor towards the 1PFW across all aspects. However, using 3PFW was also preferred for a

minority of participants in the occupant side. A Friedman test indicated that there were significant differences in ranking between the three conditions for all questions except Understanding a partner's instruction/task on the designer side. (Designer- Confident at giving/receiving instruction:  $\chi^2(11) = 6.00$ ,  $p = .050$  and Completing the task:  $\chi^2(11) = 8.33$ ,  $p = .004$ ; Occupant- Confident at giving/receiving instruction:  $\chi^2(11) = 9.50$ ,  $p = .009$ , Understanding partner's instruction/task:  $\chi^2(11) = 6.50$ ,  $p = .039$ , and Completing the task:  $\chi^2(11) = 9.50$ ,  $p = .009$ ).

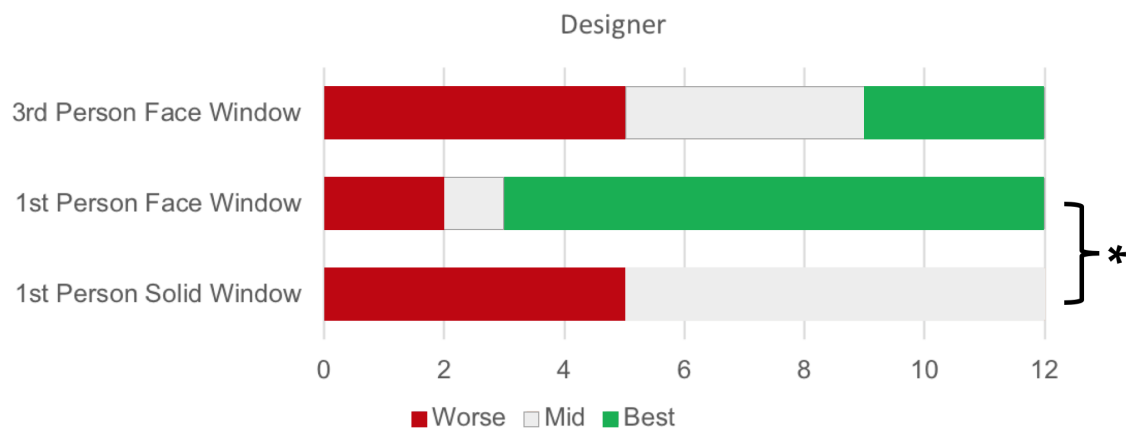


Figure 6.17: User preference-based ranking results on the designer side against the number of participants.

A separate question was also given to the participants to rank the conditions in the order of their overall preferences, as illustrated in Figure 6.17 and 6.18. In general, the majority of the participants on the designer side had selected the 1PFW as their top-ranked condition while their preferences were split between the mid and the worst in the 3PFW and 1PSW conditions. On the occupant side, participants had a similar trend of having the best ranking for 1PFW, but the majority of them also showed a strong sentiment towards the 1PSW being the worst preference. A Friedman test was conducted in each role between the three conditions and the result found a significant difference in both groups (Designer:  $\chi^2(11) = 6.50$ ,  $p = .039$ ; Occupant:  $\chi^2(11) = 6.00$ ,  $p = .050$ ). Furthermore, post hoc tests using the Wilcoxon Signed-Rank test with Bonferroni cor-

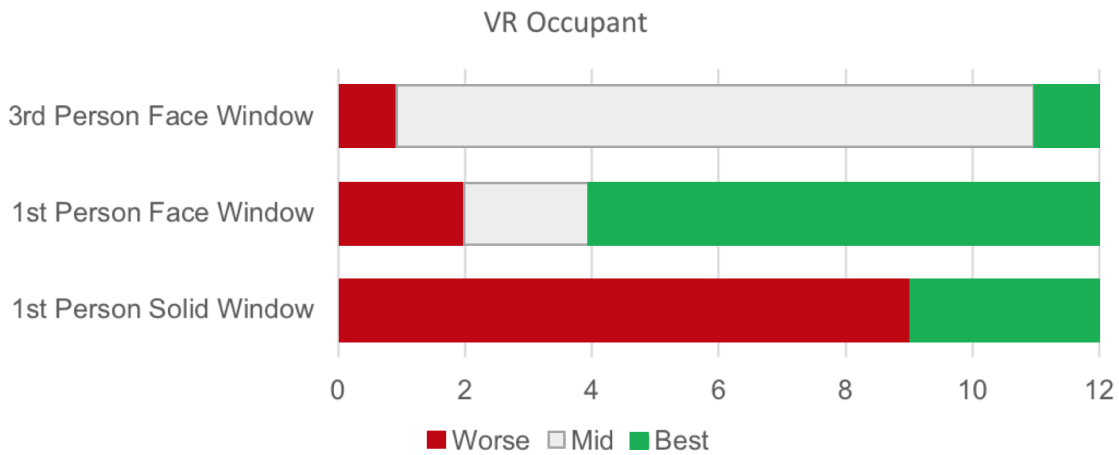


Figure 6.18: User preference-based ranking results on the occupant side.

rection ( $\alpha = .0167$ ) also showed that the participants in the designer side significantly preferred the 1PFW over the 1PSW condition ( $Z = -2.443$ ,  $p = .015$ ). However, there was no significant difference reported on the 1PFW against the 3PFW, and the 1PSW against the 3PFW, and all of the compared pairs in the occupant side.

## 6.7 Discussion

The experimental results show that participants have different preferences for the three conditions because of different reasons. In the following sections, the results are interpreted based on the characteristics of the face window as well as the differences between the first-person and third-person modes, mainly supported by qualitative feedback from the participants.

### 6.7.1 Face Window

By comparing the results of the 1PSW and 1PFW conditions, the study analyzed the significance of the face window feature for the OmniEyeballVR system. In the 1PSW condition, users found that it was not easy to inform their partner about the elevation

information. Although the designer could map the solid window directly above the target by facing the target object column, only three groups found strategies to inform the occupant whether the target was the top one or the bottom one. Among them, one group informed the top/bottom direction by turning clockwise or counterclockwise to the target column. The other two groups inform top/bottom by placing the solid window offset to the left or right of the target column. Therefore, the results in the correct rate were affected by that, while the other three groups didn't find a way to inform the top/bottom information, so they had to guess either top or bottom in the first answer. While in the 1PFW condition, participants could use gestures or gaze to indicate the top/bottom target. It is believed that in more complicated scenarios, the face window could convey more and richer information through visual cues.

As for the subjective measures, the statistical analysis of the occupant's side results showed that the face window provided comparable user experience, usability, and with its benefits, including less mental effort, a more substantial spatial presence, and a better understanding of the partner's focus. Such findings were supported by participant's comments, such as (P4) *"I was able to see where my partner was looking as well as the partner. So the task was easy."* (P3) *"I could see the gesture of my partner"*, and (P8) *"see partner's face, not being alone in my world."*

### 6.7.2 Visual Feedback of Face Window

As for the results comparing between 1PSW and 1PFW of the designer's side, since both were in the first-person mode, the main difference between the two conditions was reflected in the visual feedback window of the face window on OmniEyeball. Therefore, the difference in the results could be due to the difference in the visual feedback design to some extent. It seems that by showing a face window feedback on OmniEyeball, the participants on the designer side had better confidence at giving instructions as they showed a stronger spatial presence, better user experience, and a higher system usability score. Some participants mentioned that (P3) *"with face window, I can tell where and*

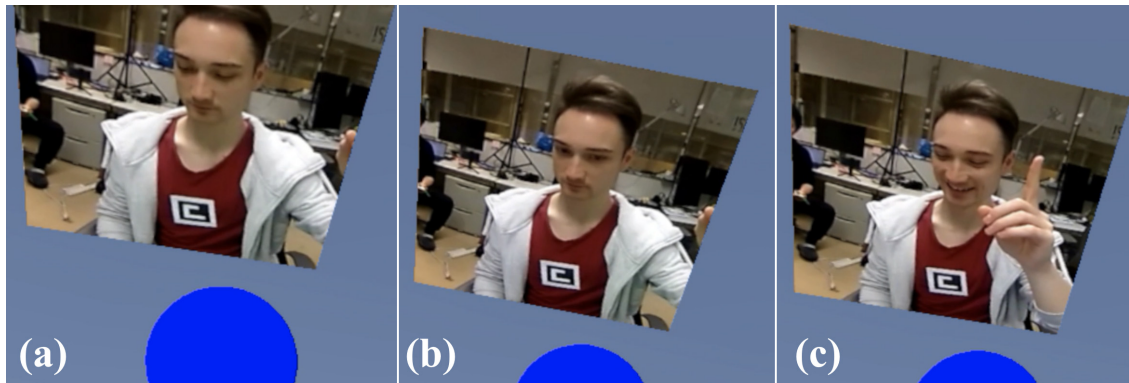


Figure 6.19: The participant(P12) used the visual feedback to correct his hand gesture shown in the face window.

