博士論文

Optical Design of Geometrically Superimposing Layered Mid-air Images onto Physical Objects for Glasses-Free Mixed Reality Interactions

（裸眼複合現実感インタラクションのための多層空中像の幾何学的重畳を可能にする光学系の設計）

金　ハンヨウル

Abstract

Mixed reality (MR) concept has been proposed to extend the features of everyday objects with the visual images. The goal of MR interactions is to provide high sense of reality from a seamless connection between visual images and physical objects. Despite the progress of display technologies, however, in current MR interactions, it is still a big challenge to realize glass-free systems for multiple users that allows users to manipulate physical objects superimposed with visual images. Therefore, this thesis focuses on visual presentation in the 3D space and implementation of MR interactions in the real world space.

From the property of MR interactions, this thesis gives importance to manipulation for physical objects and the number of viewing zones of visual images as two criteria for designing MR interactions. Based on these two criteria, a 2×2 matrix for MR interactions is proposed with providing four kinds of MR interactions: MR showcases for a single user, MR showcases for surrounding users, MR interface for a single user, and MR interface for surrounding users. The goal of this thesis is to design and implement optical systems that form layered mid-air images in 3D space and superimpose them directly onto physical objects. The proposed optical designs are applied to three glasses-free MR interactions to complete the matrix. The main contribution of this thesis is to implement and validate the optical designs through three glasses-free MR interactions.

In Chapter 3, MRsionCase is proposed as an MR showcase for surrounding users, which can superimpose two layered mid-air images onto a physical exhibit. The use of DCRA solves the limitations of virtual images by forming a real mid-air image immediately next to the physical object. Its symmetric optical design can provide visual images with four separate viewing zones so that users can see the image from different directions, front, back, left, and right. In each direction, two layers of mid-

air images are formed and sandwich the physical object to express correct occlusion between the images and the physical object.

In Chapter 4, MARIO is implemented as an MR interface for a single user, which enables a direct interaction between physical objects and a mid-air image. From the combination of a real imaging optics and linear actuator, the optical design of MARIO can form a mid-air image in the range of 350 (W)×300 (D)×250 (H) mm. The position of mid-air image can be moved back and forth along the depth direction. An artificial shadow is cast below the mid-air image and provides a high sense of reality. Thus, users can see and interact with the mid-air character in this 3D space without any interruption.

In Chapter 5, HoVerTable is developed as MR interface for surrounding users, which combines horizontal image projection and dual-sided vertical images. Two layers of vertical mid-air images are formed on the tabletop surface and superimposed onto physical objects as MRsionCase. Each vertical image layer provides view-dependent images to two facing users. The optical design combines real imaging optics with diffusion control film to provide vertical and horizontal images with a compact design. With vertical mid-air images and horizontal image projection, HoVerTable extends the display area of conventional tabletop displays to vertical direction.

The optical designs proposed in this thesis can overcome the limitations in display area of current displays with superimposing mid-air images onto physical objects for glasses-free systems. Such glasses-free MR interactions will enable people to more intuitively access visual information and finally enrich our information usage in everyday life.

Acknowledgements

Foremost, I would like to appreciate my adviser Prof. Takeshi Naemura for supporting my research and providing me with the opportunity to complete my PhD dissertation at the University of Tokyo. It was a great luck for me to conduct PhD course under his guidance. His insightful advice had shed light on me when I confronted problems.

Beside my supervisor, I want to express my sincere thanks to the rest of committee: Assoc. Prof. Masashi Toyoda, Prof. Tohru Asami, Prof. Kiyoharu Aizawa, Assoc. Prof. Takefumi Ogawa, and Assoc. Prof. Kouta Minamizawa, for their thoughtful comments and encouraging suggestion on this thesis.

I also want to express my sincere thanks to the staffs in Naemura Laboratory. Dr. Naoya Koizumi helped me to conduct research with his knowledge and experiences on user experience design. Dr. Shogo Fukushima provided me with fresh perspectives on my research through discussion. Dr. Rei Kawakami gave me kind and appropriate comments when I wrote manuscripts for papers or articles.

My sincere thanks also go to Dr. Satoshi Maekawa for allowing me to use DCRAs with no limits. With these novel optical devices, he motivated me to start research on mixed reality interactions.

I want to express my appreciation to the members of Naemura Laboratory. Especially, Dr. Youngah Seong, Junghyun Kim, and Dr. Ryo Nakashima kept motivating and stimulating me with their dedication to research. Shun Nagao for helping the implementation of MRsionCase with his programming skills and ideas on sound presentation. Issei Takahashi and Hiroki Yamamoto showed their hard work during the implementation of MARIO. Yasuko Mori allowed me to use the Hiyoko character in MARIO. Boyoung Kim, Mikyung Seo, Takahiro Tsujii, and Takefumi Hiraki assisted the demonstrations of MRsionCase and MARIO. Takayuki Kai produced a great video clip for MARIO. I also thank Hajime Kajita for doing a lot of tedious chores. Yuri

Tsuchida always helped me when I purchased supplies using research budget with handling lots of administrative work.

I would like to express my gratitude to the staffs in Miraikan who operated the exhibition of MARIO. With their dedication to daily report and maintenance, I could successfully complete the public exhibition.

I specially appreciate the MEXT program that provided me with the opportunity of higher education without expenses. As a scholarship student, I could fully focus on my research with no concern about finance. I also thank to the Global Creative Leader (GCL) program with offering me the opportunity of research assistant. The research fund enabled me to conduct research with good equipment and to participate in international conferences.

Outside academic life, my friends in Japan and Korea gave me a lot of motivation. I particularly thank to Duseok Jeong, Jeeeun Kim, Jongsun Park, and Junia Roh. They gave me comfort with understanding and sharing my worries and anxieties in the doctoral course. Dr. Donghyun Kim, Euyhwan Park, and Gyeongcheol Choi encourged me with warm words.

Lastly, I reserve the final but deepest thanks for my family. Words cannot express my gratitude to my beloved wife, Hyoju Park. She was a good listener, a sharp critic, an inquisitive questioner, and a soulmate for me. Mother, father, mother-in-law, and father-in-law were the anchor of my life. With their trust and great support, I never lost my goal during the doctoral course.

This thesis is dedicated to all the people that I listed above with my sincere thanks.

12th June 2015

Hanyuool Kim

The best way to predict the future is to invent it.

Dr. Alan Kay

Contents

Abstract iii

Acknowledgements v

List of Tables xii

List of Figures xiii

1 Introduction **1**

1.1 Background 1

1.2 Mixed Reality Interactions in the Real World Space 3

1.2.1 Manipulation of physical objects 4

1.2.2 The number of viewing zone(s) 5

1.3 Goal 7

1.4 Contribution 7

1.5 Overview 9

2 Related Work **12**

2.1 Mixed Reality for Augmenting Physical Objects 12

2.1.1 Concept 12

2.1.2 Three consistencies: time, space, and illumination 14

2.1.3 Displays used in MR interactions for image superimposition . 15

2.2 Mid-air Images for Glasses-free MR Interactions 19

2.2.1 Real and virtual images 19

2.2.2 Optics for mid-air images 21
2.3 Examples of MR interactions 29
2.3.1 MR showcases for a single user 29
2.3.2 MR showcases for surrounding users 31
2.3.3 MR interfaces for a single user 32
2.3.4 MR interfaces for surrounding users 38
2.4 Summary 46

3 MRsionCase: Multi-directionally Viewable MR Showcase 48
3.1 Introduction 49
3.2 Proposal 49
3.2.1 Spatial consistency between physical object and information . 51
3.2.2 Multi-directional viewability 51
3.2.3 Targeted information presentation 51
3.2.4 Device-free usage 52
3.3 System Design 52
3.3.1 Multi-layered mid-air images 52
3.3.2 Virtual imaging system (IS1) 53
3.3.3 Combined imaging system (IS2) 54
3.3.4 Implementation 56
3.4 Results 59
3.4.1 Optical evaluation 59
3.4.2 User study 61
3.5 Discussion and Conclusion 66

4 MARIO: Mid-air Augmented Reality Interaction with Objects 69
4.1 Introduction 70
4.2 Proposal 72

4.2.1 Glasses-free mid-air imaging display in a 3D space 72
4.2.2 Enhanced sense of reality from a mid-air image 72
4.2.3 Interaction design for practical usage 73
4.3 System Design 73
4.3.1 Mid-air imaging system 74
4.3.2 Object detection 75
4.3.3 Shadow projection 76
4.3.4 Implementation 77
4.4 Results 79
4.4.1 Optical evaluation 79
4.4.2 User study 83
4.4.3 Findings from exhibitions 86
4.5 Discussion and Conclusion 87

5 HoVerTable: Combining Dual-sided Vertical Mid-air Images with a Horizontal Tabletop Display 92

5.1 Introduction 93
5.2 Proposal 94
5.2.1 Combined display area in 3D space 94
5.2.2 Selective information sharing 95
5.2.3 Augmentation of physical objects 95
5.3 System design 96
5.3.1 Vertical mid-air images on tabletop surface 96
5.3.2 Horizontal image projection 100
5.3.3 Implementation 101
5.4 Results 102
5.4.1 Optical evaluation 102
5.4.2 User study 109

5.5 Discussion and Conclusion 112

6 Conclusion 116

6.1 Summary 116

6.2 Future work 119

A Exhibitions 123

A.1 MRsionCase 123

A.1.1 Digital Contents Expo 2012 123

A.1.2 ACM SIGGRAPH Asia 2012 124

A.2 MARIO 125

A.2.1 12th Media Laboratory Exhibition at Miraikan 125

A.2.2 Innovative Technologies 2013 126

A.2.3 ACE 2013 Creative Showcase 128

A.2.4 Japan Expo 2014 129

A.2.5 CEATEC 2014 131

B Sound Presentation from MRsionCase 133

B.1 Auditory System Design 134

B.2 Results 134

C Calibration between Mid-air Images and Real Space 136

C.1 Coordinate System Setting 136

C.2 Calibration Method 137

References 139

List of Publications 145

List of Tables

1.1 Matrix of glasses-free MR interactions in this thesis. 6

2.1 Properties of mid-air imaging optics. 28

2.2 Comparison of ExFlasion, Virtual Showcase and MRsionCase. 33

2.3 Overview of the MR interactions proposed in this thesis. 47

3.1 Survey items. 64

5.1 Expected optical property of visual images produced using diffusive films (+ represents positive effects and − for negative ones.) 101

List of Figures

1.1 Examples of MR interactions. 2

1.2 The composition of MR interactions. 3

1.3 Two criteria for MR interaction in the real world space. 4

1.4 Examples of angle and the number of viewing zones. 5

1.5 Overview of this thesis. 11

2.1 Virtual continuum [2] 13

2.2 Displays used in mixed reality interactions [19] 16

2.3 Examples of head-attached displays. 17

2.4 Examples of hand-held displays. 17

2.5 Projectors in MR applications. 18

2.6 Imaging mechanism for real and virtual images. 20

2.7 The real image formed by concave mirror. 23

2.8 Optical configuration of AIRR [29]. 24

2.9 Dihedral corner reflector array (DCRA) by Maekawa et al. 25

2.10 Comparison between the mid-air images formed by a planar mirror and DCRA [30]. 26

2.11 AIP (Aerial Imaging Plate). 27

2.12 Magic vision application for museum exhibitions [36]. Visual images are superimposed onto the real diorama. 30

2.13 ExFloasion by Nakamura et al. [37] 31

2.14 Virtual showcase by Bimber et al. [38]. 32
2.15 HoloDesk [41]. 34
2.16 RePro3D by Yoshida et al. [42]. 35
2.17 Volume display by Smoot et al. [43]. 36
2.18 Fuwa-vision by Nii et al. [44]. 37
2.19 HaptoMirage by Ueda et al. [45]. 39
2.20 Volume slicing display by Markon et al. [46]. 40
2.21 MirageTable by Benko et al. [47]. 41
2.22 Deskrama by Nagakura and Oishi [48]. Vertical images are shown from a upright display and horizontal images are displayed on the desktop surface. 41
2.23 Tablescape Plus by Kakehi et al. [49]. Characters are projected to tiny upright screens on the table. Horizontal images are displayed from the tabletop surface. 42
2.24 Lumisight Table by Kakehi et al. [50]. 43
2.25 FloasionTable by Wada et al. [52]. 44
2.26 fVisiOn by Yoshida et al. [53]. 45

3.1 Illustration of MRsionCase. (a) Viewers can see exhibit and superimposed images without wearing a special device. (b)–(e) Views from front, left, back, and right. 50
3.2 Virtual imaging system with HSMs (IS1). 53
3.3 Combined imaging system using DCRA (IS2). 55
3.4 Implementation of MRsionCase. 58
3.5 Resulting images formed by IS1. 59
3.6 A resulting mid-air image formed by IS2. In both images, arrow is floating immediately next to hands. 60
3.7 Measurement of viewing angle. 61

3.8 Demonstration of MRsionCase. 62

3.9 Spider figure and superimposed images. 63

3.10 User survey results. 65

4.1 The MARIO system. Users can directly interact with physical objects and mid-air images. A virtual character (Hiyoko, a chick) is displayed in mid-air on wooden blocks. A coordinate system is defined as x for width, y for height, and z for depth. 70

4.2 Overview of MARIO. 74

4.3 Configuration of display and AIP in mid-air imaging display. 75

4.4 Shadow projection scheme. 76

4.5 Implemented MARIO system. 78

4.6 Stereo view of mid-air image. Each figure shows a view 5° to the left or right of the center line. The stereoscopic view makes the resulting image appear to be floating in mid-air. 80

4.7 Mid-air images formed at different depths. The change in the depth direction is apparent by comparing the focuses of the images. 81

4.8 Impression given by projected shadow. 81

4.9 Viewing angle in horizontal and vertical direction versus imaging position. 82

4.10 Hiyoko jumps on blocks at different depths along a parabolic path. The character moves energetically together with animation effects. (This figure was reconstructed by overlaying frames extracted from a video clip.) 84

4.11 The MARIO system at public exhibitions. In Miraikan, kids played with the Hiyoko character by manipulating wooden blocks and having the image hop above them. 84

4.12 Accumulative percentage of movements by time interval between movements. 86

4.13 Example of MARIO being used during an exhibition. Users played with the character in unique ways that we had not anticipated in our interaction scenario. 88

5.1 The proposed “HoVerTable” system 94
5.2 Steps in optical design of HoVerTable 97
5.3 Interaction distance (I) and minimum AIP size (L_{min}) from Eqs. (5.1) and (5.2) according to viewing height (T). 98
5.4 Implemented HoVerTable system 103
5.5 Vertical mid-air image floating on physical card placed on table surface. Images were taken from left, front, and right with three different Ts. . 104
5.6 Blur in mid-air images by diffusive films. These images were taken with same camera settings (1/15 sec, F6.3). 105
5.7 Measurement of image luminance on HoVerTable 106
5.8 Luminance of mid-air image and horizontal image projection on Lumisty and semi-transparent film 106
5.9 Effect of Lumisty film on horizontal image projection 107
5.10 Dual-faced mid-air images can provide different visual images to two facing users. 108
5.11 Settings of vertical and horizontal images for user study 110
5.12 Example images used in user study. Images were taken with same camera parameter settings (1/20 sec, F8). 111
5.13 Results of user study 113
5.14 Example of application for entertainment 114

A.1 Exhibition of MRsionCase at Digital Contents Expo 2012 (Tokyo, Japan).124
A.2 Exhibition of MARIO at 12th Media Laboratory Exhibition in Miraikan (Tokyo, Japan). 126

A.3 Exhibition of MARIO at Innovative Technologies 2013 (Tokyo, Japan). 127
A.4 Exhibition of MARIO at ACE 2013 (Twente, The Netherlands). . . . 128
A.5 Exhibition of MARIO at Japan Expo 2014 (Paris, France). 130
A.6 Exhibition of MARIO at CEATEC 2014 (Chiba, Japan). 132

B.1 Vertical sound separation system. Two sets of HDLs are installed with different reflecting angles. 134
B.2 Sound pressure level distribution by height. 135

C.1 Calibration of coordinate systems in MARIO system. 137

Chapter 1

Introduction

1.1 Background

The advances of computers and ubiquitous access to the Internet have changed our usage of information. Especially, computers have changed their shapes and are widespread in everyday life. Smartphones such as iPhone [1] have allowed people to access the Internet and to obtain information anywhere and anytime. Novel display technologies also supported this change of computing experiences with presenting visual images in various ways. People can see volumetric images as 3D objects with stereoscopic displays. Wearable displays enable users to immerse in virtual environment that is constructed with visual images. In this way, visual information has been closely connected with the real world.

From this background, mixed reality (MR) concept has been proposed to extend the features of everyday objects with the visual images [2]. Figure 1.1 shows examples of MR interactions. In MR interactions, visual images are displayed with physical objects and provide additional information. Practical MR interactions have been proposed and widely used for various purposes including navigation, instruction, and entertainment. With the aid of visual information, for example, people can easily navigate their way and play a game with virtual characters in the real world.

(a) A navigation application (iOnRoad) [3]. (b) A game application (AR Cards by Nintendo) [4]

Figure 1.1: Examples of MR interactions.

The goal of MR interactions is to provide high sense of reality from a seamless connection between visual images and physical objects. To provide high sense of reality, three consistencies in time, space, and illumination are required to be implemented [10]. Among three consistencies, this thesis specifically focuses on geometrical consistency that visual images are superimposed onto physical objects and displayed in the same space.

Despite the progress of display technologies, however, in current MR interactions, it is still a big challenge to realize glass-free systems for multiple users that allows users to manipulate physical objects superimposed with visual images. Since the positions of visual images are confined inside the physical displays or surfaces, the images cannot exist in 3D space beyond the display. Therefore, this thesis focuses on visual presentation in the 3D space and implementation of MR interaction in the real world space.

Head-attached displays or 3D displays can provide visual images as if they exist in 3D space, but users need to wear additional glasses to see the images. The author believes that such limitations in imaging positions and usage of 3D glasses reduce the intuitiveness and sense of reality in MR interactions. To overcome these problems, this thesis focuses on visual presentations in the 3D space and implementation of glasses-free MR interactions.

1.2 Mixed Reality Interactions in the Real World Space

Figure 1.2 shows the essence of MR interactions. MR systems present visual images along with physical objects to connect the real world and the computers' world. The details on MR (e.g. goals, important issues) will be introduced in Section 2.1. MR interactions have arisen from this combination between visual images and physical objects.

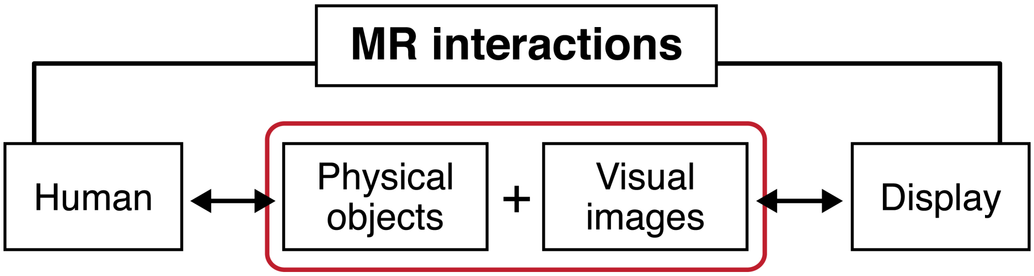


Figure 1.2: The composition of MR interactions.

In MR systems, people can interact with physical objects with various modalities. Among the modalities, this thesis specifically focuses on the manipulation of physical objects such as grab, move, and touch. This basic manipulation provides users with an intuitive and familiar way to access MR system so that the manipulation of physical objects is a critical factors in designing MR interactions.

Visual images presented in MR systems can be evaluated with several criteria such as image quality, visibility, and viewing angles. However, this thesis puts emphasis on the use of mid-air images without the need for special glasses. This enables users to intuitively see the visual images in MR systems. For mid-air images, moreover, the number of viewing zones is an important factor than others since it determines the physical range of MR interactions.

Based on these reasons, this thesis gives importance to manipulation for physical objects and the number viewing zones of visual images as two criteria for designing an MR interaction.

Figure 1.3 shows the two criteria, manipulation of physical objects and viewing zone(s) of visual images, added to the essence of MR interactions. In this thesis, MR interactions are implemented with a base on this two criteria.

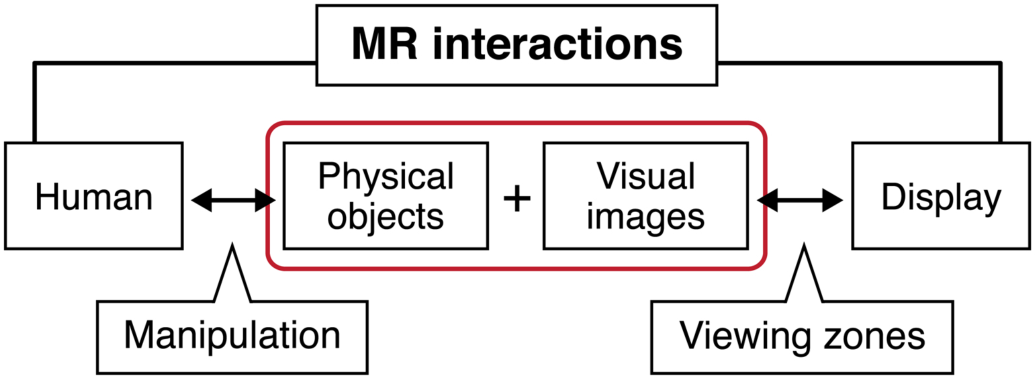


Figure 1.3: Two criteria for MR interaction in the real world space.

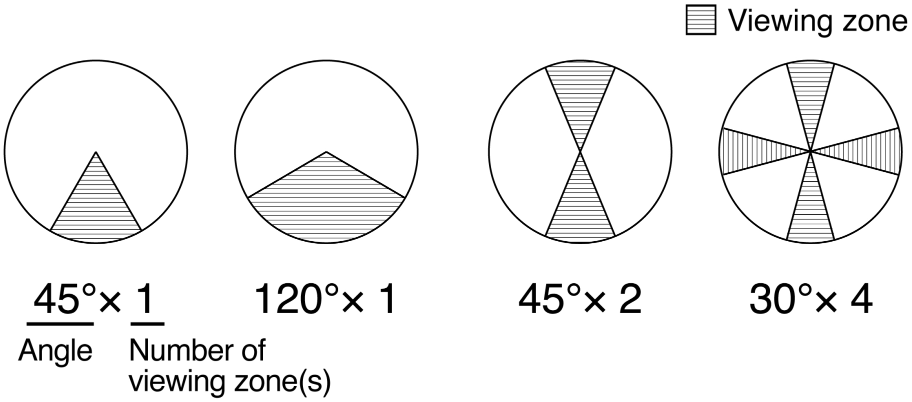
1.2.1 Manipulation of physical objects

Physical objects have provided users with an intuitive way to interact with visual images in the experience of MR applications. Users can see, point, touch, and even move the objects with MR interactions. From users' interaction with physical objects, this thesis suggests that *how users manipulate physical objects* is an important criteria for MR interaction designs.

Manipulation of physical objects are classified in two categories according to the accompaniment of physical contact: indirect, and direct manipulation. Indirect manipulation represents the situations that users interact with physical objects without physical contacts. For example, users see and/or point to physical objects without touching the objects. On the other hand, direct manipulation is for the situations that users establish a physical contact with physical objects. Moving and touching the objects can be considered as direct manipulations.

1.2.2 The number of viewing zone(s)

From the sense of providing users with visual images, MR interactions can be thought as a display. Therefore, viewing zone of visual images provides a important factor to design and evaluate MR interactions. Especially, this thesis focuses on the number of viewing zone(s) of visual images in MR interactions. This is because the number of viewing zone(s) usually determines (or sometimes limits) the number of users in MR interactions.



(a) Single viewing zone.

(b) Multiple viewing zones.

Figure 1.4: Examples of angle and the number of viewing zones.

Figure 1.4 illustrates examples of MR interactions providing different viewing zones. A viewing zone of a visual image is defined as a maximum range where a user can see the image with a continuous movement of viewing positions. Gray regions represent a viewing zone of visual images of MR interactions when visual images are displayed from the center of each circle. A viewing zone of visual images can be expressed with the angles and the number. As shown in the figure, viewing zones are classified into two categories according to their number: single and multiple. In this thesis, the number of viewing zone(s) provides another criterion for implementation of MR interactions.

Based on these two criteria, a 2×2 matrix for MR interaction is proposed as Table 1.1.

Table 1.1: Matrix of glasses-free MR interactions in this thesis.

	Single viewing zone	Multiple viewing zones
Indirect manipulation	MR showcases for a single user	MR showcases for surrounding users
Direct manipulation	MR interfaces for a single user	MR interfaces for surrounding users

As for physical objects, indirect manipulation enables users to see and point the physical object with the MR interaction. Such interactions can be realized by MR showcases which display physical exhibits. On the other hand, direct manipulation of physical objects can promote user to touch and move physical objects in the MR interaction. These interactions propose MR interfaces that users can handle visual information with physical objects.

The number of viewing zones determines the number of users of MR interactions with controlling viewing range of visual images. For example, MR interactions with a single viewing zone are more suitable for a single user's usage with providing a focused viewing range. Multiple viewing zones, on the other hands, can users to surround the MR interaction and see the visual images from multiple directions.

Therefore, in this thesis, MR interactions are classified to four categories which are MR showcases for a single user, MR showcases for surrounding users, MR interface for a single user, and MR interface for surrounding users.

MR interactions using head-attached displays or portable displays can usually realize both direct manipulation and multiple viewing zones. However, for glasses-free MR interactions, proposing each MR interaction which satisfies these criteria is a big challenge. This thesis tackles this problem with proposing optical designs for MR interactions and complete this table.

1.3 Goal

The goal of this thesis is to design and implement optical systems that form layered mid-air images in 3D space and superimpose them directly onto physical objects. This thesis completes Table 1.1 with three glasses-free MR interactions by applying the optical designs. Resulting images formed by the optical designs are evaluated through optical experiments and user study.

1.4 Contribution

The main contribution of this thesis is to implement and validate the optical designs through three glasses-free MR interactions. The mid-air images formed by the optical designs can be seen from naked eyes without wearing special glasses. From the property of MR interaction, two criteria has been proposed with manipulation of physical objects and number of viewing zone(s) that need to be considered for designing MR interactions. Based on the criteria, this thesis provided a 2×2 matrix of MR interactions and complete the matrix with proposing glasses-free MR interactions as applications of the optical designs. Three MR interactions proposed in this thesis are introduced as follows.

MRsionCase [5, 6] is an MR showcase which can superimpose two layered mid-air images onto a physical exhibit. The usage of DCRA solves the limitations of virtual images by forming a real mid-air image immediately next to the physical object. Its symmetric optical design can provide visual images with four separate viewing zones so that users can see the image from different directions, front, back, left, and right. In each direction, two layers of mid-air images are formed and sandwich the physical object to express correct occlusion between the images and the physical object. In MRsionCase, multiple users can surround the showcase and view the exhibit and superimposed images with walking around the showcase without wearing any special glasses.

MARIO [7, 8] is an MR interface for a single user which enables a direct interaction between physical objects and a mid-air image. From the combination of a real imaging optics and linear actuator, the optical design of MARIO can form a mid-air image in the range of 350 (W)×300 (D)×250 (H) mm. The position of mid-air image can be moved back and forth along the depth direction. For interaction, users can stack wooden blocks and make a physical terrain. Then, a virtual character appear in mid-air and jump around the blocks. An artificial shadow is cast below the mid-air character and provides a high sense of reality. Thus, users can see and interact with the mid-air character in this 3D space without any interruption.

HoVerTable [9] is an MR interface with a form of tabletop display which combines horizontal image projection and dual-sided vertical images. Two layers of vertical mid-air images are formed on the tabletop surface and superimposed onto physical objects as in MRsionCase. Each vertical image layer has two sides so that two facing users can see view-dependent images from the front and rear side at the same time. In the optical design, real imaging optics and diffusion control film are combined to provide vertical and horizontal images with a compact design and to reduce interferences between lights entering from different paths. Since the horizontal tabletop surface can display projected images, the tabletop surface of HoVerTable can be used as a conventional tabletop display. In HoVerTable, vertical and horizontal images are linked to enhance visual effects from the visual images. For example, when a standing character is display from a vertical image, the ground and its shadow are displayed from the horizontal surface as in MARIO. With vertical mid-air images and horizontal image projection, HoVerTable extends the display area of conventional tabletop displays to vertical direction while removing physical displays or surfaces on the tabletop surface.

1.5 Overview

This thesis is constructed as the following.

Chapter 1 Introduction

Chapter 2 Related Work

Chapter 3 MRsionCase: Multi-directionally Viewable MR Showcase

Chapter 4 MARIO: Mid-air Augmented Reality Interaction with Objects

Chapter 5 HoVerTable: Combining Dual-sided Vertical Mid-air Images with a Horizontal Tabletop Display

Chapter 6 Conclusion

Appendix A Exhibition

Appendix B Sound Presentation from MRsionCase

Appendix C Calibration between Mid-air Images and Real Space

In Chapter 2, previous work related to this thesis will be reviewed. MR concept for augmentation of physical objects will be introduced as a stating point of the main discussion. Typical displays used in MR interactions will also covered and the limitations of display area in current displays will be clarified. Then, the thesis will focus the MR interaction in the real world using mid-air images instead of physical displays. In order to start discussion on visual images beyond physical displays, imaging optics which can form mid-air images will be introduced. Then, manipulation of physical objects and the number of viewing zones are introduced as two main factors when the author construct an MR interaction in the real world. Finally, previous studies on MR interactions, which have influences on the proposed MR interactions in this thesis, will be picked up and compared with the proposed interactions.

In the following Chapters 3, 4, and 5, the details of optical designs in order to superimpose visual images onto physical objects will be introduced. Based on these optical designs, a series of glasses-free MR interactions will be implemented.

In Chapter 3, MRsionCase, an MR showcase for surrounding users, will be proposed. In MRsionCase, an optical design is devised to form two layers of mid-air images which sandwich a physical object from the front and the rear side. For the optical design, core parameters such as the size of mid-air images and exhibits will be listed and determined by considering the limitations in the implementation. Imaging positions and viewing angles of the mid-air images will be demonstrated to validate the implementation. For system evaluation, the user study result from public exhibitions will be introduced.

In Chapter 4, an MR interface for a single user, MARIO, which enables a direct interaction between physical objects and a mid-air image, will be proposed. In MARIO, an optical design is implemented to form a layer of mid-air image in 3D space, which can move back and forth with changing its depth. The details of system configuration will be introduced that consists of mid-air imaging optics, object detection, and shadow projection. Implementation results and feedbacks from in-situ demonstrations will be reported to evaluate the system.

In Chapter 5, an MR interface for surrounding users, HoVerTable, which combines vertical mid-air images and horizontal image projection, will be proposed. The optical design of HoVerTable consists of a plate-shaped mid-air imaging optics and a diffusion control film. Based on the optical parameters such as size, position and viewing range of mid-air images, the details of optical design will be discussed. The resulting visual images will be demonstrated to confirm the effectiveness of implementation. For system evaluation, a user study on comparison of vertical and horizontal images will be performed and its results will be discussed.

In Chapter 6, the discussion on superimposing mid-air images onto real object, the main topic of this thesis, will be concluded and future work and prospect are summarized with finishing the thesis.

The entire structure of this thesis is illustrated in Figure 1.5.

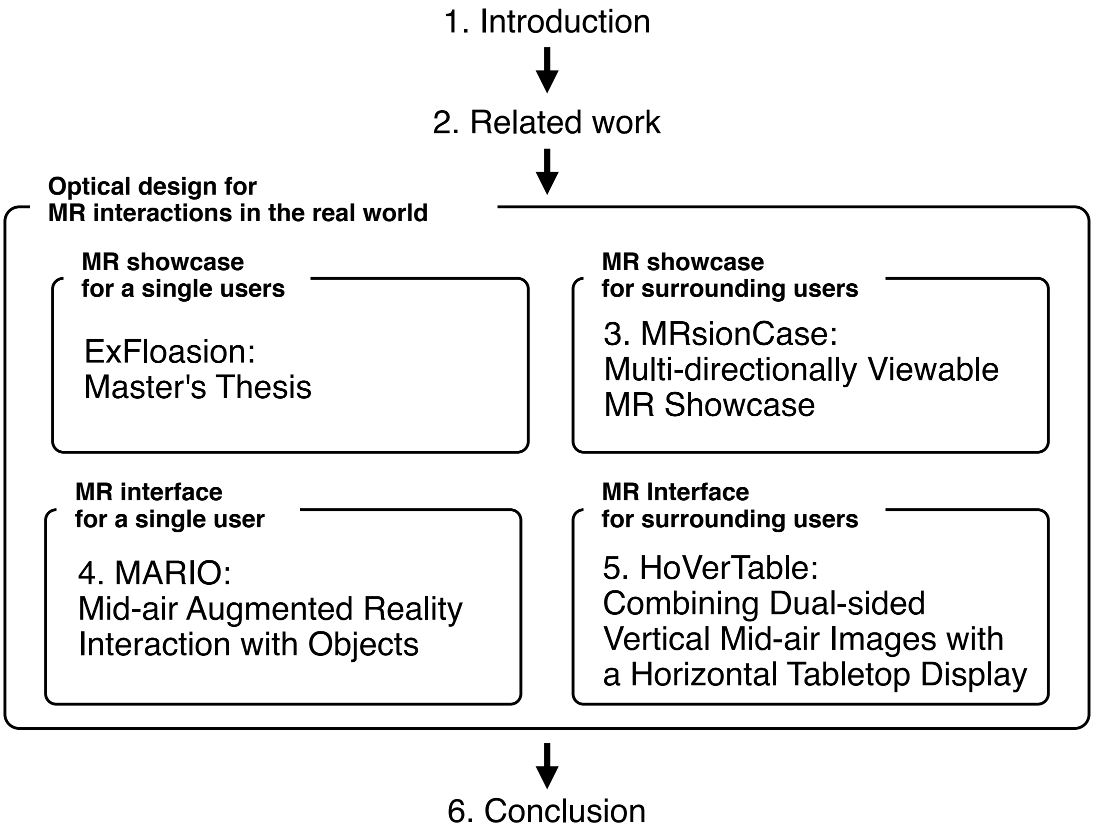


Figure 1.5: Overview of this thesis.

Chapter 2

Related Work

In this chapter, MR concept and terms used in the field of MR will be introduced. After reviewing the displays used in current MR interactions, the author will discuss the limitations of the MR interactions in terms of display area. Then, imaging optics will be surveyed that can be form visual images in 3D space. Finally, previously proposed studies on MR interactions will be reviewed as reference for system implementation.

2.1 Mixed Reality for Augmenting Physical Objects

2.1.1 Concept

From the effort to extend the features of everyday objects by connecting with visual images, Milgram and Kishino have proposed mixed reality (MR) concept [2]. In their work, a virtual continuum is defined as a continuous space where the computer world and the physical world are merged to each other as illustrated in Figure 2.1.

In the virtual continuum, there are two extrema with “pure” real and virtual environment which represent completely physical and computer world, respectively. MR is defined as anywhere between the two extrema except the both ends. This

Figure 2.1: Virtual continuum [2]

means that MR interactions are placed in hybrid environment which includes real and virtual objects at the same time.

Milgram and Kishino have also defined real and virtual objects as follows.

> ***Real objects*** *are any objects that have an actual objective existence.*

> ***Virtual objects*** *are objects that exist in essence or effect, but not formally or actually.*

Based on this definition, real objects are ordinary objects which have physical and substantial shapes, and therefore they obey physical rules. In this thesis, the term "physical objects" is used as the same meaning of the real objects defined in Milgram's work.

On the other hands, virtual objects are physically unsubstantial things such as information and concept. In this thesis, visual images which are rendered by computer graphics are mainly considered as virtual objects.

From the definitions of MR, real and virtual objects, therefore, physical objects and visual images co-exist in MR interactions. Visual images are displayed along with physical objects as an overlay onto the objects. In this way, the features of physical objects can be extended with the superimposed visual images that provides additional information. This thesis proposes MR interactions with the aim of augmenting and extending their functions with image superimposition onto physical objects.

2.1.2 Three consistencies: time, space, and illumination

When visual images are superimposed onto physical objects or the real world as a background in MR interactions, three consistencies in time, space, and illumination are required in order to provide high level of sense of reality [10].

Temporal consistency is to accord the movements of virtual objects with the time flow in the real world. In order to high level of temporal consistency, end-to-end time delay should be minimized [11]. For example, high computing power in graphics can reduce the rendering time of visual images and optical see-through displays may decrease time delay in displaying the real world background without no additional processing time.

