공학박사 학위논문

Development of Large-Scale Three-Dimensional Displays for Immersive Viewing Experiences

몰입형 시청 경험을 위한 대형 3차원 디스플레이 연구

2023년 2월

서울대학교 대학원

전기·정보공학부

조 영 진

공학박사 학위논문

Development of Large-Scale Three-Dimensional Displays for Immersive Viewing Experiences

몰입형 시청 경험을 위한 대형 3차원 디스플레이 연구

2023년 2월

서울대학교 대학원

전기·정보공학부

조 영 진

Development of Large-Scale Three-Dimensional Displays for Immersive Viewing Experiences

지도 교수 정 윤 찬

이 논문을 공학박사 학위논문으로 제출함

2023년 2월

서울대학교 대학원

전기·정보공학부

조 영 진

조영진의 공학박사 학위 논문을 인준함

2023년 2월

위 원 장: ______________________ (인)

부위원장: ______________________ (인)

위 원: ______________________ (인)

위 원: ______________________ (인)

위 원: ______________________ (인)

Abstract

Development of Large-Scale Three-Dimensional Displays for Immersive Viewing Experiences

Youngjin Jo
Department of Electrical and Computer Engineering
College of Engineering
Seoul National University

Over the past century, as display technology has evolved, people have increasingly demanded more realistic and immersive experiences. With the advancement of technology, high-definition large-scale displays are feasible, and three-dimensional (3D) displays for a more immersive experience are now emerging as the next step. However, there is a problem in achieving a 3D image from a two-dimensional (2D) panel because much more information is required than before. Since this need overwhelms the information capacity of current displays, various systematic methodologies have been defined and developed to overcome the limited information. Nevertheless, each and every 3D display technology has its own trade-offs. As a result, various limitations, such as visual fatigue, low resolution, and narrow viewing angle, have hindered the user's immersive 3D experience. In the end, with the development of hardware, a new optical methodology that can address the limitation of information capacity should be developed for a truly immersive 3D display.

In this dissertation, the author presents research conducted from a new perspective that addresses the limitations by combining the benefits of previously developed methodologies for the 3D display. Accordingly, three optical systems are newly introduced to overcome existing limitations for a more immersive experience. First, the author proposes a tomographic projection system that provides appropriate focus cues by applying a multifocal technique in the stereoscopic system. This system can potentially alleviate visual fatigue and ensure a more immersive viewing experience in a space such as a movie theater while extending the depth range of 3D objects right in front of the user. Second, the author proposes a light field projection system that can enhance 3D image quality even in a comfortable viewing environment without stereo glasses. In this system, the information between the integral imaging and multifocal display is optically converted to achieve ultrahigh-definition light fields. Finally, in the holographic display that can freely express 3D objects, the bandwidth, which is the limiting point, is effectively extended. Here, the viewing angle is widened while reducing the speckle noise by applying a multiple illumination strategy, and at the same time, the binary mask gives a degree of freedom to include more information. As a result, the author proposes a holographic system that can efficiently compress the information of holograms nine times through optimization.

Currently, ultrahigh-resolution and ultra-large displays have already become a reality with the rapid development of technology. As the virtual world is getting closer to reality via the metaverse, the author firmly believes that the next step forward will be truly 3D displays with immersive experiences. In this context, it is hoped that this dissertation will provide a new perspective and contribute to the further development of the 3D display field.

keywords: 3D display, volumetric display, stereoscopic display, light field display, holographic display, large-scale projection, immersive experience
student number: 2017-27135

Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Overview of three-dimensional displays

1.1.1 Three-dimensional space

The world is made up of three-dimensional (3D) space, and the people who live there can feel the 3D sense of space as if they were breathing naturally. With the development of technology, it is possible to provide enough information to people through various displays, from large-scale displays to ultrahigh-resolution displays. However, without 3D senses of space, it is difficult to say that display technology has reached an experience and immersion identical to reality. As the digital world gradually approaches the real world, extracting realistic 3D experiences from displays is more important than ever in today's display technology.