*what I am looking at. I can also correct myself very quickly if my reflection is wrong*". Similar comments were also given by P2, P9, P11, mentioning how the face window could be used as an individual guide. Meanwhile, an interesting case was found during the experiment with P11 and P12. When P12 (designer) gave a hint to P11 (occupant), P12 accidentally put his gesture outside the FOV of the face window (Figure 6.19 (a)). When he moved his gaze from the instruction note back to OmniEyeball (Figure 6.19 (b)), he found that his gestures were not in the face window, and quickly adjusted the position of the gesture (Figure 6.19 (c)). Therefore, it seems that the visual feedback of the face window allows a designer to reference their visual hints that are mirrored and shown on the OmniEyeball in order to ensure their instructions are delivered to the VR user appropriately.

### 6.7.3 First-person and Third-person Mode

The comparison between the results of 3PFW and 1PFW conditions helped to learn how the participants perceived the first-person and third-person views differently. All of the statistical analyses reported no significant differences between 3PFW and 1PFW on the occupant side, which is because the face window and interfaces on the occupant side are mostly the same between the two conditions. However, on the designer side, there is

.....

a significant difference in terms of the perception of participants between the two modes. The results showed that the 1PFW provided less mental effort, better spatial presence, understanding of the partner's focus, user experience, and the usability of the system compared to the 3PFW. In addition, participants (P12, P7) commented that the 1PFW was very intuitive.

One hypothesis is that visual feedback of the face window is not clearly displayed in the 3PFW condition due to the point of view (Figure 6.12d), which affected the participants' perceptions of the spatial presence and user experience in 3PFW. The comparable results of the post-experiment questionnaire seem to reveal that the first-person mode received more preferences than the third-person mode on the designer side. From the comments of the participants, some problems with 3PFW were found. For example, P5 and P9 believed that the current visual effect of the face window's feedback on OmniEyeball prevented them from knowing how they were displayed in the VR space. On the other hand, some participants said that the current 3PFW display resolution was too low (P7, P4), and this will cause problems with visibility as they said that "*the object I have to focus is far away and smaller than in the other two conditions*" (P4) and "*it's hard to see and understand the FOV of the occupant.*" (P6)

Some participants chose 3PFW as their best preference condition. For example, P3 and P9 commented that "*The 3PFW allows me to see more things, which helps to check the position of all the objects, so I can quickly find the target and give instructions to the occupant.*"

## 6.8 Limitations and Solutions

The study reports some limitations of the OmniEyeballVR. At present, there is no function designed to assist VR occupants in finding the position of the face window in VR space, so some participants indicated that once the face window disappeared out of their field-of-view, it was tough to keep looking up to search for the face window (P1, P10). Therefore, inspired by the Outside-In [48], it is interesting to implement a

.....

spatial navigation function of the face window in the occupant FOV, such as using a small translucent face window.

Meanwhile, the current spatial position of the face window cannot be arbitrarily adjusted by the user. This is limited by OmniEyeballVR's lack of robust spatial interaction methods. The participant comments implied expectations for more interaction, such as (P8, P9) "*if there is an interactive way to control the vertical position of the face window, or directly display the face window near the target object should provide better pointing ability and efficiency.*" Therefore, it is interesting in implementing a robust interactive method for OmniEyeballVR in order to support spatial input.

Another limitation was the design of the experimental task. The current user study suffered from a small number of samples (N=12) and the gender imbalance (10M, 2F). Furthermore, the system was intended for a collaboration between a single occupant with multiple designers on a design task. Yet, the experiment only investigated the communication and cooperation in a one-to-one configuration between an occupant and a designer. The task was also focused on remote-expert instruction, which only explored in a limited use case scenario. Therefore, it is interesting in investigating the system in a different task configuration in order to evaluate for other applications.

## 6.9 Summary

This chapter presented OmniEyeballVR, which is a collaborative system using the OmniEyeball (OEB2) terminal that allows a VR occupant and multiple designers to cooperate across the virtual and physical worlds. This allows designers to easily view physical assets and access the content in the VR space from any point of view. The system also supports sharing gaze awareness cues, facial expressions, and body gestures to enhance communication.

A user study was conducted to evaluate the system in three different conditions (1PSW, 1PFW, 3PFW) in a VR cooperation task where a designer gives instruction to a VR occupant to search for target objects. The results showed that participants mostly preferred

.....

the first-person view mode with the face window feature. It is quite intuitive to use, as well as provides a good spatial presence, understanding of the partner's focus and user experience. The study also discussed the advantages and limitations of the third-person mode, which helped to provide implications on how to improve such systems. It is expected that the contribution of the OmniEyeballVR will provide advice and guidelines for future telecommunication and VR collaboration system designs.

There are further opportunities to address some of the current system's limitations and expand the interactive features to make OmniEyeballVR more effective for VR prototyping. Additionally, this work enables and points to the importance of user studies with multiple participants simultaneously acting as designers/developers.

---

# Chapter 7

## Discussion

In this chapter, the insights and findings of systems and applications described in Chapters four, five, and six (OEB, OEB2, and OEBVR) as well as the user study on them will be summarized, analyzed, and discussed.

### 7.1 Research Goals

Following the research proposal, this research was proceeded based on the following subject goals.

1. The combination of the spherical display with omnidirectional cameras is a suitable choice for a 360° I/O system in panoramic video communication.
2. The invisible hemisphere issue found in the OEB experiment can be improved and resolved with the right means and design of interaction features.
3. The proposed 360° I/O device terminal not only enables panoramic video communication but also can bring possibilities to more potential scenarios about panoramic content.

## 7.2 Research Contributions

### 7.2.1 Insights and Findings

This research proposes 360° I/O systems combining omnidirectional cameras with a spherical display and apply it to panoramic video communication scenarios. The validity and usability of the system were verified through system implementation and user study. Below, conclusions are depicted with the research goals of this study.

#### **The combination of the spherical display with omnidirectional cameras is a suitable choice for 360° I/O system in panoramic video communication**

In this thesis, chapter 4 introduces the OEB, which combines a panoramic camera with a hemispherical display. Chapter 5 presents the OEB2, which combines a panoramic camera with a full-spherical display. User studies show that the combination of omnidirectional cameras with a spherical display works well in the 360° panoramic video communication scenario.

In chapter 4, by comparing 2D flat display and hemispherical display, the main findings show that a 2D flat display enables participants to see everything in one view but increases cognitive load due to the image distortion and the necessity of mental rotation. The spherical display can only show half of the environment at once but provides a better sense of space and consistency/coherency of two distant spaces. As for the affordance, the upward orientation of the 2D flat display affords participants to hold objects right above the display. The spherical display affords participants to hold objects in front of their chest naturally. The 2D flat display requires participants to look down to see the display, which is quite unnatural for the participants when they are showing objects. A similar issue was also reported by Licoppe et al. [47]. The spherical display does not have such issues. The spherical display better affords one to one conversation (side-channel communication) while the 2D flat display better affords many to many conversations (more communications without standing face-to-face happened in the 2D condition). Although the spherical display requires a participant to walk around to talk with a particular person

.....

or to see a particular object, such movement clearly indicates the participant's point of interest.

The findings of visual perception, affordance, side-channel communication, and user preference show that the spherical display is more preferred than the 2D display and implies it as a suitable choice for 360° panoramic video communication.

In chapter 5, the study of OEB2 further proved that the combination of a panoramic camera with a full-spherical display could support one-to-one users to do video communication in a more complicated remote collaboration scene.

### **The invisible hemisphere issue found in the OEB experiment can be improved and resolved with the right means and design of interaction features**

The study of OEB reported the invisible hemisphere issue. To address this issue, the OEB2 presents new hardware improvement as well as interactive features.

In chapter 5, the experimental results show that the features presented by OEB2 are effective in resolving the invisible hemisphere issue, with three main contributions: (1) Extension of the full-sphere field of view (2) OmniBar with segmentation design (3) Physical Rotation.

1. By making OEB2 a full 360° I/O device, the participants could place any objects at any places quite casually, which affords casual visual interaction that utilizes the whole 360° environment in the video communication.
2. The OmniBar is a novel feature invented for OEB2. Its novelties are that 1) it makes use of the wasted area of the OEB2 by repeating the full field of view twice, participants can get the whole 360° view from any direction, and also 2) it's a natural touch interface to enable intuitive monitoring and referencing of the remote environment.
3. The results showed that the invisible hemisphere issue was resolved by the OmniBar and mainly physical rotation, which caused less perceived workload in the

remote collaboration. Furthermore, the analysis of the feature set indicated that the participants preferred Physical Rotation more than OmniBar because of good affordance and supporting better gaze awareness. It leads to a concluded implication that more physical interaction should be supported for spherical display systems, which is also mentioned in the user study of Qoom [69].