Geometrical consistency is to coordinate the positions of virtual and real objects in the same coordinate system, or a display frame. For example, when a virtual object is overlaid on the real world inside a display, the information on the virtual object and the real world (e.g. size, position, shapes, etc.) is necessary. Virtual objects can be considered as a 3D model with a shape and position in the display coordinate. With the aid of sensing technologies, the real world is also mapped to the same coordinate. With a transformation from virtual and real space to the display coordinate, virtual and real objects can be registered in the same space and registered into the real world and they can be displayed together in the same space with expressing correct occlusions [12, 13] and collision [14].

Illumination consistency is to accord the lighting conditions of virtual objects with the actual illumination condition in the real world. Such accordance in illumination conditions between real and virtual objects can realize more natural and seamless connection when the objects construct a scene. Especially, the estimation of light sources in the real world has been studied to understand the lighting condition and to accord the illumination between virtual and real objects [15, 16]. Based on the estimated light sources, optical parameters such as shades and shadows have been implemented in virtual objects to express correct illumination expressions as real objects [17, 18].

As above, three consistencies in time, space, and illuminations have been proposed to provide high sense of reality and to realize the seamless connection between virtual and real objects. However, in conventional MR studies, there are limitations in display area of visual images. Thus, these consistencies have been discussed in MR interactions inside the frame of physical displays. Visual images can be displayed and superimposed on the real objects, but the images exist in the display coordinates not the real world coordinates.

In this thesis, the author focus the real world space out of the display surface as display area of visual images. Visual images are superimposed onto physical objects in the real space instead of inside physical displays. Same level of temporal consistency, as achieved in optical see-through MR interactions, is provided by eliminating rendering process of real and physical objects. Illumination consistency between virtual objects and the real world is provided with artificial shadows formed by the virtual objects.

2.1.3 Displays used in MR interactions for image superimposition

In order to superimpose visual images onto physical objects, various types of visual displays have been used in MR interactions. Bimber and Raskar have classified visual displays according to the position of displays with respect to the user and physical objects as shown in Figure 2.2 [19].

From a user's eye, head-attached displays are placed at the closest position since the displays are worn in the user's head. Figure 2.3 shows some examples of head-attached displays including retinal displays and head-mounted displays (HMDs). Retinal displays show visual images by directly projecting lights to the user's retinal

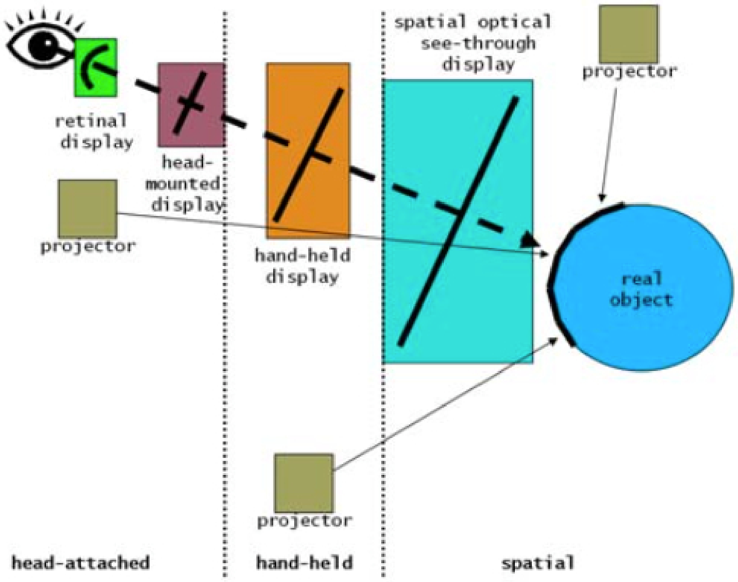
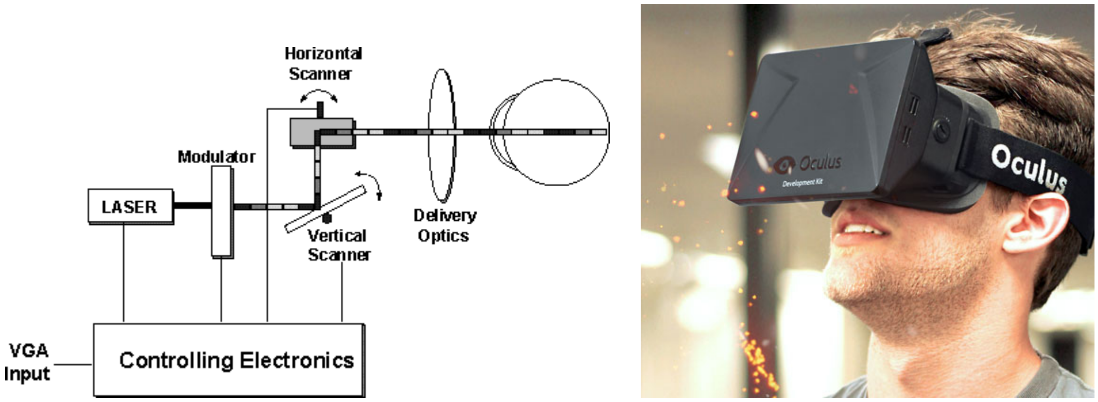


Figure 2.2: Displays used in mixed reality interactions [19]

surface. HMDs can superimpose visual images on the real world in front of the user's sight. These head-attached displays can keep the relative distance with the user's eye so that visual images can be superimposed during his/her movement. For head-attached displays, however, users need to wear the displays in order to see the visual images, and thus preparation is necessary before experiencing the MR interaction. Such need to wear displays to see visual images might make users feel uncomfortable and cumbersome during the experience [22]. This does not accord with the author's philosophy on MR interactions that walk-up-and-use is critical for MR interactions.

As portable devices advance, hand-held displays (HHDs) including mobile phones, portable game players, and personal digital assistants (PDAs), have been widely used in MR interactions, especially in applications for multiple users. These displays are usually located within arm reach from a user. Examples of hand-held displays are shown in Figure 2.4. Due to portability and high degree of freedom on positioning in 3D space, HHDs can display visual images with various perspectives. Most recent HHDs including mobile phones and tablet devices are equipped with cameras so that video see-through MR interactions are easily implemented with HHDs. Despite these



(a) The configuration of retinal display [20]

(b) A head-mounted display (Occulus VR [21]).

Figure 2.3: Examples of head-attached displays.

(a) A portable game player [23].

(b) A personal digital assistant (PDA) [24]

Figure 2.4: Examples of hand-held displays.

advantages of HHDs in MR interactions, visual images on HHDs are displayed only from the displays. In other words, HHDs cannot display visual images out of the display area. Due to the limitations of displays, visual images cannot be placed in the 3D space along with physical objects, and thus complete geometrical consistency cannot be provided with HHDs.

Projectors have been used to display visual images by projecting lights onto physical surfaces such as screens and walls. With the variety of shapes, projectors can be placed at different locations for head-attached, hand-held, or fixed usage. Such various usage of projectors enables flexible image superimposition in MR interactions as

(a) Image projection on a physical object [25]. (b) User interaction with hand-held projectors [26]

Figure 2.5: Projectors in MR applications.

shown in Figure 2.5. For example, the appearance of physical objects can be changed by the visual images projected on the surface [25]. Users can interact with visual images by moving hand-held projectors [26]. However, in order to provide visual images by projectors, a screen is necessary to diffuse the projected light so that the images cannot be superimposed onto physical objects due to the need for physical screens.

Spatial optical see-through displays, including transparent displays and mirroring optics, can superimpose visual images onto physical objects using transparency or virtual images formed by mirror reflection. Most MR interactions that use optical see-through displays can provide high level of time consistency, since physical objects and the real world background is seen through transparent windows and only visual images are rendered by computers. Moreover, optical see-through displays are usually installed onto the MR interactions so that users do not need to wear additional displays to view visual images. Thus, users can walk up and experience the MR interactions. However, visual images provide limited viewing region so that viewing range should be carefully designed with considering the interaction area and users' viewing positions.

Reminding that the main goal of this thesis is to propose optical designs for glasses-free MR interactions, this thesis adopts the advantages of optical see-through displays but also extend the display area into the 3D space in the real world. For this purpose, visual images need to be formed in mid-air space beyond physical displays using optical phenomena.

2.2 Mid-air Images for Glasses-free MR Interactions

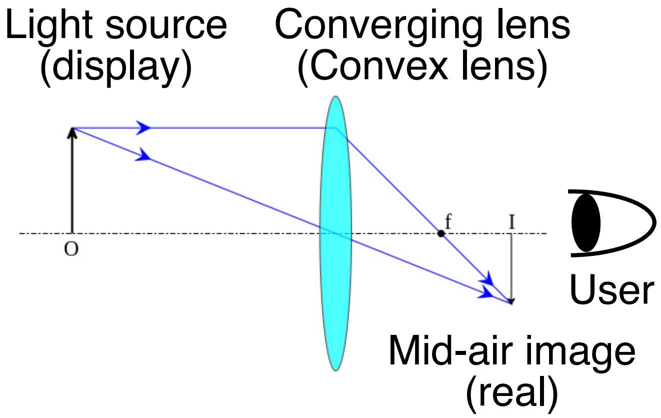
In this thesis, mid-air images stand for the visual images which are formed by optical phenomena such as reflection and refraction. Since mid-air images do not need physical displays or surfaces to show visual images, they can be placed along with physical objects. Moreover, users are not required to wear additional displays to see mid-air images. Thus, mid-air images can be used to implement glasses-free MR interactions.

2.2.1 Real and virtual images

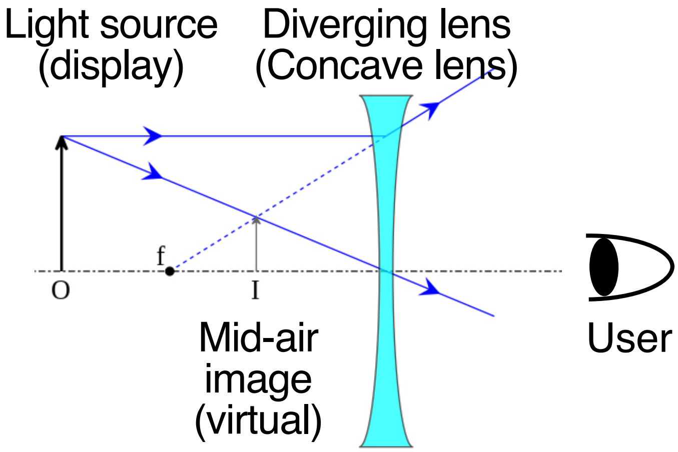
Geometrical optics can form visual images in mid-air by reflection or refraction of incident lights, which enter the optics. Based on the imaging mechanisms, mid-air images are classified into two kinds: Real and virtual images. In geometrical optics, real and virtual images are defined as the following descriptions.

> ***Real images*** *are (the images) formed by the converging lenses in that they can be made visible on a screen [27].*
>
> ***Virtual images*** *are (the images) produced with converging lenses when the object is placed between the focal point and the lens and with diverging lenses when the object is in any position [27].*

Based on these definitions, imaging mechanism for real and virtual images can be illustrated with light rays as Figure 2.6.



(a) A real image formed by a convex lens.



(b) A virtual images formed by a concave lens.

Figure 2.6: Imaging mechanism for real and virtual images.

For real images, mid-air images are formed in front of imaging optics. It is because real images are formed by converging lights and the converging point is placed in front of the imaging optics after passing through it. When a user see the mid-air image in front of optics, the user can access the mid-air image without any interruption by the imaging optics. Therefore, real images can be used for MR interactions that users need to touch and manipulate the visual images and physical objects in front of the imaging optics.

On the other hand, virtual images are always placed behind imaging optics since the origin of diverging light rays exist inside the imaging optics surface. In this propery of virtual images, when virtual images are used in MR interactions, the images are always seen through the imaging optics. Since the position of virtual images are limited to behind the imaging optics, users are hindered from approaching the mid-air images and the physical objects, which are placed behind the imaging optics. Thus, virtual images are often used in see-through MR applications such as showcases and viewing windows.

2.2.2 Optics for mid-air images

In this subsection, imaging optics which can form virtual and real images will be introduced. In specific, the optical property of half-silvered mirrors for virtual images, and curved mirrors and lenses, and retro-reflective optics for real images will be reviewed with two criteria, optical distortions and the kind of resulting images.

Half-silvered mirrors

As the simplest way to form virtual images, planar mirrors have been widely used. In Pepper's Ghost [28], for example, a plate of glass reflects the incident lights and forms a mirror image as semi-transparent mirrors using the difference of illuminance. Half-silvered mirrors (HSMs) are semi-transparent mirrors that control the transparency and reflectivity with optical coating.

This semi-transparency enables HSMs to overlay visual images onto physical objects and the real world background. When HSMs reflect incident lights, they form virtual images as ordinary planar mirrors. On the other hands, when users see through the HSMs, the scenery behind HSMs can be seen as users look through transparent glasses. The duality of transparency and reflectivity is the core feature of HSMs.

The positions of resulting image and the light source are symmetric to the HSM surface. The size of resulting image is identical to the size of the light source.

The imaging position has a linear relationship with the distance between the HSM and the light source so that the imaging position can be easily and precisely calculated before the implementation of optical design. Based on this linear relation, imaging position can be simply changed in mid-air. Moreover, the mirror surface is planar and the imaging process involves only linear reflections so that resulting images are free from optical distortion unlike curved mirrors.

Despite such advantages, HSMs have stringent limitations on the imaging position: they cannot form mid-air images in front of their surface since the images are virtual ones and can only be placed behind the HSM. Thus, HSMs are suitable for see-through applications usage only.

From this optical property of HSMs, in this thesis, HSMs are used as a distortion-free imaging optics when virtual images are necessary for a see-through usage.

Concave mirrors and convex lenses

Concave mirrors (CCMs) and convex lenses (CVLs) can overcome the limitations in imaging positions of virtual images with forming real images. By converging incident lights, these optics can form real images in mid-air and place them in front of the optics. Thus, the images formed by CCMs and CVLs can be superimposed onto physical objects placed in front of the optics. This allows people to directly access the visual images. However, unlike HSMs, the images formed by CCMs and CVLs have a nonlinear relation in the distance between the light source and the imaging optics. Figure 2.7 shows the geometrical relation between a light source, a CCM, and the resulting image. This relation can be expressed in the lens equation as the following equation.

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \tag{2.1}$$

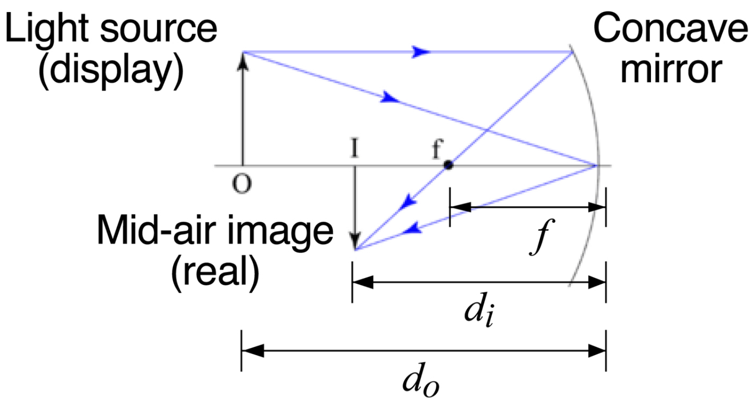


Figure 2.7: The real image formed by concave mirror.

The notion d_o is the distance from the light source (display) to the imaging optics, d_i is the distance from the imaging optics to the resulting image, and f is the focal length of the optics. As shown in the Eq. 2.1, the relation between d_o and d_i. Because of this nonlinearity, it is very complicated to calculate the position and size of the image during the optical design. In addition, CCM-CVLs cause optical distortions in resulting images due to their curved surfaces. Despite of their advantages in image positions, CCM-CVLs have limitations in dynamic change of image positions and image quality due to the nonlinearity of lens equation and the optical distortions.

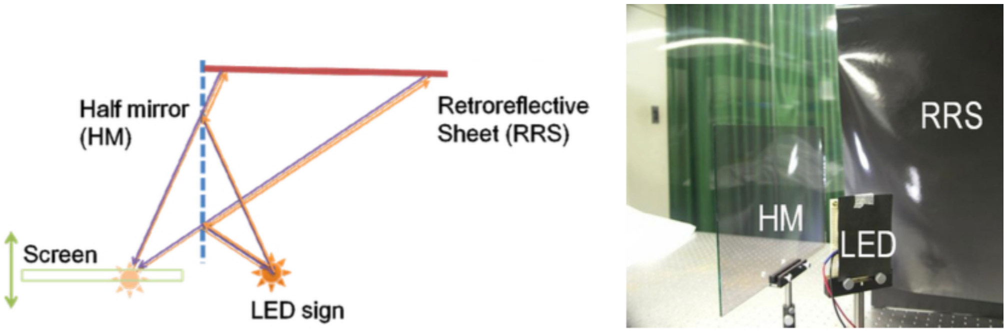
Retro-reflective imaging optics for real images

To overcome the limitations of CCM-CVLs, retro-reflective imaging optics (RIOs) have been recently proposed as combining the merits of HSMs and CCM-CVLs. The imaging mechanism of RIOs is based on reflections on orthogonal mirror pairs or retro-reflective materials (e.g. micro-beads). The resulting images are located in the symmetric position of lighting source about the optics plane. Unlike CCM-CVLs, RIOs can form real images without optical distortion in principle. In addition, the size of resulting images is identical to that of light source since only reflections are involved in the imaging process. Therefore, RIOs have similar optical properties on the position and size of mid-air image to plane mirrors and HSMs except they can form real images.

As examples of RIOs, AIRR, dihedral corner reflector array (DCRA), and aerial imaging plate (AIP) will be introduced.

AIRR [29]

AIRR [29], which consist of a HSM and reflective sheets, can form a mid-air real image by reflecting the incident lights from a light source.



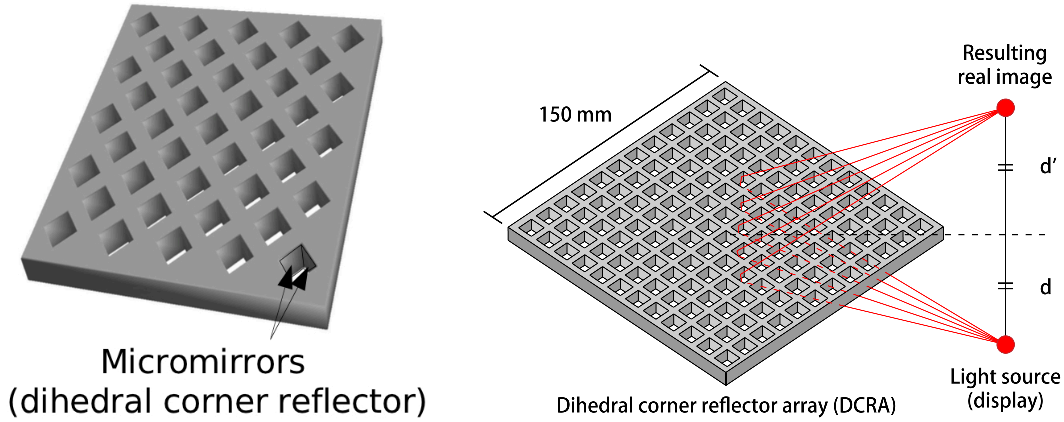
(a) Overview. (b) Experimental setup.

Figure 2.8: Optical configuration of AIRR [29].

Figure 2.8(a) illustrates the imaging process of AIRR. An LED sign is used as a light source for a mid-air image. The lights from the light source proceed to the retro-reflective sheet (RRS) after reflecting on the half mirror (HSM as the notation in this thesis). Due to the retro-reflection, the lights entered RRS are reflected at the RRS and penetrate the HSM. Finally, the retro-reflected lights from the RRS converge at the same point, which is marked as a screen in the Fig. 2.8(a). From these process, the resulting image are formed at the plane-symmetric point to the HSM. Although AIRR can form a mid-air image with a series of retro-reflections, its configuration requires large space for the installation. Therefore, more compact imaging optics are necessary for a compact optical design for MR interactions.

DCRA [30]

Maekawa et al. have proposed a plate-shaped device called DCRA, Dihedral Corner Reflector Array, which can form real images in mid-air [30]. The imaging mechanism is similar to AIRR since DCRA also employs retro-reflections of incident lights. However, DCRA has realized an imaging optics with a thin and compact size by arranging micro-sized corner reflectors in array.

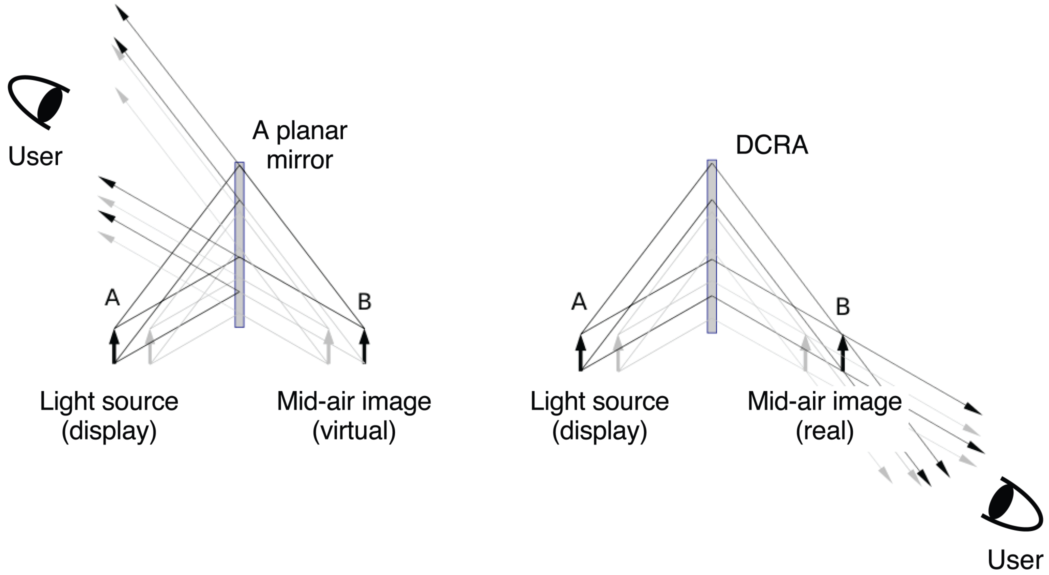


(a) Internal structure of DCRA [30].

(b) Optical property.

Figure 2.9: Dihedral corner reflector array (DCRA) by Maekawa et al.

Figure 2.9 shows the internal structure and imaging mechanism of DCRA. Inside the DCRA, there is an array of micro-mirrors (Fig. 2.9(a)). Each micro-mirror works as a reflector as the retro-reflecting sheet does in AIRR. Fig. 2.9(b) shows the imaging process in DCRA. A light source (display) is placed below the DCRA. As lights enter the DCRA, the incident lights are reflected at the corner reflectors inside the DCRA. Then, the reflected lights proceed to the other side of the light source and converge at the symmetric point of the light source to the DCRA plane. The converged lights finally form a mid-air image at the point as duplication of the light source with the same distance ($d' = d$). Since the imaging mechanism of DCRA employs only linear reflection as AIRR, the size and position of mid-air image can be easily calculated and the resulting image has no optical distortion in principle. However, the internal structure of DCRA is complex so that it is difficult to extend the size of DCRA in practical implementation. The currently available DCRA has a size with 150 (W) × 150 (D) × 5 mm (H).

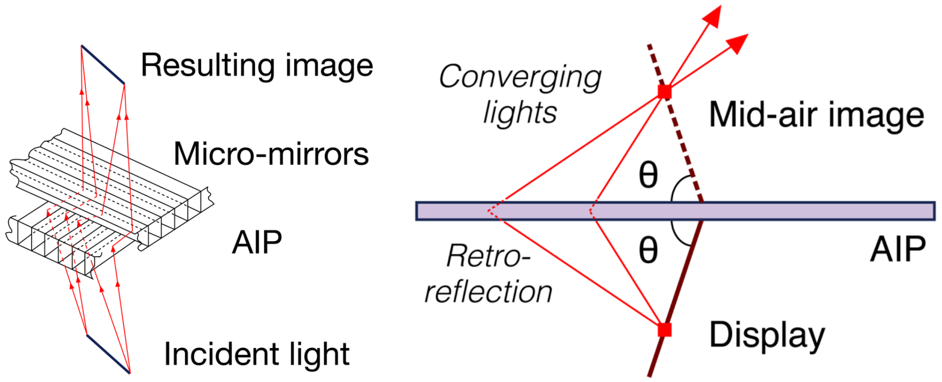


(a) A virtual image formed by a planar mirror. (b) A real image formed by DCRA.

Figure 2.10: Comparison between the mid-air images formed by a planar mirror and DCRA [30].

Figure 2.10 shows the comparison of resulting images formed by a planar mirror and a DCRA. In both images, the arrows at the point A represent the light source and the arrows at the point B are the resulting mid-air images. Planar mirrors can form only virtual images behind the mirror surface from the users' viewing position (Fig. 2.10(a)). Thus, light sources (displays) and users are placed in front of the mirror surface and resulting images are formed behind the mirror when mirrors are used in MR interactions. HSMs have also this optical property as a mirroring optics.

Unlike planar mirrors, in the case of DCRA, light sources are placed behind the DCRA and resulting images and users are located in front of the optics. With



(a) Internal structure [32]. (b) Optical property.

Figure 2.11: AIP (Aerial Imaging Plate).

DCRA, visual images can be placed along with physical objects in front of the DCRA.

AIP [31]

AIP (Aerial Imaging Plate) is another plate-shaped real imaging optics that uses retro-reflection to form mid-air images [31]. Although imaging mechanism of AIP is similar to DCRA, the internal structure of AIP is different to that of DCRA.

Figure 2.11 shows the internal structure of AIP and its optical property. As illustrated in Fig. 2.11(a), AIP consists of two layers that have micro-mirror walls in array, and the layers are crossed in 90°. When incident lights enter the AIP, the lights are retro-reflected at the micro-mirrors with penetrating the AIP. Then, the lights converge at the points which are symmetric to the AIP plane of the incident lights (Fig. 2.11(b)). As AIRR and DCRA, the imaging mechanism of AIP is based on retro-reflection. Thus, the position of mid-air images can be easily calculated as well. In addition, since the internal structure of AIP is simple and easy to extend the optics size, AIP can be implemented in a large size in principle. The maximum size of currently available AIP product is 350 (W) × 350 (D) × 5 mm (H). Therefore, AIP can form larger mid-air image than DCRA.

So far, AIRR, DCRA, and AIP have been reviewed as examples of RIOs. These imaging optics have linear relation between resulting mid-air images and light sources and thus can form mid-air images without optical distortion, so that image size and position can easily calculated for implementation. However, AIRR is not suitable for compact implementation of MR interactions since the optical setting needs a space. On the other hand, the compact shape of DCRA and AIP is helpful for implementing a compact optical design.

From the survey of imaging optics for virtual and real images, the optical properties of HSMs, CCM-CVLs, and RIOs are summarized in Table 2.1.

Table 2.1: Properties of mid-air imaging optics.

Imaging optics	**Distortion-free**	**Kind of mid-air image**
Half-silvered mirrors	✓	Virtual
Concave mirrors, Convex lenses		Real
Retro-reflective imaging optics (AIRR, DCRA, AIP)	✓	Real

Distortion-free optics have advantages in optical designs that imaging position can be easily calculated, especially for moving images. Therefore, in this thesis, only distortion-free optics are used for virtual and real images. As for the kind of mid-air images, virtual and real images are appropriately chosen according to the purpose of usage.

2.3 Examples of MR interactions

In order to complete the Table 1.1 with the MR interactions in this thesis, three usage examples including showcases, MR interfaces for a single user, and tabletop displays as MR interfaces for multiple users. MR showcases, which superimposes visual images onto physical exhibits, are required to provide visual images with multiple viewing zones, and allow users only indirect manipulation of the exhibits. MR interfaces for a single user display visual images within a single viewing zone or limited viewing angles with supporting direct manipulation of physical objects. For MR interfaces for surrounding users, tabletop displays will be reviewed since multiple viewing zones of visual images and direct manipulation of physical objects can extend the current tabletop displays.

2.3.1 MR showcases for a single user

MR interactions have been combined with museum applications with improving exhibition experiences [33, 34]. Especially, the ability of image superimposition onto physical objects has led to research on MR showcases.

Magic vision

Magic vision [35] forms a virtual images using mirror reflection in order to superimpose visual images onto physical objects.

Figure 2.12 shows an example application of Magic vision as an MR showcase for museum exhibitions. Magic vision is often used in museums to explain exhibits or dioramas with animated visual images which are displayed in mid-air. Similar to Pepper's Ghost, visual images on the display are reflected on the HSMs, and overlaid on physical objects inside a showcase. Due to the usage of HSM, imaging position can be easily calculated and resulting images are free from optical distortions. Despite

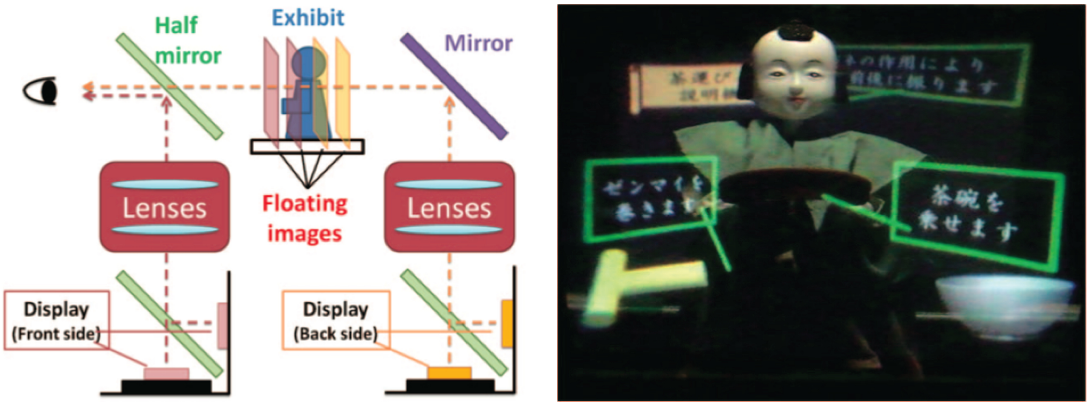
Figure 2.12: Magic vision application for museum exhibitions [36]. Visual images are superimposed onto the real diorama.

these merits, however, visual images are only superimposed from the front side of physical objects so that the occlusion between visual images and physical objects cannot be expressed when the visual images are placed behind the physical objects.

ExFloasion

ExFloasion [37] is an MR showcase which can superimpose four layers of mid-air images on physical exhibit(s) using a set of Fresnel lenses (Figure 2.13). Since the visual images are formed as mid-air images, users do not need to wear special glasses in order to see the superimposed images and the exhibit(s). The optical design of ExFloasion can form both virtual and real images so that the mid-air images can be placed immediately next to the physical exhibit inside the showcase by sandwiching the exhibit from the front and back.

Although ExFloasion could propose an optical design to superimpose real and virtual images onto a physical exhibit and realize glasses-free MR interaction for an exhibit showcase, its viewing zone is limited to one region which can be seen by only one user. However, the viewing zones of exhibit showcases need to provide visual



(a) System configuration.

(b) Mid-air images superimposed on a physical exhibit.

Figure 2.13: ExFloasion by Nakamura et al. [37]

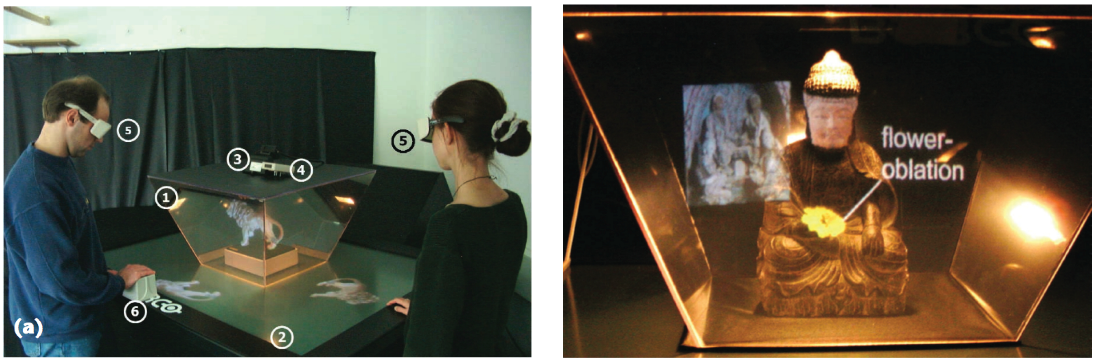
images to as wide range as possible since the showcases are usually appreciated by multiple users.

2.3.2 MR showcases for surrounding users

Virtual Showcase

Virtual Showcase [38, 39] proposes an MR showcase for surrounding users with multiple viewing zones (Figure 2.14). Using HSMs and stereoscopic projectors, stereoscopic images can be superimposed onto physical objects. Its symmetric design enables viewers to see the showcase from lateral directions with four discrete viewing zones. Although stereoscopic images enable richer depth perceptions, viewers are required to wear 3D glasses to see the superimposed images. In order to handle occlusion problems between physical exhibit(s) and visual images, Virtual Showcase features a computational solution that limits the optimal viewing position to the position of 3D glasses [40].

From the merits of ExFloasion and Virtual Showcase, this thesis proposed MRsionCase, an MR showcase for surrounding users in Chapter 3. Giving the priority of implementation on glasses-free usage and natural occlusions between visual images



(a) System configuration.

(b) Mid-air images superimposed on a physical exhibit.

Figure 2.14: Virtual showcase by Bimber et al. [38].

and physical objects, in MRsionCase, this thesis proposed an optical design with imaging optics to form two mid-air image layers. As ExFloasion did, the image layers consists of both virtual and real images and sandwich the exhibit from the front and the back. For real imaging optics, DCRA is used instead of Fresnel lens due to the merits of no optical distortion and compact size. The combination of HSM and DCRA forms distortion-free mid-air images with a linear relation between imaging position and light source. As for depth perception, although 3D glasses can provide richer stereoscopic perception than two mid-air layers, MRsionCase can be extended to provide richer depth conception when auto-stereoscopic displays are used in the future.

The comparison of ExFloasion, Virtual Showcase and MRsionCase is summarized in **Table 2.2**.

2.3.3 MR interfaces for a single user

MR interfaces have been proposed to superimpose visual images with physical objects using virtual and/or real images. In specific, single user MR interfaces have displayed visual images within limited viewing angle for user interactions that suitable for a single user usage. User interaction of MR interfaces can be determined by the way

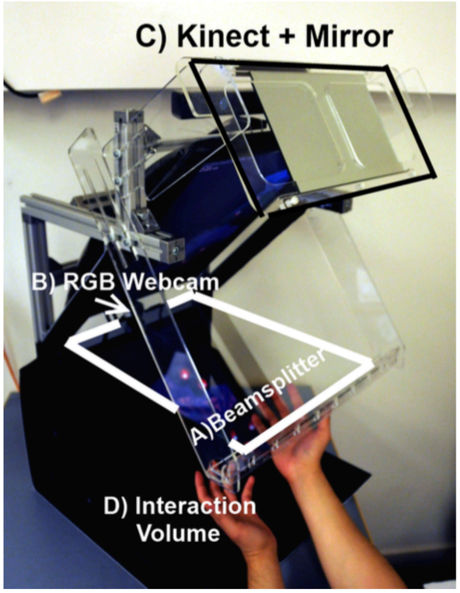
Table 2.2: Comparison of ExFlasion, Virtual Showcase and MRsionCase.

Criteria	**ExFloasion**	**Virtual Showcase**	**MRsionCase**
Solution for occlusion problem	Optical (natural)	Computational	Optical (natural)
Optimal viewing position	Not limited	Limited to one position	Not limited
Number of viewing zones	Single	Multiple (four)	Multiple (four)
3D perception	Normal (four layers)	Rich (stereoscopic)	Simple (two layers)

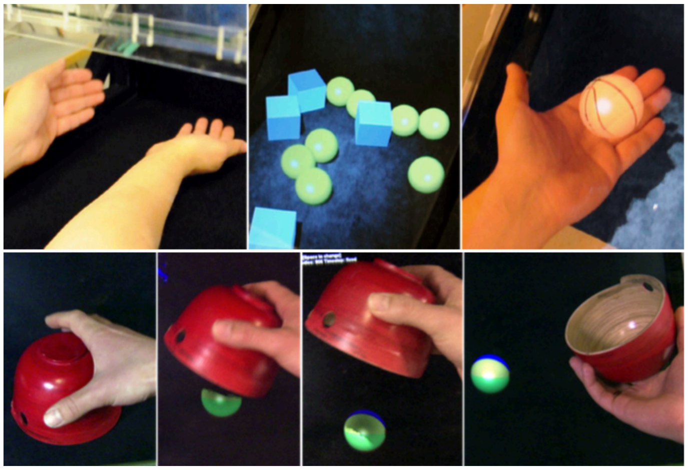
to display visual images. For example, systems using virtual images with mirror reflection provide see-through applications due to the imaging positions. On the other hands, MR displays using real imaging optics enable users to directly access the visual images.