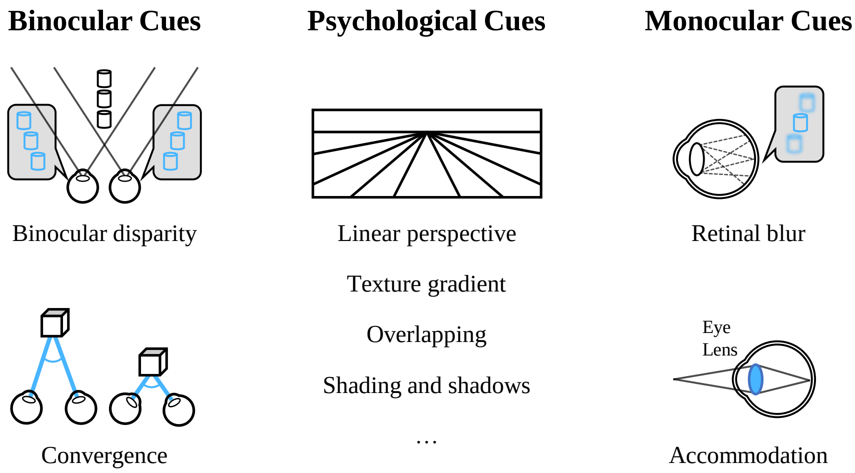

Figure 1.1 Various depth cues for three-dimensional vision.

To convey a sense of depth in a display, one must first know how a person perceives depth. For living things, estimation of the object's depth was an essential element for survival, and the human visual system has evolved to perceive the depth using physiological and psychological depth cues from the obtained images through the eyes [4]. If there is a point in space (also called a voxel) in the real environment, both eyes acquire the light emitted from the point and recognize it as being in space via a number of properties, as shown in Fig. 1.1. First, humans have two eyes approximately 63mm apart from each other [5], which makes the images received from each eye dissimilar. This binocular disparity determines the degree of convergence, the movement of the two eyes to form a fused image. Second, as the light rays enter the eye's pupil, each eye tunes the focal power of the eye lens to form clear retinal images. This focal power adjustment of the eye lens, called accommodation, corresponds to the monocular focus cues. These two oculomotor movements of accommodation and convergence are closely related and hard to be separated independently. Specifically, when one is stimulated, the other is induced and stimulated together via cross-links (AC/A and CA/C ratios) [6].

In addition to the physiological cues described above, there are various psy-

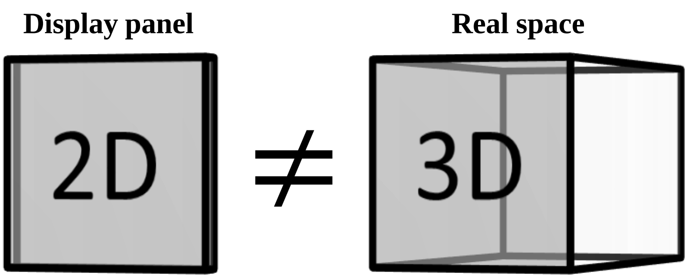

Figure 1.2 Difficulty in depth representation. Extending a 2D display into a 3D volume usually requires more information.

chological cues, such as linear perspective, texture gradient, overlapping, shading, and shadow [7]. When all these depth cues come to the user harmoniously and naturally, as in the real world, we can feel the real depth through the display and get an truly immersive visual experience. Among these cues, psychological depth cues can be easily supported through computer graphics on display. On the other hand, in the case of physiological depth cues, there is a critical limitation because the display panel we watch is 2D rather than 3D, as shown in Fig. 1.2. Thus, previous research in the 3D display field is a history of how these problems have been addressed and overcome [8]. Therefore, it is important to outline the conventional 3D displays in terms of the limitations when experiencing immersive 3D.