In chapter 6, the OEBVR system provides an additional solution to the invisible hemisphere issue. By utilizing the VR fish-tank technology [18], the issue is naturally resolved in the 3rd-person view mode. Therefore, it is interesting to find out that the invisible hemisphere issue of the spherical display system can be addressed by the right means and design of interaction.

**The proposed 360° I/O device terminal not only enables panoramic video communication but also brings possibilities to more potential scenarios about panoramic content**

When there is only one application scenario for an interactive terminal device, it dramatically limits and reduces the sustainability of this terminal device, e.g., OmniEyeball (including OEB and OEB2) can only be used as a device for video communication.

Fortunately, chapter 6 succeeds in portraying a new application scenario that makes the OEB2 more than just a communication device that connects through two real spaces. It also can be a communication device that connects through real and virtual spaces. As a unique 360° I/O device, it provides a novel collaboration experience in VR and 360° content prototyping that is never provided by the previous works. Its presentation also hints at the possibilities that the combination of omnidirectional cameras with a spherical display can bring new interactions in more areas.

### 7.2.2 Innovation

Besides the findings from the user study, the contributions of this research also come from the novelties of the systems, which are shown as follows:

1. The OEB supports the first study of how omnidirectional cameras support symmetric 360° panoramic video communication and evaluate it in a collaboration setting.
2. The OEB2 is the first interactive system to combine a full-spherical display with the omnidirectional camera. In addition, it is the first to design interactive functions for such a 360° I/O device to support the symmetric 360° panoramic video communication experience.
3. By utilizing spherical displays in a VR collaboration scene, the OEBVR is the first collaboration tool to simultaneously provide a diversified viewpoint (active first-person view and third-person view) as well as face-to-face communication in VR collaboration prototyping.

## **7.3 Design Implication of 360° I/O Systems**

### **7.3.1 The Synchronized Rotation to Support Gaze Awareness**

The user study of OEB2 and VR shows that the physical rotation feature is an effective design for a 360° I/O device. Not only does it effectively and intuitively resolve the invisible hemisphere issue, but it also provides good gaze awareness [83] (Figure 7.1), which is named as “Pseudo Gaze Leading Effect”. Pseudo gaze leading effect indicates that a video communication system should provide users an intuitive way to lead a remote partner’s gaze to the intended orientation and also presenting the user’s gaze orientation to the remote partner. Utilizing symmetric spherical displays with physical rotation is one of the effective solutions.

The reason the physical rotation feature proposed in this thesis is possible with the Pseudo Gaze Leading Effect is because it is a synchronized rotation. That is, since the camera and the spherical display are orientately bounded, the camera will also follow the rotation synchronously when the user tries to rotate the spherical display. However, physical rotation proposed in this thesis has its limitations, e.g., physically rotating the whole device may cause messy outside of the system. One possible solution is to implement the

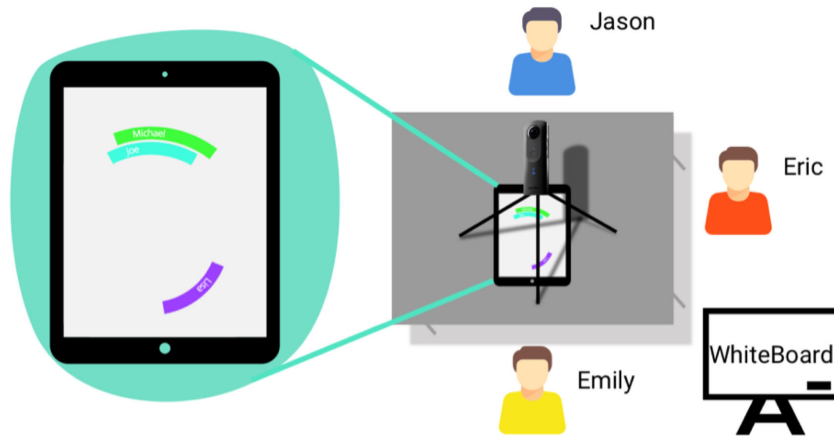


Figure 7.1: Attention from Afar [83] proposed by Xu et al..

rotation platform based on a better mechanical structure. If designers plan to implement a similar physical rotation in the digital approach, the user can digitally rotate the display by touching and swiping on the screen. The possible solution for the digital way may be to use image processing to make the image of the remote display synchronically follow the rotating input of the user.

### 7.3.2 Improvement for Reversed Space

Although the OmniEyeball is implemented based on the teleconference metaphor so that the reversed space issue exists in these two systems. However, both the system's user study did not show this issue affected the system performance or user interaction. In particular, OEB2's experiment task includes an arrangement task that requires the participant to understand the spatial orientation of the remote space and give instructions based on the map and observed image. Nevertheless, participants gave no reports about the reversed space issue.

The experimental results show that the reversed space issue is less likely to affect the remote collaboration on panoramic video communication negatively. However, there is no guarantee that it won't appear in all the remote collaboration scenes. According to

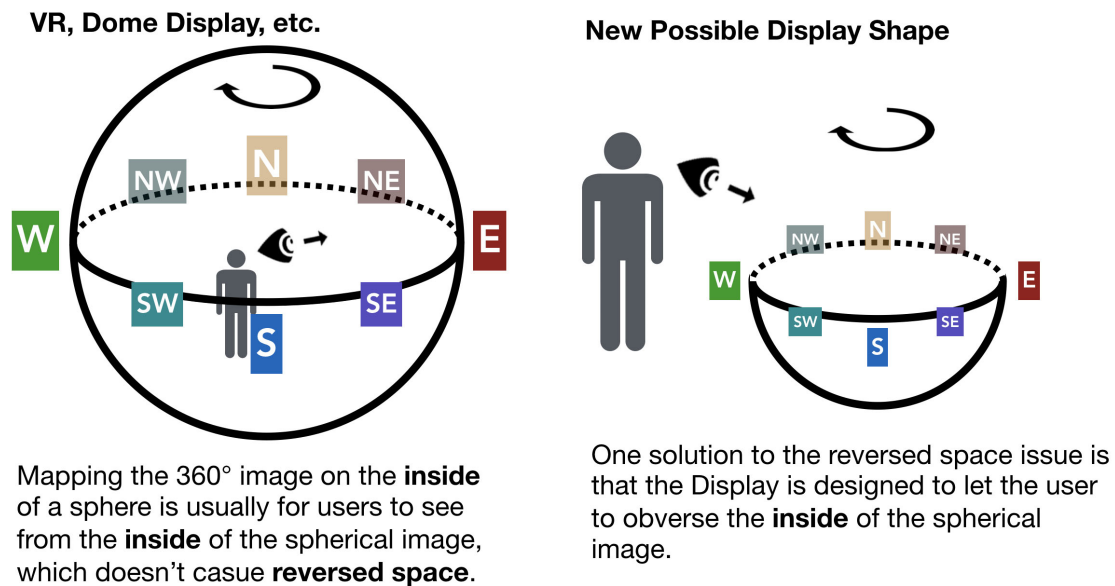


Figure 7.2: One solution to the reversed space issue.



Figure 7.3: DisplayBowl [54] proposed by Miyafuji et al..

chapter 4.3.3, the reversed space happens because the user observes from the outside of a 360° spherical images. As shown in Figure 7.2, if a hemispherical display can be designed as a bowl-shaped display so that users are able to observe the inside of the spherical image even they are standing outside of the sphere, the reserved issue can be resolved in this display setup.

.....

Fortunately, displays of this form have begun to appear in the latest studies. The DisplayBowl [54] proposed by Miyafuji et al. is designed for remote robot operation (see Figure 7.3). Combining this type of display with a 360° camera may open up new possibilities for panoramic video communication.

## 7.4 Limitations

In addition to the limitations for each of the three systems (OEB, OEB2 and OEBVR) discussed in the respective chapters, several common limitations to 360° video I/O systems are discussed here.

### 7.4.1 Vertical Gaze Misalignment

Gaze misalignment is a common problem in video communication. It is caused by a difference in the position of the camera and the display, causing a mismatch between the direction of observation of the person in the video and the actual direction of observation of that person. The OEB2 applies a Ricoh Theta camera standing at the top of the spherical display. The vertical distance between the camera and display is not small, which causes an unavoidable vertical gaze misalignment.

However, in the actual user experiments, few results showed that the vertical gaze misalignment caused distress to user interaction and communication. This makes it reasonable to believe that vertical gaze misalignment is not as important in the common video-based remote collaboration.