HoloDesk [41]

HoloDesk [41] is an MR interface which enables a 3D user interaction between visual images and physical objects (Figure 2.15). For visual images, in HoloDesk, virtual images are formed by HSM and a 3D display. Images of 3D display are reflected on HSM, and virtual images are formed behind the HSM without distortion as a result of mirror reflection. However, the position of the virtual images are fixed since the HSM and 3D display are fixed to the system. For stereoscopic depth perception, users are required to 3D glasses when they see the visual images. Although HoloDesk can form visual images in mid-air and display them in 3D space with stereoscopy, the display area of visual images are limited only to behind the HSM due to the limitations of virtual images, and thus the interaction area as well. Therefore, HoloDesk can support only see-through MR interactions, and users' access to the interaction area



(a) HoloDesk configuration.



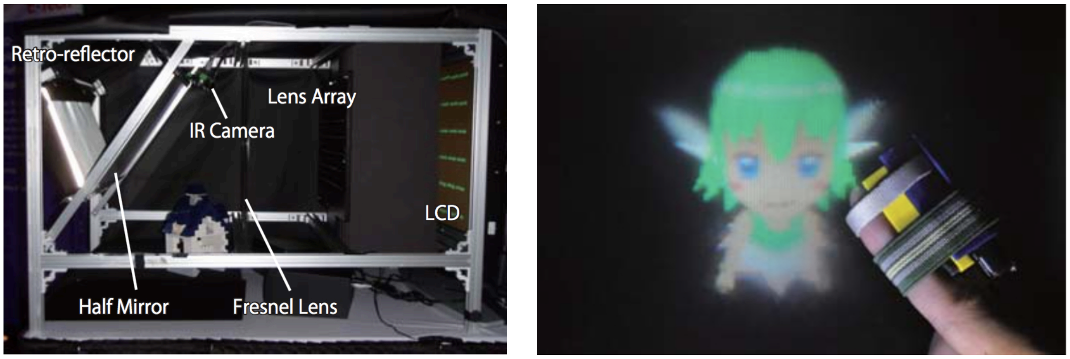
(b) User interaction with physical objects and visual images.

Figure 2.15: HoloDesk [41].

is hindered due to the HSM which is placed in front of the interaction area as an imaging optics.

RePro3D [42]

RePro3D [42] enables an interaction between visual images and physical objects with tactile feedbacks (Figure 2.16). Using projector arrays and retro-reflective material, RePro3D can control light rays and provide visual images to the user's eye as if they



(a) System configuration.

(b) User interaction with visual images.

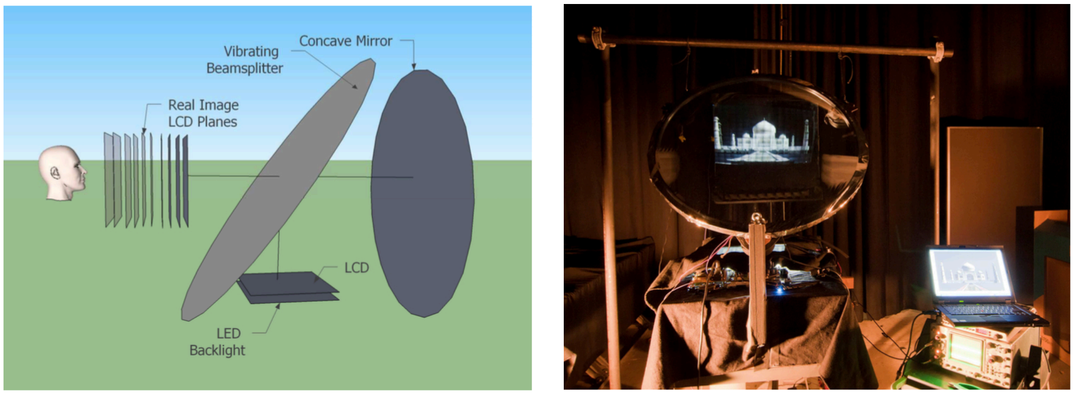
Figure 2.16: RePro3D by Yoshida et al. [42].

are floating in mid-air. The visual images can be superimposed onto the physical objects or users' fingers in 3D space. With RePro3D, users can interact with a character shown in mid-air by touching and poking it or moving a see-saw to make the character jump on it. However, the imaging position is fixed so that the visual images can be superimposed on only specific positions. In addition, due to the optical design of RePro3D, users cannot directly access to the interaction area since HSM is placed between visual images and the users.

Volumetric display by vari-focal beam splitter and concave mirror [43]

Smoot et al.'s work [43] presents multiple mid-air image layers in different depths by using a CCM and a vari-focal beam splitter (Figure 2.17). In this system, imaging position, especially in a depth direction, is determined by the distance between a light source (an LCD) and a concave mirror so that a vari-focal beam splitter with a dynamic control of focal length can form image layers in mid-air with different depths about 30-cm range. Since the resulting mid-air images are real ones, the mid-air images can be placed in front of imaging optics. Thus, the interaction area is not blocked or divided by imaging optics. Users can place physical objects along with visual images in this area without any interruption.

The number of viewing zones in this system is limited to one, which is enough and suitable for a single user. However, the resulting images have intrinsic optical distortion due to the use of the CCM. In addition, the relation between image size and position has non-linearity so that complex calculation is required for optical design and change in imaging positions.

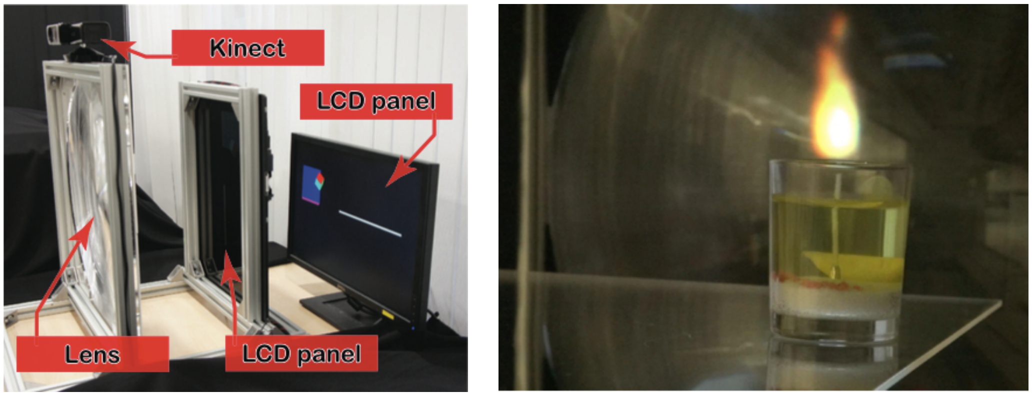


(a) System configuration. (b) User interaction with visual images.

Figure 2.17: Volume display by Smoot et al. [43].

Fuwa-vision [44]

Fuwa-vision [44] has proposed an auto-stereoscopic MR interface using a Fresnel lens Figure 2.18. Adaptive parallax barrier system is devised with mixed usage of a transparent LCD and head tracking so that the visual images can provide rich depth perception as 3D glasses. Imaging position is set to the front of imaging optics so that visual images can be superimposed onto physical objects. The applications implemented in Fuwa-vision shares the main goals discussed in this thesis with a large part. Users can interact with physical objects and visual images without any disturbance caused by imaging optics. However, in Fuwa-vision, the movement of imaging positions in a depth direction depends on only binocular disparity. Since the convergence angle of human eyes is limited, depth range of visual images also has limitation. Especially, when the interaction is performed within arm's reach, visual images have little flexibility in the change of depth direction. Moreover, since the depth movement of visual image is implemented by adaptive parallax barrier, which is based on head tracking, robust and precise head tracking is highly required to provide visual images in exact depth positions where the images are supposed to be displayed.



(a) System configuration. (b) User interaction with visual images.

Figure 2.18: Fuwa-vision by Nii et al. [44].

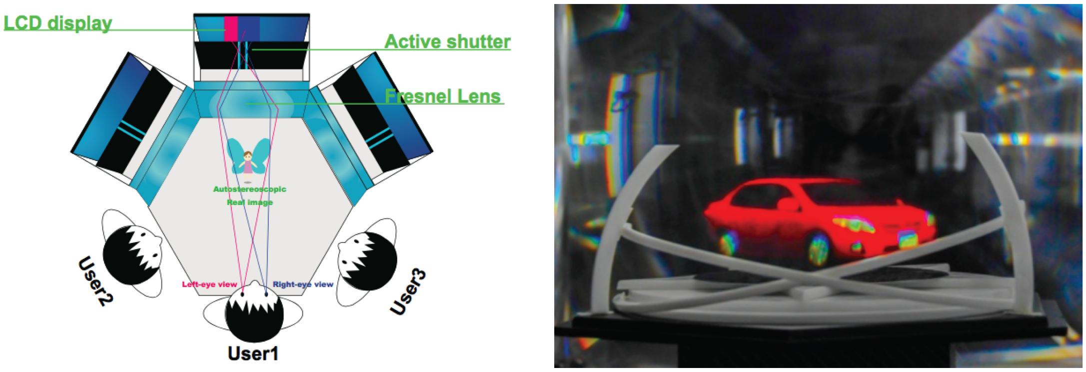
From the review of the previous MR interfaces for a single user, core features can be summarized in terms of manipulation of physical objects and viewing zones of visual images. In order to enable users to access and manipulate physical objects, the interaction area needs to be located in front of imaging optics without interruption or blockage by the imaging optics. Meanwhile, as for the viewing zone of visual images, one viewing zone is enough for a single user interaction. But rather, the author have focused on imaging position which is more critical than viewing zone since visual images are displayed along with physical objects in 3D space. If users move physical objects in 3D space, visual images can change the position according to the position of physical objects in order to provide high level of geometrical consistency. Therefore, display area and interaction area should be set to the identical area, and they are placed in front of imaging optics for direct manipulation and enhanced spatial connection between physical objects and visual images. For this purpose, in Chapter 4, this thesis proposed a new MR interaction, MARIO, which allows direct manipulation of physical objects and movement of mid-air image position. If the visual images can be formed under the linear relation between the position of mid-air images and light sources, the position of mid-air image can also be changed in more simple way. Thus, the optical design of MARIO adopted AIP as a real imaging optics. By changing the distance between display and AIP using a motorized actuator, a layer of mid-air image can be moved in 30-cm depths. The author believes that such a simple change of image depths is effective not only for a single user but also for multiple users in an MR display.

2.3.4 MR interfaces for surrounding users

For MR interfaces for surrounding users, interaction and display area need to be extended to multiple directions and wider viewing zones.

HaptoMirage [45]

HaptoMirage [45] provide a mid-air image with wider viewing range by connecting three units of Fresnel lens (Figure 2.19). Each set of imaging optics is based on Fuwa-vision [44], which consists of a Fresnel lens, an active shutter with a transparent LCD, and an LCD display for a light source. Since the Fresnel lenses are disconnected at the boundary, viewing zones of mid-air images are also divided into three regions. These three viewing zones enable three users to see the visual images at the same time. In order to propose an MR interface for surrounding multiple users with providing three viewing zones of visual images, HaptoMirage requires additional sensors to detect users' head positions since the light rays for mid-air images are controlled by active shutters. Moreover, in HaptoMirage, visual images are displayed from vertical layer only so that horizontal tabletop area is not used for user interaction. However, the author considers that tabletop surfaces provide horizontal space for sharing visual images among multiple users and improve the spatial link between mid-air images and the real world.



(a) System configuration.

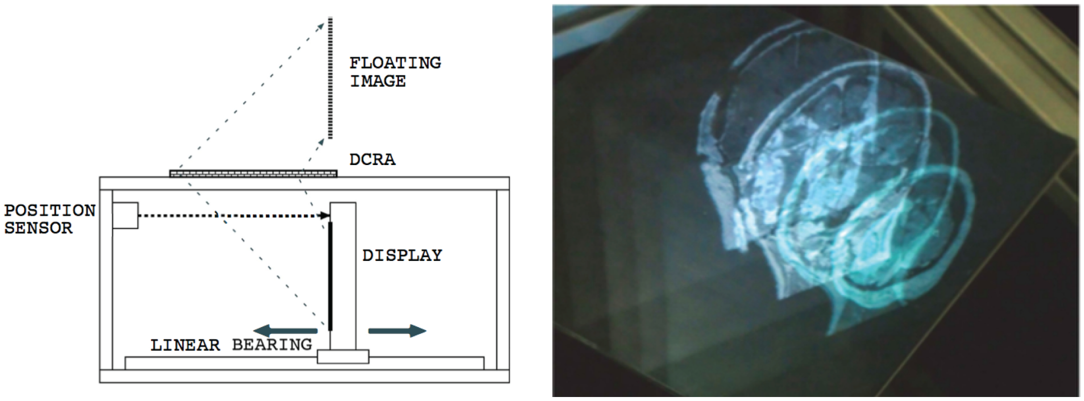
(b) A visual image (car) is superimposed onto physical objects (a table).

Figure 2.19: HaptoMirage by Ueda et al. [45].

Volume slicing display with DCRA [46]

Markon et al. [46] proposes a new optical configuration with DCRA, which shows possibility on the usage of vertical mid-air image on a tabletop display (Figure 2.20). In the optical design of Markon's work, DCRA is placed as a horizontal plane and forms a vertical mid-air image on its surface. The vertical mid-air image can move along one direction (back and forth) and be superimposed onto physical objects since the image has no physical shape. Although this system provides only vertical mid-air images without horizontal images, the optical configuration has provided good reference when the author devised an optical design of HoVerTable, a tabletop display which can combine vertical mid-air images with a horizontal tabletop surface. In HoVerTable, AIP is used as a tabletop surface in order to form vertical mid-air images without optical distortion as reported in Markon et al.'s work [46].

In HaptoMirage and Markon et al.'s work, visual images are only displayed in mid-air as floating objects. However, the author believes that the mixed usage of vertical mid-air images and horizontal images enhances the visual expression in MR



(a) System configuration. (b) Mid-air images floating on the DCRA.

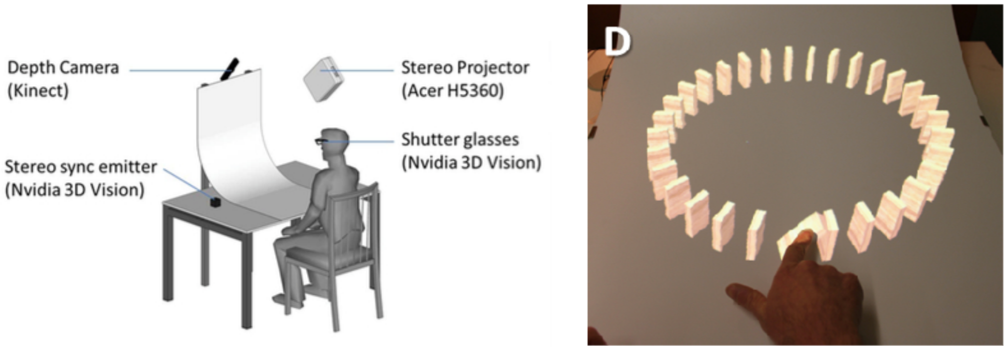
Figure 2.20: Volume slicing display by Markon et al. [46].

interfaces. For example, shadows can be cast on the horizontal surface below mid-air images.

As reference for the usage of horizontal surface as displays, studies on tabletop displays which aim at the combination of horizontal and vertical images will be reviewed.

MirageTable [47]

MirageTable [47] has combined vertical and horizontal image planes with using a desk and a curved screen (Figure 2.21). In MirageTable, a desk is seamlessly connected to a upright screen as a curved surface so that users can place physical objects on the desk and interact with visual images in front of them. Since visual images are projected from a stereoscopic projector, 3D depth perception can be realized with the aid of 3D glasses. Although MirageTable can combine vertical and horizontal images, the position of visual images is fixed on the desktop surface. Moreover, the system mainly focuses on the single user due to its desktop configuration.



(a) MirageTable configuration.

(b) User interaction with visual images

Figure 2.21: MirageTable by Benko et al. [47].

Deskrama [48]

Deskrama [48] also proposes combined usage of vertical and horizontal images by placing a upright display (LCD) on the desk (Figure 2.22). Using both vertical and horizontal image planes, 3D objects can be illustrated with multiple images viewed from two perspectives. Users can move the upright display on the table and explore the imaging space on the table with viewing the visual image from the display. However, visual images cannot be directly superimposed onto physical objects since the visual images are displayed on the display panel.

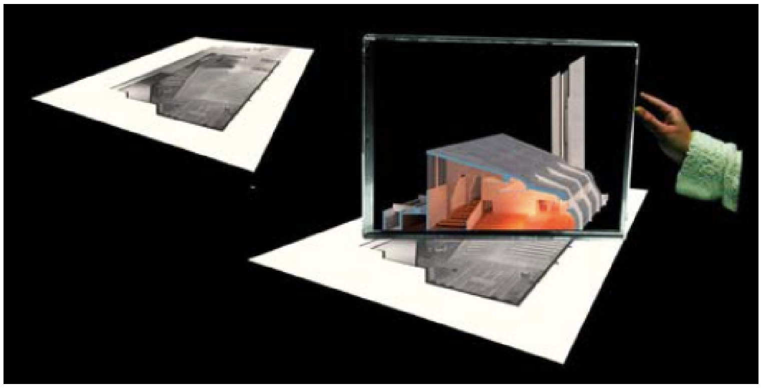


Figure 2.22: Deskrama by Nagakura and Oishi [48]. Vertical images are shown from a upright display and horizontal images are displayed on the desktop surface.

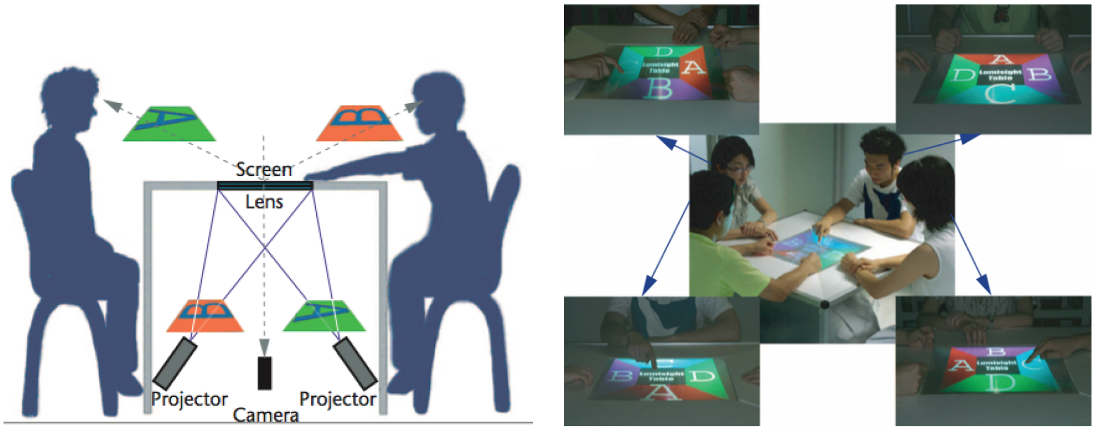
Tablescape Plus [49]

Tablescape Plus [49] suggests a similar approach to Deskrama using a tiny vertical screen on a tabletop display (Figure 2.23). For vertical images, a tiny and upright semi-transparent paper is placed on the tabletop surface and used as a screen for image projection. The upright screen can be moved by users with changing image positions during user interactions. Horizontal images are also changed according to the vertical images by enhancing the link between visual images. In front of the table, multiple users can interact with visual images by moving the vertical screens and by manipulating physical objects. Due to the usage of rear projection, however, the interaction and viewing area is limited to the front side only.

Figure 2.23: Tablescape Plus by Kakehi et al. [49]. Characters are projected to tiny upright screens on the table. Horizontal images are displayed from the tabletop surface.

Lumisight Table [50]

When multiple users surround a tabletop display and view the displayed visual images, there is needs that the contents of visual images are selectively provided to the users according to their viewing positions. Lumisight Table [50] proposes a tabletop display which can selectively display visual images to each users with a special diffusive film (Figure 2.24). A special film, Lumisty [51], is used to display horizontal visual images to only specific range. The Lumisty film have an unique optical property that the film diffuses the incident lights with a specific angle range. Therefore, incident lights entering from this range are diffused and the others only pass through the film as transparent glass sheet. Although Lumisight Table can provide only horizontal visual images on the tabletop surface, the usage of Lumisty film can increase the number of viewing zones to four directions and display different sets of visual images to surrounding users according to their viewing positions.



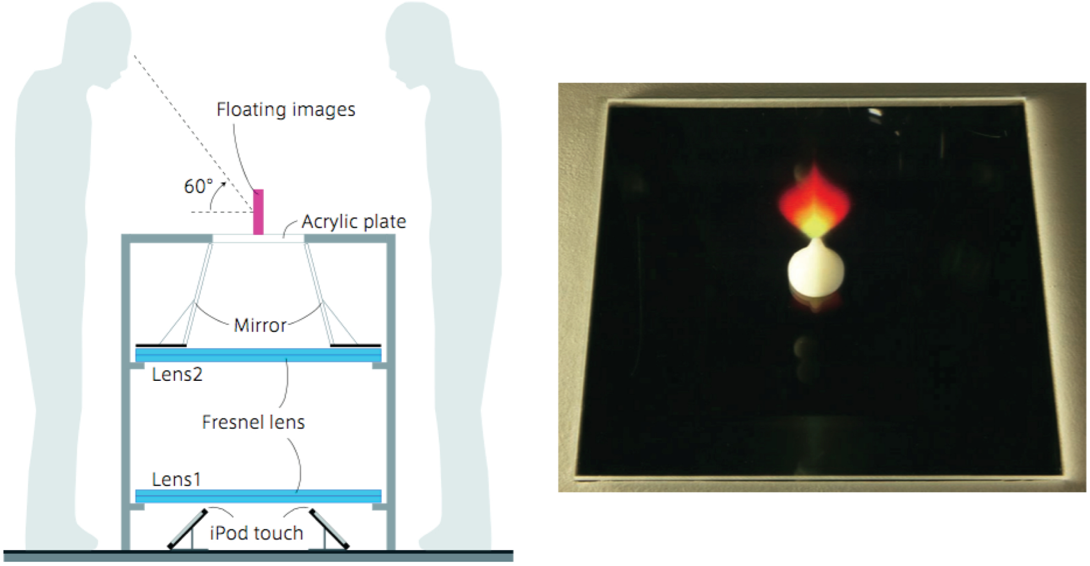
(a) System configuration.

(b) Visual images viewed from each user.

Figure 2.24: Lumisight Table by Kakehi et al. [50].

FloasionTable [52]

In the attempt of extending horizontal tabletop displays to the vertical direction, mid-air images have been formed with standing on tabletop displays (Figure 2.25. In FloasionTable [52], upright mid-air images are formed by Fresnel lenses to provide vertical images to each viewing zone. Since the vertical images can be seen from four viewing zones, four users can view the visual images at the same time. The position of vertical images are set to the center of the tabletop surface so that the vertical images can express four different views from the front, left, right, and back side. Although each layer of vertical mid-air images is two-dimensional, 3D-like objects can be displayed by constructing 2D vertical image layers which are standing on the table surface. Due to the usage of mid-air images, FloasionTable can display vertical images without the need for physical displays on the tabletop surface so that users can place physical objects on the table and freely access the tabletop space. However, the positions of mid-air images are fixed and cannot be flexibly changed due to its use of Fresnel lenses in the optical design.



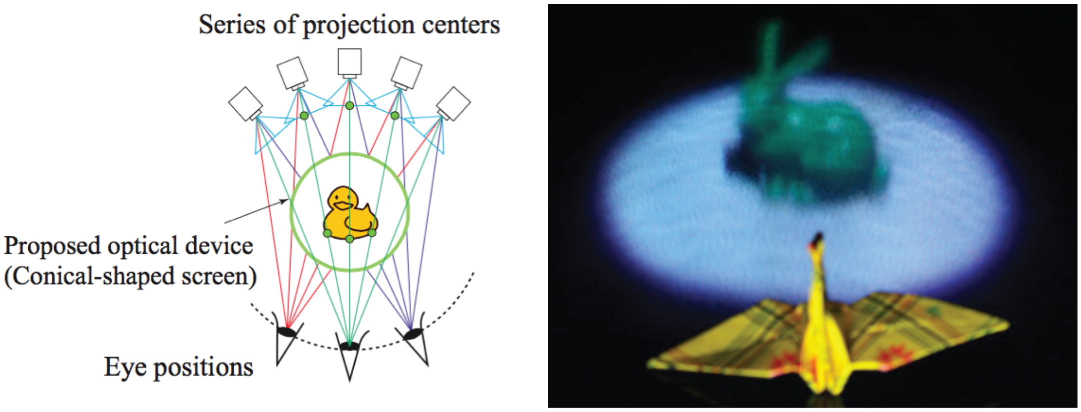
(a) System configuration.

(b) A mid-air image (flame) is superimposed onto a physical object (candle) placed on the tabletop surface.

Figure 2.25: FloasionTable by Wada et al. [52].

fVisiOn [53]

fVisiOn [53] also extends the display area of an MR interface by providing a vertical mid-air image with increased viewing range. Vertical images are formed on a tabletop surface by projector arrays (Figure 2.26). Multiple projectors construct a vertical mid-air image with multiple images viewed from various perspectives. In addition to the vertical image, a horizontal image is projected to the tabletop surface which is below the vertical image. Both horizontal and vertical images are linked in applications so that vertical images seem “really” standing on the horizontal surface. For example, when a character, which is shown from a vertical image, is dancing on a tabletop display, the shadow of the character is also displayed and moving on the table surface with accordance to the character’s movement. Since the visual images do not need physical displays, the image can be superimposed onto physical objects as well. Despite such features in fVisiOn, the imaging position is fixed and immovable due to the internal structure of optical design.



(a) System configuration to reproduce a light field (top view).

(b) The resulting image (Stanford bunny) is shown along with a real paper origami (crane).

Figure 2.26: fVisiOn by Yoshida et al. [53].

Based on the survey, the author proposed a tabletop display HoVerTable in Chapter 5, which can support MR interactions for surrounding multiple users. HoVerTable has several core features as design requirements as follows. For visual images, not only horizontal images but also vertical images are displayed and both images are combined to enhance the link between them. Vertical images are formed as mid-air images instead of using physical displays or screens so that users can freely access the tabletop surface and place physical objects on it without any interruption. Moreover, visual images provide view-dependent appearance to viewers according to their viewing positions. This enables selective sharing of visual images under the situation that the display is shared by multiple users. As for imaging position, the vertical mid-air image can change the position and move on the tabletop surface for interactive applications. In addition, multiple layers of vertical images are formed to broaden the range of visual expression in MR interactions.

2.4 Summary

In this chapter, MR concept is introduced to share the core goal with the readers of this thesis that visual images are seamlessly merged with physical objects. Among MR interactions, the author has aimed to realize MR interactions in the real world space. Visual images are formed in the real world space beyond physical displays, and thus they can be directly superimposed onto the physical objects. As a foundation of main discussion, optical properties and geometric optics for mid-air images have been reviewed. From the optical property of real and virtual images, it is confirmed that real images can be formed in front of imaging optics so that the real images can be superimposed onto physical objects without any see-through window. On the other hands, virtual images are suitable for see-through applications such as showcases due to the imaging position. As for imaging optics, distortion-free optics, which uses only linear reflection in imaging processes, such as HSMs for virtual images and DCRA and AIP for real images enables simple optical design with removing complex calculation for the position and size of resulting images.

Table 2.3 shows the overview of the MR interactions which are proposed in this thesis. The main discussion starts from an MR showcase for a single user, including Magic mirror [35] and ExFloasion [37]. With providing viewing zones to four directions, MRsionCase is proposed in Chapter 3. As removing the front HSM from Magic mirror and ExFloasion and forming two layers of mid-air images along with the physical objects, MARIO is implemented in Chapter 4. Finally, as eliminating

Table 2.3: Overview of the MR interactions proposed in this thesis.

	Single viewing zone	Multiple viewing zones
Indirect manipulation	Magic mirror [35] ExFloasion [37], (Related work)	Ch. 3: MRsionCase: Multi-directionally Viewable MR Showcase [5, 6]
Direct manipulation	Ch. 4: MARIO: Mid-air Augmented Reality Interaction with Objects [7, 8]	Ch. 5: HoVerTable: Combining Dual-sided Vertical Mid-air Images with a Horizontal Tabletop Display [9]

all HSMs (i.e. show windows) from MRsionCase and extending the viewing angle of MARIO to multiple directions, HoVerTable can be implemented in Chapter 5.

Chapter 3

MRsionCase: Multi-directionally Viewable MR Showcase

This chapter addresses the challenging problem on the optical design that implements indirect manipulation of physical objects and that provides mid-air images with multiple viewing zones. In specific, the author aims to propose an MR showcase that superimposes visual images onto physical exhibits. Several mixed reality (MR) showcases have recently been proposed to augment physical exhibits with visual images. However, most of these showcases require viewers to wear a special device to see the superimposed images. To overcome the problems, in this chapter, a glasses-free MR showcase, MRsionCase, is proposed. MRsionCase provides spatially consistent visual images along with physical exhibits. A DCRA is used to place mid-air images inside the showcase without optical distortions. The symmetric design of MRsionCase displays visual images with multiple viewing zones. Thus users can view the exhibit with surrounding and walking around the showcase. Results of a user survey demonstrated the effectiveness of MRsionCase as an MR showcase.

3.1 Introduction

Showcases have long and widely served as an effective means of display in museums and advertisements. Recently, mixed reality (MR) showcases that augment physical objects by superimposing images to provide viewers with an enhanced exhibition experience have been proposed.

Since most MR showcases require viewers to wear a special device to see the superimposed images, the viewing area and the number of viewers are limited in comparison with conventional showcases.

The objective of the work reported here is to develop an MR showcase that overcomes these limitations and that meets the following four requirements:
(1) superimposing easily understandable information onto a physical exhibit,
(2) motivating intellectual exploration of various aspects of the exhibit,
(3) providing information targeted at different groups of viewers to attract their attention, and
(4) enabling walk-up-and-use for viewers without preparation.

The proposed MR showcase is called MRsionCase, which stands for "Mixed Reality + emerSION of mid-air images + showCASE." MRsionCase features spatially consistent visual and auditory information that facilitates the viewer's understanding of the exhibit. Its symmetric design enables viewing of the exhibit from four directions, as shown in Figure 3.1. MRsionCase also provides visual information targeted at different groups of viewers with gestural interaction. Most importantly, viewers do not need to wear a special device for experiencing an MR exhibition.

3.2 Proposal

Through MRsionCase, it is aimed to create a new paradigm for MR showcases that fulfills four main goals: 1) spatial consistency between physical object and informa-

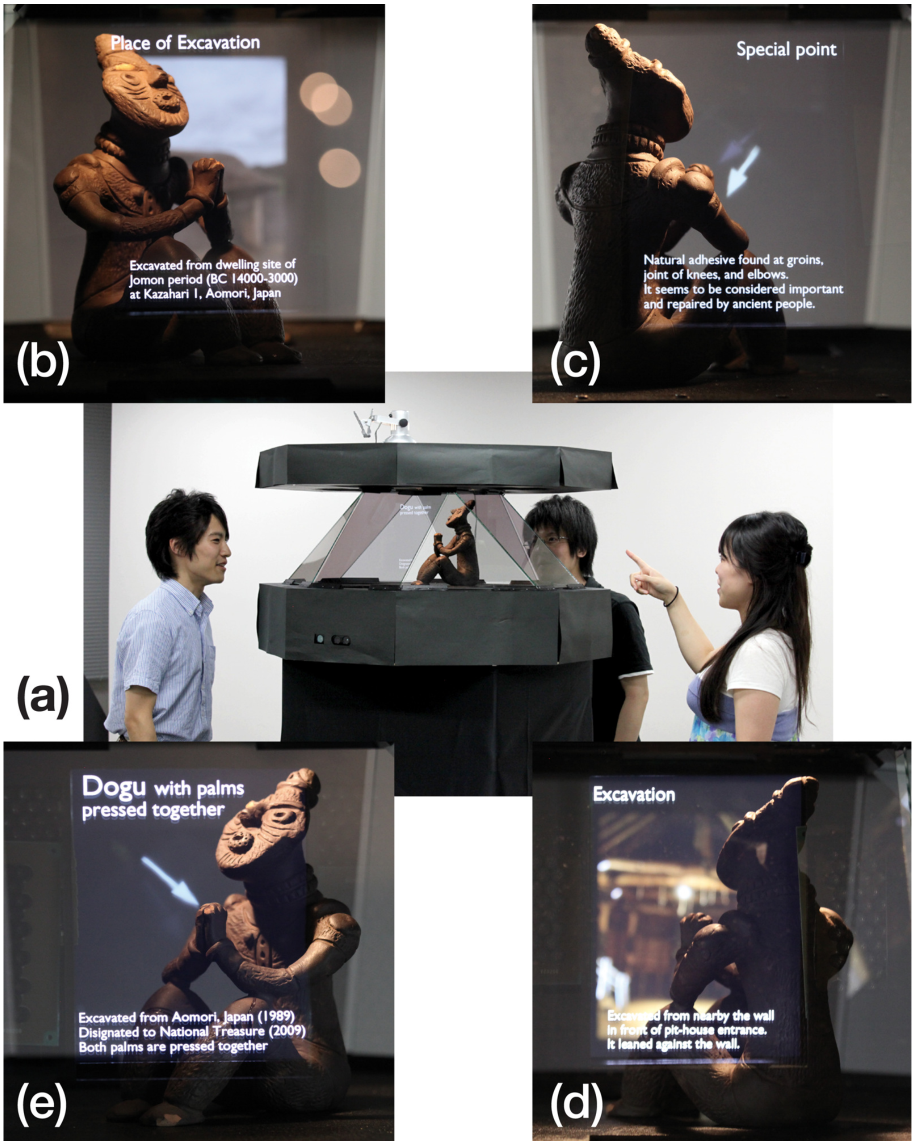


Figure 3.1: Illustration of MRsionCase. (a) Viewers can see exhibit and superimposed images without wearing a special device. (b)–(e) Views from front, left, back, and right.

tion, 2) multi-directional viewability, 3) targeted information presentation, and 4) device-free usage. These four goals correspond to the requirements of MRsionCase.

3.2.1 Spatial consistency between physical object and information

A core requirement is that information be displayed with a high spatial consistency with the target exhibit to strengthen the link between them. This will help viewers deepen their understanding of the exhibit in an easy and intuitive way. Thus a system is designed to place mid-air images inside the showcase along with physical exhibit(s). Specifically, a physical exhibit is sandwiched by two mid-air images formed at both the front and rear sides. Front image layers are suitable for providing descriptive text since they are not occluded by the exhibit. Rear image layers provide background images, naturally occluded by the exhibit, or direct annotation right next to the exhibit. Directional sounds convey auditory information from the showcase to the viewer.

3.2.2 Multi-directional viewability

Viewability from multiple directions is an essential feature of showcases, especially for volumetric exhibits. Therefore the system is designed so that the optical images and auditory guides are displayed around a showcase. This enables viewers to move around the showcase and enjoy the visual images and audio guides from each side.

3.2.3 Targeted information presentation

Showcases need to be appealing to a wide variety of viewers, so the focus of our design was to provide information targeted at different groups. For example, a different set of visual images can be provided with multiple languages by enabling users to choose the visual images through user interaction. Such user-adaptive information presentation enhances the practicality of MR showcases.

3.2.4 Device-free usage

Some of the ideas introduced above have been realized in other MR systems, but users were required to wear special devices such as 3D glasses or headphones. In this thesis, the author have firmly believed that showcases should be easily accessible, meaning that viewers do not have to wear a special device. The use of mid-air images and directional sound could eliminate the need for a special device. MRsionCase thus enables viewers to participate in exhibitions without preparation.

3.3 System Design

To superimpose mid-air images onto physical objects, MRsionCase has two types of imaging systems, virtual imaging system and combined imaging system.