1.1.2 Bottlenecks of conventional 3D display

Stereoscopic display

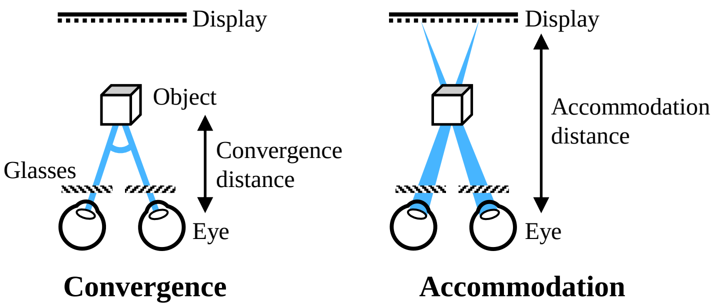

Figure 1.3 Convergence and accommodation distance in the stereoscopic display.

The first limitation to consider is that a 2D flat display located at a certain depth cannot provide well-matched focus cues. The stereoscopic display, the most popular type of 3D display, uses an optical filter in the form of glasses in front of the eyes to provide different images to each eye, creating binocular disparity. On the other hand, the focus cue is fixed at a certain depth where the display or screen is positioned, as shown in Fig. 1.3. Therefore when attempting to perceive an object with a different depth from the display, the two physiological depth cues do not adequately match in the stereoscopic environment, and the intrinsic cross-linked behavior between convergence and accommodation is decoupled. Consequently, viewers may experience a conflict between convergence and accommodation, which can cause double vision or blurred images. Furthermore, uncomfortable experiences such as dizziness and nausea may occur if this conflict is significant [9–11]. Therefore, with a single focus cue of the current stereoscopic display, only 3D contents with a limited depth range near its focal plane can be expressed [12]. This unpleasant viewing experience and narrow depth range potentially limit the immersion.

Autostereoscopic display

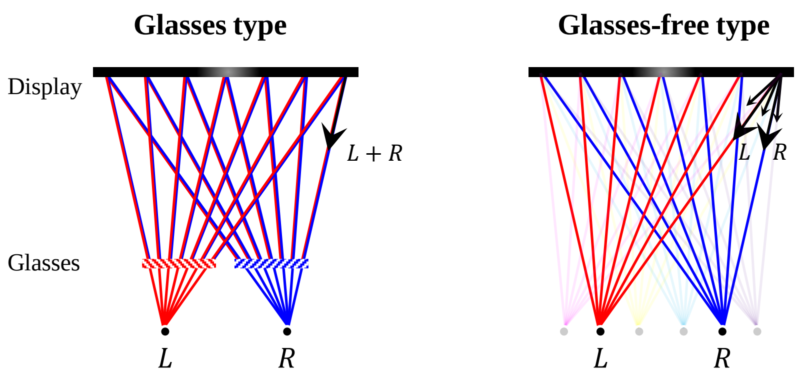

Figure 1.4 Illustration of stereoscopic 3D displays with or without glasses. Both systems deliver depth information via binocular disparity. Compared to glasses type displays, glasses-free type displays should reconstruct light field or holographic wavefront to provide binocular disparity.

Second, as one dimension expands from 2D to 3D, the amount of information to be displayed increases dramatically. Unless the glasses are positioned as a filter, the 3D display should produce several independent images for different viewing angles to provide the binocular disparity, as shown in Fig. 1.4. This is called an autostereoscopic display, and it should be able to express not only spatial resolution like 2D imagery but also angular resolution in which information is different for each angle. To achieve this, methods for forming an angular resolution through the barrier [13, 14], lenses [15, 16], stacking of panels [17, 18], and interference of light [19, 20] have been proposed. However, since the information capacity of the display is limited by spatial resolution, a reduction in the spatial resolution is inevitable to produce the angular resolution. Thus, in the process of dividing the display's available resources as efficiently as possible, the viewing area is limited, or the image quality is lowered. These challenges are another barrier to an immersive visual experience. In short, taking off the glasses ensures a comfortable viewing environment for 3D displays, but the system must be able to express a huge amount of information, which has been a difficult challenge to solve.