At the same time, face feature point detection and semantic segmentation are becoming more mature as the field of computer vision develops. If the vertical gaze misalignment generates a negative impact on the 360° video communication in a particular scene, it is possible to resolve this issue by correcting the gaze alignment using image processing methods.

### 7.4.2 Privacy Issue

Privacy protection has also been a common issue in video communication. Especially when the field of view for video communication is expanded to 360°, the entire space is transmitted unreservedly to the other side. It can render some of the user's familiar means of privacy protection ineffective, such as making the device's front camera shoot only primarily themselves.

Since the user study in this research focused primarily on remote collaboration, the privacy protection of users in daily use was not much explored. However, in the video communication tool, which has been heavily used in the recent Covid-19 epidemic, available methods are found that can be applied to 360° systems. That's the background context replacement. With the further development of computer vision and image processing, a stable and efficient way of replacing backgrounds will be a simple but significant solution to video communication's privacy issues.

### 7.4.3 Limited Use Scenario

Cutting-edge technologies often encounter problems with restricted application scenarios. Usually, due to price, the devices used in cutting-edge technologies (such as panoramic cameras) do not have a sufficient user base to achieve the interaction and application revolution at the current point of time.

The proposed OmniEyeball (including OEB, OEB2) encountered a similar problem in this study. Due to the lack of popularity of panoramic cameras, the current commercial panoramic camera is basically only used for taking photos and videos. And in the first half of the exploration of this study, it wasn't easy to find any other application scenarios other than video chat. This lack of an application scenario came to an end until a combination of OmniEyeball and VR applications was discovered.

Therefore, for the similar 360° I/O devices, in addition to the interaction and functional improvements needed for panoramic video communication application, researchers should look more towards the VR and MR fields. It is believed that the ability of

panoramic cameras to capture the entire space may open up more possibilities for VR application. The relevant discussion is written in the next subsection.

#### **7.4.4 Communications among Three Spaces**

The OmniEyeball (including OEB, OEB2) applications and experiments in this thesis are based on symmetric video communication between two spaces. Due to the limitation of the number of OmniEyeball system terminals (only two terminals were implemented), this thesis does not study and explore video communication among more than three spaces. Therefore, the 360° video communication application connecting more than three spaces is one of the limitations in this thesis.

For cases where more than three spaces are connected by video communication, the configuration of such a system application can be designed based on the thesis's experiments and findings. The findings of this research imply that one OmniEyeball terminal is a natural and intuitive representative of one remote space. Therefore, when using the OmniEyeball system to connect three spaces, two OmniEyeball terminals can be set in each space to represent two remote spaces respectively. In such a case, six OmniEyeball terminals are needed to enable the application for three spaces. Although the device cost is high, such an application can tightly connect the three spaces, which may bring benefits for specific remote collaboration scenarios.

### **7.5 A Vision for the Future**

#### **7.5.1 Commercial Prospects**

As discussed in chapter 1, more digital I/O devices support video communication within three decades. Such as laptops, smartphones, or tablets, combined the 2D display with small embedded cameras together as image input and output. As 360° panoramic cameras become more and more accepted and used by the general public, it is envisioned that 360° image I/O terminal devices will also be available to the public.

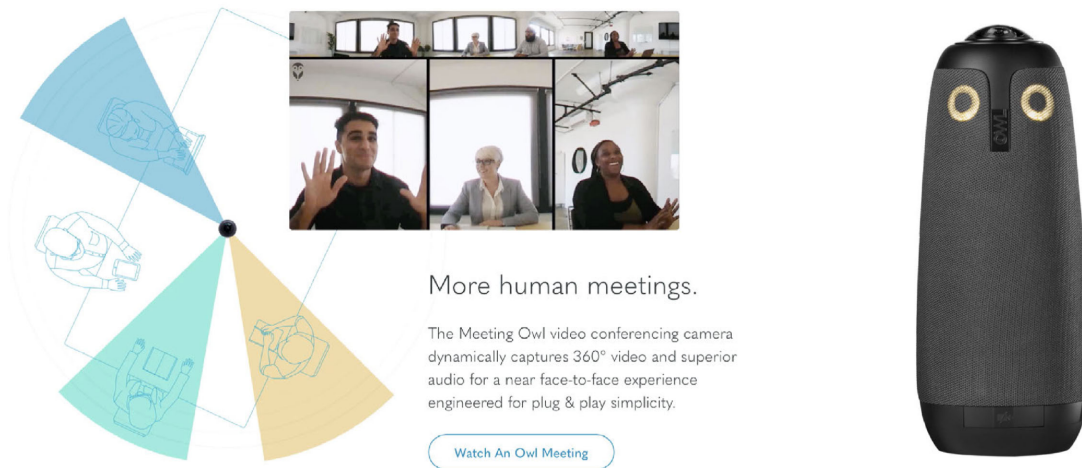


Figure 7.4: The Meeting Owl produced by Owl Labs.

Recently, some latest video conference cameras, such as Meeting Owl <sup>1</sup> (Figure 7.4) or Kandao Meeting <sup>2</sup>, embed 360° cameras to support 360° panoramic remote conference experience. However, in order to control costs and selling prices, these commercial products have compromised on the display side. In order to adapt to the big monitors that have been adapted in most meeting rooms, they improve some of the drawbacks of 360° panoramic images displayed on a 2D screen with a proper 2D UI design.

The first is that because the camera and monitor are placed separately (the camera is in the center of the conference table, while the monitor hangs on one side of the wall), the user's gaze will not be able to properly convey the message when making a video communication this way. For example, when a user is looking at the remote partner, his gaze appears to be looking not at the remote partner but to the side.

The second reason is that the 360° panoramic image actually contains very rich orientation information. By combining with a 360° display (e.g., a spherical display), this orientation information is possible to get a very intuitive expression and application, such as the "Pseudo Gaze Leading Effect". Therefore, to meet the needs of the increasingly

<sup>1</sup><https://www.owllabs.com/meeting-owl>

<sup>2</sup><https://www.kandaovr.com/kandao-meeting/>

.....

diverse applications of remote collaboration scenarios in the future, it is still quite important to explore the possibility of a commercial combination of the 360° panoramic camera with a 360° stereo display. The OEB2 presented in this research shows one possibility to that future.

As this paper is being written, the world of 2020 is becoming different than ever before because of Covid-19. Remote-education, teleconferencing, and video communication applications are being diffusely used to increase social distance. In particular, two teleconferencing software programs, Zoom and DingTalk, are widely used by schools and companies. If the 360° I/O devices proposed in this paper could be easily adapted to these popular video communication software, it would be very beneficial to the promotion of 360° video devices. Currently, my colleagues and I are trying to implement APIs for 360° video devices, so that the next generation of OmniEyeball can be easily used through the Zoom/Skype connection.

### 7.5.2 Mobile 360° Video Communication Application

The previous subsection mainly discusses the prospects of a desktop-based 360° panoramic video communication device; this subsection focuses on the mobile site and discusses the mobile 360° video communication application. The main interaction devices on mobile are still smartphones and smart tablets. Currently, there are already many 360° cameras that support wireless or port connection to the smartphone. Then, how can the mobile smart devices plus 360° cameras support mobile 360° video communication may become a good research topic.

Insta360 once proposed a 360° video chat application by utilizing their Insta360 Nano S camera<sup>3</sup> with smartphones (Figure 7.5). However, their video chat application configuration is uni-directional, which means that the 360° camera is only used at one side. Therefore, it seems that there are several future research topics, to support symmetric mobile 360° video communication by using mobile smart devices with 360° cameras,

---

<sup>3</sup><https://www.insta360.com/product/insta360-nanos>



Figure 7.5: The 360° video chat application enabled by Insta360 Nano S camera.

The first question is that since 360° video communication enables actively observing any side of the remote space, how can a user know which side the remote person is looking at? In the case of using the OmniEyeball system, the user can intuitively know it by checking the gaze or facing direction of the remote person's image on the spherical display. However, since the mobile smart devices usually only equip a 2D display, this will be an issue to solve.

The second question is that, by using the Pseudo gaze leading effect introduced in OEB2, the OmniEyeball system user can easily guide the remote person to see any local direction by physical rotation. In the case of conventional mobile video chat, the user can guide the remote person by moving and rotating the smartphone because the back camera of the smartphone can only capture in one direction. However, when combining a smartphone with 360° cameras, the directionality of the 360° camera becomes confusing. So I believe this may also become an issue to solve in mobile 360° video communication.