3.3.1 Multi-layered mid-air images

Mid-air images in MRsionCase are formed at two different depths. These images sandwich a real exhibit from the front and rear sides. Superimposing multiple images at different depths has a merit that no additional calculation is not needed to express occlusion between exhibits and superimposed images. When a mid-air image with one layer is superimposed onto a real exhibit, occlusion can be a problem [40]. In specific, when the one-layered mid-air image is placed behind the exhibit, the intersection between the superimposed image and the exhibit should be excluded. Only after this process, correct occlusions between real exhibits and superimposed images can be expressed. Despite such occlusion handling, however, it is optically impossible for the one-layered mid-air image to be placed behind the exhibit. Therefore, the optical design of MRsionCase has focused to form two-layered mid-air images which can sandwich an exhibit. These two-layered mid-air images do not need additional

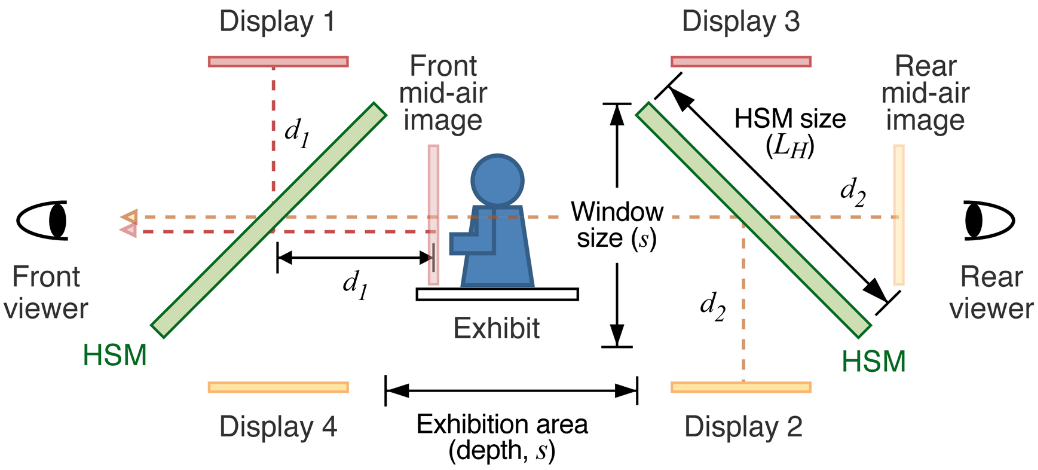


Figure 3.2: Virtual imaging system with HSMs (IS1).

calculation for occlusion expression. A front image layer will be placed in front of the exhibit and a rear image layer will show the rear part behind the exhibit.

3.3.2 Virtual imaging system (IS1)

To present visual images that can be seen without wearing a special device, we use optical devices to form images in mid-air. First, a virtual imaging system (IS1), which consists of HSMs and displays in a simple structure, has been devised. Due to their transparency and reflectivity, HSMs can function as windows as well as mirrors.

Figure 3.2 illustrates the virtual imaging system (IS1) with optical parameters. As a showcase, the physical exhibit is placed at the center in the exhibition area. The size of exhibition area is assumed as a cubic area with a side of s. In front of and back of the exhibition area, a pair of HSMs is placed at 45° and 135° from the perpendicular, respectively. Since the height of exhibition area is s, the size of HSM (L_h) is equal to $\sqrt{2}s$. Then, a display is placed above and below each HSM. From the size of HSM and exhibition area, the maximum size of display is limited to s (W) $\times$ s (H). Among the four displays in Figure 3.2, for the front viewer, Displays 1 and 2 form virtual images at the front and rear of the exhibit by reflection on the HSMs.

The distances between Display 1, 2 and each corresponding HSM are marked as d_1 and d_2, respectively. This distance will determine the position of mid-air images. For the front mid-air image, the imaging position should be inside the exhibition area to superimpose visual images onto the exhibit. Therefore, the maximum value of d_1 is set to $\frac{3}{2}s$. Unlike the front image, d_2 can be flexibly determined since the rear image is placed outside the showcase. For rear viewers, Displays 3 and 4 perform the same process.

IS1 could provide multiple viewing zones and device-free usage features with a simple design. However, there is a problem that the rear image layer cannot be formed inside the showcase due to the limitation of the mirror reflection. Such limitations reduce the flexibility of imaging position and thus hinder the complete achievement of spatial consistency between physical objects and visual images in an MR showcase.

3.3.3 Combined imaging system (IS2)

In order to place rear images immediately next to physical exhibits, a combined imaging system (IS2) that can form both virtual and real images is devised. Figure 3.3 shows the detailed design of IS2 for the front viewer. This IS2 consists of a DCRA, a mirror, and a display, which are added below the rear HSM of the IS1. The size of mirror DCRA and mirror is defined as L_D and L_M, respectively. The rear image layer is formed with a real image by the DCRA in front of the rear HSM after reflection. The distances from displays to the mirror, and from the mirror to DCRA, DCRA and HSMs, and DCRA and the mid-air image can be marked as d_1, d_2, d_3, and d_4, respectively. From the optical properties of the DCRA, $d_1 + d_2$ is equal to $d_3 + d_4$. Thus, the position of the mid-air image (d_4) can be easily determined to any depth inside the showcase with adjusting d_1. For a rear viewer, DCRA, a mirror, and a display needed to be added below the front HSM.

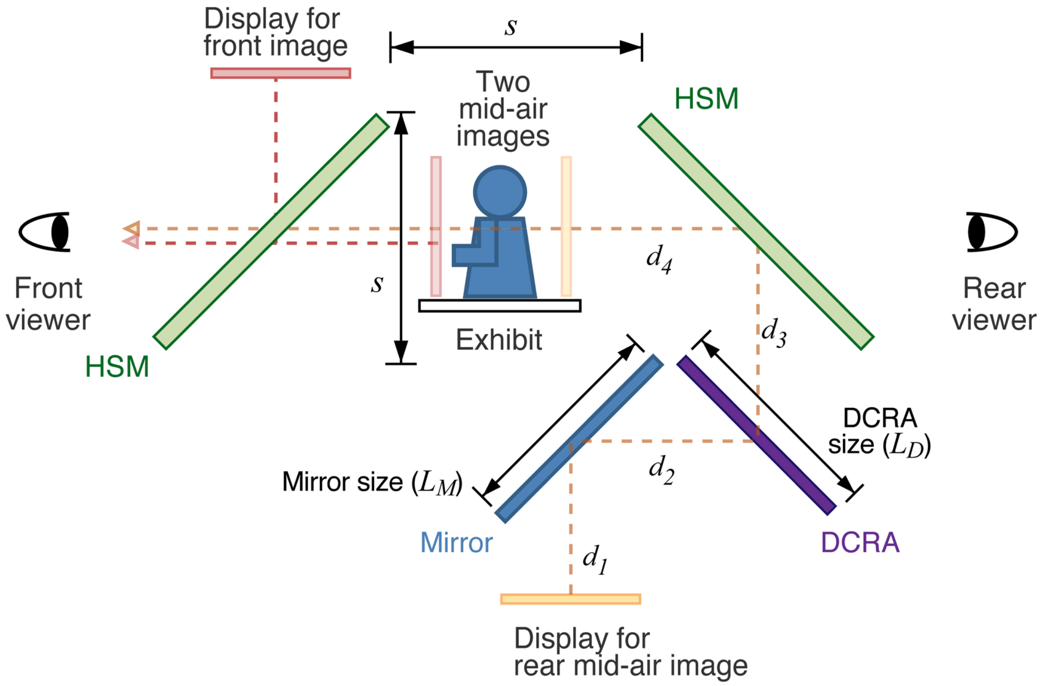


Figure 3.3: Combined imaging system using DCRA (IS2).

In order to express the size of IS2 with parameters, the exhibition room is assumed as a cube with the sides of s as in the case of IS1. Then, the maximum HSM size is set to s (W) × $\sqrt{2}s$ (H). Below two HSMs and exhibition area, two sets of mirror and DCRA are placed. From this relation, L_D and L_M can be calculated from the following equation.

$$2\left(\frac{L_D}{\sqrt{2}} + \frac{L_M}{\sqrt{2}}\right) = 3s \tag{3.1}$$

$$\therefore\ L_D + L_M = \frac{3\sqrt{2}}{2}s \tag{3.2}$$

$$L_D = L_M = \frac{3\sqrt{2}}{4}s\ (where\ L_D = L_M) \tag{3.3}$$

Thus, the maximum size of the mirror and DCRA will be s (W) × $\frac{3\sqrt{2}}{4}s$ (H). The maximum size of displays is also determined as s (W) × $\frac{3}{4}s$ (H) from L_M. From the calculation, the maximum size of IS2 can be calculated as a function of s. In order to provide visual images to four viewing zones, two sets of IS2 need to be crossed. Thus, the entire size will be $3s$ (W) × $3s$ (D) × $\frac{13}{4}s$ (H), or $\frac{117}{4}s^3(\sim 29.3s^3)$, which is a function of s. This means the volume of MRsionCase becomes approximately 30 times larger compared to the exhibition area in the center of the showcase.

By introducing a DCRA into the optical design, IS2 realizes the core requirements that MRsionCase has aimed for: multi-directional viewability, device-free usage, and high spatial consistency between physical objects and visual images.

During the design process of imaging systems, resolution of mid-air images has been an important issue. Virtual images formed by HSM have almost the same resolution to that of image source (display). Real images formed by DCRA, however, have a little lower resolution than image sources, due to discrete reflector arrays of DCRA and more complicated imaging mechanism including reflection and convergence. According to the preliminary study on mid-air image quality [54], it has been confirmed that viewers can easily recognize letters in a real image with 5 × 7 mm in size, and larger one.

3.3.4 Implementation

Figure 3.4 shows the details of MRsionCase implmenetation. In the prototype, IS1 and IS2 are combined with crossing 90° to each other as shown in Figure 3.4(a). With IS1 and IS2, users can compare the imaging positions of real and virtual images. In both imaging systems, two layers of mid-air images are formed in front of and the rear of the exhibit and thus the exhibit is sandwiched by the visual images.

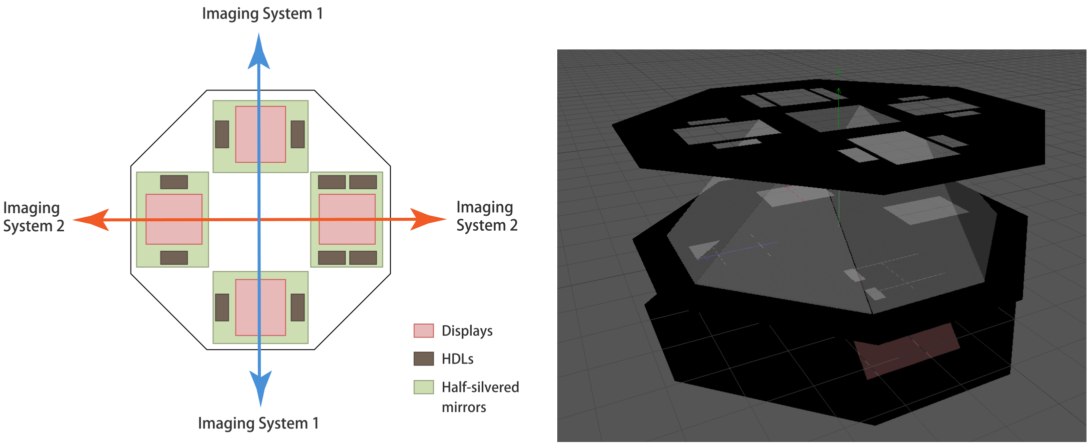
During the implementation, s and L_H are the most critical factors that affects in the system scale. Although exhibition room with larger s can accommodate larger exhibits, large HSMs (over 300 mm) are expensive and difficult to find from off-the-shelf products. However, s should be at least 200 mm, since the exhibits for the application (i.e. a clay doll and figure model) requires this room size for accommodation. From this requirement, four sheets of HSMs, produced by Edmund Optics [55], with a size of 254 (W) × 356 (H) mm are chosen due to the large size and cost effectiveness. This means $s = 254\ mm$ in the current implementation of MRsionCase.

The size of mirrors used in MRsionCase is 230 (W) × 230 mm (H). The DCRA, placed with a rotation of 45°, has a size of 213 mm (W) × 213 mm (H). This size of mirrors and DCRA used in the implementation complies to the maximum size of mirrors and DCRA, 254 (W) × 269 mm (H) calculated in Section 3.4.2. For displays, iPad2 tablet devices (197 (W) × 148 (H) mm) are used as a light source of mid-air images due to its compact size and versatile usage.

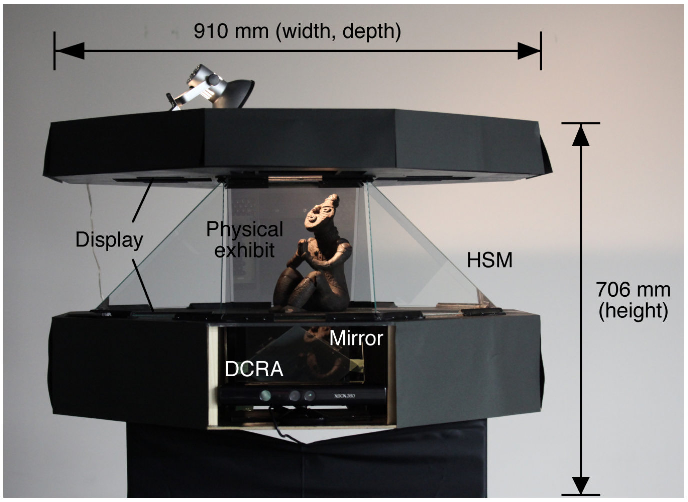
To determine the position of mid-air image layers, distance parameters, d_1 and d_2 in IS1 and d_4 in IS2, need to be considered. In IS1, d_1 was set to 196 mm since the front image needed to be placed 70 mm behind the HSM surface due to the shape of the figure model.

For rear mid-air image layers, the position of displays should be thought separately in IS1 and IS2 due to the difference of imaging mechanisms. In IS1, since the rear images cannot be placed inside the exhibition room, the images need to be placed as near positions as optically possible. Thus, Display 2 and 4 are placed just below the HSM with minimizing the distance between displays and HSMs, or $d_2 = 127$ mm. On the other hand, in IS2, the rear image layer can be formed inside the exhibition room, so that the visual images can directly be superimposed onto the exhibit. Considering the shape of exhibits, d_4 needs to be set to 207 mm, and thus d_1 results in 230 mm.

After the size determined, the 3D overview is made by CAD (computer-aided design) tools as illustrated in Figure 3.4(b). Showing windows consist of plate-shaped HSMs, and the space between HSMs are covered by triangle-shaped glasses. The space above and below the HSMs are reserved for displays and optics for real images including DCRA and mirrors.



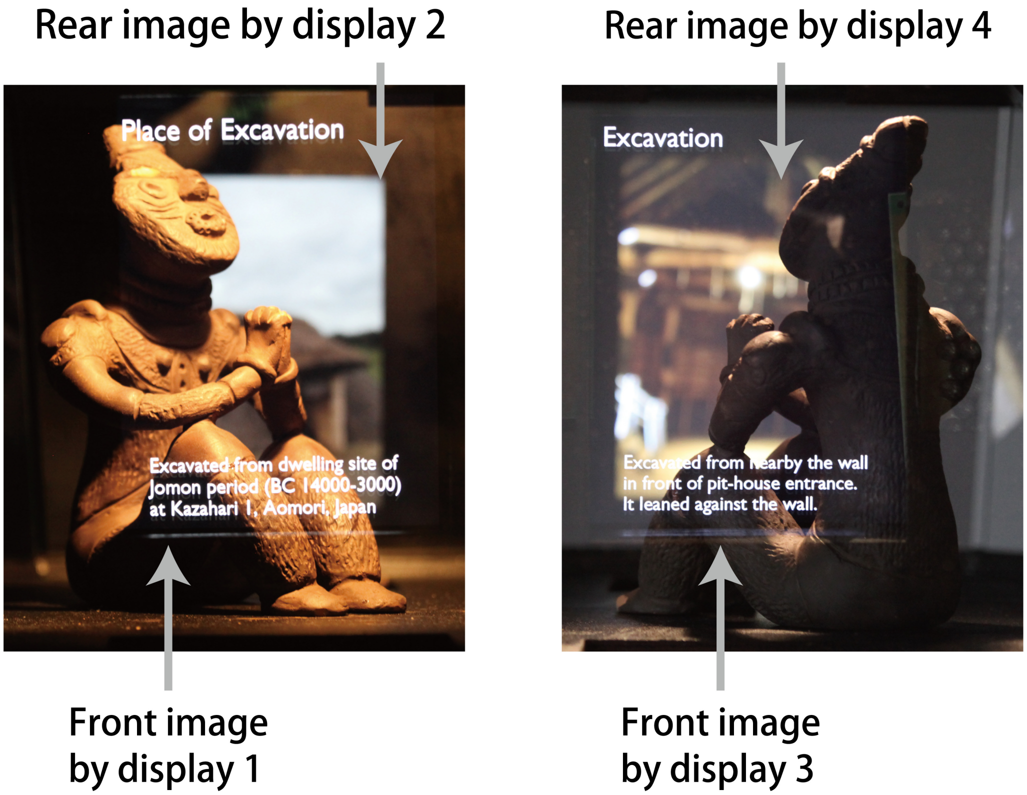
(a) Topview of MRsionCase implementation (sketch). (b) 3D overview of MRsionCase in CAD design.



(c) Front view of implementation.

Figure 3.4: Implementation of MRsionCase.

Figure 3.4(c) shows the actual implementation of MRsionCase. Housing cases for displays and imaging optics are made by wooden plates. The frames for fixation of HSMs and glass panels are designed with CAD software and originally made by 3D printers. After the complete assembly, MRsionCase has a size of 910 (W) × 706 (H) × 910 (D) mm.



(a) View from front side.

(b) View from rear side.

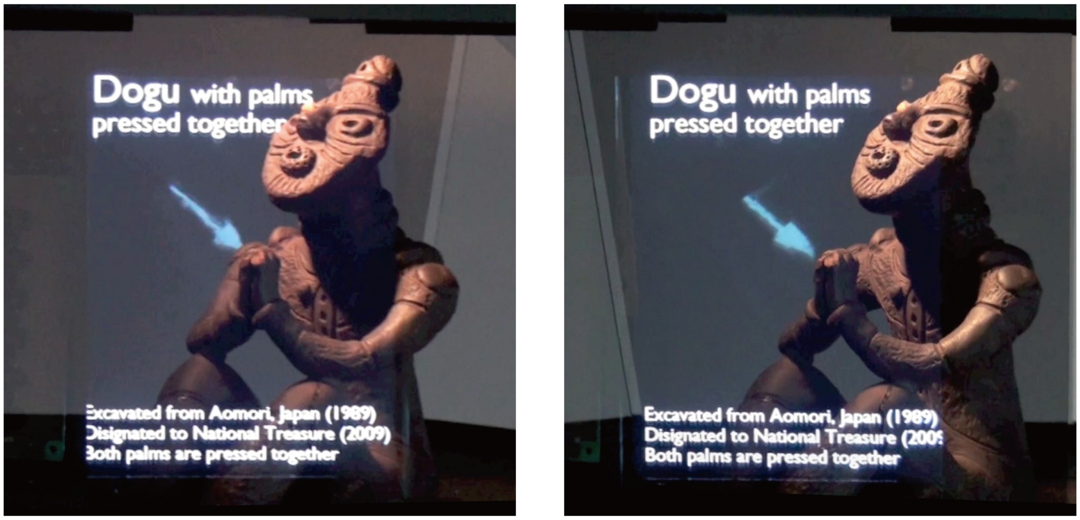
Figure 3.5: Resulting images formed by IS1.

3.4 Results

3.4.1 Optical evaluation

The resulting mid-air images formed by IS1 are shown in Figure 3.5. The descriptive text comes from the front image layer and the background comes from the rear image. Rear images are occluded by the exhibit from their physical positions.

The images formed by IS2 are shown in Figure 3.6, which are seen from the left and right, displaced 5° from the front, respectively. The description text is shown from the front of the exhibit as a virtual image formed by HSM. The arrow in mid-air is a real image formed using the DCRA from the rear of the exhibit. It is placed immediately next to the hands as a direct annotation.

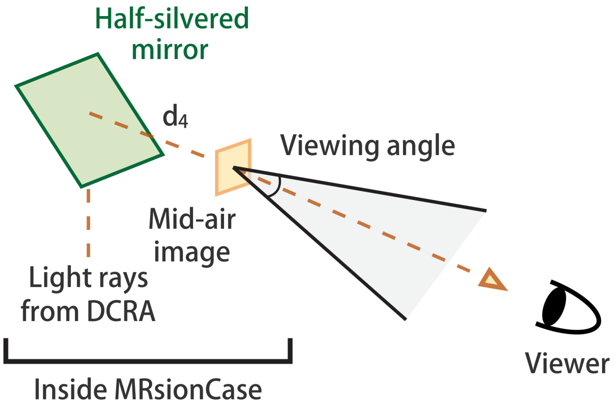


(a) A mid-air image (an arrow) viewed from 5° left from front.

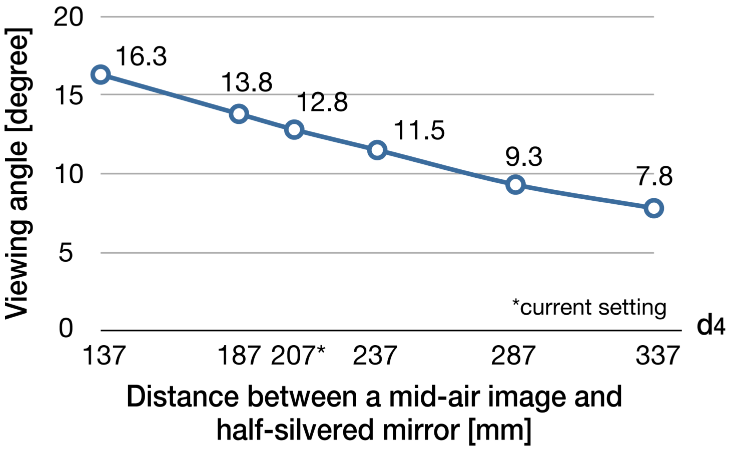
(b) View from 5° right from front.

Figure 3.6: A resulting mid-air image formed by IS2. In both images, arrow is floating immediately next to hands.

In IS2, however, there is an issue with the viewing area due to the size and shape of the DCRA. We therefore measured the maximum viewing angle of the mid-air image in order to confirm the viewing area of the real image formed by IS2 (Figure 3.7(a)). Maximum viewing angles are measured by changing the distance between a mid-air image and the HSM (d_4). As shown in Figure 3.7(b), the viewing angle becomes smaller as the mid-air images are formed with larger d_4 due to the relationship between the size of the DCRA and the mid-air image. Since d_4 is set to 207 mm at the current setting, the viewing angle of a mid-air image formed by IS2 is 12.8°. This means that the real images can be seen from a 35.2 cm range when a viewer stands 1 m away from the front HSM of MRsionCase.



(a) Viewing angle of a real image (rear image layer) formed by IS2.



(b) The measured viewing angle with different imaging positions (d_4).

Figure 3.7: Measurement of viewing angle.

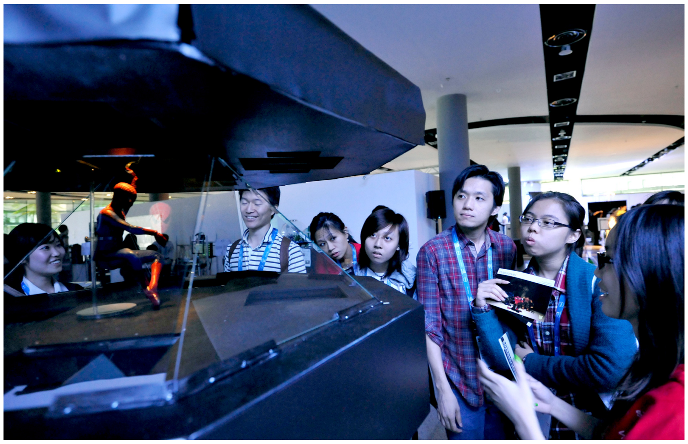
3.4.2 User study

Demonstration

MRsionCase has been presented through open exhibitions at Digital Contents Expo (DCExpo) 2012 [56] and ACM SIGGRAPH Asia 2012 Emerging Technologies [57] (Figure 3.8).

At DCExpo, we displayed a museum exhibit in MRsionCase. As shown in Figure 3.1, a *dogu*, an ancient Japanese clay doll, was placed inside MRsionCase, and related information was presented through mid-air images and directional sounds. De-

(a) Exhibition at DCExpo 2012.



(b) Exhibition at ACM SIGGRAPH Asia 2012.

Figure 3.8: Demonstration of MRsionCase.

tailed explanations of the exhibit were given in front mid-air images. The rear image layers provided direct and ambient information. An arrow pointing directly at a specific part of the exhibit was displayed in mid-air to attract attention, and background scenes were displayed to enhance the viewer's sense of reality (Figure 3.1(b)–(e)). Along with the visual images, an audio guide was presented in each direction. The sound was spatially linked with the exhibit, so viewers could easily understand the exhibit by relating the auditory information to the exhibit.

At SIGGRAPH Asia, using a Spiderman figure as an exhibit, we superimposed relevant video clips on the object and played songs through speakers (Figure 3.9).

(a) View from front side. (b) View from rear side.

Figure 3.9: Spider figure and superimposed images.

Dynamic and animated sets of visual images and auditory sounds could enhance the entertainment factor as well.

User survey

A user survey was conducted at both exhibitions for system evaluation. We focused on the effect MRsionCase might have on a person's interest in visiting a museum and his/her appreciation of the MR showcase, the readability of the mid-air images, the spatial consistency between the physical object and the superimposed images, the sound quality, and the hearing range. The items in the survey are listed in Table3.1. Items 1 to 8 are required for all respondents. Four additional optional questions were asked to obtain viewers' opinions and suggestions. The items fall into three categories: overall system (1–2), mid-air images (3–5), and directional sound (6–8). Five (1, 3,

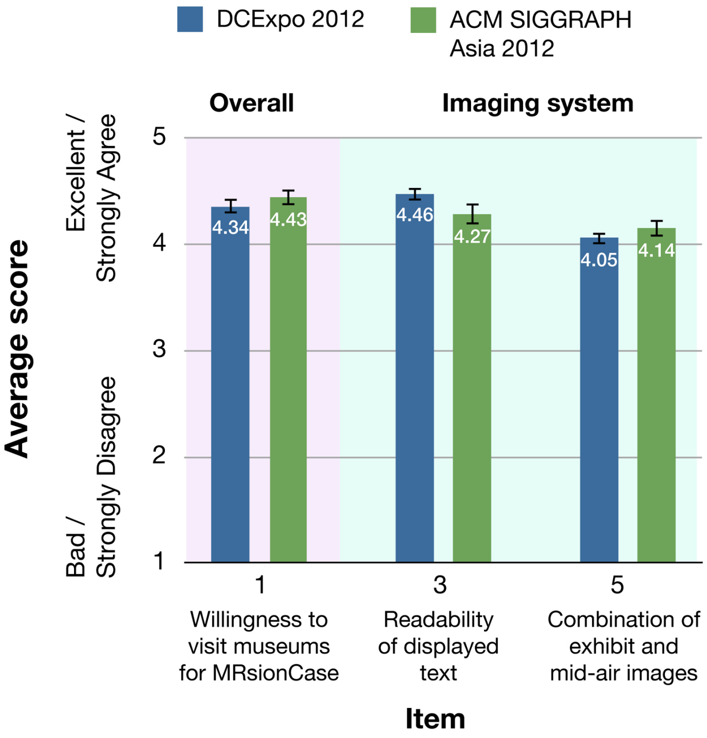
Table 3.1: Survey items.

Number	**Statements**
1	If the system is installed in a museum, do you want to go and see?
2	In MRsionCase, there are four sides (front, right, left, back) to see superimposed information. Which of them, I saw the MRsionCase from ( ) side(s) out of 4. (Please write a number.)
3	How easy or hard was it to read the letters?
4	Please give us the reason why you chose the answer for item 3.
5	Were the exhibit and mid-air images spatially well combined?
-	What makes you come here to see MRsionCase? System's appearance, Sound from this system, Crowded with people, Poster, Lab homepage, Conference map, Other (please specify)
-	Please let us know your overall impression on this system?
-	If you can use this system, what would you want to exhibit with it?
-	If you have any suggestions, feel free to tell us.

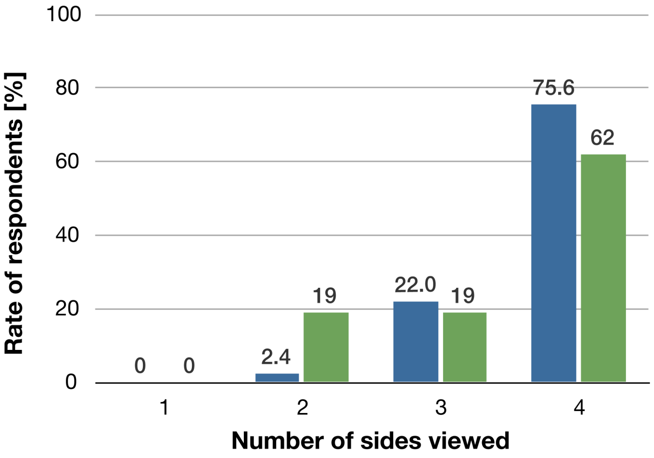
and 5–7) were based on a five-level assessment from 1 (bad / strongly disagree) to 5 (excellent / strongly agree). Item 2 required an integer number (0 to 4), and 4 and 8 required a free description.

The surveys were completed by 41 people at DCExpo and 22 people at SIGGRAPH Asia. Their age range was relatively broad (20 to 61 years of age). All were taller than 160 cm, putting them in the upper hearing range of the auditory system.

Results of required items are shown in Figure 3.10. The results from the free description items (items 4 and 8) are not plotted in the graph but are referred to during the discussion process.



(a) Results for 5-level evaluation items (item 1, 3, and 5).



(b) Number of sides viewed (item 2).

Figure 3.10: User survey results.

Figure 3.10(a) shows the average points for the 5-level based questions (items 1, 3, and 5) along with standard error bars. For item 1, the average score exceeded 4 (agree) for both events. The respondents generally agreed that they were likely to visit a museum at which MRsionCase was installed. This suggests that MRsionCase could encourage people to visit museums to enjoy an MR experience.

Items 3 and 5 focused on the display of the mid-air images. In specific, item 3 is mainly targeted at front images formed by HSM since text information is provided from front images only. The average score exceeded 4 (good) for both events. This shows that the respondents were able to easily read the text information displayed in the mid-air images. Several viewers commented that the letter size and contrast were such that it was easy to read the text. It also confirms that mid-air images formed by HSM are appropriate to display text information due to their higher resolution as we expected in section 3.4 (system design). With item 5, we aimed to evaluate overall combination of MRsionCase between physical exhibit and mid-air images, including both real and virtual ones. The average score for item 5 was ∼4 (good). This indicates that the respondents felt the exhibit and mid-air images, shown in the front and rear of the exhibit, were well combined. In other words, MRsionCase thus could provide visual information spatially well-combined with a physical object in an MR environment.

Figure 3.10(b) shows the results for item 2. In both events, most viewers (97.6 and 81.0%, respectively) viewed MRsionCase from more than two sides. The average was 3.73 and 3.43, respectively. Moreover, no viewers viewed MRsionCase from only one side. This confirms that most viewers viewed the exhibit from multiple sides by moving around MRsionCase. Such multi-directional viewability of MRsionCase might thus encourage viewers to actively enjoy museum exhibitions.

3.5 Discussion and Conclusion

It was interesting that ∼43% of the respondents (27 out of 63) mentioned “system appearance” as the factor that attracted them to MRsionCase. This means that its unique appearance in addition to the mixed reality experience might motivate people to approach MRsionCase and view the exhibit. Several respondents offered ideas for potential applications, such as digital signage to display pieces of jewelry and artwork.

The observation of viewers at both demonstrations showed us a unique thing: viewers felt it was interesting to see other viewers through the showcase. Especially, when their friends or family members were seen, viewers reacted more joyfully. It is natural that MRsionCase shows other viewers through its showcase since HSMs are transparent. Although we did not perform any quantitative analysis on this issue, it might provide viewers with motivation to communicate with other viewers during museum visits, which could have potential educational benefits.

The current design of MRsionCase has several limitations that were found independently of the user survey. First, due to the size of the DCRA device, real images formed by IS2 had a limited viewing angle (Figure 3.7(b)). Although sufficient for one viewer, a wider viewing angle is preferred for real museum environments so that more viewers can see the image at the same time. We expect to solve this problem as the process of manufacturing DCRAs improves.

In addition, since 2D displays (iPad2) are used as a light source of imaging systems, the resulting mid-air images are also only 2D images, which include only one depth layer. Changing to a 3D light source such as integral photography (IP) displays or auto-stereoscopic 3D displays would enhance the depth perception of mid-air images.

The fixed imaging position in mid-air is another issue. We are currently implementing a moving structure with a linear actuator to dynamically control the imaging position. The next version of MRsionCase will support the real-time change of imaging positions inside the showcase.

Resolution of mid-air images has revealed an important issue in MRsionCase. Virtual images formed by HSM have almost the same resolution to that of image source (display). Real images formed by DCRA, however, have a little lower resolution than image sources, due to discrete reflector arrays of DCRA and more complicated imaging mechanism including reflection and convergence. According to the preliminary study on mid-air image quality [54], it has been confirmed that viewers can easily recognize letters in a real image with 5 × 7 mm in size, and larger one. It is expected that DCRA with finer corner reflectors can improve the resolution of resulting real images.

Our device-free mixed reality showcase, MRsionCase, features a new paradigm for MR showcases by realizing four core objectives. Two mid-air image layers sandwich a physical exhibit and thus provide natural occlusion. A real image formed by a DCRA can be placed immediately next to the physical exhibit to enhance the spatial consistency between the two. From a user survey, most viewers (over 80%) saw MRsionCase from more than 3 directions (out of 4), indicating that MRsionCase would probably encourage viewers to actively enjoy museum exhibitions. Mid-air images obtained a good evaluation (∼4 out of 5) in readability and spatial combination with the physical exhibit. Most importantly, viewers do not need to wear a special device, so people feel more free to stop and enjoy an MR exhibition even in an unplanned encounter. As future work, the author will intend to improve MRsionCase by solving the aforementioned limitations. Finally, the author hopes to provide an exhibition with MRsionCase in a real museum environment.

Chapter 4

MARIO: Mid-air Augmented Reality Interaction with Objects

This chapter addresses the challenging problem on the optical design that enables direct manipulation of physical objects and that provides mid-air images with a single viewing zone. Unlike MRsionCase, physical objects can be accessed and moved by users as a tangible and interactive tool. According to users' manipulation, visual images are formed in mid-air and along with the physical objects. The visual images are required to be placed in front of imaging optics and thus formed as a real image. In order to display visual images in 3D space without special glasses, a layer of real image dynamically changes its imaging position. Therefore, the optical design needs to be able to change the imaging positions without optical distortions with providing users with enough viewing zone. The position and shape of physical objects are also detected from sensors and aligned with that of visual images to provide high geometrical consistency. The spatial registration between visual images and physical objects is critical for correct expression of geometrical relations. In order to provide a high sense of reality from an MR interface, optical consistency between mid-air images and the real world needs to be achieved. For this purpose, an artificial shadow

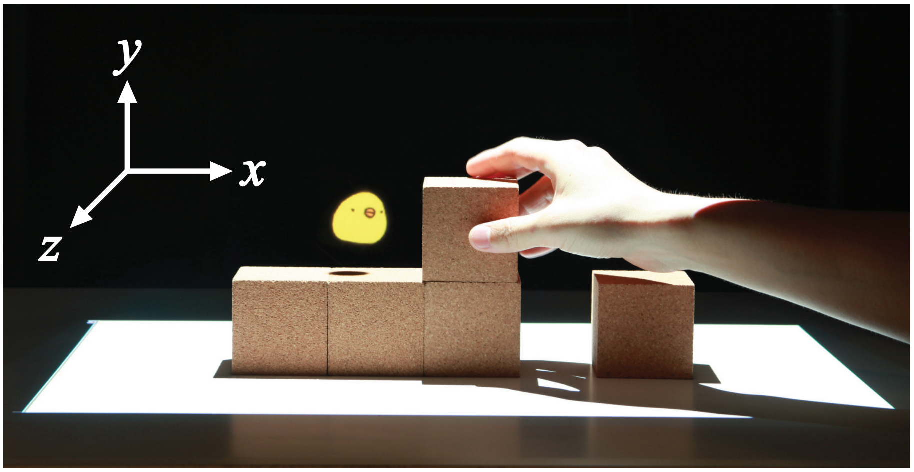


Figure 4.1: The MARIO system. Users can directly interact with physical objects and mid-air images. A virtual character (Hiyoko, a chick) is displayed in mid-air on wooden blocks. A coordinate system is defined as x for width, y for height, and z for depth.

is projected below a mid-air image as if the shadow is formed by the mid-air image. In this chapter, the author aims to propose an MR interface for a single user that the user can interact with a mid-air image by manipulating physical objects in 3D space. As a usage example, an entertainment system, MARIO, is implemented and evaluated the effectiveness of system design through optical experiments and public exhibitions.

4.1 Introduction

What if a game player could make Mario [58] run a course of his or her own design made by arranging wooden blocks in the real world and what if they could control Mario's movements through means other than game pads? If such features could be realized, game players could immerse themselves in the game space and thus enjoy interactions with game characters in more intuitive ways.

Aiming at an attractive game experience, the author has imagined a scenario where users can play with physical objects and characters moving through the air.