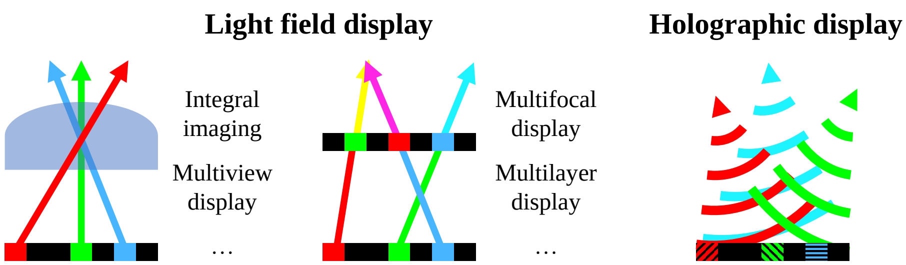

Figure 1.5 Various autostereoscopic displays.

Integral imaging display

As shown in Fig. 1.5, light field display can represent spatial-angular information of light rays in space [21]. One of the promising autostereoscopic displays, integral imaging display, utilizes a lens array to realize the light field of 3D objects. If the rays of the light field are dense enough, the viewer can obtain both monocular and binocular cues simultaneously with continuous parallax [22]. In addition, increasing the display size is easily possible with projection, and much research has been conducted to enhance the viewing angle or image resolution through the spatial-temporal multiplexing method [23–25]. However, the light field is defined as a one-to-one relationship between a ray and a pixel, so there is a trade-off between spatial and angular resolution. As a result, the image resolution of integral imaging displays is much lower than that of a flat panel display. Furthermore, with the use of lens optics, the expressible depth of 3D images is restricted around the central depth plane [26, 27].

Multifocal display

Another light field display, the multifocal display, positions focal planes in space [28]. Because the focal plane is physically or optically located in real space, it can support monocular focus cues. Furthermore, unlike the lens-based light field display, it has the representative advantage of being able to compress the light fields using the correlation between the focal planes [29–31]. This approach allows for more efficient use of limited information and a relatively high-resolution 3D image. However, the multifocal display is difficult to implement on a large scale because the actual focal plane should be constructed in space [32]. Furthermore, low light efficiency and diffraction effect become problems for the method of stacking multiple panels [33]. Even in the method of sequentially floating the focal planes, a device capable of high-speed operation is essential, and it suffers from losses in terms of bit depth or frame rate [34, 35]. In addition to this, the depth range of multifocal displays is also limited to near the focal planes, such as the limitation of integral imaging.

Holographic display

In this regard, the holographic display is the most promising technology as it can produce high-resolution 3D objects in a wide depth range even without glasses [36]. Moreover, as the holographic display modulates the wavefront of light using a spatial light modulator (SLM), it can be inherently free from the aforementioned ray-based limitations [37]. However, since the interference and diffraction of coherent light are used to form a wavefront, the hologram can be modulated within the diffraction range of the SLM's pixel [38]. So, an extremely small pixel pitch is essential for a sufficient viewing angle. At the same time, it requires a huge amount of information and computational resources to realize a display size similar to that of a conventional 2D display [39]. Therefore, the holographic display has great potential for small-scale applications such as near-eye displays for augmented and virtual reality [40, 41], whereas replacing larger displays requires significant hardware advances from the current level of technology. Furthermore, speckle noise generated from a coherent laser source degrades the image quality, which is another problem that must be addressed in holographic displays.

1.2 Motivation and purpose of this dissertation

Method	Focus cues	Depth range	Image resolution	Scale enlargement	Viewing angle	Issue
Stereoscopic display	No	Narrow	High	Easy	Wide	Glasses
Integral imaging display	Yes	Moderate	Low	Easy	Moderate	
Multifocal display	Yes	Moderate	Moderate	Moderate	Moderate	
Holographic display	Yes	Wide	High	Difficult	Narrow	Speckle noise

Figure 1.6 Comparison of current 3D displays with their characteristics. Limitations to be addressed in this dissertation are marked with red boxes.