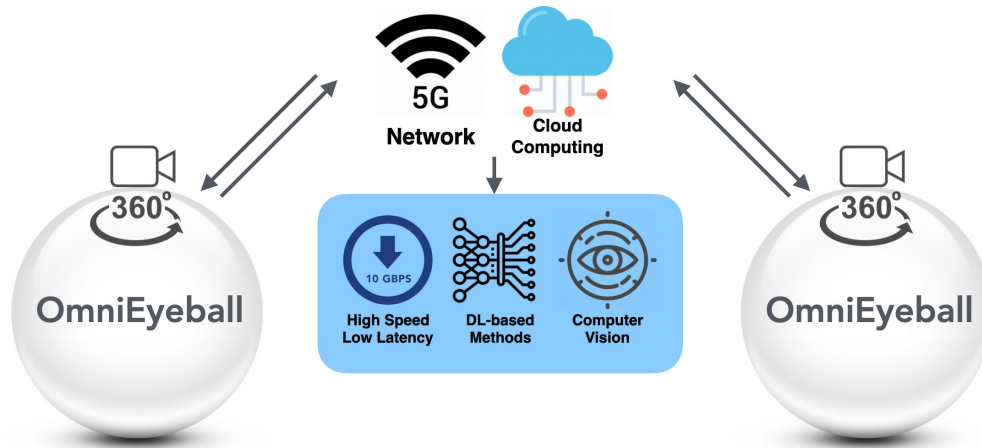


Figure 7.6: DL-based computer vision with 5G cloud computing support new OmniEyeball.

### 7.5.3 Computer Vision with 5G Cloud Computing

The OmniEyeball, as an interactive device equipped with a 360° camera, is bound to be tightly integrated with computer vision. Different from traditional hand-held smart-devices, the OmniEyeball system needs a robust user detection method to provide a proper user interface for multiple users. In the current version, only DL-based face detection is used. However, limited by the device's computing power, current face detection algorithms do not provide highly accurate real-time face detection. Besides, applying gaze, hand, and pose detection may also support diversifying user interaction.

Enabling many vision-based algorithms will put a huge computational cost on local devices. The latest 5G network with cloud computing technology may provide a solution to this situation. The 5G mobile network may provide 10Gbps data up/downloading speed, which can transmit the high-resolution live streaming with low latency. In addition, since the live video must be transmitted to the cloud to be sent to the remote site, utilizing cloud computing to process the vision-based algorithm seems quite reasonable. Therefore, I believe the 5G technology with cloud computing can play an essential role in the high-resolution 360° video communication (Figure 7.6).

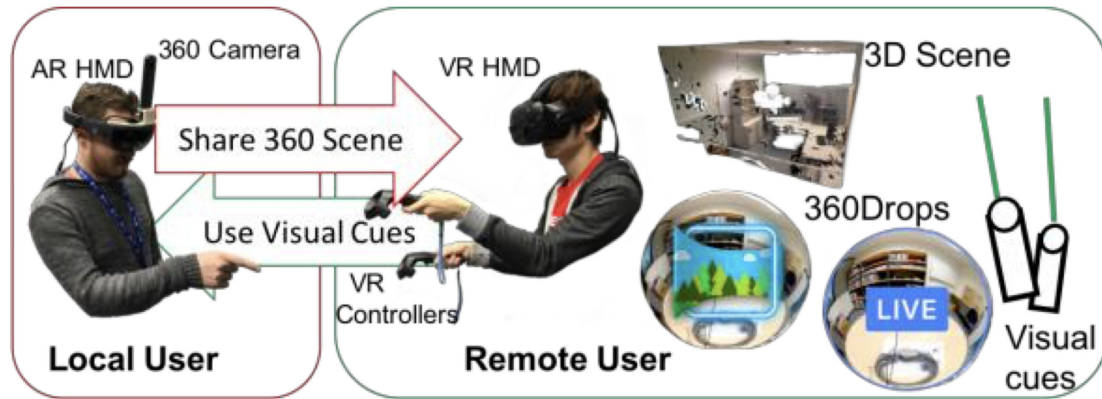


Figure 7.7: The 360Drops system proposed by Teo et al..

#### 7.5.4 VR Communication Scenario

The OmniEyeballVR makes the OmniEyeball system more than just a communication device that connects through two real spaces, and it also can be a communication device used for the VR field. The implementation of the OmniEyeballVR application shows the possibility to connect through real and virtual spaces. However, is there a possible scenario where communication between the two virtual spaces needs to be implemented?

The 360Drops [75] proposed by Teo et al. showed that space could be represented as a virtual sphere in VR (see Figure 7.7) for communication and collaboration. Then, one conceivable scenario is that in the future, when two people in their respective virtual spaces, if they want to just make video communication rather than meeting one virtual space, could they do such video communication by using two virtual OmniEyeball? It is expected that the insights and implications of this work can contribute to the research about VR-to-VR communication.

---

## Chapter 8

### Conclusion

This thesis introduced full 360° image I/O systems, OmniEyeball (including OEB, OEB2), which combines an omnidirectional camera with a spherical display to support the symmetric 360° panoramic video communication. The system overview is presented in detail, which makes it solid reproducible. The experiment was conducted to prove such 360° panoramic video communication can support multiple users for remote collaboration. Furthermore, a VR application tool named OmniEyeballVR was introduced, which supports communication and cooperation in VR prototyping and design.

Chapter One introduces the background of video communication systems and existing issues, narrow field-of-view. Then, the research motivation of the proposed 360° panoramic I/O system in this thesis is presented.

Chapter Two presents related works of the 360° video communication system, which consists of two parts: video communication systems and spherical display systems. Then, the existing limitations of the previous researches are listed.

Chapter Three explains the research proposal in detail, which presents how the proposed system can solve the existing issue in video communication systems. Meanwhile, a prototype system is introduced, which helps to confirm the implementation method for the OmniEyeball system.

Chapter Four introduces the OmniEyeball system. It supports to explore the possibilities of using 360° hemispherical video communication systems to assist symmetrical remote collaboration. The experimental results showed that the proposed systems enabled two pairs of people to communicate over the distance simultaneously from anywhere

.....

around the devices. However, subjective results indicated that participants preferred the spherical display (the OmniEyeball condition) over the 2D flat display. This could possibly be due to better image quality (less distortion) and a better sense of three-dimensional space, directional relation, and telepresence.

Chapter Five firstly analyzes the limitations of the OmniEyeball system. Then, to address these limitations, an improved symmetric full 360° video communication system, OmniEyeball2, is presented. The OmniEyeball2 is designed to have four novel features to support remote observing and increased shared awareness. Based on the evaluation of the comparative experiments, the four features were effective in reducing the inconvenience mainly caused by the partial invisibility of the spherical display. The results also indicate how the proposed features supported the remote space awareness and gazed leading intuitively and effectively.

Chapter Six presents an MR cooperation application of the whole OmniEyeball project, named as OmniEyeballVR. It utilizes the OmniEyeball2 to allow a VR occupant user and multiple designers to cooperate across the virtual and physical world. This allows designers to easily view the physical assets and access the content in the VR space from any point of view. The system also supports sharing gaze awareness cue, facial expression, and body gestures to enhance communication. A user study was conducted to evaluate the system. The results showed that participants mostly preferred the first-person view mode with face window feature as it is quite intuitive to use, as well as providing a good spatial presence, understanding of the partner's focus, and user experience.

Chapter Seven presents the insights and findings of systems and applications described in Chapters four, five, and six (OmniEyeball, OmniEyeball2, and OmniEyeballVR). The user studies on them are summarized, analyzed, and discussed. Furthermore, it discusses how the findings and contribution of the whole OmniEyeball video communication systems (including OEB, OEB2), as well as its VR application, can provide advice and guidelines for future telecommunication and VR collaboration system designs.

A conclusion graph of the whole thesis is shown in Figure 8.1.