A user comes to our system with no preparation. This is very similar to a situation in which people play with game machines in game arcades. Blocks are placed on the game table, instead a controller. On this table, the user freely arranges the blocks and makes a 3D structure as he/she likes. Since the blocks are completely ordinary ones, the user does not have any difficulty in handling them. The shape, position, and orientation of the blocks, which the user has arranged, are detected by a sensor and reflected in the game world. A game character appears in mid-air and jumps on the blocks. The character will bounce off the surface of the block and the orientation of the blocks's surface controls the direction of the bounce.

On the basis of this scenario, a novel MR interface called "Mid-air Augmented Reality Interaction with Objects (MARIO)" is proposed. MARIO realizes a direct user interaction between visual images and physical objects beyond the confines of a physical display (Figure 4.1). An image is displayed and moves in 3D (xyz) space. Users of MARIO are freed of the need for physical displays during their interaction with mid-air images. Moreover, MARIO does not require users to wear additional devices, making it very easy to use. As in most MR systems, it also maintains high levels of temporal, spatial, and optical consistency between images and physical objects, and users can experience a strong sense of reality and immersion from the game.

The contributions in MARIO are as follows:

(1) a new mid-air imaging display shows images that freely move in a mid-air in a 3D space measuring 350 (W) × 300 (D) × 250 (H) mm without the need for special glasses;

(2) an intuitive MR application enables users to control mid-air images with everyday objects or their hands. Shadows are cast by the images to give viewers the sense of reality;

(3) the popularity and effectiveness of MARIO as an entertainment system was proven during a six-month public exhibition.

4.2 Proposal

MARIO is aimed to suggest a new MR interface for a single user that achieve three core goals: 1) glasses-free mid-air imaging display in a 3D space, 2) enhanced sense of reality from a mid-air image, and 3) interaction design for practical usage. These three goals accord with the requirements of MARIO.

4.2.1 Glasses-free mid-air imaging display in a 3D space

Since the MARIO supports direct manipulation of physical objects in 3D space, the visual images also need to be placed in 3D space. Moreover, the position and shapes of the physical objects is changed with users' manipulation so that the position of mid-air images also needs to be changed. Stereoscopic displays can display visual images in 3D space with binocular parallax. However, this method has limitations that are not suitable for MARIO: Users need to wear 3D glasses to see the visual images and the position of visual images are far from the users' eye. The author chose a 2D image layer that is formed by imaging optics instead of 3D displays. The layer of 2D image moves in mid-air along with a depth direction. Although 3D depth perception cannot be implemented, this approach can display a 2D visual image in 3D space without the need for 3D glasses. The author believes this approach enables intuitive experience and dynamic visual expression of MR interfaces.

4.2.2 Enhanced sense of reality from a mid-air image

The author believes that mid-air images can provide higher sense of reality when they have a close relation with the real world. Especially, the accordance of position and illuminating conditions between mid-air images and physical objects is critical and effective. The position of a mid-air image formed by imaging optics needs to be understood in the real world space through coordinate transformations. As for the illumination, a controlled light source is used and an artificial shadow is cast based on the calculation with the positions of mid-air image and the light source. The mid-air images can provide users with high sense of existence and thus users may treat the mid-air images as physical objects. This enhanced sense of reality from a mid-air images leads users to more immerse into MR interfaces without the need to distinguish visual images from the real world.

4.2.3 Interaction design for practical usage

The MARIO is aimed to provide a practical usage example that people can readily experience. For this purpose, MARIO is designed not to require users any preparation before participating the interaction. Users do not need to wear 3D glasses to see the visual images. Since the position and shape of physical objects are detected with an infrared-based sensor, most everyday objects can be used in general. The contents of visual images are carefully selected with an animated character that can draw users' attention and motivate the participation. The combination of proper visual contents and the merits of optical design enhances the practicality of user interaction in MARIO.

4.3 System Design

Figure 4.2 illustrates an overview of MARIO. The entire system consists of three parts: mid-air imaging display, object detection, and shadow projection. The mid-air imaging display, which combines a display, linear actuator, and real imaging optics, forms an image by using geometrical optical phenomena such as reflection and con-

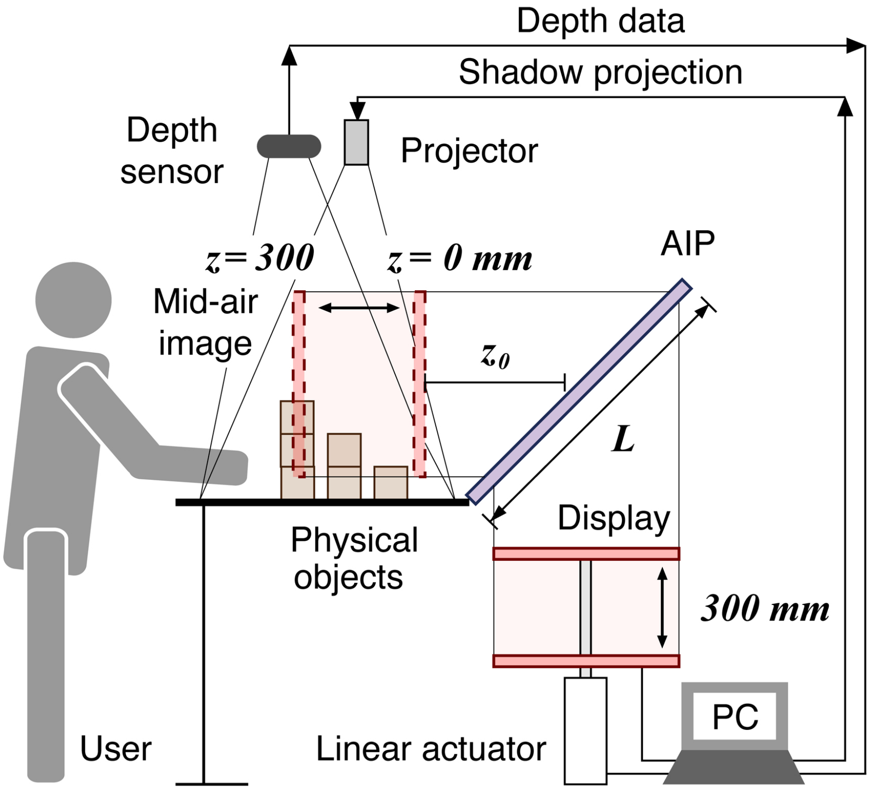


Figure 4.2: Overview of MARIO.

vergence of light rays. Object detection involves constructing a terrain model based on the shapes and orientations of everyday objects that users can manipulate and arrange in the real world. Shadow projection from a projector casts an artificial shadow of mid-air images to convey a sense of reality to viewers.

4.3.1 Mid-air imaging system

The priority of making a walk-up-and-use MR interface and displaying images in 3D space led us to devise a mid-air imaging display combining a LCD monitor, linear actuator, and real imaging optics.

As discussed in Section 2.2.2, an AIP is used because of its distortion-free imaging and advantage in positioning the image. Based on the optical property of AIP, the configuration of display and AIP is shown as Figure 4.3. To form an image in the

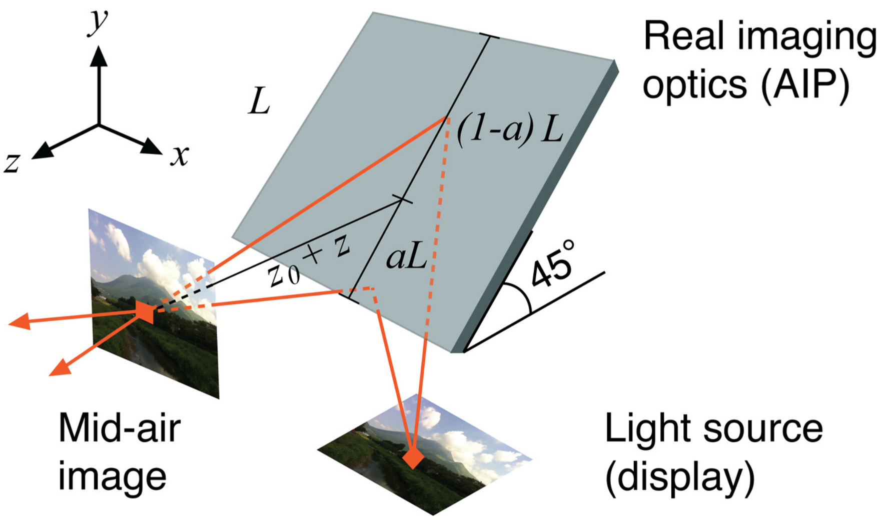


Figure 4.3: Configuration of display and AIP in mid-air imaging display.

xy-plane, the AIP device is obliquely placed at a 45° angle from the table's surface. The position of the mid-air image is determined by the distance between the light source and AIP. A display is placed below the AIP as a light source.

Since the AIP does not have a fixed focal length, the imaging position can easily be moved by changing the distance between the light source and AIP without changing the size of the resulting images. This means that the imaging position can be changed in xyz-space when the display moves to different heights.

4.3.2 Object detection

To enable people to use everyday objects as controllers, the exact status of those objects has to be detected. In MR applications, various sensing devices, from cameras to laser scanning sensors, have been used to detect the shape or orientation of physical objects. In particular, infrared depth sensors such as Microsoft's Kinect [59] have been used to detect physical objects in the real world.

Due to its accessibility and low cost, in MARIO, a Kinect was chosen as a depth sensor for object detection. Raw depth images are acquired using a Microsoft Kinect

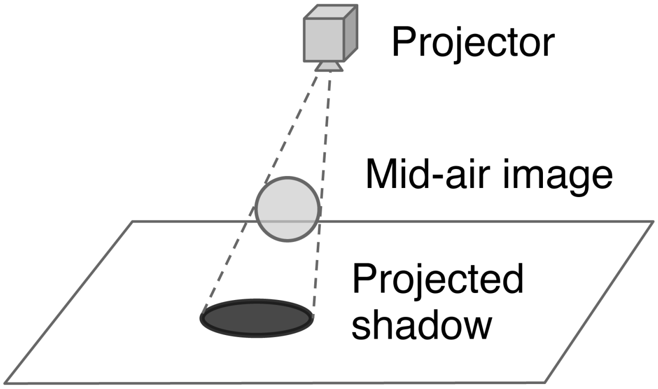


Figure 4.4: Shadow projection scheme.

SDK. Each pixel of the raw images represents the distance from the sensor to the object in millimeters. From the depth image, the highest point was detected and sent to the mid-air imaging display to form the mid-air image.

4.3.3 Shadow projection

Since the mid-air image formed by the AIP has no physical shape, users had difficult perceiving the exact positions of the image with it. In particular, most viewers pointed out that the depth (on the z-axis) of the mid-air images was more difficult to perceive than the other directions. From this feedback, the author searched for a way to give users a clue that would enable them to easily recognize the positions of mid-air images.

In virtual reality (VR) and AR studies, a shadow is often used to give images a sense of reality. Naemura et al. proposed the virtual shadow concept wherein projectors project an artificial shadow on images [17]. Sugano et al. confirmed the effectiveness of such projected shadows in helping people to perceive the depth and height of images overlaid on the real world [60]. Referring these studies, the author decided that the mid-air images should cast an artificial shadow so that their positions could be more easily recognized. This artificial shadow also enhances the optical consistency between the mid-air images and the real world.

For an artificial shadow, the author used an overhead projector with projecting an image (a black dot) below the mid-air image. As shown in Figure 4.4, the projector was considered to be a point-like light source. This helps the projected shadow have consistency with natural shadows. The shadow was calculated by blocking light from the projectors with the cross-section of the mid-air image. The 2D image from the projector cast an artificial shadow on the top of any physical object regardless of the shape of the projection surface.

4.3.4 Implementation

Figure 4.5 shows the implementation of MARIO. The author aimed to implement an application in entertainment purposes with MARIO that draws the attention from children. In specific, the children with a height between 120 to 140 cm was targeted as the main users of MARIO. Based on the users' height, the system size is determined as follows.

In the front, a table (880 (W)$\times$600 $\times$900mm (D)) is prepared as interaction area with enabling users to manipulate physical objects on it. The AIP (AIP-350, 350 (W) $\times$ 350 (D) $\times$ 3 mm (H)) is placed behind the table and forms a mid-air above the table. As a light source, a 19-inch color LED-backlit display (EIZO S1903-TBK) is placed below the AIP device. The maximum size of mid-air image that can be formed in MARIO is 350 (W)$\times$ 247 mm (H). This image size can provide enough viewing range to the targeted users.

As for the depth movement, a mid-air image is designed to move in 300-mm range. Considering the arm reach of the target users, the author considered this range provide an enough range in depth direction. The imaging position can be changed when the distance between the AIP and the display. To move the display in the vertical direction, a linear motor actuator (RCP4-RA5C) is connected to the display. The maximum moving stroke of the linear actuator is 300 mm to enable a

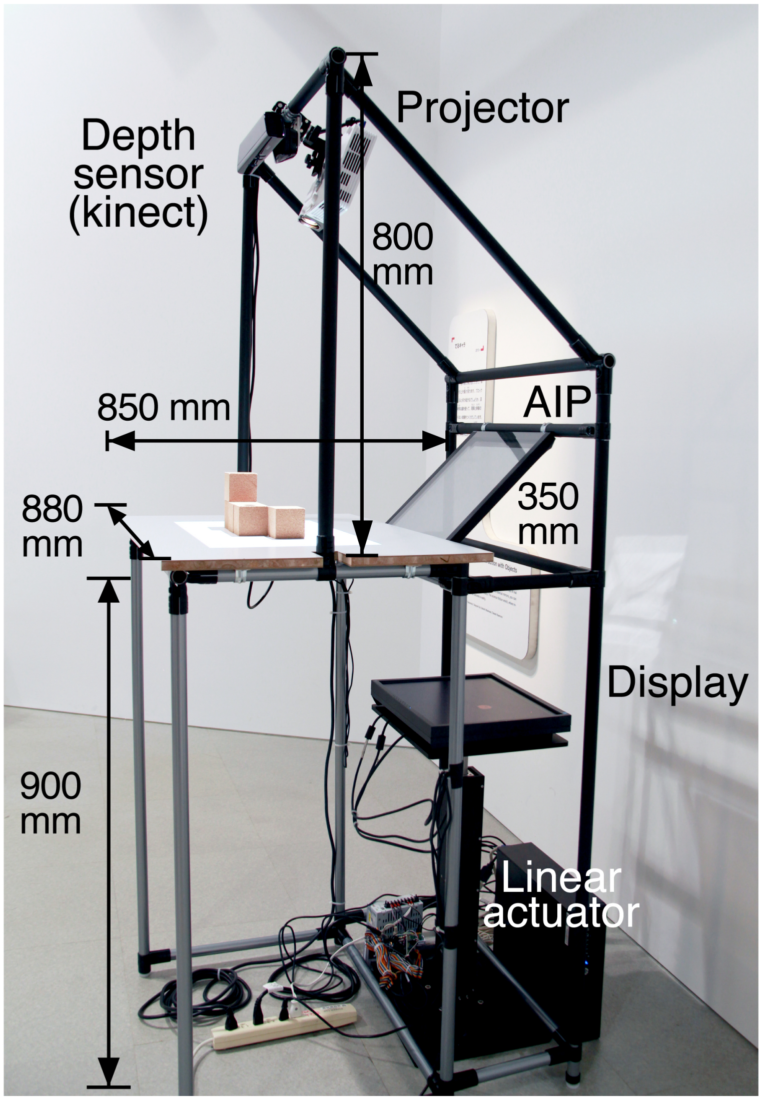


Figure 4.5: Implemented MARIO system.

mid-air image to move 300 *mm* in the depth direction (z-axis). The maximum speed of the actuator is 450 mm/s, but the actual moving speed is set to 300 mm/s to prevent physical shocks to the display and to reduce vibration during movement. At this speed, it approximately takes only one second for the actuator to move through the maximum stroke (300 *mm*). The error in the position during movement is much less than 1 *mm* (0.01–0.02 *mm*).

Although the combination of a display, linear actuator, and AIP can effectively form a mid-air image in 3D space, occlusions can be a problem when physical objects are placed between the images and the AIP. Since the mid-air image is formed by converging light coming from the AIP, it is impossible to place it in front of a physical object which blocks the light path. As a practical solution to this problem, the author developed a software adjustment by imposing a rule on the imaging positions: mid-air images are placed at the highest point of the physical objects.

For sensing the shape and position of physical objects on the table, the depth sensor (Kinect) is suspended 800 mm above the table. The object detection area is the same size as the display area. If objects or hands enter this area, their shapes and orientations can be estimated from the depth data.

For shadow projection, a LED projector (BenQ GP10) was installed at the top of the system along with the Kinect sensor. The projector was used as the sole illuminating source in MARIO system, since the exterior case of the system blocks light from outside (e.g. environmental light). Thus, objects in the interaction area (table) also had their own natural shadow generated by the light from the projector.

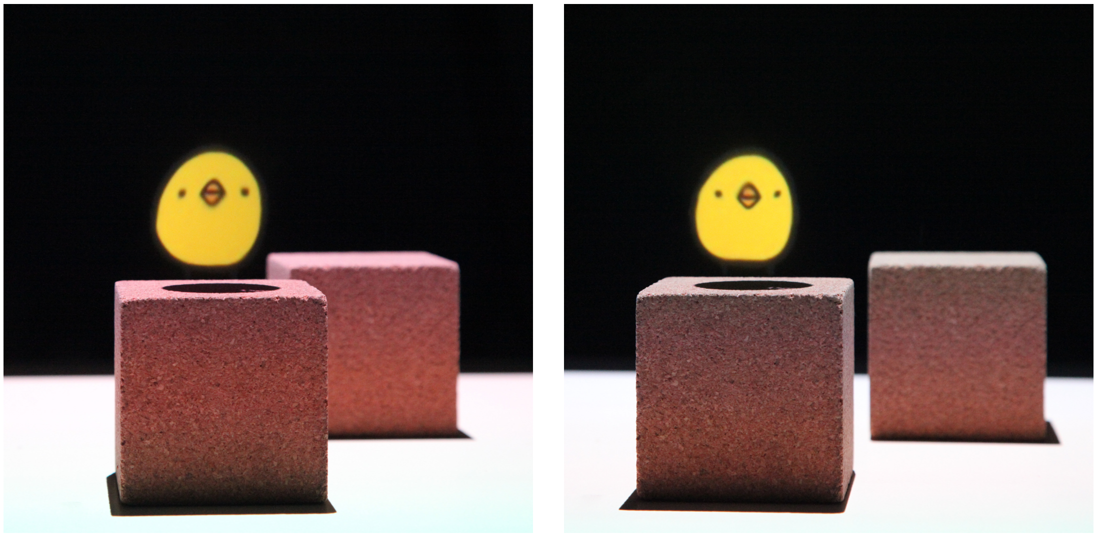
4.4 Results

Several experiments were performed on the major features of the MARIO system: the mid-air images, depth movement, shadow projection, and viewing angle.

4.4.1 Optical evaluation

Mid-air images

Figure 4.6 shows a mid-air image formed by the display. The images are those seen 5° to the left and right of the center line and they confirm that the image appears to be floating in mid-air. Binocular parallax gives the image a particular depth in mid-air.



(a) Left view (b) Right view

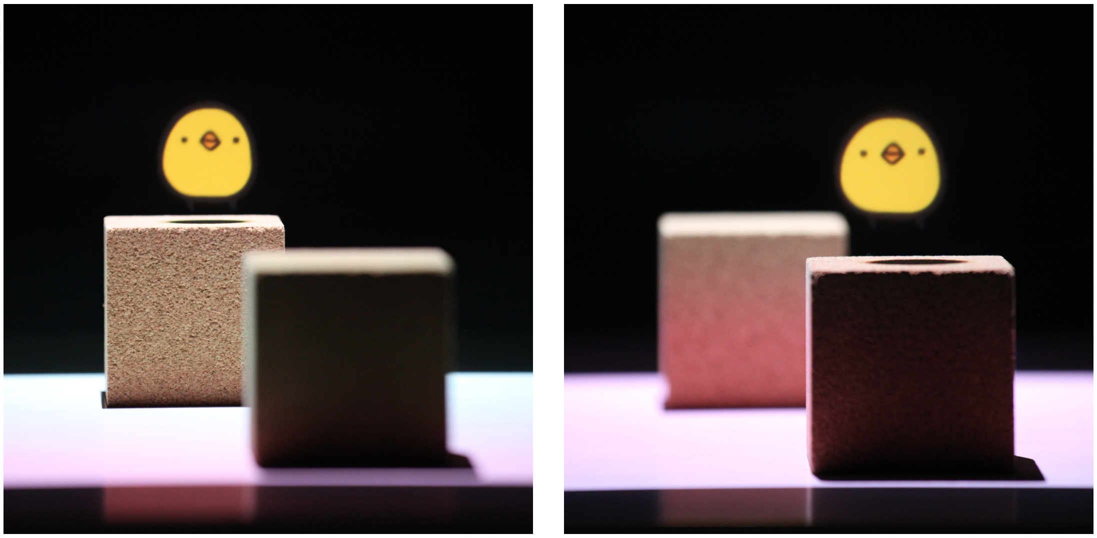
Figure 4.6: Stereo view of mid-air image. Each figure shows a view 5° to the left or right of the center line. The stereoscopic view makes the resulting image appear to be floating in mid-air.

Depth movements

The mid-air image also can change its 3D coordinates (xyz-space). Figure 4.7 shows such a change in the depth direction (along the z-axis). The image moves from the back ($z = 10\ cm$) to the front ($z = 30\ cm$). The mid-air images were taken by changing the focus of the camera. Comparing the block and mid-air image, one can see that the mid-air image is formed at different depth positions.

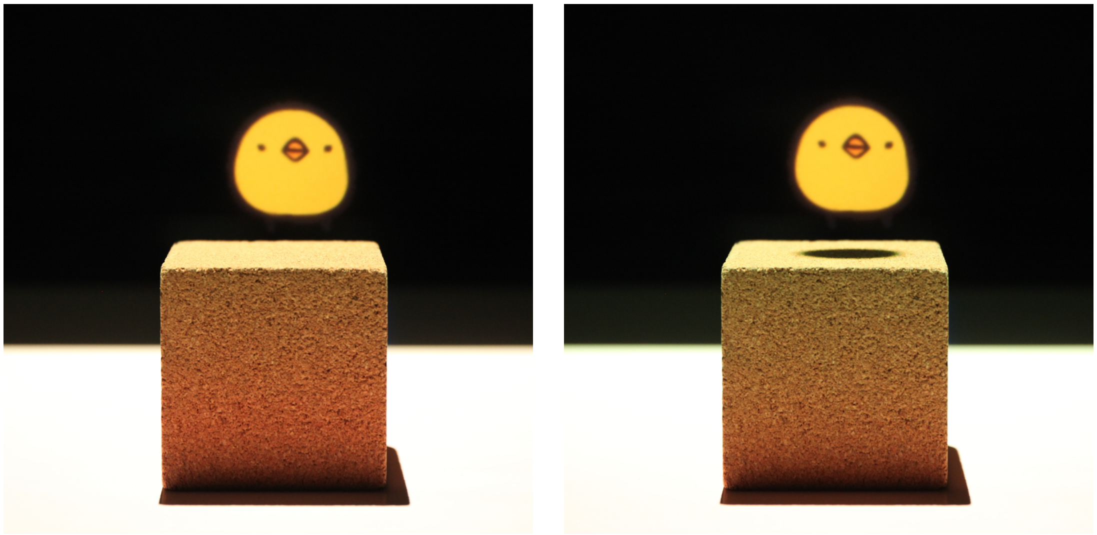
Shadow

Figure 4.8 shows the results of the shadow projection. Comparing the with and without cases, it becomes apparent that the mid-air image with the projected shadow seems more natural, and it is as if the image is really above the block. Since the illuminating conditions for the wooden blocks and mid-air images are identical, the resulting shadows are consistent with each other.



(a) $z = 10\ cm$ (b) $z = 30\ cm$

Figure 4.7: Mid-air images formed at different depths. The change in the depth direction is apparent by comparing the focuses of the images.



(a) No shadow projection (b) With shadow projection

Figure 4.8: Impression given by projected shadow.

Viewing angle

From the geometric relation between display and AIP (see Figure 4.2 and 4.3), viewing angle can be expressed as an arctangent of the ratio of the AIP width (L) to the distance from the mid-air image to AIP ($z_0 + z$) as Equation 4.1 and 4.2.

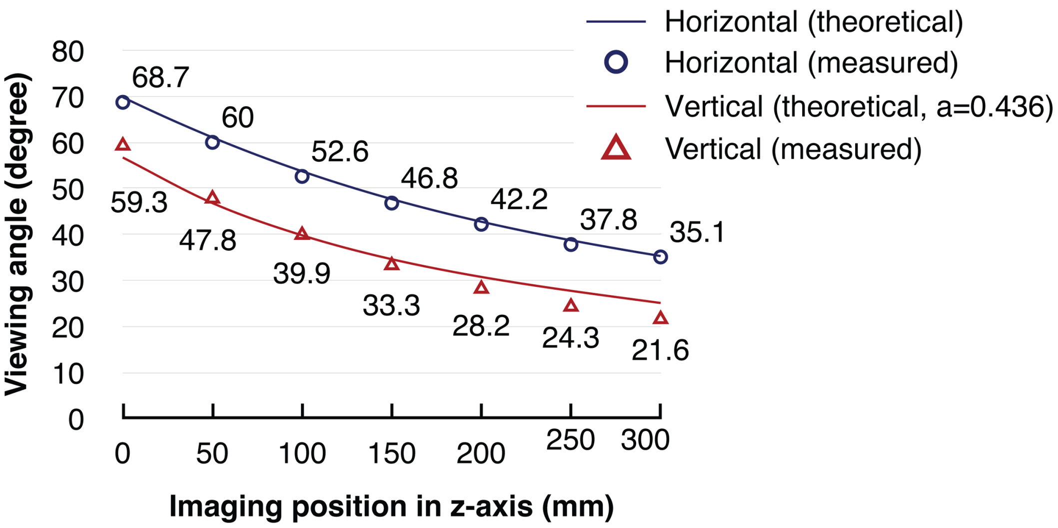


Figure 4.9: Viewing angle in horizontal and vertical direction versus imaging position.

$$horizontal\ viewing\ angle = 2\arctan(\frac{\frac{L}{2}}{z_0 + z}) \tag{4.1}$$

$$vertical\ viewing\ angle = \arctan(\frac{\frac{aL}{\sqrt{2}}}{z_0 + z - \frac{aL}{\sqrt{2}}}) + \arctan(\frac{\frac{(1-a)L}{\sqrt{2}}}{z_0 + z + \frac{(1-a)L}{\sqrt{2}}}) \tag{4.2}$$

Thus, it could be expected that viewing angle decreases with larger distance (z). To confirm the viewing range of MARIO system, the viewing angle of mid-air images was measured with different z values.

Figure 4.9 shows the measured viewing angles of mid-air images. The image of Hiyoko (30 mm) can be seen within 35.1–68.7° in the horizontal direction and 21.6–59.3° in the vertical direction. These results indicate that a mid-air image formed at $z = 150\ mm$ can be seen inside a 799 (W) × 552 (H) mm viewing range by a viewer standing 600 mm from the MARIO system. This result also indicates the MARIO system provides a wider viewing area to viewers standing farther from it.

4.4.2 User study

A user study was conducted through public exhibitions to evaluate the MARIO system as an entertainment system.

Application

Through MARIO system, the author created an interactive application in which users make physical terrain with wooden blocks and a character appears in mid-air jumping around the blocks. The application was mainly intended to show the feasibility of our user scenario. A yellow chick character, named "Hiyoko" ("Chick" in Japanese), appears in mid-air to run across and jump over wooden blocks stacked by the user as physical terrain (Figure 4.1). The character appears to move back and forth at different depths when the imaging position is changed with an actuator (Figure 4.10). The movements and positions of the character change in accordance with the shape of the terrain, so users can play with Hiyoko simply by arranging the blocks. To avoid occlusions when the character could not be formed in front of the wooden blocks, it can jump to the highest point on the table. If the terrain is changed, Hiyoko jumps to the next point along the path of a parabola connecting the current and next point. The animation sequences of Hiyoko are designed so that they look smooth, natural and charming. In specific, the durations for each jump sequence are carefully tuned, since it looks boring when the duration is too long. From several trials, every duration was set within 4 seconds.

Exhibition: in-situ evaluation

To evaluate the effectiveness of MARIO as an entertainment system, it was publicly exhibited at different venues: a tech show (Innovative Technologies 2013 [61]), an academic conference (ACE 2013 Creative Showcase [7]), and a public museum (The National Museum of Emerging Science and Innovation, Miraikan [62]).

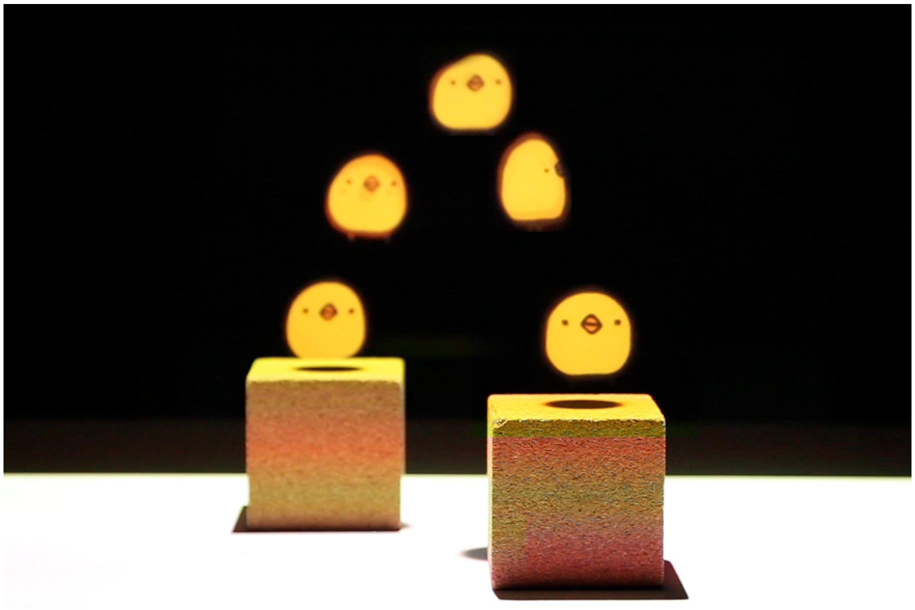
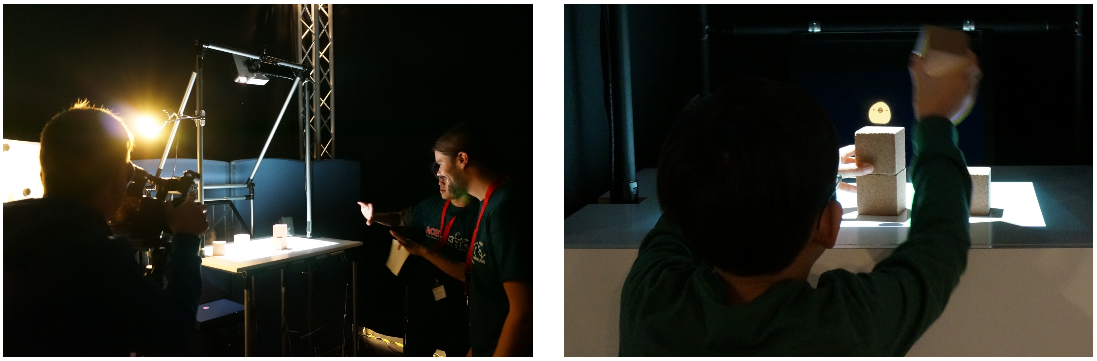


Figure 4.10: Hiyoko jumps on blocks at different depths along a parabolic path. The character moves energetically together with animation effects. (This figure was reconstructed by overlaying frames extracted from a video clip.)



(a) Demonstration at ACE 2013 creative showcase.

(b) Exhibition in National Museum Emerging Science and Innovation (Miraikan).

Figure 4.11: The MARIO system at public exhibitions. In Miraikan, kids played with the Hiyoko character by manipulating wooden blocks and having the image hop above them.

Innovative Technologies is an event, held by the Japanese Ministry of Economy, Trade and Industry, to introduce emerging technologies. Visitors mainly have engineering and technology backgrounds. ACE is an international academic conference that focuses on entertainment-related interfaces between computers and people. Most ACE participants have academic backgrounds and are interested in factors that affect entertainment. On the other hand, Miraikan is a popular public museum of science and technology. It attracts people, including small kids, students, adults, and elderly people, from all over Japan. Our exhibition at Miraikan lasted for six months (171 days), and over 129,000 people visited it.

Analysis of system usage

Since the author thought the frequent usage is very important factor for entertainment systems, the usage of MARIO system was analyzed from the public demonstration at Miraikan. Position data was logged that represented the location of the mid-air images in 3D space during the users' interactions. The logged data contained position values in millimeters (x, y, and z-axis) together with timestamps. The range of imaging positions was the same as the display area (x=0–350 mm, y=0–250 mm, and z=0–300 mm). Each position data was automatically saved when the mid-air image made an action such as moving or jumping. Data was collected for 27 days (10 am–5 pm, 2013/12/11 to 2014/1/13 except holidays).

Due to the noise in the depth data acquired by the Kinect sensor, the character sometimes jumped in the same place or moved slightly (a few millimeters) even if there was no user input. Thus, a threshold was set to determine the effective movements made by users. Since the Hiyoko character measured about 30 mm, it was assumed that only movements over half of its size (15 mm) were intended by users. After filtering, the total number of effective movements during the 27 days was 72,051, and the average distance of each movement was 106.2 mm.

In specific, the user's interaction pattern was estimated by focusing on the time intervals between movements. Figure 4.12 illustrates the accumulative percentage of movements by time interval. It shows that over the half of the total interactions (52.3 %) occurred within 4 seconds. Moreover, the time intervals of 89.6 % of the movements were less than 15 seconds. These results indicate that most user inter-

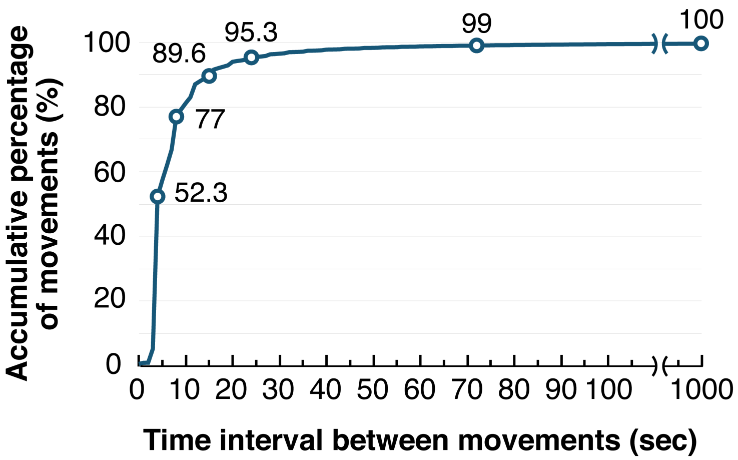


Figure 4.12: Accumulative percentage of movements by time interval between movements.

actions occupied a short time interval. Considering that the animation sequences of jump is up to 4 seconds, the author believes that these short time intervals indicate the popularity of the MARIO system since users continuously played with the system during most of the exhibition time.

4.4.3 Findings from exhibitions

Most viewers could see the Hiyoko character floating in mid-air, and also perceive the exact imaging position. Moreover, when viewers noticed a mid-air character could jump to different depths, they tried to arrange blocks at the front and the rear of the display area to confirm the change in depth themselves.

Meanwhile, as for the shadow projection, many viewers seemed to think the artificial shadow was a natural shadow. Only a few people noticed that the projected shadow was not natural since the mid-air image did not have a physical shape. The author demonstrated two situations in which the shadow projection was on and off by turning the projector on and off. Although a quantitative analysis was not performed, it was empirically confirmed that the projected shadow helped users to detect mid-air images from the comments of the viewers. After experiencing MARIO system, many viewers commented that “thanks to the (projected) shadow, the mid-air character looked as if it really existed as a living thing.” and that “projecting a shadow was a clever idea.” In the future, the author planes to perform a more quantitative experiment to confirm the effect of projected shadows on the perception of mid-air images.