The previous section identified the technical background and limitations of existing 3D display technology. There are many factors that need to be met for a true 3D display. Among them, immersive 3D experiences can be realized via monocular focus cues and wide depth range, as well as high resolution, wide viewing angle, and larger screen size for high-quality images. However, as shown in Fig. 1.6, each 3D display technique suffers from inherent drawbacks such as the absence of focus cues, the need for glasses, low image quality, limited viewing angle, and depth range.

Current issues of 3D displays mainly originate from the difficulty of filling 3D space with the information capacity of flat-panel displays. In other words, with enough information, it is possible to realize perfect 3D displays on a large scale. However, even if a display with ultrahigh-resolution and a very high frame rate is developed, related problems would arise in other fields, such as cost, rendering time, and shortage of transmission bandwidth. Therefore, this dissertation does not rely on improving hardware but on developing an information-efficient system through a new optical design for a realistic and immersive 3D display along with the existing technologies.

In this dissertation, the author presents an optical-based solution that is less limited in information capacity. Each chapter addresses the most critical issues for the realization of an immersive, large-scale 3D display. The purpose of this dissertation is to give an optical system with a new perspective that overcomes the limitations by combining the benefits of different techniques and methodologies in the 3D display. First, in the stereoscopic display, monocular focus cues are provided. It allows displaying of 3D objects right in front of the eyes with a wide depth range, providing a more immersive viewing experience. Then, after taking off the 3D glasses that cause inconvenience, the trade-off relationship in the autostereoscopic display is addressed. The proposed optical design eliminates the inefficient use of information, enabling ultrahigh-resolution 3D experiences without glasses. Finally, for an advanced viewing experience with a wider depth range, a holographic display is studied. At this time, the author introduces the concept of extending the narrow bandwidth, which is a bottleneck for realizing a large screen, while alleviating the challenge of speckle noise. Figure 1.7 visualizes the relative values of each factor in the form of a hexagon. The direction and goals to be achieved in this dissertation are expressed with red arrows and hatched areas.

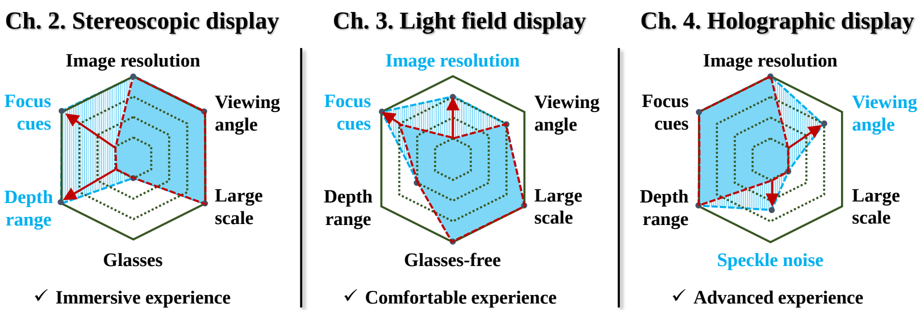

Figure 1.7 Overview of chapters in the dissertation.