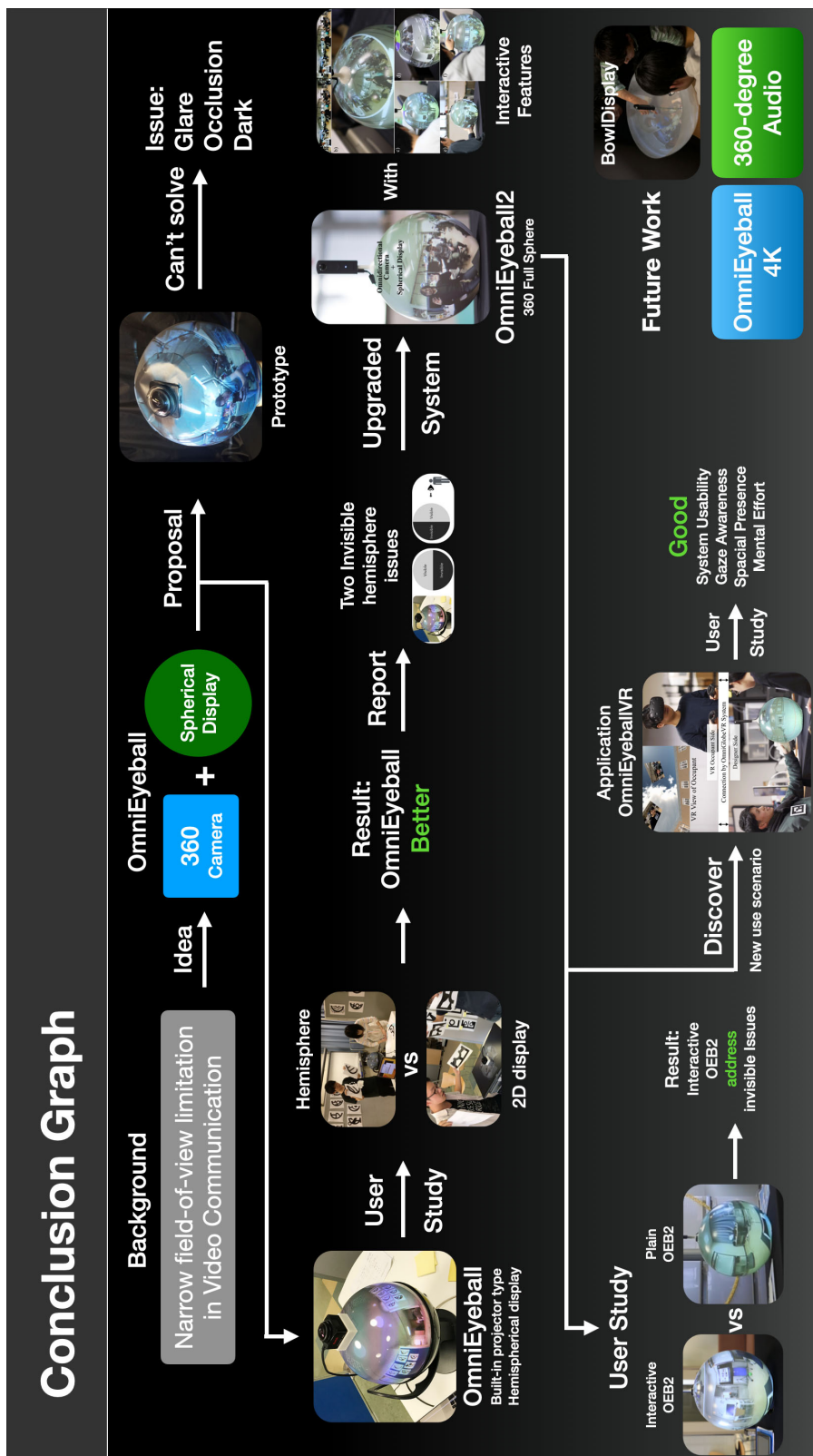


Figure 8.1: A conclusion graph of the whole thesis.

## Acknowledge

First and most importantly, I want to thank Professor Hideki Koike for the patient instruction and thoughtful guidance in both academics and life through these five years of my master and doctor degree.

I also want to thank the professors in the co-authors' list, especially Professor Hideaki Kuzuoka, who helped and gave me advice for three papers.

Then, I want to thank the co-authors and friends: Shio Miyafuji, Erwin Wu, Theophilus Teo, and Zhaoce Liu. Thank you for working together with me so that we made these academic achievements.

After that, I would like to thank all professors in my thesis defense committee: Professor Koichi Shinoda, Professor Takenobu Tokunaga, Professor Masamichi Shimosaka and Professor Yuta Itoh. Thank you for discussing and giving comments to me so that I can improve my doctor thesis.

Then, I thank all Koike lab members for these five-year memories, studying, and working together.

Finally and also importantly, I would like to thank my parents, Li Xihai and Zhao Xiuyan, for supporting me in finishing my doctor degree in both mental and financial perspectives.

I also want to say thank you to my fiancée, Huang Jianmin, who took care of me and cooked meals for me every day during this Covid-19 tough year so that I could finish my doctor thesis.

Thank you very much.

## Bibliography

- [1] Sang-Gyun An, Yongkwan Kim, Joon Hyub Lee, and Seok-Hyung Bae. Collaborative experience prototyping of automotive interior in vr with 3d sketching and haptic helpers. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '17, pages 183–192, New York, NY, USA, 2017. ACM.
- [2] O. R. Belloc, M. R. Nagamura, D. Fonseca, A. Rodrigues, D. A. R. Souza, C. S. Kurashima, M. M. Almeida, E. Z. Borba, R. D. Lopes, and M. K. Zuffo. Orbevr: A handheld convex spherical virtual reality display. In *ACM SIGGRAPH 2017 Emerging Technologies*, SIGGRAPH '17, New York, NY, USA, 2017. Association for Computing Machinery.
- [3] Hrvoje Benko and Andrew D. Wilson. Multi-point interactions with immersive omnidirectional visualizations in a dome. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, page 19–28, New York, NY, USA, 2010. Association for Computing Machinery.
- [4] Hrvoje Benko, Andrew D Wilson, and Ravin Balakrishnan. Sphere: multi-touch interactions on a spherical display. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 77–86. ACM, 2008.
- [5] Francois Berard and Thibault Louis. The object inside: Assessing 3d examination with a spherical handheld perspective-corrected display. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 4396–4404, New York, NY, USA, 2017. ACM.
- [6] Mark Billinghurst and Hirokazu Kato. Collaborative augmented reality. *Commun. ACM*, 45(7):64–70, July 2002.

- [7] John Bolton, Kibum Kim, and Roel Vertegaal. Privacy and sharing information on spherical and large flat displays. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, CSCW '11*, pages 573–574, New York, NY, USA, 2011. ACM.
- [8] John Brooke et al. Sus-a quick and dirty usability scale. usability evaluation in industry 189, 194 (1996), 4–7. *This page intentionally left blank*, 1996.
- [9] Marcio Cabral, F Ferreira, Olavo Belloc, Gregor Miller, C Kurashima, R Lopes, Ian Stavness, J Anacleto, Sidney Fels, and M Zuffo. Portable-spheree: A portable 3d perspective-corrected interactive spherical scalable display. In *2015 IEEE Virtual Reality (VR)*, pages 157–158. IEEE, 2015.
- [10] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman, editors. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.
- [11] Li-Wei Chan, Yi-Fan Chuang, Meng-Chieh Yu, Yi-Liu Chao, Ming-Sui Lee, Yi-Ping Hung, and Jane Hsu. Gesture-based interaction for a magic crystal ball. In *Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology, VRST '07*, pages 157–164, New York, NY, USA, 2007. ACM.
- [12] Liwei Chan and Kouta Minamizawa. Frontface: Facilitating communication between hmd users and outsiders using front-facing-screen hmds. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '17*, New York, NY, USA, 2017. Association for Computing Machinery.
- [13] Huiwen Chang and Michael F. Cohen. Panning and zooming high-resolution panoramas in virtual reality devices. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology, UIST '17*, page 279–288, New York, NY, USA, 2017. Association for Computing Machinery.

- .....
- [14] Thomas Crespel, Patrick Reuter, and Xavier Granier. 43-4: A low-cost multitouch spherical display: Hardware and software design. In *SID Symposium Digest of Technical Papers*, volume 48, pages 619–622. Wiley Online Library, 2017.
- [15] Thomas Crespel, Brett Ridel, Clara Rigaud, Anke M. Brock, and Patrick Reuter. Code the globe: Interactive content for spherical multi-touch displays with simple webpages. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays*, PerDis '17, New York, NY, USA, 2017. Association for Computing Machinery.
- [16] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. Surround-screen projection-based virtual reality: The design and implementation of the cave. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, page 135–142, New York, NY, USA, 1993. Association for Computing Machinery.
- [17] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. The cave: Audio visual experience automatic virtual environment. *Commun. ACM*, 35(6):64–72, June 1992.
- [18] Dylan Fafard, Ian Stavness, Martin Dechant, Regan Mandryk, Qian Zhou, and Sidney Fels. Ftvr in vr: Evaluation of 3d perception with a simulated volumetric fish-tank virtual reality display. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 533:1–533:12, New York, NY, USA, 2019. ACM.
- [19] Dylan Brodie Fafard, Qian Zhou, Chris Chamberlain, Georg Hagemann, Sidney Fels, and Ian Stavness. Design and implementation of a multi-person fish-tank virtual reality display. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, VRST '18, New York, NY, USA, 2018. Association for Computing Machinery.