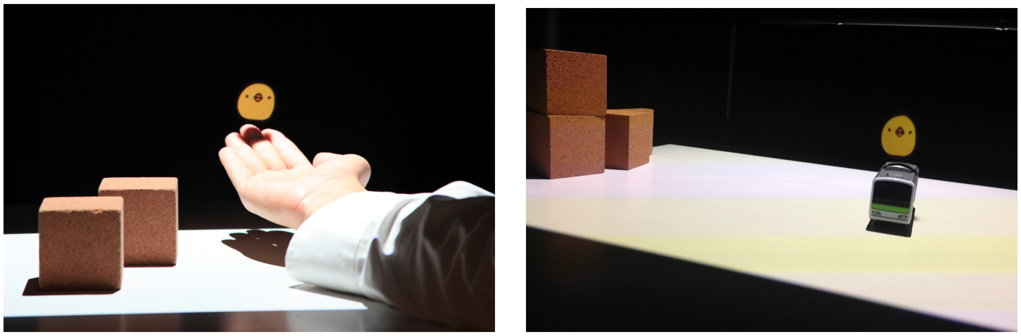
It was also found that viewers had little problem viewing mid-air images with the proposed viewing angle of MARIO system. They were able to see the mid-air character by adjusting the positions of their heads into the viewing range.

During the six-month public exhibition at Miraikan, a lot of feedback was collected from users, including people of different professions, ages, and nationalities. In addition, Miraikan’s professional staff, who operated the exhibition, made daily reports during the whole exhibition period (171 days).

These reports showed that MARIO was very popular with age groups ranging from young kids to elderly people (60–70’s). There were often waiting lines, sometimes lasting a couple of hours for using our system. The exhibition staff had to put up partitions or change the route of the waiting line to manage the large number of visitors. In specific, it was the popularity and waiting lines of MARIO system that the exhibition staff most frequently reported (54 out of 171 days, 31.6%). These results correspond with the results from time interval analysis.

4.5 Discussion and Conclusion

MARIO system was aimed at providing a new entertainment experience through a virtual character, which appears in the real space not inside the display. The author believes that the distance between the character and users should be shorter for closer and more intimate interactions.



(a) A user plays with the Hiyoko character on his hand.

(b) A kid makes Hiyoko take a ride on his toy train.

Figure 4.13: Example of MARIO being used during an exhibition. Users played with the character in unique ways that we had not anticipated in our interaction scenario.

Some kids were witnessed saying “bye bye” and waving their hands at Hiyoko after their play. These actions did not seem forced or taught by their parents or other instruction; rather they seemed natural as if they had just said goodbye to their friends. Some people even flicked it off when the character had jumped on their hands or arms.

In addition, many people had unique ways of interacting with the MARIO system, ones which we could not expect from the user scenario. Some users stretched their hands to grab or touch the mid-air character. One kid even put his own toy train on the table and let Hiyoko ride on it (Figure 4.13). Using everyday objects as interaction keys increases the flexibility of the user interaction and increases the variety and novelty of actions.

Moreover, many kids tried to “smash” the character by swinging blocks at it. However, the character always escaped and jumped to a higher point (e.g. on the blocks or player’s hands). Although Hiyoko was never caught or smashed, its ability to escape seemed to intensify the kids’ motivation to chase and catch it. Seeing a kid continuously chase and try to catch a mid-air character was really interesting to the author.

Many viewers could actively interact with Hiyoko character in very unique and friendly manners. Thus, the author believes that characters outside the display are effective as an entertainment element in other character-based entertainment systems.

After the exhibition at Miraikan, there was an opportunity to hear some opinions of the Miraikan staff about our system. The chief director of the exhibition, in particular, pointed out that MARIO did not need verbal explanations, and because of that, many foreign visitors who did not understand Japanese could easily experience MARIO for longer times than they could with other exhibitions.

The MARIO system seems complicated because of its combination of imaging optics, depth sensor, and projector. However, the user interaction is straightforward: arrange wooden blocks, then see and play with a character appearing in mid-air. The author believes such intuitiveness is essential when people play with an MR interface with an entertainment purpose. MARIO provides an intuitive interaction and its exhibitions confirmed that it is easy to use.

After all of the exhibitions had completed, the error was investigated between the position data obtained by the depth sensor and the actual imaging position in order to see how precisely the mid-air images were formed in the 3D space. The error was measured at 125 points within a cubic volume measuring $200 \times 200 \times 200\ mm$. The average error was $1.54\ mm$ in total, and $1.19\ mm$ at the 27 center points. This result shows that the spatial precision of mid-air images is high enough for our application, since the mid-air images are $30\ mm$ in size and the wooden blocks are $65\ mm$ cubes. The coordinate settings and calibration process in MARIO system is described in Appendix C.

The experiments and user study also revealed some limitations of the current MARIO system. Especially during the public exhibitions, many viewers pointed out issues with mid-air images regarding depth perception, viewing angle, and spatial density.

First of all, a 2D flat panel display was used as a light source to form the mid-air image. Therefore, the mid-air images were 2D images as well. In contrast, 3D mid-air images would have conveyed a stronger sense of reality. Autostreoscopic displays or integral photography (IP) 3D displays could be used to maintain the walk-up-and-use property.

Moreover, MARIO used a single display, so that only one image layer could be shown in mid-air at one time. Having a more independently controllable display such as InFORM [63] would improve the spatial density and allow multiple characters to appear together in mid-air.

During the exhibition at Miraikan, many viewers and staff pointed out that the dependence of the viewing angle on the viewer's height was a big problem. Since our design of the mid-air imaging display was only suitable for forming an image in an xy-plane along the z-axis, the best viewing position was limited to the front. If the angle of the AIP and display were to adaptively vary according to viewer's height, the viewing angle issue will be alleviated.

MARIO implemented a new MR interface by connecting mid-air images and physical objects in 3D space. Its optical design can form a mid-air image in 3D space that is 350 (W) $\times$ 300 (D) $\times$ 250 (H) mm in size. Through the use of a display, linear actuator, and real imaging optics, MARIO can make mid-air images freely move in the real world. Since the real imaging optics are free of distortion, the mid-air images had positioning errors of less than 2 mm. An artificial shadow was also projected to give a sense of reality to the mid-air images. Moreover, these features do not require the viewer to wear any special devices, and this allows them to intuitively interact with mid-air images.

The results of the user study on the public exhibitions confirmed the popularity and effectiveness of the MARIO system as an entertainment system. MARIO was popular with people of all ages, gender, and nationalities. The interaction log revealed that most user interactions were continuously concentrated within a short time interval. This also indicates that users played with our system during most of the exhibition period.

In addition, the chief director of Miraikan commented that the MARIO system did not need a verbal explanation to use. Users could immediately use MARIO system and interact with mid-air images simply by arranging blocks or moving their hands. The walk-up-and-use factor and the simple interaction requiring no explanation reinforce our belief that the system does not disturb players with features which are irrelevant to the core experience.

Meanwhile, the user study clarified the limitations of the current MARIO system. Many viewers pointed that mid-air image could not be seen at certain heights. Restrictions on the viewing angle of the mid-air images mean that our display is only suitable for viewing from the front. The author supposes that appropriate wider viewing area can be provided to viewers by using imaging optics that adaptively change angle according to the user's height.

As a future study, the author plans to improve the depth perception and spatial density of the mid-air images. It is expected that autostereoscopic displays can form 3D mid-air images that are viewable to the naked eye. Increasing the number of display–actuator pairs will enable MARIO to provide multiple mid-air images at the same time. In addition, a quantitative study is necessary to evaluate the effects of the projected shadow on the viewer's perception of the mid-air image. For example, it can be studied whether an artificial shadow can help users to recognize mid-air images more quickly or with high precision.

Chapter 5

HoVerTable: Combining Dual-sided Vertical Mid-air Images with a Horizontal Tabletop Display

This chapter addresses the challenging problem on the optical design that enables direct manipulation of physical objects and that provides mid-air images with multiple viewing zones. For direct manipulation, the optical design need to allow users to freely access physical objects in 3D space as implemented in MARIO. In order to provide an MR interaction to multiple users, the optical design is required to provide mid-air images with multiple viewing angles as in MRsionCase. The author believes that the values from direct manipulation and multiple viewing zones accord to the goals of tabletop displays, and thus can suggest a tabletop display as an MR interface. Therefore, in this chapter, the author aims to overcome the limitations of display area in conventional tabletop displays by implementing direct manipulation and multiple viewing zones. The advantages of mid-air images, physical objects, and even horizontal images on the tabletop surface are combined and finally a new MR interface for surrounding users is implemented.

5.1 Introduction

Tabletop displays provide a co-located workspace for multi-user interactions. With tabletop displays, people gather around and share ideas with others. In addition, horizontal displays on a tabletop not only display visual images but also support intuitive manipulation with multi-touch interactions. However, conventional tabletop displays have limitations in terms of visual presentation: The display area is limited to the horizontal surface, and all the users surrounding the table see an identical visual image, which limits the users' viewing direction.

In order to overcome these limitations, the author focused on the space "above" and "around" the tabletop surface. For these purposes, in this chapter, a tabletop display called "HoVerTable" is proposed, which stands for "**Ho**rizontal and **Ver**tical image presentation on a **Table**top display" [9]. HoVerTable was designed to satisfy the following three requirements:

(1) extending the display area of tabletop displays above the horizontal surfaces without the need for special glasses,

(2) providing view-dependent appearance of visual images to multiple users around a table, and

(3) achieving a mixed reality showcase with readable text annotation

Figure 5.1 shows the HoVerTable system. Vertical floating images, the labels "Winglet" and "Cockpit," are formed in mid-air and superimposed onto a physical object, the airplane model. The images can be seen with the naked eye without the need for 3D glasses. The horizontal images on the tabletop surface are shared among all users around the table.

In this chapter, the optical design of HoVerTable is described that combines plate-shaped imaging optics and diffusion control film. The imaging optics form mid-air images in the vertical direction and used as the tabletop surface. The diffusion control

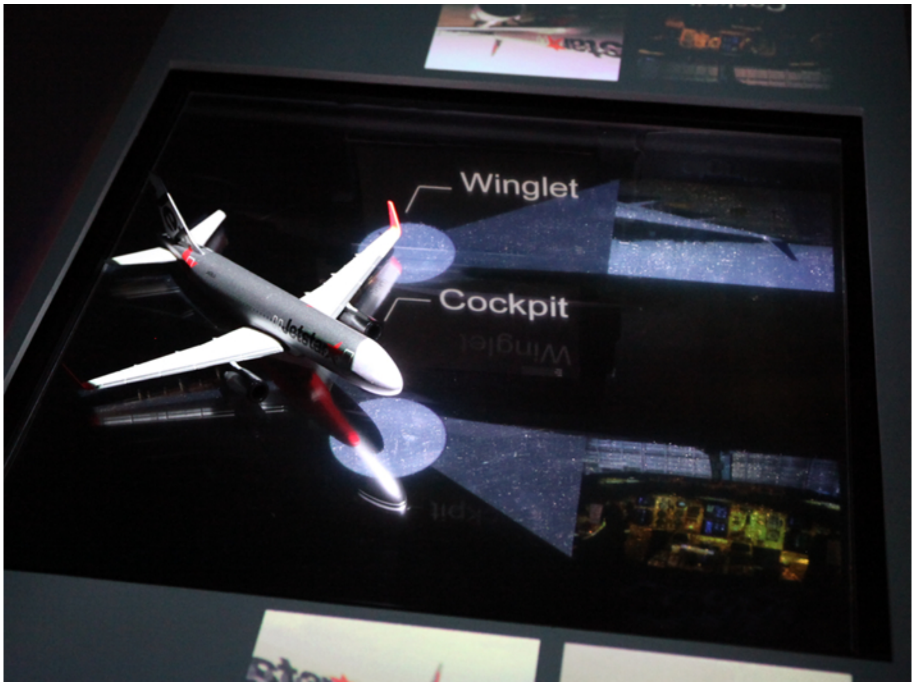


Figure 5.1: The proposed “HoVerTable” system

film enables horizontal image projection on the imaging optics. This combined usage makes compact design of HoVerTable possible.

5.2 Proposal

Through HoVerTable, it is aimed to suggest a new standard for MR tabletop displays provides three core values: 1) combined display area in 3D space, 2) selective information sharing, and 3) augmentation of physical objects. These three values correspond to the requirements of HoVerTable.

5.2.1 Combined display area in 3D space

A core feature of HoVerTable is to provide visual images in 3D space with combining vertical and horizontal images. Multi-layered vertical images are formed above the tabletop surface and located at different positions. Horizontal images are displayed on the table surface as conventional tabletop displays. These vertical and horizontal images are spatially and temporally linked to each other. For example, horizontal images can be dynamically changed according to the position and contents of vertical images. Such combination of visual images in 3D space enriches the visual expressions of tabletop displays.

5.2.2 Selective information sharing

Sharing visual images among multiple users is an essential feature of tabletop displays as a co-located and collaborative workspace. Especially, the author focuses on selective visual presentation according to each user's need with providing different visual images. The viewing zones of vertical images can be controlled by imaging optics so that the images can be displayed to targeted users. Each user can obtain different visual information according to his/her perspective while surrounding the same table. For example, different cards can be provided to players or text information can be provided in different languages. Such selective visual presentation helps tabletop displays to satisfy the needs from multiple users.

5.2.3 Augmentation of physical objects

As an MR tabletop display, HoVerTable is required to augment the features of physical objects with visual representations. Physical objects, such as mockups or prototype models, are often placed on tabletop displays as an intuitive tool. The author believes that the core features of HoVerTable introduced above, 3D display area and selective visual presentation, can suggest another physical object augmentation on tabletop displays with providing high geometrical connection between physical objects and visual images. Visual images can be superimposed onto the physical object and users can see different visual images according to their position. From the advantages of these visual representations, novel applications are implemented such as showcases and entertainment interactions with extending the features of everyday objects.

5.3 System design

In the system design, the author focused on combining vertical mid-air images and horizontal image projection. Vertical mid-air images are formed on the tabletop surface with dual sides to extend the display area above the tabletop surface and provide different visual images to multiple users around the tabletop surface. Horizontal images are projected onto the tabletop surface and shared among these users.

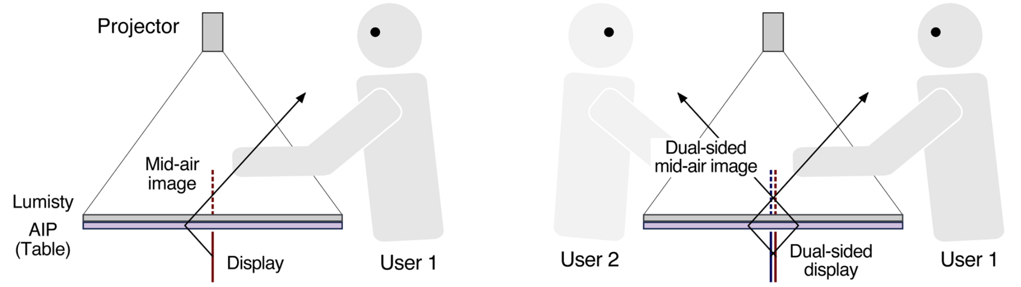
5.3.1 Vertical mid-air images on tabletop surface

Requirements for vertical images

Since text annotation is considered as a main use of vertical mid-air images, the images need to have sufficient resolution and size for users to read. To superimpose visual images onto different positions, at least two sets of vertical images are formed on the tabletop surface. All the vertical images should provide enough viewing range to users sitting in front of the table.

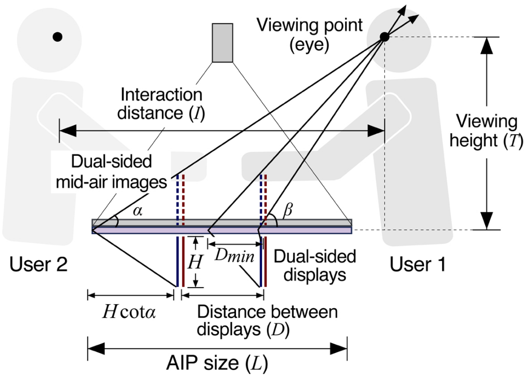
Optical principle

To form vertical mid-air images on a tabletop surface, a plate-shaped imaging optics called the aerial imaging plate (AIP) was used [32]. As shown in Fig. 2.11(a), when a display is placed below the AIP at an angle of θ, a mid-air image is formed on the AIP surface with the same angle (θ) by retro-reflection. Figure 2.11(b) illustrates the retro-reflection process inside the AIP. Since the AIP consists of two layers of micro-mirror arrays crossed at 90°, when light enters the AIP, it is reflected twice at



(a) A mid-air image formed above the table for User 1

(b) A dual-sided mid-air image for the users around the table (User 1 and 2)



(c) Optical design for two dual-sided mid-air images with design parameters

Figure 5.2: Steps in optical design of HoVerTable

the mirror arrays and converged at the same point symmetric to the AIP plane. This converged light forms an image in mid-air. Since the imaging process involves only linear reflections, the AIP can form distortion-free mid-air images, unlike Fresnel lenses and concave mirrors. However, the mirror arrays are discrete so that the resolution of the resulting images has limitations in principle. The readability of text captions formed with the AIP to the horizontal projected images will be compared in Section 5.4.2.

Optical design: step-by-step

Figure 5.2 shows the optical design of HoVerTable with three steps. In the first step, a single-sided mid-air image is formed above the AIP surface for User 1 (Fig. 5.2(a)).

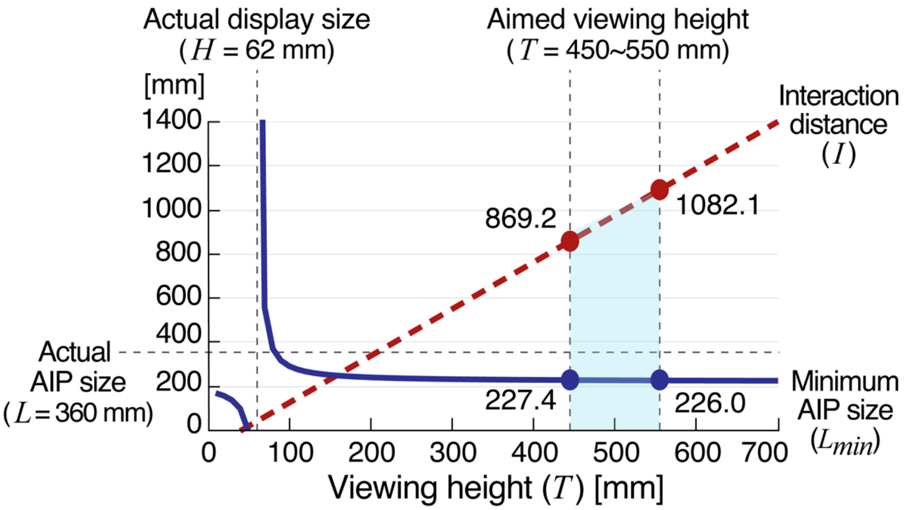


Figure 5.3: Interaction distance (I) and minimum AIP size (L_{min}) from Eqs. (5.1) and (5.2) according to viewing height (T).

To make the mid-air image stand on the tabletop surface, a display was placed with upright postures under the AIP, as Markon et al. reported [46], which means θ=90° in Fig. 2.11(a). This enables the author to use the AIP as imaging optics as well as the tabletop surface for a compact optical design. For horizontal image projection, a diffusive film (Lumisty) is placed on the AIP as a screen, and a projector is fixed above the table.

In the second step, a dual-sided mid-air image is formed on the AIP surface to provide view-dependent appearance of visual images to two users around the AIP (Fig. 5.2(b)). Each user can view each side of the mid-air image.

In the final step, a dual-sided display is added to form two sets of dual-sided mid-air images on the AIP (Fig. 5.2(c)).

Parameter settings

For forming two sets of vertical images on the AIP, design parameters are defined for system implementation, as shown in Fig. 5.2(c).

The notation T is viewing height (vertical distance from the tabletop surface to a user's eye) and I is interaction distance (distance between the eyes of two users facing one another). The notation H is display height (the vertical size of mid-air image), D is the distance between displays, and L is AIP size. The notation α is the angle between the AIP surface and the end of the far vertical image, β is the angle formed by a user's eye, the bottom of the near vertical image, and the AIP. From an experiment, it was found that the viewable range of a mid-air image is 30°–60°. Thus, α should be equal to or larger than 30°, and β should be equal to or less than 60°. From Fig. 5.2(c), I can be derived as Eq. (5.1).

$$I = T(\cot\alpha + \cot\beta) - H\cot\alpha \tag{5.1}$$

When the author forms two sets of vertical images on the AIP, space is necessary between the displays to avoid occlusions. The minimum distance between displays (D_{min}) can be calculated from the homologous triangles in Fig. 5.2(c) as the following equation.

$$D_{min} = \frac{TH}{T-H}\cot\beta$$

To form a vertical image at a certain H, the AIP should be located $H\cot\alpha$ behind the image. The minimum size of the AIP (L_{min}) can be calculated by Eq. (5.2).

$$L_{min} = 2H\cot\alpha + D_{min} \tag{5.2}$$

Figure 5.3 illustrates the I and minimum AIP size (L_{min}) from Eqs. (5.1) and (5.2) at T. The H is set to 62 mm with actual display size and the α and β are set to 35° and 55°, respectively. Actual AIP size (360 mm) is also plotted as a dashed line. For T, the author mainly targeted the 450–550 mm range since the author expects most users belong to this height range. In this range, the L_{min} should be 226–227.4 mm.

The I is determined between 869.2 and 1082.1 mm. The author consider this distance appropriate for face-to-face communication and direct access to the tabletop surface for two users. The D_{min} is also calculated as ranging from 48.9 to 50.3 mm. Based on these calculation results, the optical system of HoVerTable was implemented.

5.3.2 Horizontal image projection

Requirements for horizontal screen

For horizontal images, overhead projection was chosen from the top of the tabletop surface instead of rear projection to avoid cross contamination between the light sources, the projector and displays below the AIP. Since the AIP only passes the light from the projector, a diffusive screen is necessary on the AIP surface for horizontal image projection. In HoVerTable, a diffusive screen should satisfy two requirements: (1) penetrating light from the displays below the AIP as a transparent layer for clear vertical images, and (2) diffusing the projected light from the overhead projector for bright horizontal images.

Possible options for a horizontal screen

The author surveyed two diffusive films as possible diffusive screens: Lumisty and a semi-transparent screen.

Lumisty (MFX-1515) is a film that diffuses the entering light at an incident angle between 75° and 105° from the film surface. However, for the light coming from outside this range, Lumisty works as a transparent sheet. Due to this selective diffusion control by incident angles, it is expected that Lumisty would barely affect vertical images. Unlike the studies discussed in the Section 2, the use of Lumisty is for sharing horizontal images rather than providing view-dependent visual images.

Table 5.1: Expected optical property of visual images produced using diffusive films (+ represents positive effects and − for negative ones.)

	Vertical	Horizontal
No film (AIP only)	+++++	−−−−−
Semi-transparent film	+	+++++
Lumisty	+++	+++

A semi-transparent screen (CF-500-1525) [64] is also considered feasible for horizontal image projection due to its transparency and diffusiveness. Since this semi-transparent screen is mainly used for rear projection, it can be used for overhead projection. In addition, according to the catalogue specifications, the viewing angle is about 150° from the center of the screen. This viewing angle can cover the viewing angle of both vertical and horizontal images on HoVerTable. However, unlike Lumisty, this screen does not change its diffusion by incident angles; thus, vertical images may be blurred by diffusion.

Based on the optical properties of these films, Table 5.1 summarizes the assumptions on the visual images. For vertical images, the author expects the AIP without any film forms the clearest images since the light is not affected by diffusive films and Lumisty provides clearer images than semi-transparent film due to the selective diffusion control by incident angles. A semi-transparent film may provide a wider viewing range and brighter horizontal images because the diffusion does not depend on viewing angles. In Section 5.4.1, the detailed results of visual images produced using diffusive films will be described.

5.3.3 Implementation

Figure 5.4 shows the current implementation of HoVerTable. As shown in Fig. 5.4(a), HoVerTable consists of four parts: A projector, AIP with Lumisty film, and dual-sided displays. The size of HoVerTable is 830 (W)×795 (D)×750 mm (H). For horizontal image projection, an LED projector (BenQ GP20) is used due to the compact size

(387 (W) × 247 (D) × 111 mm (H)) and brightness (700 lm). The projector is fixed at 700 mm above the AIP, the lowest position where the projected image can cover the tabletop surface. The AIP is horizontally placed at the center of HoVerTable as a tabletop surface. The size of the AIP is 360 (W)×360 (D)×5 mm (H), which is larger than L_{min} (227.4 mm). A diffusive film (Lumisty) covers the AIP surface as a horizontal projection screen. Below the tabletop surface (the AIP), there are dual-sided displays.

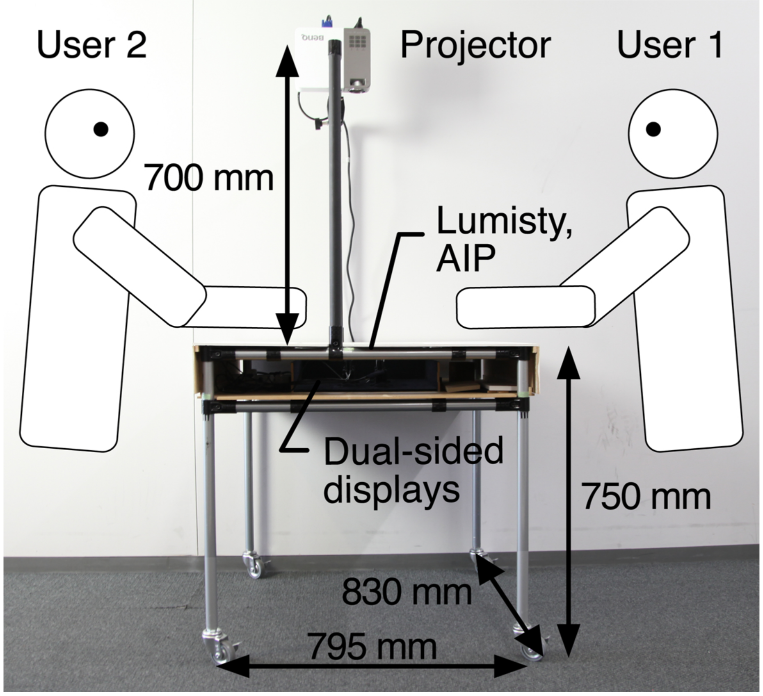
Figure 5.4(b) shows the details of the dual-sided displays. The display of a 5-inch smartphone device (Nexus 5) is used as the light source for vertical images. Each display forms a vertical mid-air image on the tabletop surface after passing through the AIP. A total of four layers of vertical images are formed, and each user can see the two image layers. The distance between the display sets was 85 mm, which is larger than D_{min}, to avoid occlusion between the display sets. The size and resolution of the display were 111×62 mm and 1920×1080 pixels, respectively. The space around the displays was covered with a light-absorbing sheet to prevent unnecessary reflections. All displays were connected to a computer via Wi-Fi as a sub-display through an Android application.

5.4 Results

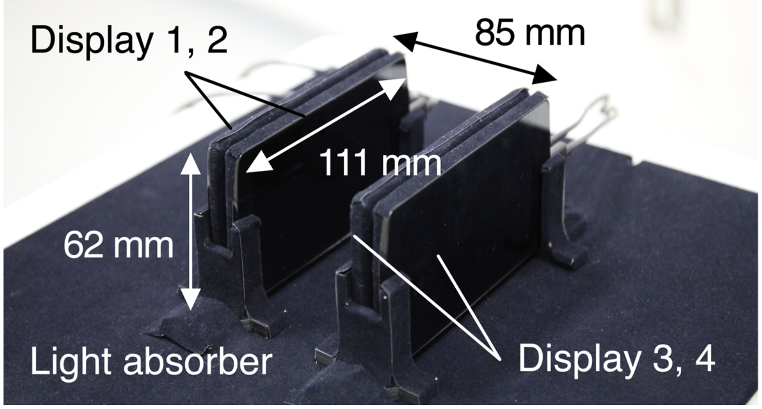
5.4.1 Optical evaluation

Vertical floating images

Figure 5.5 shows the results of a vertical floating image on the HoVerTable. Each image was taken from the left (-50 mm), front, and right (+50 mm) of center with different Ts (T=450, 500, and 550 mm). For comparison, a physical card was horizontally placed on the table surface below the vertical mid-air image. It is confirmed



(a) Side view



(b) Details of dual-sided displays

Figure 5.4: Implemented HoVerTable system

that HoVerTable can form a vertical mid-air image on the tabletop surface and users can see the image without the need for special glasses.

Diffusive films for horizontal image projection

To investigate the effects of diffusive films on horizontal image projection, the author examined the blur of vertical images and luminance of horizontal images.

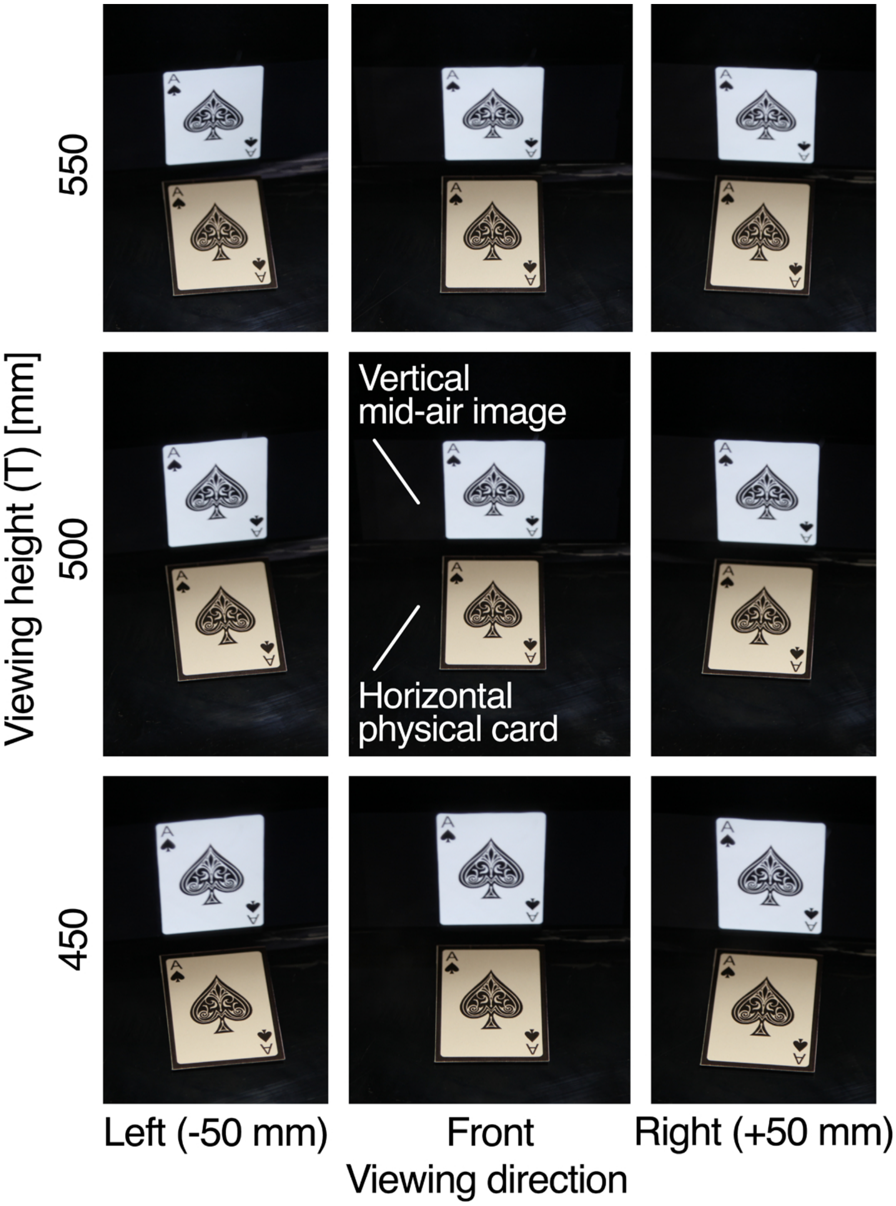
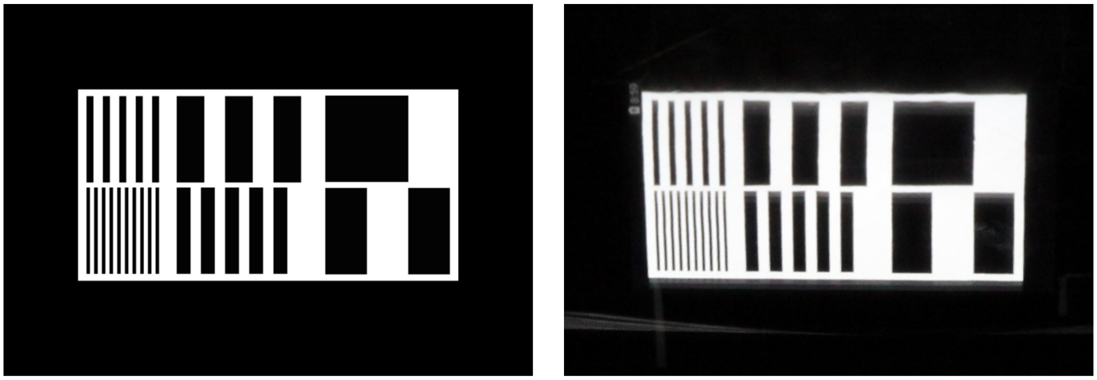


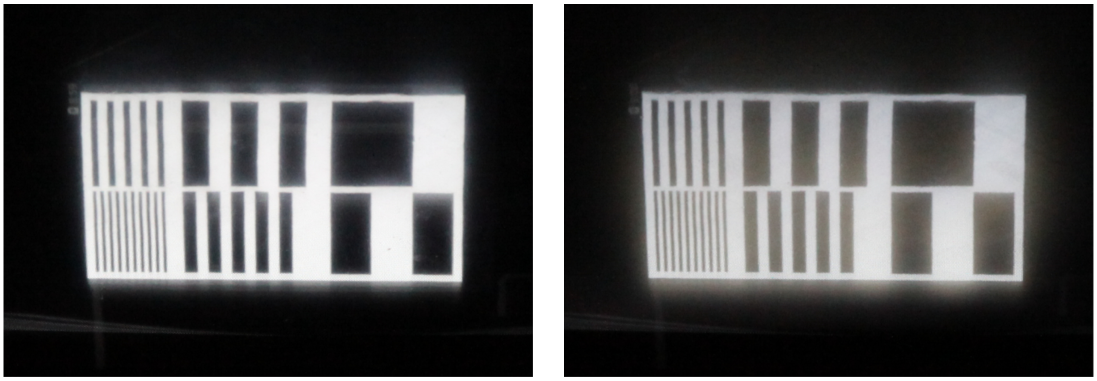
Figure 5.5: Vertical mid-air image floating on physical card placed on table surface. Images were taken from left, front, and right with three different Ts.

Blur in vertical mid-air images

The effects of diffusive films on vertical images were compared. The source and resulting vertical images are shown in Fig. 5.6. A vertical image of stripe patterns with different widths was made as a resolution test chart (Fig. 5.6(a)). As expected, the image formed without any diffusive film was the clearest (Fig. 5.6(b)). Lumisty



(a) A source image (b) No diffusive film



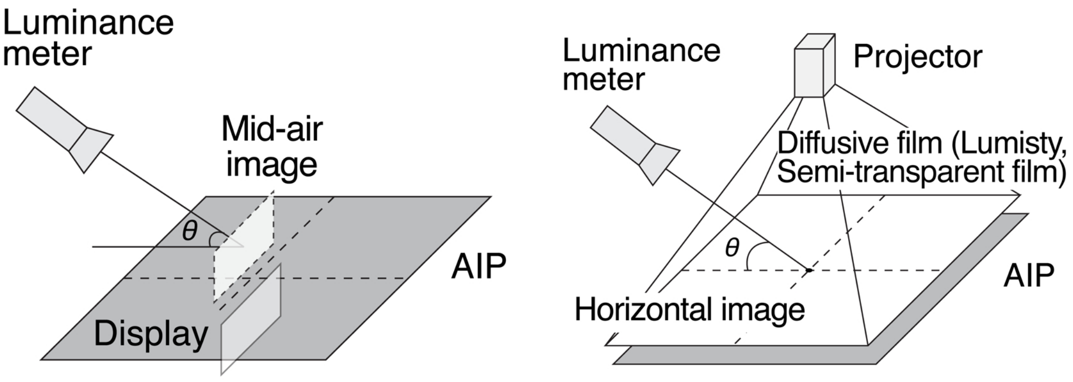
(c) Lumisty film (d) Semi-transparent film

Figure 5.6: Blur in mid-air images by diffusive films. These images were taken with same camera settings (1/15 sec, F6.3).

caused a slight blur in the mid-air image (Fig. 5.6(c)), and the semi-transparent film caused a severe blur in the mid-air image (Fig. 5.6(d)). Thus, the author concluded that Lumisty provides higher quality mid-air images than semi-transparent film.