1.3 Scope and organization

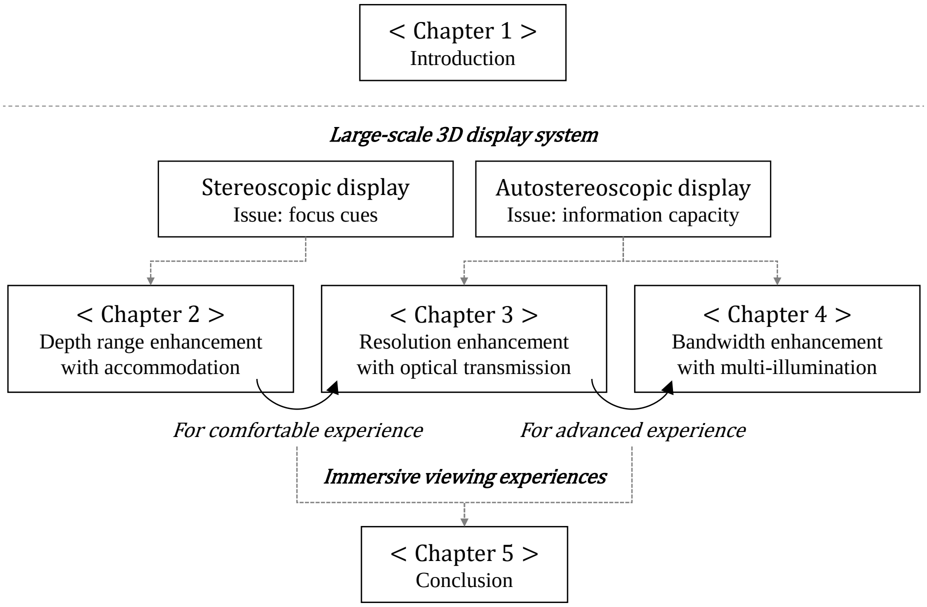

Figure 1.8 Organization of the dissertation.

In Chapter 1, the essential factors in achieving a realistic 3D experience are introduced, and the current issues in research on 3D displays are described based on two perspectives of information capacity and monocular focus cues. Then, the author briefly describes the dissertation and the approaches to reaching large-scale 3D displays for a more immersive experience. The organization of this dissertation is described in a block diagram shown in Fig. 1.8. Each chapter contains the concept idea, the principle of the optical system, and experimental results that support the idea with a thorough analysis.

In Chapter 2, the author proposes a large-scale stereoscopic display featuring a wide depth range with focus cues. The stereoscopic can easily provide binocular disparity and high-resolution images on a large screen with 3D glasses that serve as filtering. However, another depth cue, accommodation cues, has not been considered in stereoscopic systems despite being essential for a comfortable and immersive 3D experience. The author introduces a tomographic projection system capable of providing focus cues by forming multiple focal planes. The presented system can reconstruct 60 focal planes without image quality degradation. With a continuous range of focal planes from 25 cm to optical infinity, 3D objects can be placed right in front of the user's eyes. Consequently, Experiments and analysis demonstrate that the proposed system enables a more immersive 3D experience in stereoscopic environments such as movie theaters. However, the need for glasses is another issue that has to be addressed for a more comfortable viewing experience.

The research in Chapter 3 lays a technological foundation for realizing an ultrahigh-definition 3D display without 3D glasses. To take off the glasses for a comfortable viewing environment and improve 3D image quality, the author introduces a novel optical design that combines two 3D technologies: integral imaging and multifocal display. Optically connecting these two technologies enables efficient use of information and addresses the trade-off between spatial and angular resolution in the integral imaging display. As a result, light field projection on a large scale is realized, with the effective resolution increased by 36 times. However, since the proposed method is based on integral imaging, which utilizes additional optics such as a lens array, the resolution of 3D objects decreases as the object is far from the in-focus plane.

In Chapter 4, the author introduces a holographic display with enhanced information capacity. Holographic displays have competitive advantages of wide depth range, natural parallax, and high image quality, which can support a more advanced 3D experience. However, limitations such as narrow viewing angle and small display size have hindered its practical use, as the amount of information required overwhelms the hardware of the display. To overcome this, two approaches are introduced: multi-illumination strategy and binary mask. The author applies multiple illuminations to extend the narrow diffraction angle of the holographic display while reducing speckle noise. Moreover, using the binary mask mitigates information correlation and effectively compresses the information. This perspective of overcoming the hardware-limited performance could inspire further developments for large-scale holographic displays.

Finally, Chapter 5 concludes the dissertation by summarizing the contributions of the previous chapters. Several research topics to improve the proposed methods are also discussed.