- .....
- [20] F. Ferreira, M. Cabral, O. Belloc, G. Miller, C. Kurashima, R. de Deus Lopes, I. Stavness, J. Anacleto, M. Zuffo, and S. Fels. Spheree: A 3d perspective-corrected interactive spherical scalable display. In *ACM SIGGRAPH 2014 Posters*, SIGGRAPH '14, pages 86:1–86:1, New York, NY, USA, 2014. ACM.
- [21] Robert S. Fish, Robert E. Kraut, and Barbara L. Chalfonte. The videowindow system in informal communication. In *Proceedings of the 1990 ACM Conference on Computer-supported Cooperative Work*, CSCW '90, pages 1–11, New York, NY, USA, 1990. ACM.
- [22] Susan R. Fussell, Robert E. Kraut, and Jane Siegel. Coordination of communication: Effects of shared visual context on collaborative work. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, CSCW '00, pages 21–30, New York, NY, USA, 2000. ACM.
- [23] Athanasios Gaitatzes, Georgios Papaioannou, Dimitrios Christopoulos, and Gjergji Zyba. Media productions for a dome display system. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 261–264, 2006.
- [24] William W. Gaver. The affordances of media spaces for collaboration. In *Proceedings of the 1992 ACM Conference on Computer-supported Cooperative Work*, CSCW '92, pages 17–24, New York, NY, USA, 1992. ACM.
- [25] William W. Gaver, Abigail Sellen, Christian Heath, and Paul Luff. One is not enough: Multiple views in a media space. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, CHI '93, pages 335–341, New York, NY, USA, 1993. ACM.
- [26] William W. Gaver, Gerda Smets, and Kees Overbeeke. A virtual window on media space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, pages 257–264, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

- [27] Glomal350. Glomal350 spherical display. <http://glomal.jp>. Accessed: 07-01-2017.
- [28] Markus Gross, Stephan Würmlin, Martin Naef, Edouard Lamboray, Christian Spagno, Andreas Kunz, Esther Koller-Meier, Tomas Svoboda, Luc Van Gool, Silke Lang, and et al. Blue-c: A spatially immersive display and 3d video portal for telepresence. *ACM Trans. Graph.*, 22(3):819–827, July 2003.
- [29] Markus Gross, Stephan Würmlin, Martin Naef, Edouard Lamboray, Christian Spagno, Andreas Kunz, Esther Koller-Meier, Tomas Svoboda, Luc Van Gool, Silke Lang, Kai Strehlke, Andrew Vande Moere, and Oliver Staadt. Blue-c: A spatially immersive display and 3d video portal for telepresence. *ACM Trans. Graph.*, 22(3):819–827, July 2003.
- [30] Tovi Grossman, Daniel Wigdor, and Ravin Balakrishnan. Multi-finger gestural interaction with 3d volumetric displays. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology*, UIST '04, pages 61–70, New York, NY, USA, 2004. ACM.
- [31] Jan Gugenheimer, Evgeny Stemasov, Harpreet Sareen, and Enrico Rukzio. Facedisplay: Towards asymmetric multi-user interaction for nomadic virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, 2018. Association for Computing Machinery.
- [32] Georg Hagemann, Qian Zhou, Ian Stavness, Oky Dicky Ardiansyah Prima, and Sidney S. Fels. Here's looking at you: A spherical ftvr display for realistic eye-contact. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ISS '18, page 357–362, New York, NY, USA, 2018. Association for Computing Machinery.
- [33] Zhenyi He, Karl Toby Rosenberg, and Ken Perlin. Exploring configuration of mixed reality spaces for communication. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2019.

- .....
- [34] Christian Heath and Paul Luff. The naturalistic experiment: Video and organizational interaction. *Organizational Research Methods*, page 1094428117747688, 2017.
- [35] David Holman and Roel Vertegaal. Organic user interfaces: Designing computers in any way, shape, or form. *Commun. ACM*, 51(6):48–55, June 2008.
- [36] Hikaru Ibayashi, Yuta Sugiura, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Masaaki Mochimaru, and Takeo Igarashi. Dollhouse vr: A multi-view, multi-user collaborative design workspace with vr technology. In *SIGGRAPH Asia 2015 Posters, SA '15*, pages 24:1–24:1, New York, NY, USA, 2015. ACM.
- [37] Craig Jin, Abhaya Parthy, and Andr  van Schaik. Measured and theoretical performance comparison of a broadband circular microphone array. In *Audio Engineering Society Conference: 31st International Conference: New Directions in High Resolution Audio*, Jun 2007.
- [38] Steven Johnson, Irene Rae, Bilge Mutlu, and Leila Takayama. Can you see me now?: How field of view affects collaboration in robotic telepresence. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, pages 2397–2406, New York, NY, USA, 2015. ACM.
- [39] Brennan Jones, Anna Witcraft, Scott Bateman, Carman Neustaedter, and Anthony Tang. Mechanics of camera work in mobile video collaboration. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, pages 957–966, New York, NY, USA, 2015. ACM.
- [40] Abhijit Karnik, Archie Henderson, Andrew Dean, Howard Pang, Thomas Campbell, Satoshi Sakurai, Guido Herrmann, Shahram Izadi, Yoshifumi Kitamura, and Sriram Subramanian. Vortex: Design and implementation of an interactive volumetric display. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems, CHI EA '11*, pages 2017–2022, New York, NY, USA, 2011. ACM.

- .....
- [41] Shunichi Kasahara and Jun Rekimoto. Jackin head: Immersive visual telepresence system with omnidirectional wearable camera for remote collaboration. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology, VRST '15*, page 217–225, New York, NY, USA, 2015. Association for Computing Machinery.
- [42] Stefanie Kettner, Christopher Madden, and Remo Ziegler. Direct rotational interaction with a spherical projection. In *Creativity & Cognition Symposium on Interaction: Systems, Practice and Theory*, pages 473–490, 2004.
- [43] Davis E. King. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758, 2009.
- [44] Gun A. Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. Shared-sphere: Mr collaboration through shared live panorama. In *SIGGRAPH Asia 2017 Emerging Technologies, SA '17*, New York, NY, USA, 2017. Association for Computing Machinery.
- [45] Gun A Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. A user study on mr remote collaboration using live 360 video. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 153–164. IEEE, 2018.
- [46] Zhengqing Li, Shio Miyafuji, Toshiki Sato, and Hideki Koike. Omnieyeball: Spherical display embedded with omnidirectional camera using dynamic spherical mapping. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology, UIST '16 Adjunct*, pages 193–194, New York, NY, USA, 2016. ACM.
- [47] Christian Licoppe, Paul K. Luff, Christian Heath, Hideaki Kuzuoka, Naomi Yamashita, and Sylvaine Tuncer. Showing objects: Holding and manipulating artefacts in video-mediated collaborative settings. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17*, pages 5295–5306, New York, NY, USA, 2017. ACM.

- .....
- [48] Yung-Ta Lin, Yi-Chi Liao, Shan-Yuan Teng, Yi-Ju Chung, Liwei Chan, and Bing-Yu Chen. Outside-in: Visualizing out-of-sight regions-of-interest in a 360° video using spatial picture-in-picture previews. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, page 255–265, New York, NY, USA, 2017. Association for Computing Machinery.
- [49] Paul Luff, Christian Heath, Hideaki Kuzuoka, Jon Hindmarsh, Keiichi Yamazaki, and Shinya Oyama. Fractured ecologies: Creating environments for collaboration. *Hum.-Comput. Interact.*, 18(1):51–84, June 2003.
- [50] Paul Luff, Naomi Yamashita, Hideaki Kuzuoka, and Christian Heath. Hands on hitchcock: Embodied reference to a moving scene. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 43–52, New York, NY, USA, 2011. ACM.
- [51] Tamotsu Machida. Geo-cosmos: World's first spherical display. In *ACM SIGGRAPH 2002 Conference Abstracts and Applications*, SIGGRAPH '02, pages 189–189, New York, NY, USA, 2002. ACM.
- [52] Gloria Mark. Extreme collaboration. *Commun. ACM*, 45(6):89–93, June 2002.
- [53] Akira Matsuda and Jun Rekimoto. Scalablebody: A telepresence robot supporting socially acceptable interactions and human augmentation through vertical actuation. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16 Adjunct, pages 103–105, New York, NY, USA, 2016. ACM.
- [54] Shio Miyafuji, Florian Perteneder, Toshiki Sato, Hideki Koike, and Gudrun Klinker. A bowl-shaped display for controlling remote vehicles. In *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*, VRCAI '19, New York, NY, USA, 2019. Association for Computing Machinery.
- [55] Shio Miyafuji, Masato Sugasaki, and Hideki Koike. Ballumiere: Real-time tracking and spherical projection for high-speed moving balls. In *Proceedings of the 2016*

.....