Brightness of horizontal mid-air images

To examine the brightness of projected images, their luminance was measured. The experimental setup is shown in Fig. 5.7. The author projected a white image onto each film and measured the luminance at the center of the projected image from the angles of 30°–60° with a luminance meter (Minolta CS-100A). The luminance of a white vertical image was also measured for reference. Figure 5.8 shows the measured luminance of the vertical and horizontal images by viewing angle. The



(a) Vertical image (b) Horizontal image

Figure 5.7: Measurement of image luminance on HoVerTable

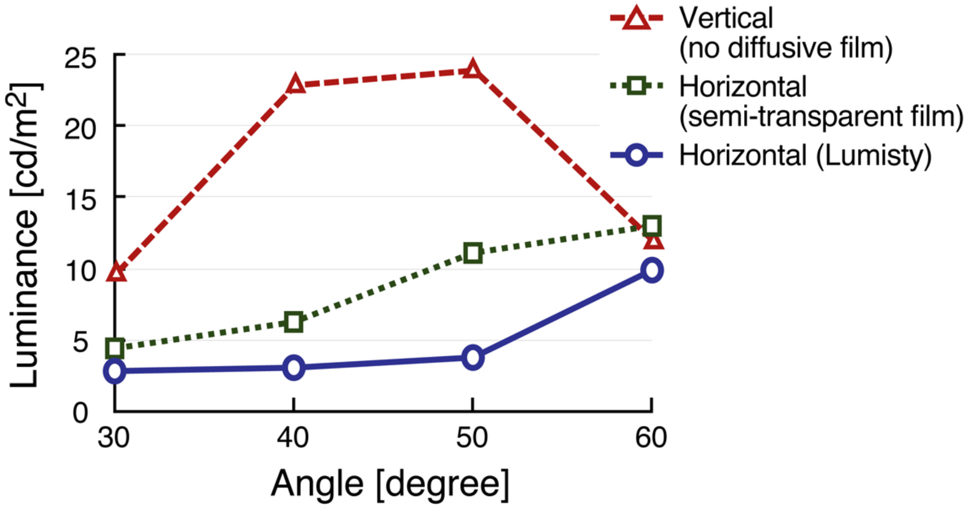
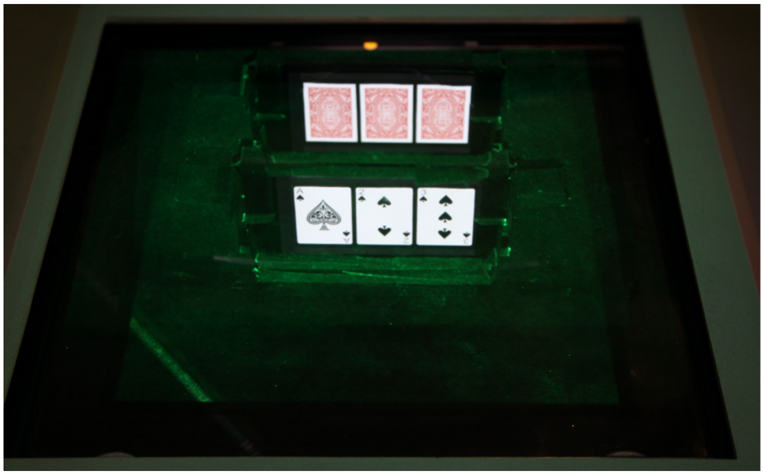


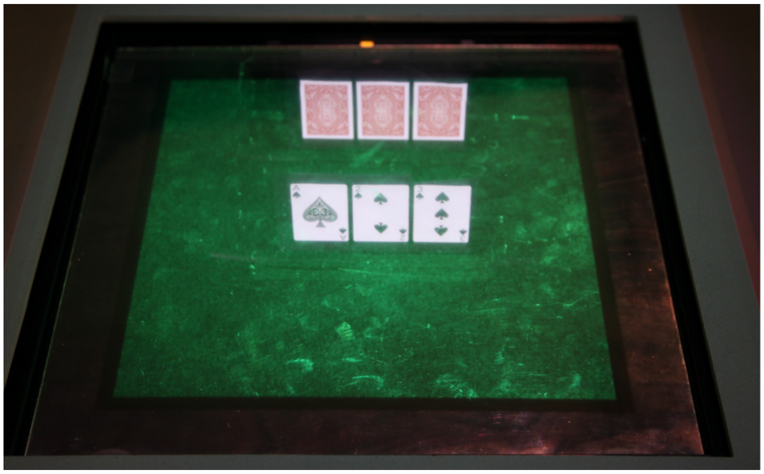
Figure 5.8: Luminance of mid-air image and horizontal image projection on Lumisty and semi-transparent film

semi-transparent film provided brighter horizontal images than Lumisty from these viewing angles.

In the 30°–60° range, regarding the brightness of horizontal images, the semi-transparent film provided brighter images than Lumisty film. On the other hand, Lumisty provided clearer vertical images with less blur. The difference in image brightness was too small to sacrifice the clear images formed by Lumisty film. Therefore, Lumisty film was chosen as the diffusive film for horizontal image projection with the emphasis on the combination of vertical and horizontal images.



(a) No Lumisty

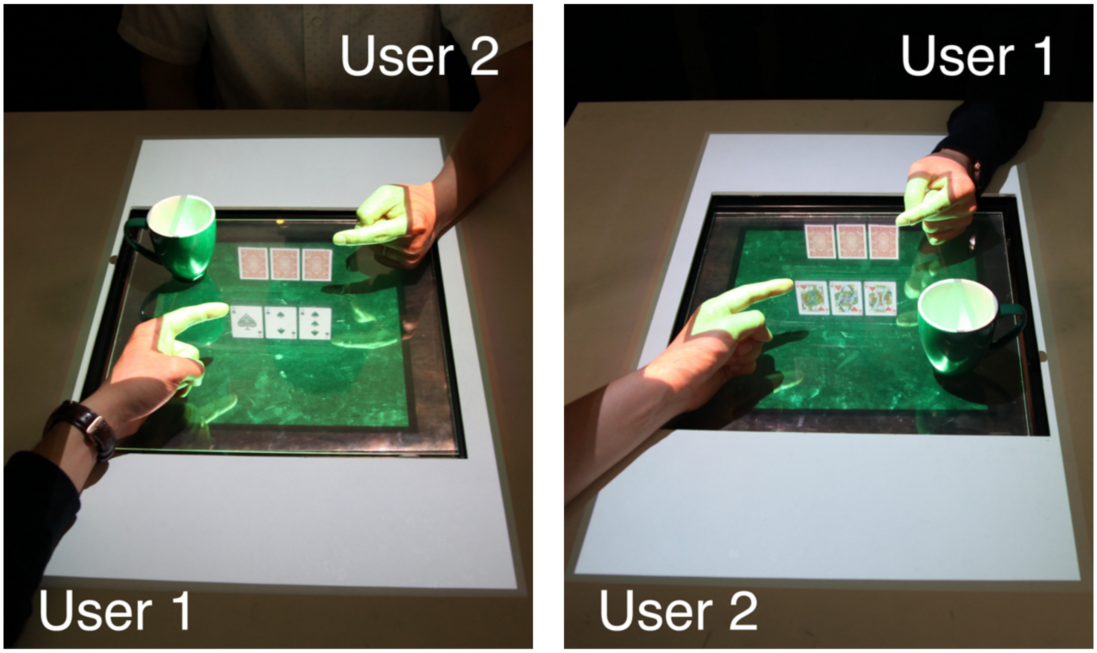


(b) With Lumisty

Figure 5.9: Effect of Lumisty film on horizontal image projection

Horizontal image projection

Figure 5.9 shows the effect of Lumisty film on horizontal image projection. Without Lumisty film, horizontal images cannot be clearly displayed due to the transparency of the AIP (Fig. 5.9(a)). On the other hand, when Lumisty film is placed on the AIP, a projected image (a green mat) can be displayed from the tabletop surface (Fig. 5.9(b). From these results, it was confirmed that Lumisty film is effective as a projection screen on the AIP.



(a) View from User 1 (b) User 2

Figure 5.10: Dual-faced mid-air images can provide different visual images to two facing users.

View-dependent appearance for two facing users

To provide view-dependent visual images from HoVerTable, mid-air images was created with two sides. Figure 5.10 shows the mid-air images viewed by two users facing one another, Users 1 and 2. From User 1's viewpoint, the cards with spades and the back of User 2's cards could be seen. On the other hand, User 2 could see different cards and the back of User 1's cards. This result confirms that dual-sided mid-air images can provide view-dependent visual images in two directions.

Usage example: A mixed reality showcase with text annotation

Based on the benefit of vertical and horizontal images, the author implemented a usage example: A mixed reality (MR) showcase that superimposes text annotation onto a physical model, is shown in Fig. 5.1. When users place a physical object (an airplane model) on the HoVerTable, text labels are shown in mid-air and directly superimposed onto the model. Horizontal images are displayed from the tabletop surface to be shared among all users. Since the vertical images have two sides, different visual images can be displayed to each user. For example, English captions are provided to User 1 and Japanese ones to User 2.

5.4.2 User study

Goal

Since vertical images were chosen for text annotation instead of horizontal ones, the readability of vertical images were critical for reliability of the usage example. Compared to horizontal images, vertical images have the advantage of good readability with higher brightness, as shown in Fig. 5.8. As mentioned regarding the optical properties of the AIP, however, there is a disadvantage in that the readability of vertical images is reduced: the mid-air images formed using the AIP have limitations in resolution due to its discrete mirror arrays. Moreover, HoVerTable can provide the view-dependent appearance of visual images to two users facing one another. It is expected that text captions can be displayed with appropriate orientations to each user according to their viewing positions.

In particular, the author investigated two assumptions through the user study: With HoVerTable, (1) text captions displayed in vertical images are as easy to read as horizontal projected images despite the differences in brightness and resolution, and (2) text captions with correct orientation from the user's viewpoint are easier for a user to read.

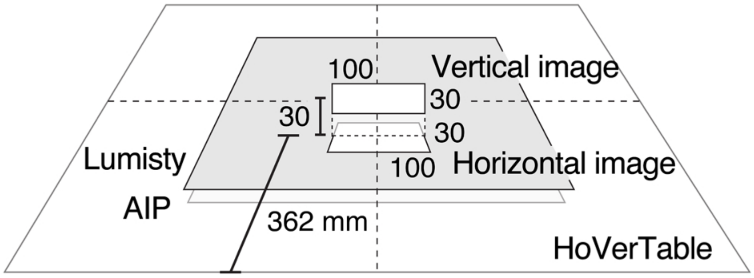


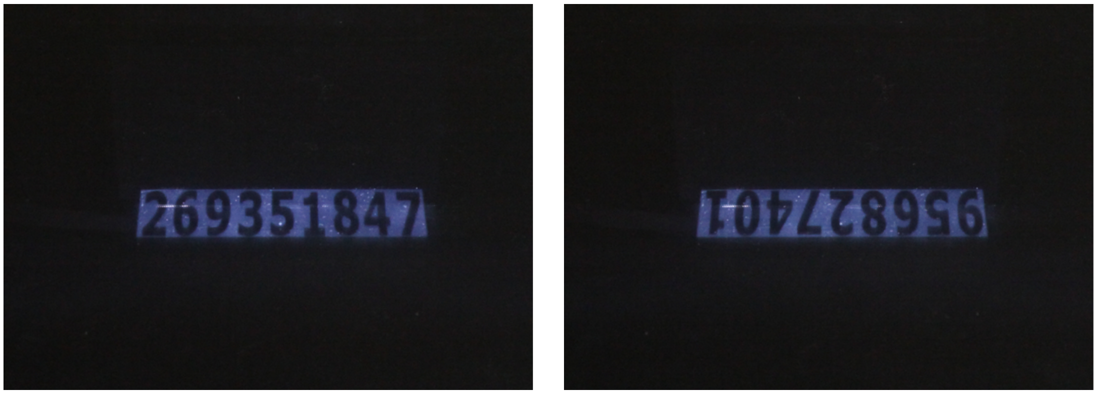
Figure 5.11: Settings of vertical and horizontal images for user study

Settings and procedures

Figure 5.11 illustrates the settings of the horizontal and vertical images. Horizontal images were placed 362 mm from the end of the HoVerTable surface and vertical images were placed 30 mm above the center of horizontal images. The size of all visual images were 100 (W)×30 mm (H).

A 9-digit random number was used as the text caption. The reason we chose numbers is that they have only ten elements with a distinct orientation. Figure 5.12 shows the example images used in the user study. In each task, a 9-digit number randomly appeared on HoVerTable under one of the following four conditions: (a) a horizontal image with correct orientation, (b) horizontal image with upside-down orientation, (c) vertical image with correct orientation, and (d) vertical image with reflected orientation. When a number appeared on HoVerTable, participants were required to input the number using keypads.

An experimental set consisted of 20 input tasks, which included five tasks from each condition. Each participant conducted three sets of experiments in total. In the first set, all participants were guided in a practice tutorial. After each set was completed, participants took at least a one-minute break. While participants observed the provided images, head movement was allowed for adjusting their viewpoint. The time taken for each task and the numbers input by the participants were recorded. For statistical analysis, only the time taken for the tasks with correct input was used.



(a) Horizontal image with a correct orientation (b) Horizontal image with an upside-down orientation

(c) Vertical image with a correct orientation (d) Vertical image with a reflected orientation

Figure 5.12: Example images used in user study. Images were taken with same camera parameter settings (1/20 sec, F8).

Twelve participants (ten males, two females), who were recruited from our university, participated in the experiment. Their ages ranged from 23 to 38 (median=24). They had normal eyesight including some with corrected vision. The Ts of all participants were in the targeted range from 450–550 mm. The illuminance of the experiment room was set to 28.6 lux so that participants could see the vertical and horizontal images on HoVerTable and the keypads for the input task.

Results

Figure 5.13 shows the average input times for each condition of visual image presentation. For assumption (1), the average input time for vertical images was significantly shorter than that for horizontal images ($F(1, 11) = 14.51$, $p < .01$). This result confirms assumption (1) and indicates that vertical images on HoVerTable are easier to see than horizontal images despite the differences in brightness and resolution. Two possible explanations can be suggested for this result: the differences in the brightness and effective size of the horizontal and vertical images. As mentioned in Section 4.2, with HoVerTable, vertical images are brighter than horizontal ones: In the user study, the average luminance of the vertical images (7.42 $\mathrm{Cd/m^2}$) was higher than that of the horizontal images (1.53 $\mathrm{Cd/m^2}$). Moreover, participants' viewing angle was less than 45° so that vertical images had a larger effective size, the size viewed from a participant's perspective, than horizontal images, although their sizes were identical.

For assumption (2), text annotation provided in the correct orientation to participants' viewing direction were associated with significantly shorter input times on average than those with upside-down and reflected orientations ($F(1, 11) = 25.65$, $p < .01$). This result confirms that providing view-dependent appearance with dual-sided vertical images is an effective method for text annotation with HoVerTable, as expected. During the experiment, most participants commented that they found it difficult to differentiate the numbers "6" and "9" due to their reversed shapes.

5.5 Discussion and Conclusion

The resulting images and user study results revealed limitations of HoVerTable.

The current implementation can support a usage scenario for only two users facing one another as a minimum configuration of multiple user interaction. However, many tabletop displays can support more than two users and various viewing positions. It is necessary to increase the number of users and expand the viewing directions of vertical images by improving the optical design.

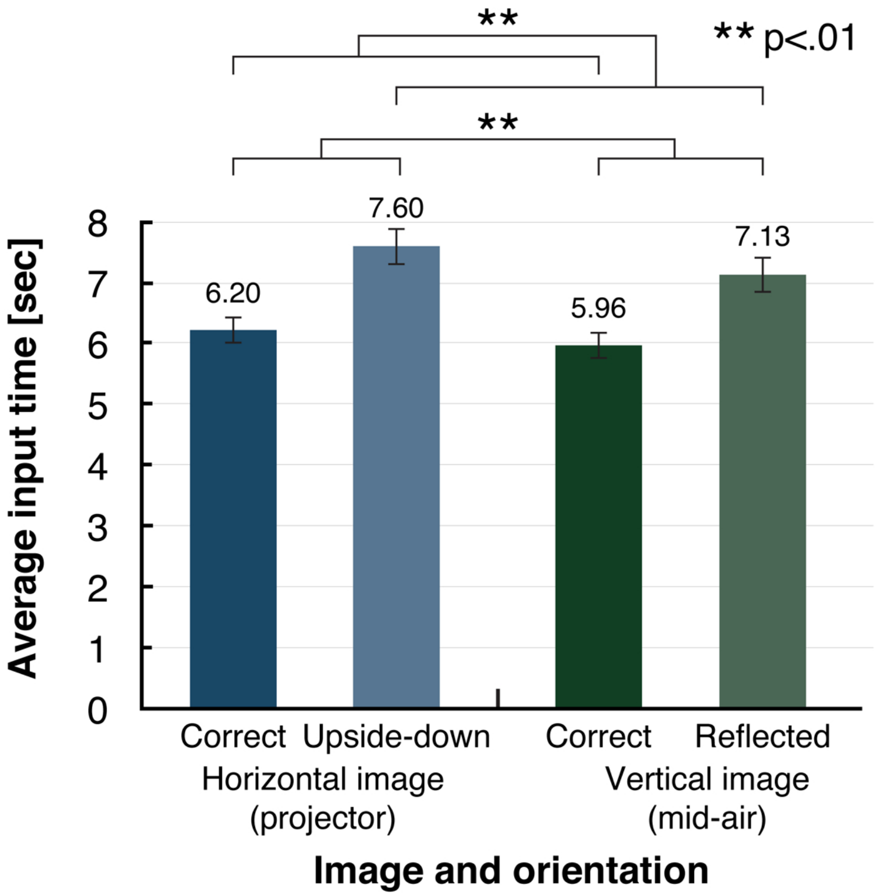


Figure 5.13: Results of user study

Since the author focused on text annotation as a main function of vertical images, a user study on readability with numeric strings was conducted. However, the vertical and horizontal images on HoVerTable can display other kinds of visual information such as photos, figures, and illustrations. In the future, the author plans to study the usability of HoVerTable with the view-dependent appearance of these images.

Although the use of the AIP as a tabletop surface has enabled a compact optical design, there is an occlusion problem when physical objects are placed on the AIP and block the light reflected by the AIP. This problem can be avoided by considering the shape and orientation of physical objects using object detection. Moreover, the position of vertical images is fixed in the current implementation. In the future,

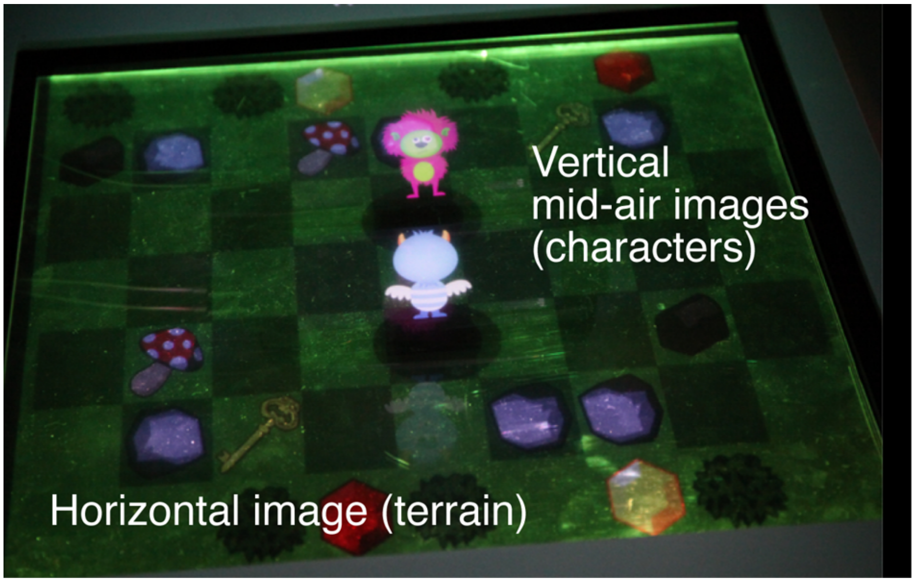


Figure 5.14: Example of application for entertainment

displays will be moved using mechanical actuators and dynamic visual expressions can be provided. The author expects that changing the vertical image positions will also solve the occlusion problem.

The brightness of visual images needs further improvement. The vertical and horizontal images on HoVerTable exhibit low luminance compared to ordinary displays such as LCD monitors or projectors. Although the displays used for vertical images (Nexus 5) had a luminance of 429.6 Cd/m^2, the highest luminance of the resulting vertical images was 24.0 Cd/m^2. Horizontal projected images showed much lower luminance, with an average of 4.88 Cd/m^2 at viewing angles between 30° and 60°. To enable the use of HoVerTable under ordinary lighting conditions (e.g., office environment), it is necessary to improve the brightness of the visual images. Changing the angle of the vertical images may be a solution. To provide bright vertical images, it is preferable for a vertical image to be perpendicular to a user's viewing angle. If the mid-air images were formed obliquely to the AIP surface instead of the upright postures, the vertical images would appear brighter.

The author discussed the optical design of “HoVerTable” that combines vertical and horizontal images on the tabletop surface. The optical design formed dual-sided vertical mid-air images using plate-shaped imaging optics (the AIP). Users could see the vertical mid-air images without the need for special glasses. For horizontal image projection, Lumisty was chosen as a diffusive screen due to less blur in vertical mid-air images. Dual-sided vertical images provided text annotation with correct orientation to two users facing one another. From these advantages, an MR showcase was implemented that superimposed text captions onto an airplane model. The user study on text readability confirmed the effectiveness of text annotation with vertical images or with a correct orientation for the MR showcase. However, the current implementation has several limitations due to the occlusion problem, fixed image positions, and low brightness. In the future, the author plans to implement more usage examples with various visual images. A card battle game can be considered as an example of entertainment use, as shown in Fig. 5.14. Users can play a battle game with the characters summoned from the physical cards onto the table. The author believes this combination of vertical and horizontal images with HoVerTable enables new user experiences with tabletop displays.

Chapter 6

Conclusion

6.1 Summary

In order to overcome the limitations in display area of current displays, in this thesis, optical designs, which can form layered mid-air images and superimposes them onto physical objects, are proposed. Based on the proposed optical designs, novel glasses-free mixed reality interactions are developed.

MRsionCase

MRsionCase formed two layers of mid-air images that sandwich a physical object between the mid-air images. The optical design consists of half-silvered mirrors and a real imaging optics (DCRA). MRsionCase has realized three main functions as a new MR showcase.

1) Hybrid usage of virtual and real mid-air images enables the mid-air images to be located at the very next to the physical objects. This gives a high level of spatial consistency between physical object and visual images. Users can see both a physical exhibit and superimposed images at the same sight.

2) The imaging system of MRsionCase is designed with a symmetry to support multi-directional viewability. Users can appreciate the exhibit and visual images by walking around the showcase. In live demonstrations, it is confirmed that most users viewed MRsionCase from more than three directions.

3) All these features do not require users to wear special devices such as 3D glasses. Users of MRsionCase could walk up and participate the interaction without any preparation. I believe device-free usage motivate users to experience MR interactions.

MARIO

The optical design of MARIO enabled users to access visual images and manipulate physical objects during interaction. The combined usage of AIP and a linear actuator can change positions of a mid-air image in the depth direction. MARIO have three contributions as the follows.

1) The optical design of MARIO formed a mid-air image moving inside a 3D space of 350 (W)×300(D)×250 (H) mm. This 3D space could be used for interaction area as well as display area.

2) With the aid of object detection, interactive applications using everyday objects are developed. Users could make a terrain for a mid-air character with stacking and manipulating wooden blocks.

3) A high level of optical consistency between a mid-air image and the real world has realized by an artificial shadow projected an overhead projector. The cast shadow helped users to perceive the positions of the mid-air image and provide higher sense of reality during user interaction. It was witnessed that users voluntarily said “bye-bye” to the mid-air character after their experiencing MR interaction in MARIO system.

HoVerTable

HoVerTable provided multi-layered mid-air image to two opposite directions and enabled direct manipulation of physical objects on the tabletop surface. In HoVerTable, the optical design was focused to form each mid-air image with dual sides and to combine them with horizontal image projection.

1) The combination of vertical mid-air images and horizontal images on the tabletop surface extended the display area of tabletop displays. For this purpose, a novel combination of AIP and a diffusive control film (Lumisty) was implemented. The vertical images are formed by imaging optics so that no physical displays or surfaces that disturbs face-to-face are needed. Users could communicate with each other in face-to-face as the case they sat around ordinary tables.

2) Dual-sided mid-air images provided view-dependent images to two facing users. Two sides of mid-air images were formed at the same depth, but controlled as an independent image layer. Providing different images according to users' viewpoints broadened the visual expression.

3) Physical objects can be placed on the tabletop surface and augmented by vertical and horizontal images. Without see-through displays or optics between users and physical objects, the imaging system is complete inside a tabletop display.

All optical designs in the proposed systems are based on geometrical optics. The mid-air images formed by these applications are not limited to the current applications but also can be applied to future possible applications for glasses-free MR interactions. Therefore, the findings and discussion of this thesis will be helpful to further implementation of optical designs for glasses-free MR interactions.

6.2 Future work

The results of discussion on this thesis revealed the limitations of current implementations and showed a direction for future work.

Scalability in implementation

To support various usage of mid-air images, high scalability is required in visual images and optical design in the implementation processes. However, current imaging optics have only few options in size and shape. Although the system size was theoretically calculated in various scale, the size of actual implementation was limited to certain scale. Moreover, imaging optics (i.e. HSMs, DCRA, and AIP) were made with glass so that it was so hard to change their shape according to various requirements on implementation. If imaging optics (i.e. HSMs and RIOs) can be made with DIY-able material such as resin or cuttable sheet and provide a variety of sizes and shapes, imaging optics can provide a variety of sizes and shapes. Such DIY-able imaging optics will enable more rapid prototyping of optical designs and improve the scalability of MR interactions.

Enhancement of visual expression

Visual expression of mid-air images also needs improvement in several points.

Fixed imaging positions suggest a problem to be solved in the future work. To provide a dynamic MR interaction, mid-air images need to be placed in various positions with movements. In the current optical design of MRsionCase and HoVerTable, if displays move and change the position, mid-air images also move to the corresponding positions. Especially, in HoVerTable, moving displays can easily move the mid-air images on the tabletop surface. The moving images will enrich visual expression on tabletop displays with extending the display area of mid-air images to complete 3D space. Moreover, the direct link between the positions of display and mid-air images will provide more interactive applications that encourage users' participation. For example, users can directly move a character, shown from mid-air images, with manipulating the display instead of a gamepad.

In MARIO and HoVerTable, horizontal and vertical images are combined and linked to each other with providing a high sense of reality from mid-air images. However, the horizontal images were limited to the physical surface due to the use of image projection. If the horizontal images are presented with mid-air images and move in vertical direction, two layers of mid-air images can be displayed with enhancing degree of freedom in movement over 3D space completely beyond the limitations of physical displays or surfaces.

Along with image superimposition by mid-air images, projection mapping will enable more direct visual presentation onto physical objects. For example, a static exhibit can provide dynamic visual expression with the projection of moving images: The color and apperance of the exhibit can be changge. This dynamic visual effects will help users to interactively appreciate the exhibit from various perspectives.

Another issue is occlusion expression when multiple layers of mid-air images are overlaid. The optical design of MRsionCase and HoVerTable could provide two layers mid-air images. When these mid-air image layers are overlaid, the lights of the overlaid images are added to each other since the mid-air images are transparent. This makes the overlaid images brighter, and thus the visibility of the images decreases. To solve this problem, the addition of lights between the overlaid images needs to be controlled with adjusting the transparency of the mid-air images.

Feedbacks from mid-air images to the real world will be a challenging issue for a sense of reality. As introduced, the goal of MR interactions is to provide a high sense of reality from visual images with a seamless connection between the images and physical objects. In MARIO and HoVerTable, visual images are displayed as an output to users' actions. To establish a complete connection between visual images and physical objects, the feedback from mid-air images to the physical objects also needs to be implemented. Although artificial shadows on the real blocks seemed effective, more active feedback such as tactile stimuli and physical forces will improve the sense of reality with mid-air images. For example, a mid-air image might move a physical block place on the table or a mid-air character bounces on a user's hand with a weight. The author believes that such bi-directional feedback between mid-air images and physical objects will enhance a sense of reality with removing the border between the real and virtual world.

Use of various imaging sources

This thesis used ordinary displays as a light source for mid-air images. However, the application of mid-air images can be diversified when more various lights are used as a lighting source. The DCRA and AIP can form a mid-air "image" with reflecting and converging other electromagnetic waves even with infrared rays. Thus, users can feel heat from mid-air images if infrared ray sources are used along with displays.

Moreover, additional information (e.g. 3D position data) can be encoded into mid-air images. For example, if a mid-air image consists of a pattern of high-frequency flickering lights, the pattern can convey additional information along with the mid-air image. Since human's eye cannot detect the high-frequency patterns, additional information can be provided without visual disturbance. Moreover, if the position data can be combined with mid-air images, the MR system does not need additional sensors for object detection so that more various interactions using physical objects will be possible. In this way, the use of other imaging sources provides various information from mid-air images.

Toward wide and everyday usage

There are practical limitations for everyday usage of the proposed optical designs for mid-air images. Most of all, the cost for imaging optics, especially novel RIOs, is much higher (∼\$10k) than ordinary displays since the production of the imaging optics are at an early stage. Moreover, the size of mid-air images are small for practical use in comparison with ordinary displays such as LCD monitors or TV screens. The size of DCRA and AIP was limited to 150 (W)×150 mm (H) and 360 (W)×360 mm (H), respectively. Since the size of mid-air images depends on that of imaging optics, larger imaging optics are necessary to form large mid-air images. For example, in order to form a 2-m mid-air image for a life-sized figure, the size of RIO should be at least 2 m to 2.8 m. These limitations of RIOs will be solved as the production process matures. Cost-effective optical design and larger size of mid-air images will be used in practical applications such as digital signage systems and visual effects in performance stages. Such everyday usage of mid-air images will seamlessly connect the physical and information world, and enables people to interact with visual images as they do with ordinary objects by making the two worlds indistinguishable.

Appendix A

Exhibitions

Proposed MR interactions in this thesis have been demonstrated from exhibition venues.

A.1 MRsionCase

A.1.1 Digital Contents Expo 2012

Digital contents Expo (DCExpo) is an annual event to introduce media and contents technology. DCExpo 2012 was held at the National Museum of Emerging Science and Innovation (Miraikan) from October 25th to 27th in 2012. Since DCExpo is an open event, guests at DCExpo come from various fields of background such as art, research, entertainment and broadcasting.

In DCExpo, MRsionCase was presented as a museum showcase (Figure A.1). The exhibit was a dogu, which is an ancient Japanese clay doll. In order to provide viewers related information about the exhibit, its history, place of excavation and special points are superimposed onto it.

(a) Users look into the implementation of MRsionCase.

(b) Users see the exhibit with surrounding MRsionCase.

Figure A.1: Exhibition of MRsionCase at Digital Contents Expo 2012 (Tokyo, Japan).

A.1.2 ACM SIGGRAPH Asia 2012

ACM SIGGRAPH Asia is an academic conference which mainly focuses on computer graphic and interactive techniques held in Asia-pacific region.

SIGGRAPH Asia 2012 was held in Singapore Expo from November 28th to December 1st, 2012. MRsionCase is selected for live demonstration in Emerging Technologies site, oral presentation in Technical Briefs, and poster presentation in Poster sessions.

In the demonstration, MRsionCase was introduced as a digital signage showcase. The exhibit was a spider-man figure. Video clips and text information were superimposed onto the figure to provide relevant information with the figure and the story of movie. In addition to superimposed images, narrations and theme songs were played from hyper directional loudspeakers for auditory information.

A.2 MARIO

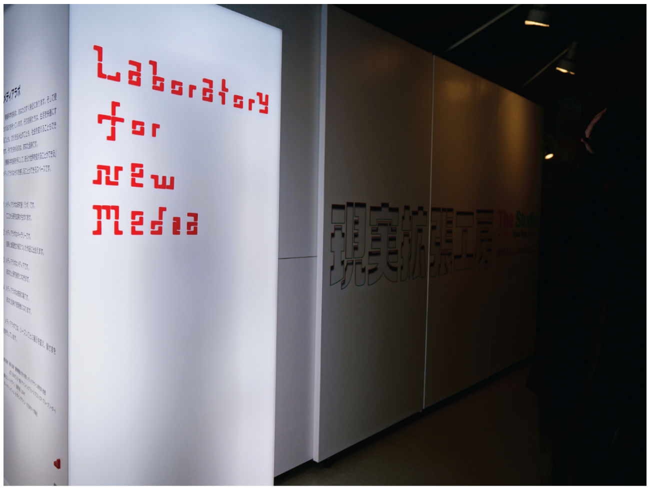
A.2.1 12th Media Laboratory Exhibition at Miraikan

MARIO was exhibited at a permanent exhibition, 12th Media Laboratory Exhibition, in National Museum of Emerging Science and Innovation (Miraikan) as shown in Figure A.2. The exhibition was lasted for six months from July 1st, 2013 to January 14th, 2014. During the exhibition, over 110,000 people came to the exhibition booth. MARIO system was experienced by a lot of guest from various age groups, including kids to elderly people.

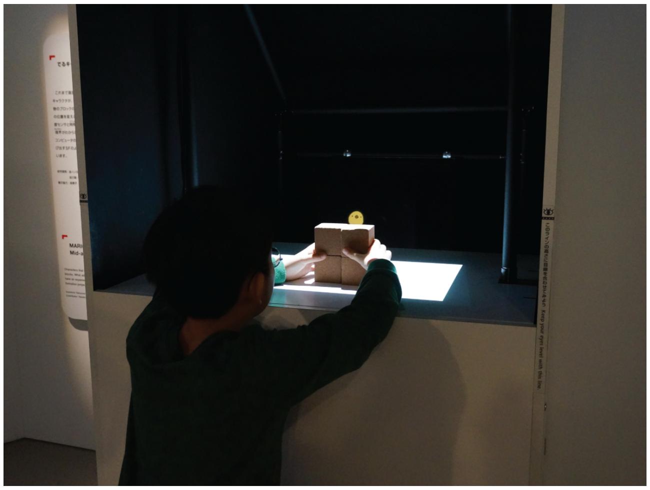
In this exhibition, children was considered as a main target of guests. Thus, the height of interaction table was set to 120 cm from the floor. The character was carefully selected with a chick character having pretty appearance and somewhat clumsy movement.

The operation of exhibition was supported by science communicators, technical staffs, and volunteers in Miraikan. Operation staffs recorded a daly brief after finishing everyday's exhibition. The record was a good reference when analyzing users' behavior during their usage of MARIO.

Despite the limited viewing angle of MARIO system, it was frequently observed that many pairs of guests played with MARIO with each other by moving wooden blocks and looking for the Hiyoko character in mid-air. Collaborative work performed in MARIO system by multiple users will be an interesting theme for further study.



(a) Exhibition booth.



(b) A kid plays with MARIO.

Figure A.2: Exhibition of MARIO at 12th Media Laboratory Exhibition in Miraikan (Tokyo, Japan).

A.2.2 Innovative Technologies 2013

Innovative Technologies is a Japanese annual event to pick up innovative technologies in Japan and introduce them to domestic and foreign people.

MARIO was honorably selected as one of the Innovative Technologies in 2013 by Japanese Ministry of Economy, Trade, and Industry (Figure A.3). Selected technologies were accompanied with public demonstration during the event. The event was held in Miraikan from October 24th to 27th, 2013.

(a) Exhibition booth.

(b) A user interacts with the mid-air character with a hand.

Figure A.3: Exhibition of MARIO at Innovative Technologies 2013 (Tokyo, Japan).

Preparing for a demonstration at Innovative Technologies, second version of MARIO system was made. Compared to Media Laboratory exhibition, which mainly focuses on children, the main target of Innovative Technologies is adults. Thus, the second system was designed a little bit higher than its prior version. The projector was also installed at higher place to improve object detection precision.

However, parts such as a projector, a monitor, and a linear actuator were identical to those used in the previous system. Sharing the parts, a high level of compatibility between two MARIO systems could be maintained.

During the exhibition, MARIO was featured as a near-future display and interaction technology by television and internet media. Most guests asked a question about the practical usage of MARIO system, for example, as an entertainment or digital signage system.

A.2.3 ACE 2013 Creative Showcase

MARIO was accepted from ACE 2013 Creative Showcase. ACE 2013 is an international conference which mainly focuses on entertainment applications using computer and human interactions. Thus, participants of ACE have a deep interest in entertainment and interactive technologies.

Figure A.4: Exhibition of MARIO at ACE 2013 (Twente, The Netherlands).

Creative showcase session includes on-site demonstration (Figure A.4). The demonstration was held in November 14th, 2013 at University of Twente, Netherland. Since ACE was a single-tracked conference, all participants could see and experienced the demonstration systems.

In Miraikan exhibition, MARIO system was housed in the wooden case. The case could isolate illumination condition inside the system from the outside illuminations. However, the system exhibited in ACE did not have no outer housing for higher transportation efficiency. Thus, MARIO was placed in a dark area of the demonstration floor in order to enable viewers to see a mid-air character more clearly.