*ACM on Interactive Surfaces and Spaces*, ISS '16, pages 33–37, New York, NY, USA, 2016. ACM.

- [56] Alessandro Mulloni, Hartmut Seichter, Andreas Dünser, Patrick Baudisch, and Dieter Schmalstieg. 360° panoramic overviews for location-based services. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, page 2565–2568, New York, NY, USA, 2012. Association for Computing Machinery.
- [57] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. Collavr: Collaborative in-headset review for vr video. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pages 267–277, New York, NY, USA, 2017. ACM.
- [58] James Norris, Holger Schnädelbach, and Guoping Qiu. Camblend: An object focused collaboration tool. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, page 627–636, New York, NY, USA, 2012. Association for Computing Machinery.
- [59] James Norris, Holger M. Schnädelbach, and Paul K. Luff. Putting things in focus: Establishing co-orientation through video in context. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1329–1338, New York, NY, USA, 2013. ACM.
- [60] Oyewole Oyekoya, William Steptoe, and Anthony Steed. Sphereavatar: A situated display to represent a remote collaborator. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2551–2560, New York, NY, USA, 2012. ACM.
- [61] Ye Pan and Anthony Steed. Preserving gaze direction in teleconferencing using a camera array and a spherical display. In *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, 2012, pages 1–4. IEEE, 2012.

- [62] Ye Pan, William Steptoe, and Anthony Steed. Comparing flat and spherical displays in a trust scenario in avatar-mediated interaction. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 1397–1406. ACM, 2014.
- [63] Thammathip Piumsomboon, Arindam Day, Barrett Ens, Youngho Lee, Gun Lee, and Mark Billinghurst. Exploring enhancements for remote mixed reality collaboration. In *SIGGRAPH Asia 2017 Mobile Graphics and Interactive Applications*, SA '17, New York, NY, USA, 2017. Association for Computing Machinery.
- [64] Thammathip Piumsomboon, Gun A. Lee, Jonathon D. Hart, Barrett Ens, Robert W. Lindeman, Bruce H. Thomas, and Mark Billinghurst. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, 2018. Association for Computing Machinery.
- [65] Thammathip Piumsomboon, Gun A. Lee, Andrew Irlitti, Barrett Ens, Bruce H. Thomas, and Mark Billinghurst. On the shoulder of the giant: A multi-scale mixed reality collaboration with 360 video sharing and tangible interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 228:1–228:17, New York, NY, USA, 2019. ACM.
- [66] Jun Rekimoto, Keishiro Uragaki, and Kenjiro Yamada. Behind-the-mask: A face-through head-mounted display. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, AVI '18, New York, NY, USA, 2018. Association for Computing Machinery.
- [67] Ricoh Company, Ltd. Ricoh Theta S 360 camera. <https://theta360.com/en/>. Accessed: 07-01-2017.
- [68] Jeff Sauro and Joseph S. Dumas. Comparison of three one-question, post-task usability questionnaires. In *Proceedings of the SIGCHI Conference on Human Fac-*

- .....
- tors in Computing Systems*, CHI '09, page 1599–1608, New York, NY, USA, 2009. Association for Computing Machinery.
- [69] Toshiki Sato Hideki Koike Shio Miyafuji, Zhengqing Li. Qoom: An interactive omnidirectional ball display. In *Proceedings of the 30th annual ACM symposium on User interface software and technology*. ACM, 2017.
- [70] Nikita Soni, Sayli Bapat, Schuyler Gleaves, Alice Darrow, Carrie Schuman, Hannah Neff, Peter Chang, Kathryn A Stofer, and Lisa Anthony. Towards understanding interactions with multi-touch spherical displays. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2019.
- [71] Quentin Stafford-Fraser. On site: The life and times of the first web cam. *Communications of the ACM*, 44(7):25–26, 2001.
- [72] Kenji Tanaka, Junya Hayashi, Masahiko Inami, and Susumu Tachi. Twister: An immersive autostereoscopic display. In *IEEE Virtual Reality 2004*, pages 59–278. IEEE, 2004.
- [73] Kenji Tanaka, Junya Hayashi, Yutaka Kunita, Masahiko Inami, Taro Maeda, and Susumu Tachi. Twister: A media booth. In *ACM SIGGRAPH 2002 conference abstracts and applications*, pages 81–81, 2002.
- [74] Anthony Tang, Omid Fakourfar, Carman Neustaedter, and Scott Bateman. Collaboration with 360° videochat: Challenges and opportunities. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, DIS '17, pages 1327–1339, New York, NY, USA, 2017. ACM.
- [75] Theophilus Teo, Gun A. Lee, Mark Billingham, and Matt Adcock. 360drops: Mixed reality remote collaboration using 360 panoramas within the 3d scene\*. In *SIGGRAPH Asia 2019 Emerging Technologies*, SA '19, page 1–2, New York, NY, USA, 2019. Association for Computing Machinery.

- [76] Theophilus Teo, Gun A. Lee, Mark Billinghurst, and Matt Adcock. Investigating the use of different visual cues to improve social presence within a 360 mixed reality remote collaboration\*. In *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*, VRCAI '19, New York, NY, USA, 2019. Association for Computing Machinery.
- [77] Soichiro Toyohara, Toshiki Sato, and Hideki Koike. Balloonygen: Extended tabletop display embedded with balloon-like deformable spherical screen. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, VRST '18, New York, NY, USA, 2018. Association for Computing Machinery.
- [78] Peter Vorderer, Werner Wirth, Feliz R Gouveia, Frank Biocca, Timo Saari, Futz Jäncke, Saskia Böcking, Holger Schramm, Andre Gysbers, Tilo Hartmann, et al. Mec spatial presence questionnaire (mec-spq): Short documentation and instructions for application. *Report to the European community, project presence: MEC (IST-2001-37661)*, 3, 2004.
- [79] Julie R Williamson and Daniel Sundén. Enter the circle: Blending spherical displays and playful embedded interaction in public spaces. In *Proceedings of the 4th International Symposium on Pervasive Displays*, pages 195–200, 2015.
- [80] Julie R Williamson, Daniel Sundén, and Jay Bradley. Globalfestival: evaluating real world interaction on a spherical display. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1251–1261, 2015.
- [81] Julie R. Williamson, Daniel Sundén, and Keith Hamilton. The lay of the land: Techniques for displaying discrete and continuous content on a spherical display. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays*, PerDis '16, page 38–44, New York, NY, USA, 2016. Association for Computing Machinery.

- .....
- [82] Julie R. Williamson, John Williamson, Daniel Sundén, and Jay Bradley. Multi-player gaming on spherical displays. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '15*, page 355–358, New York, NY, USA, 2015. Association for Computing Machinery.
- [83] Bin Xu, Jason Ellis, and Thomas Erickson. Attention from afar: Simulating the gazes of remote participants in hybrid meetings. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, pages 101–113. ACM, 2017.
- [84] Wataru Yamada, Kazuhiro Yamada, Hiroyuki Manabe, and Daizo Ikeda. isphere: Self-luminous spherical drone display. pages 635–643, 10 2017.
- [85] Naomi Yamashita, Keiji Hirata, Shigemi Aoyagi, Hideaki Kuzuoka, and Yasunori Harada. Impact of seating positions on group video communication. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, CSCW '08*, pages 177–186, New York, NY, USA, 2008. ACM.
- [86] Naomi Yamashita, Katsuhiko Kaji, Hideaki Kuzuoka, and Keiji Hirata. Improving visibility of remote gestures in distributed tabletop collaboration. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, CSCW '11*, pages 95–104, New York, NY, USA, 2011. ACM.
- [87] Yuko Yanagawa, Kaori Ikematsu, Chihiro Suga, Mana Sasagawa, Yasushi Matoba, and Itiro Siio. Anamorphicons: An extended display utilizing a cylindrical mirror widget. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction, OZCHI '17*, page 457–461, New York, NY, USA, 2017. Association for Computing Machinery.
- [88] Tomohiro Yendo, Naoki Kawakami, and Susumu Tachi. Seelinder: the cylindrical lightfield display. In *ACM SIGGRAPH 2005 Emerging technologies*, pages 16–es. 2005.

- .....
- [89] Qian Zhou, Georg Hagemann, Sidney Fels, Dylan Fafard, Andrew Wagemakers, Chris Chamberlain, and Ian Stavness. Coglobe: A co-located multi-person fivr experience. In *ACM SIGGRAPH 2018 Emerging Technologies*, SIGGRAPH '18, pages 5:1–5:2, New York, NY, USA, 2018. ACM.
- [90] Ferdinand Rudolf Hendrikus Zijlstra. Efficiency in work behaviour: A design approach for modern tools. 1995.