Since most viewers in ACE were expected as adults, a higher version of MARIO system was exhibited at ACE, which was the same one exhibited at Innovative Technologies. However, some viewers were too tall beyond the expectation. They should adjust their head position to see the mid-air character. Some viewers found out that the mid-air image were easily seen from further position from the system due to the limitation on viewing angle of a mid-air image.

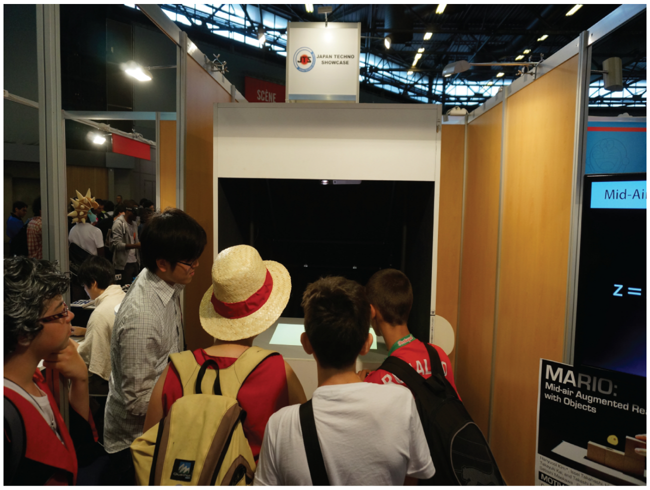
During the demonstration, most viewers showed their surprise after interacting with a character appeared in mid-air. Some visitor expressed his impression with saying “It is a crazy display.”

All participants of ACE had a vote for selecting best demonstration. MARIO was honorably chosen as an awardee of Best Demo Gold Award by participants’ vote. Through MARIO, a game stage was extended to the real world with a mid-air image coming out of the boundary between computer monitor and users’ world. The author considers that, in ACE conference, such “mixed” quality of MARIO system that could contribute to a new entertainment application was highly praised.

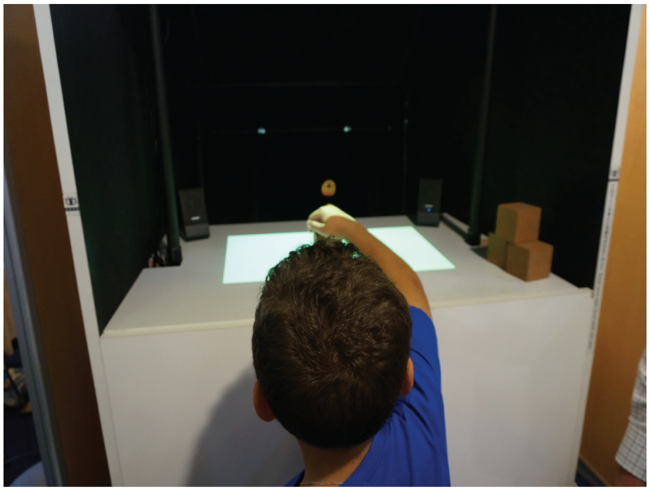
A.2.4 Japan Expo 2014

Japan Expo is an annual event to introduce Japanese culture and technologies to overseas. The events are held twice a year in Europe and USA.

MARIO was invited to Japan Expo, held in Paris, as a part of technical showcase (Figure A.5). Visitors of Japan Expo come from various backgrounds. However, most of them have a common interest on Japanese culture.



(a) A lot of visitors experienced MARIO.



(b) A kid plays with the mid-air character on this hand.

Figure A.5: Exhibition of MARIO at Japan Expo 2014 (Paris, France).

Actually, most visitors came to the event with even wearing a costume of “anime” characters. They seemed very familiar with Japanese culture. Some of them could speak and understand Japanese as well.

Unlike to Dutch people, French people were little bit shorter. Thus, visitors had little problem to see a mid-air character despite its limited viewing angle.

It was very interesting to see users' reactions to MARIO system from the exhibition. Some users yelled at the character by showing their surprise when the character jumped onto their arm. Other users even flicked off the character with an action of big surprise. Many people commented that MARIO system was "really weird" with a positive meaning. This "weird" experience was the MARIO system aimed by superimposing a mid-air character onto the real world.

During the exhibition, it was often observed that French people tend to use their hands for interacting with Hiyoko character than blocks. Most visitors showed interest in grabbing and touch the character by hands. Only few people moved the wooden blocks for interaction. Although it was an empirical observation, this is the opposite to the case of Japanese visitors in Miraikan. The characteristics of users might bring a different way of usage on the same system.

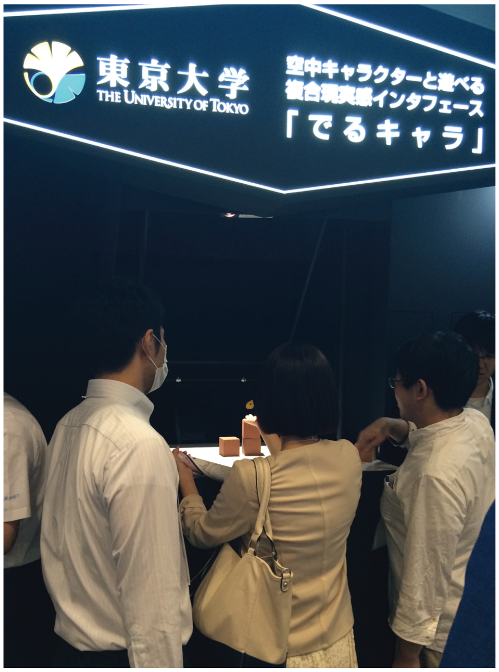
A.2.5 CEATEC 2014

CEATEC is the biggest technical showcase event in Japan. CEATEC 2014 was held in Makuhari Messe from October 7th to 11th, 2014. Many visitors came from IT companies, and thus they had professional background and interest in engineering, technologies, and business.

MARIO was invited as a near-future entertainment application using mid-air images instead of 2D displays (Figure A.6). The system was the same version presented at Innovative Technologies and ACE 2013, a higher one, since the main target of visitors were considered as adults.

Most people could experience the interaction with Hiyoko character by moving blocks or using their hands. Although it was informal observation, women and children more likely to use hands and grasp the mid-air character compared to men. Especially, many women visitors enjoyed the interaction with saying the character was cute.

(a) Exhibition site of CEATEC 2014 (Makuhari Messe)



(b) A user experiences MARIO with manipulating physical blocks.

Figure A.6: Exhibition of MARIO at CEATEC 2014 (Chiba, Japan).

After exhibition, some company had contacted with a future usage of MARIO system. They showed a deep interest in utilizing MARIO as a digital signage system for advertisement and promotion events.

Appendix B

Sound Presentation from MRsionCase

This chapter will introduce directional sound presentation in MRsionCase from the collaborative work with Shun Nagao.

As a complement to visual information, the use of sound presentation in MR systems has been proposed [65, 66]. For example, LISTEN [67] provides an adaptive audio guide with 3D sound in exhibitions. The 3D effect improves the spatial consistency between sound and exhibits. However, a viewer must wear headphones during the exhibition. To eliminate this requirement and to provide spatially consistent sounds, hyper directional loudspeakers (HDLs) [68] can be used. Invoked computing [69] places a sound source on physical objects by reflecting the directional sound from HDLs onto the objects. A rotating platform installed on a ceiling enables the directional sound to be reflected and directed to any point as a sound source. This approach could create a sound field over a large area (e.g., a room). However, in MRsionCase, we installed HDLs inside showcases for compact system implementation.

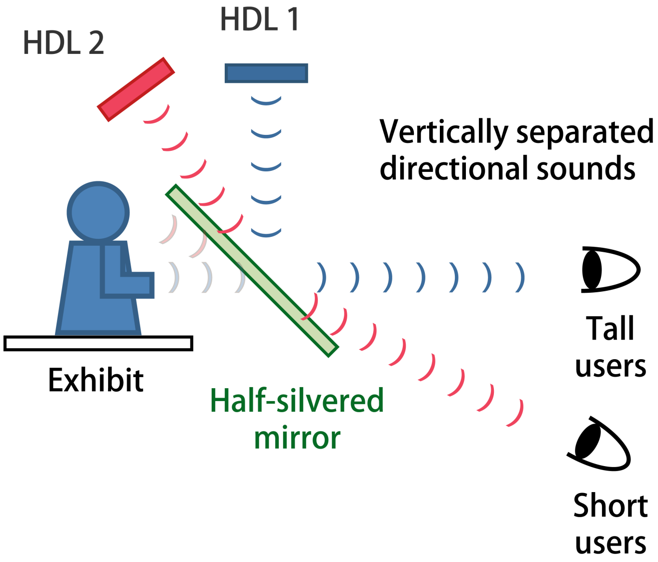


Figure B.1: Vertical sound separation system. Two sets of HDLs are installed with different reflecting angles.

B.1 Auditory System Design

An auditory system is designed for vertical sound separation to provide different auditory information in different directions and heights while making the sound appear to come from the exhibit itself. In our auditory system, we add a new function to HSMs: reflecting directional sound on the HSM surface.

Figure B.1 shows the auditory system design. Since two sets of HDLs are installed above the HSM with different reflecting angles, directional sounds are directed to different heights. Assuming that parents and children will be viewing the showcase at the same time, different audio guides will be heard in accordance with the viewer's height. As a result, this auditory system design realizes targeted information presentation and device-free usage.

B.2 Results

To confirm the vertical sound separation, the sound pressure level every 10 cm in the vertical range from 100 to 220 cm was measured at 1 m from MRsionCase. As

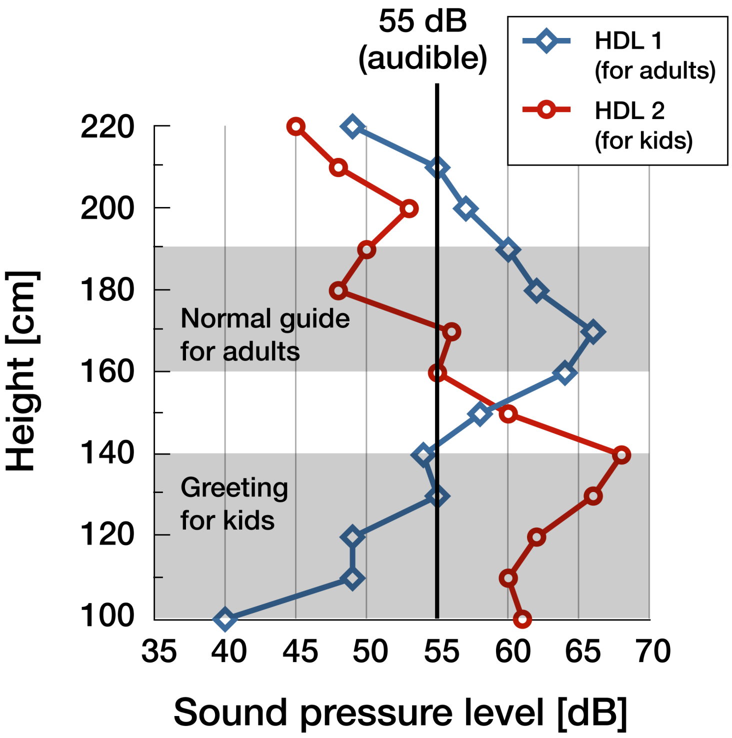


Figure B.2: Sound pressure level distribution by height.

a sound source, 1-kHz white noise from each of the HDLs is used. A Brüel & Kjær 2250 sound level meter was used as the measuring instrument. The results of sound pressure level are plotted in Figure B.2. The red line represents the pressure level of sound from HDL 2 in Figure B.1, which is intended for short people (i.e., children), and the blue line represents that from HDL 1 in Figure B.1, which is intended for taller people (i.e., adults). The red line has a distinct peak at 140 cm, and the blue one has a distinct peak at 170 cm, with the difference over 10 dB to each other. Furthermore, only one directional sound is dominant to its counterpart in each region (100 to 140 and 160 to 190 cm, respectively) as audible sound over 55 dB. This means that two different audio guides are separated to be distinguished from each other. Thus, different auditory content can be delivered to short and tall viewers.

Appendix C

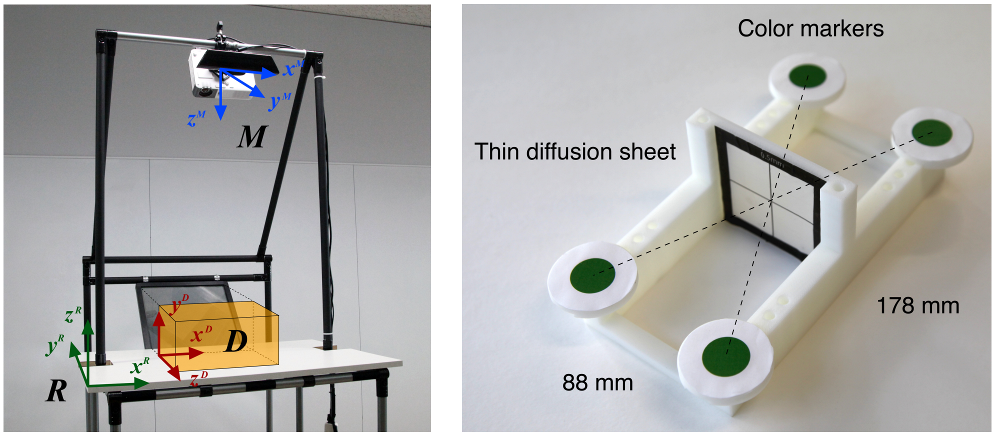
Calibration between Mid-air Images and Real Space

This chapter will introduce a calibration method between mid-air images and real space used in MARIO from the collaborative work with Issei Takahashi.

In MARIO system, there are different coordinate systems for display, depth sensor and real space. Thus, calibration between these coordinate systems is required to form mid-air images with a high level spatial consistency.

C.1 Coordinate System Setting

As for mid-air images, we set three coordinate systems for display (D), depth sensor (M) and real space (R) as shown in Figure C.1(a). D is defined as the space in which mid-air images can be formed (an orange box in Figure C.1(a)). Its origin is located at the left rear corner of the bottom of the imaging space. x- and y-axis are set for horizontal and vertical directions respectively, and z-axis for depth direction. M is defined as the measurable space of the depth sensor. Its origin is at the IR-camera of the depth sensor. z-axis is along with the perpendicular line to the table, which is the same direction of depth measurement. R is for real space. Its origin is defined at



(a) Coordinate system setting. (b) A pointing tool for a mid-air image.

Figure C.1: Calibration of coordinate systems in MARIO system.

the left front corner of the table. xy-plane represents the table surface and z-axis for the height from it.

The goal of this calibration is to calculate a transformation matrix from M to D (T_{DM}).

C.2 Calibration Method

When considering this calibration, however, we faced a problem: mid-air images cannot be detected by an IR-based depth sensor since they have no physical shape. In order to overcome this problem, we devised a new pointing tool which enables a depth sensor to detect the position of mid-air images (Figure C.1(b)). The pointing tool has four markers at the corners which can be detected by RGB-camera of the depth sensor. In the center, there is an upright screen with semi-transparent diffusion sheet. Thus, we can find the imaging position when the screen and the mid-air image are fitted in at the same depth. Since we designed the center of the screen to be identical to the center of four markers, the depth sensor can find the position of the mid-air image by detecting the pointing tool. The coordinates of RGB- and IR-camera were aligned before calibration process.

Calibration process is performed as follows:

1) Place the pointing device in R with given position.

2) Record the position of the pointing device from the depth sensor in M ($\boldsymbol{p_i^M}$).

3) Arrange a mid-air image to the center of the pointing device by adjusting the imaging position, and then record the position of the mid-air image in D ($\boldsymbol{p_i^D}$).

4) Repeat the step 1) to 3) N times at different positions in R.

$\boldsymbol{p_i^M}$, $\boldsymbol{p_i^D}$ $(i = 1, 2, ..., N)$ and T_{DM} are defined as C.1. Since the values in $\boldsymbol{p_i^M}$ and $\boldsymbol{p_i^D}$ include errors, we applied singular value decomposition (SVD) method to calculate T_{DM} which minimizes squared differences using C.2 and C.3.

$$\boldsymbol{p_i^M} \equiv \begin{bmatrix} x_i^M \\ y_i^M \\ z_i^M \\ 1 \end{bmatrix}, \ \boldsymbol{p_i^D} \equiv \begin{bmatrix} x_i^D \\ y_i^D \\ z_i^D \\ 1 \end{bmatrix} (i = 1, 2, ..., N), \ \mathrm{T}_{DM} \equiv \begin{bmatrix} a & d & g & t_x \\ b & e & h & t_y \\ c & f & i & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{C.1}$$

$$\boldsymbol{p_i'^D} \equiv \mathrm{T}_{DM} \boldsymbol{p_i^M} \tag{C.2}$$

$$\underset{\mathbf{T}_{DM}}{\arg\min} \sum_{i=1}^{N} \|\boldsymbol{p_i'^D} - \boldsymbol{p_i^D}\|^2 \tag{C.3}$$

At this time, we used 125 points in 200 × 200 × 200 mm-cube for calibration ($N = 125$).

As for shadow projection, the projector can be assumed as a virtual light source in D from its position measured in R. Then, a shadow will be projected according to the mid-air image positions in D, which is sent to the projector from the mid-air imaging display.

References

[1] Apple Inc.: “iPhone,” `https://www.apple.com/iphone/`

[2] P. Milgram and F. Kishino : “A taxonomy of mixed reality visual displays,” IEICE Trans. on Information and Systems, E77-D, 12, pp. 1321–1329 (1994)

[3] iOnRoad, `http://www.ionroad.com/`

[4] Nintendo Co.Ltd., AR Cards, `http://www.nintendo.com/3ds/ar-cards/`

[5] H. Kim, S. Nagao, S. Maekawa, and T. Naemura: “MRsionCase: A Glasses-free Mixed Reality Showcase for Surrounding Multiple Viewers,” ACM SIGGRAPH ASIA 2012, Technical Briefs, Article No.10 (2012).

[6] H. Kim, S. Nagao, S. Maekawa, and T. Naemura: “MRsionCase: A Glasses-free Mixed Reality Showcase for Surrounding Multiple Viewers,” ITE Trans. on Media Technology and Applications, vol.2, no.3, pp.200-208 (2014).

[7] H. Kim, I. Takahashi, H. Yamamoto, T. Kai, S. Maekawa, and T. Naemura, “MARIO: Mid-Air Augmented RealityInteraction with Objects,” Advances in Computer Entertainment pp.560-563 (2013).

[8] H. Kim, I. Takahashi, H. Yamamoto, S. Maekawa, and T. Naemura, MARIO: Mid-air Augmented Reality Interaction with Objects, Elsevier Entertainment Computing, vol.5, no.4, pp.233-241 (2014).

[9] H. Kim, H. Yamamoto, N. Koizumi, S. Maekawa, and T. Naemura: “HoVerTable: Dual-sided Vertical Mid-air Images on Horizontal Tabletop Display,” In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA ’15), pp.1115-1120 (2015).

[10] I. Sato, Y. Sato, and K. Ikeuchi: “Acquiring a Radiance Distribution to Superimpose Virtual Objects onto a Real Scene,” IEEE Transactions on Visualization and Computer Graphics, Vol. 5, No. 1, pp.1–12, (1999)

[11] R. T. Azuma : “A survey of augmented reality,” Presence-Teleoperators and Virtual Environments, 6, 4, pp. 355–385 (1997)

[12] K. Hayashi, H. Kato, and S. Nishida: "Occlusion detection of real objects using contour based stereo matching," In Proceedings of the 2005 international conference on Augmented tele-existence (ICAT '05), ACM, New York, NY, USA, pp.180–186 (2005).

[13] T. Fukiage, T. Oishi, and K. Ikeuchi, "Reduction of contradictory partial occlusion in mixed reality by using characteristics of transparency perception," In Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR'12), pp.129–139 (2012).

[14] D. E. Breen, R. T. Whitaker, E. Rose, and M. Tuceryan, "Interactive Occlusion and Automatic Object Placement for Augmented Reality," Computer Graphics Forum, Vol.15, No.3, pp.11–22 (1996).

[15] Y. Liu, X. Qin, S. Xu, E. Nakamae, and Q. Peng, Qunsheng: "Light source estimation of outdoor scenes for mixed reality," The visual computer, Vol.25, No.5-7, pp.637–646 (2009).

[16] M. Knecht, G. Tanzmeister, C. Traxler, and M. Wimmer: "Interactive BRDF Estimation for Mixed-Reality Applications," Journal of WSCG, Vol.20, No.1, pp.47–56 (2012).

[17] T. Naemura, T. Nitta, A. Mimura, and H. Harashima: "Virtual Shadows in Mixed Reality Environment Using Flashlight-like Devices," Transactions of the Virtual Reality Society of Japan, Vol.7, No.2, pp.227–237 (2002).

[18] T. Kakuta, T. Oishi, and K. Ikeuchi: "Shading and Shadowing of Architecture in Mixed Reality," In Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '05). pp.200–201 (2005).

[19] O. Bimber and R. Raskar: "Modern approaches to augmented reality," In ACM SIGGRAPH 2006 Courses (SIGGRAPH '06). ACM, Article 1 (2006)

[20] E. Viirre, H. Pryor, S. Nagata, and T. A. Furness: "The Virtual Retinal Display: A New Technology for Virtual Reality and Augmented Vision in Medicine," In Proceedings of Medicine Meets Virtual Reality, San Diego, California, USA, pp.252–257, Amsterdam: IOS Press and Ohmsha (1998).

[21] Oculus VR: "Oculus Rift," `http://www.oculus.com/rift/`

[22] R. Pausch, M. A. Shackelford, and D. Proffitt: "A user study comparing head-mounted and stationary displays," In Proceedings of Symposium on Research Frontiers in Virtual Reality, pp.41–45 (1993).

[23] Sony Computer Entertainment: "AR Play," `http://us.playstation.com/psvita/apps/psvita-app-ar.html`

[24] D. Wagner, T. Pintaric, F. Ledermann, and D. Schmalstieg: "Towards Massively Multi-user Augmented Reality on Handheld Devices," Pervasive Computing, LNCS Vol.3468, pp.208–219 (2005).

[25] R. Raskar, G. Welch, K.-L. Low, and D. Bandyopadhyay: "Shader Lamps: Animating Real Objects With Image-Based Illumination," In Proceedings of the 12th Eurographics Workshop on Rendering Techniques, Steven J. Gortler and Karol Myszkowski (Eds.). Springer-Verlag, pp.89–102 (2001).

[26] K. D.D. Willis, I. Poupyrev, S. E. Hudson, and M Mahler: "SideBySide: adhoc multi-user interaction with handheld projectors," In Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11), ACM, pp.431–440 (2011).

[27] F. A. Jenkins and H. E. White: "Fundamentals of Optics," McGraw-Hill Kogakusha, Tokyo, p.65 (1976).

[28] B. Kattelman: Spectres and Spectators: "The Poly-Technologies of the Pepper's Ghost IllusionTheatre," In Performance and Analogue Technology, K. Reilly (Eds.), Palgrave Macmillan, Ch. 10 (2013).

[29] H. Yamamoto and S. Suyama: "Aerial 3D LED display by use of retroreflective sheeting,," In Proceedings of SPIE Stereoscopic Displays and Applications XXIV, Vol.8648, No.86480Q (2013).

[30] S. Maekawa, K. Nitta, and O. Matoba: "Transmissive optical imaging device with micromirror array," In Proceedings of SPIE, Vol.6392, No.63920E (2006).

[31] ASUKANET, AERIAL IMAGING Technologies [in Japanese] `http://aerialimaging.tv` (Accessed: 1 April 2014)

[32] ASUKANET, Aerial Imaging Optics, Japan patent No.P5437436 [in Japanese]

[33] B. Brown, I. MacColl, M. Chalmers, A. Galani, C. Randell, and A. Steed : "Lessons from the lighthouse: collaboration in a shared mixed reality system," SIGCHI Conference on Human Factors in Computing Systems 2003, pp. 577–584 (2003)

[34] T. Naemura, Y. Kakehi, T. Hashida, Y. Seong, D. Akatsuka, T. Wada, T. Nariya, T. Nakashima, R. Oshima, and T. Kuno : "Mixed reality technologies for museum experience," 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry, pp. 17–20 (2010)

[35] T. Sugiura: "The Principle and Effects of Magic vision," ITEJ Technical Report Vol.8, No.25, pp.37–42 (1984) [in Japanese].

[36] City of Ashibetsu: "Information of Permanent Exhibitions," `http://www.city.ashibetsu.hokkaido.jp/hyakunenkinenkan/hkanri/kinenkan_info.html` [in Japanese] [Accessed: 1May 2015].

[37] T. Nakashima, T. Wada, and T. Naemura : "ExFloasion: Multi-layered floating vision system for Mixed reality Exhibition," 16th International Conference on Virtual Systems and Multimedia, pp.95–98, 2010.

[38] O. Bimber, B. Frohlich, D. Schmalsteig, and L. M. Encarnação : "The Virtual Showcase," IEEE Computer Graphics and Applications, 21, 6, pp. 48–55 (2001)

[39] O. Bimber, S. M. Gatesy, L. M. Witmer, R. Raskar, and L. M. Encarnação : "Merging fossil specimens with computer-generated information," IEEE Computer, 35, 9, pp. 25–30 (2002)

[40] O. Bimber and B. Frohlich : "Occlusion shadows: using projected light to generate realistic occlusion effects for view-dependent optical see-through displays," International Symposium on Mixed and Augmented Reality 2002, pp. 186–195, 319 (2002)

[41] O. Hilliges, D. Kim, S. Izadi, M. Weiss, A. Wilson, HoloDesk: direct 3d interactions with a situated see-through display, In Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI'12), pp.2421–2430, 2012.

[42] T. Yoshida, K. Shimizu, T. Kurogi, S. Kamuro, K. Minamizawa, H. Nii, and S. Tachi, "RePro3D: full-parallax 3D display with haptic feedback using retroreflective projection technology," In Proceedings of IEEE International Symposium on VR Innovation (ISVRI'11), pp.49–54 (2011).

[43] L. Smoot, Q. Smithwick, D. Reetz, A volumetric display based on a rim-driven varifocal beamsplitter and LED backlit LCD, In ACM SIGGRAPH 2011 Emerging Technologies (SIGGRAPH'11), Article 22, 1 page (2011)

[44] H. Nii, K. Zhu, H. Yoshikawa, N. L. Htat, R. Aigner, R. Nakatsu, Fuwa-vision: an auto-stereoscopic floating-image display. In SIGGRAPH Asia 2012 Emerging Technologies (SA'12), Article 13, 1 page (2012)

[45] Y. Ueda, K. Iwazaki, M. Shibasaki, Y. Mizushina, M. Furukawa, H. Nii, K. Minamizawa, and S. Tachi: "HaptoMIRAGE: mid-air autostereoscopic display for seamless interaction with mixed reality environments," In ACM SIGGRAPH 2014 Emerging Technologies (SIGGRAPH '14), Article 10, 1 page (2014)

[46] S. Markon, S. Maekawa, A. Onat, and H. Furukawa, Interactive medical visualization with floating images, In Proc. of International Conference on Complex Medical Engineering 2012 (ICME'12), pp.20–23, 2012.

[47] Benko, H., Jota, R., Wilson, A.: MirageTable: freehand interaction on a projected augmented reality tabletop; In Proc. of CHI 2012, pp.199-208 (2012).

[48] Nagakura, T., Oishi, J.: "Deskrama," In Proceedings of ACM SIGGRAPH 2006 Emerging technologies, Article 6, 1 page (2006).

[49] Kakehi, Y., Naemura, T., Matsushita, M.: Tablescape Plus: Interactive Small-sized Vertical Displays on a Horizontal Tabletop Display; In Proc. of TABLETOP 2007, pp.155-162 (2007).

[50] Kakehi, Y., et al.: Lumisight Table: Interactive View-Dependent Tabletop Display Surrounded by Mutiple Users, IEEE CG&A, **Vol. 25**, No.1, pp 48-53 (2005).

[51] Lintec: Lumisty film; `http://www.lintec.co.jp`.

[52] Wada, T., Naemura, T.: FloasionTable : Multidirectional Tabletop Floating Vision System [in Japanese]; In Technical report of IEICE-MVE, **Vol.109**, No.466, pp.29-34 (2010).

[53] Yoshida, S., Yano, S., Ando H.: Implementation of a tabletop 3D display based on light field reproduction. In Proc. of SIGGRAPH 2010 Posters, Article 61, 1 page (2010).

[54] H. Kim, S. Maekawa, and T. Naemura: "Study on Real Imaging Optical Systems for Multi-layered Floating Images in Mixed Reality Exhibition System MRsionCase," IEICE technical report, Image engineering, 111, 478, pp. 151–156 (2012) (in Japanese)

[55] Edmund optics, 254 x 356mm, 50R/50T, Plate Beamsplitter:
`http://www.edmundoptics.com/optics/beamsplitters/plate-beamsplitters/plate-beamsplitters/46584/`

[56] Digital Contents Expo Official Website [Online]. Avaliable: http://www.dcexpo.jp/ (Accessed: 1 December 2013)

[57] H. Kim, S. Nagao, S. Maekawa, and T. Naemura : "MRsionCase: A Glasses-free Mixed Reality Showcase for Surrounding Multiple Viewers," SIGGRAPH Asia 2012 Emerging Technologies, Article 17, 4 pages (2012)

[58] Nintendo, Super Mario Bros.
`http://mario.nintendo.com` (Accessed: 1 June 2015)

[59] Microsoft, Kinect.
`http://www.xbox.com/en-US/kinect` (Accessed: 1 June 2015)

[60] N. Sugano, H. Kato, K. Tachibana: "The Effects of Shadow Representation of Virtual Objects in Augmented Reality," In Proc. of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR'03), pp.76–83, 2003.

[61] MARIO: Mid-air Augmented Reality Interaction with Objects, Innovative Technologies 2013, Ministry of Economy, Trade and Industry, Japan, 2013.

[62] DeruChara, The 12th MediaLab exhibition, The National Museum of Emerging Science and Innovation (Miraikan), 2013.
`http://miraikan.jp/medialab/en/12.html` (Accessed: 1 April 2014)

[63] S. Follmer, D. Leithinger, A. Olwal, A. Hogge, H. Ishii: “inFORM: dynamic physical affordances and constraints through shape and object actuation,” In Proc. of the 26th annual ACM symposium on User Interface Software and Technology (UIST ’13), pp.417–426, 2013.

[64] Theater house, Co. Ltd.: Semi-transparent projection screen [in Japanese]; `http://theaterhouse.co.jp/p_rear/item_clear_film.html` (Accessed: 2015/6/1).

[65] D. E. Hughes: “Defining an Audio Pipeline for Mixed Reality,” Human Computer Interfaces International 2005 (2005)

[66] K. Higa, T. Nishiura, A. Kimura, F. Shibata, and H. Tamura : “A Two-by-Two Mixed Reality System That Merges Real and Virtual Worlds in Both Audio and Visual Senses,” 6th IEEE/ACM International Symposium on Mixed and Augmented Reality, pp. 203–206 (2007)

[67] A. Zimmermann and A. Lorenz: “LISTEN: a user-adaptive audio-augmented museum guide,” User Modeling and User-Adapted Interaction, 18, 5, pp. 389–416 (2008)

[68] P. J. Westervelt: “Parametric Acoustic Array,” Journal of the Acoustical Society of America, 35, 4, pp. 535–537 (1963)

[69] A. Zerroug, A. Cassinelli, and M. Ishikawa: “Invoked computing,” Virtual Reality International Conference 2011 (2011)

List of Publications

Journal articles (referred)

[1] H. Kim, H. Yamamoto, N. Koizumi, S. Maekawa, and T. Naemura: "HoVerTable: Combining Dual-sided Vertical Mid-air Images with a Horizontal Tabletop Display," Journal of Human Interface Society [submitted].

[2] H. Kim, I. Takahashi, H. Yamamoto, S. Maekawa, and T. Naemura: MARIO: Mid-air Augmented Reality Interaction with Objects, Elsevier Entertainment Computing, vol.5, no.4, pp.233–241 (2014).

[3] H. Kim, S. Nagao, S. Maekawa, and T. Naemura: "MRsionCase: A Glasses-free Mixed Reality Showcase for Surrounding Multiple Viewers," ITE Trans. on Media Technology and Applications, vol.2, no.3, pp.200–208 (2014).

Conference proceedings (referred)

[4] H. Yamamoto, H. Kajita, H. Kim, N. Koizumi, and T. Naemura: "Mid-air Plus: A 2.5 D Cross-sectional Mid-air Display with Transparency Control," ACM SIGGRAPH 2015, Posters (2015) [Accepted].

[5] H. Kim H. Yamamoto, N. Koizumi, S. Maekawa, and T. Naemura: "HoVerTable: Dual-sided Vertical Mid-air Images on Horizontal Tabletop Display," In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15), pp.1115–1120 (2015).

[6] H. Kim, I. Takahashi, H. Yamamoto, T. Kai, S. Maekawa, and T. Naemura: "MARIO: Mid-Air Augmented RealityInteraction with Objects," Advances in Computer Entertainment pp.560–563 (2013).

[7] H. Kim, S. Nagao, S. Maekawa, and T. Naemura: "MRsionCase: A Glasses-free Mixed Reality Showcase for Surrounding Multiple Viewers," ACM SIGGRAPH ASIA 2012, Technical Briefs, Article No.10 (2012).

[8] H. Kim, S. Nagao, S. Maekawa, and T. Naemura: "MRsionCase: A Glasses-free Mixed Reality Showcase for Surrounding Multiple Viewers," ACM SIGGRAPH ASIA 2012, Emerging Technologies, Article No.17 (2012).

[9] H. Kim, S. Nagao, S. Maekawa, and T. Naemura: “MRsionCase: A Glasses-free Mixed Reality Showcase for Surrounding Multiple Viewers,” ACM SIGGRAPH ASIA 2012, Posters, Article No.25 (2012).

Conference proceedings (non-referred)

[10] H. Yamamoto, H. Kim, N. Koizumi, and T. Naemura: “Horizontal and Vertical Mid-air Images Inside and Outside of a Mirror for Mixed Reality,” 3D Image Conference 2015 (2015) [in Japanese, to be presented].

[11] I. Takahashi, H. Kim, N. Koizumi, and T. Naemura: “Registration of mid-air images and physical objects for MR applications,” IPSJ SIG technical reports, Vol.2014–EC–32, No.28, 6 pages (2014) [in Japanese].

[12] H. Yamamoto, H. Kim, N. Koizumi, and T. Naemura: “A Proposal of Shadow Projection onto Mid-air Images for Mixed-Reality System,” 3D Image Conference 2014, Article No.4–3 (2014) [in Japanese].

[13] H. Kim, S. Maekawa, and T. Naemura: “Study on Real Imaging Optical Systems for Multi-layered Floating Images in Mixed Reality Exhibition System MRsionCase,” Technical report of IEICE–MVE, Vol.111, No.479, pp.151–156 (2012) [in Japanese].

[14] H. Kim, S. Maekawa, and T. Naemura: “MRsionCase: Multidirectionally Viewable MR-Showcase using Multi-layered Floating Images,” In Proceedings of VRSJ the 16th Annual Conference, 22B–4 (2011) [in Japanese].

Exhibitions

[15] “MARIO: Mid-air Augmented Reality Interaction with Objects,” CEATEC 2014, Chiba, Japan (Sep. 2014).

[16] ”MARIO: Mid-air Augmented Reality Interaction with Objects,” Japan EXPO 2014, Paris, France (Aug. 2014).

[17] “MARIO: Mid-air Augmented Reality Interaction with Objects,” Advances in Computer Entertainments 2013(ACE ’13), Twente, The Netherlands (Nov. 2013).

[18] “MARIO: Mid-air Augmented Reality Interaction with Objects,” Innovative Technologies 2013, Tokyo, Japan (Oct. 2013).

[19] “DeruChara (Mid-air Augmented Reality Interaction with Objects),” Miraikan 12th Media Laboratory Exhibition, Tokyo, Japan (Jul. 2013–Jan. 2014).

[20] “MRsionCase: A Glasses-free Mixed Reality Showcase for Surrounding Multiple Viewers,” ACM SIGGRAPH ASIA 2012 Emerging Technologies, Singapore (Nov. 2012).

[21] “MRsionCase,” Digital Contents EXPO 2012, Tokyo, Japan (Oct. 2012).

[22] “MRsionCase: Mixed Reality + emerSion of image + showcase,” ConTex Digital Contents EXPO 2011, Tokyo, Japan (Oct. 2011).

Awards and honors

[23] Best Demo Gold Award, “MARIO: Mid-Air Augmented RealityInteraction with Objects,” Creative Showcase, Advances in Computer Entertainment (2013).

[24] Innovative Technologies, “MARIO: Mid-Air Augmented RealityInteraction with Objects,” Digital Contents EXPO, Ministry of Economy, Trade and Industry (2013).