Chapter 2. Depth range enhancement of stereoscopic display by providing focus cues

2.1 Introduction

2.1.1 Current issues in stereoscopic 3D for large scale

As displays continue to evolve, people have increasingly sought out more realistic and immersive experiences on a large scale. Accordingly, movie theaters have evolved from the silent film era to present-day 3D cinema in an effort to offer audiences a fuller viewing experience. Recently, stereoscopic 3D movies have been introduced in many theaters, allowing audiences to perceive depth in the imagery. Stereoscopic 3D movies enable audiences to observe different images with binocular disparity using their two eyes [42]. With binocular depth cues and high-resolution images, stereoscopic 3D movies could deliver an unprecedented immersive viewing experience. The stereoscopic 3D movie industry has already proven its commercial value, notwithstanding the inconvenience of wearing glasses for stereoscopic viewing. In 2009, the 3D movie 'Avatar' has witnessed tremendous success and is still ranked as one of the most profitable movies. However, several concerns have been reported since stereoscopic 3D movies were commercialized. Some people had uncomfortable experiences such as dizziness and nausea after watching a stereoscopic 3D movie [43].

As described in Chapter 1, one of the most convincing causes of those uncomfortable experiences is visual fatigue [10]. This visual fatigue is caused by the discrepancy between observing stereoscopic 3D imagery and a real volumetric object. Because stereoscopic 3D systems can only provide binocular dis-

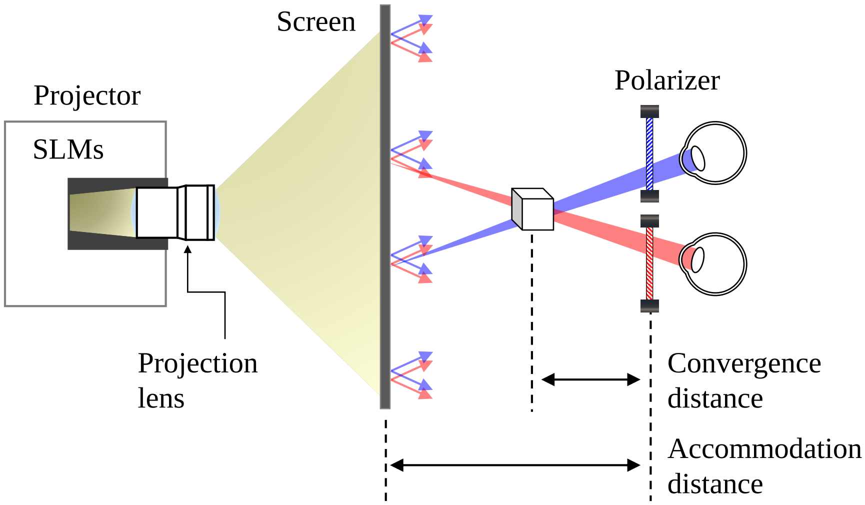

Figure 2.1 Projection-type stereoscopic display system.

parity among the physiological cues, the absence of appropriate focus cues may lead to visual fatigue in some circumstances [44]. In stereoscopic 3D theaters, binocular images stimulate the eyes to converge at various depths in order to perceive the 3D effect. However, the focus cue is fixed at the screen depth where eye lenses should focus for clear observation, as shown in Fig. 2.1. In that case, viewers are unconsciously confused because the depth information corresponding to the binocular stimulation differs from that of the monocular accommodation cue [45]. Because convergence and accommodation cannot be set independently with its cross-links, each physiological depth cue repeatedly affects another one. This can lead to a vergence-accommodation conflict (VAC) [46].

2.1.2 Requirements for stereoscopic displays

Since the conflict must be mitigated for comfortable viewing, the importance of providing focus cues has been consistently emphasized in many research articles. Especially providing focus cues has been thoroughly discussed in applications such as augmented reality (AR) or virtual reality (VR), where viewers interact with objects at close distances. As near-eye displays for AR/VR primarily deliver depth information through binocular disparity, it is essential to provide focus cues simultaneously in order to alleviate the VAC problem.. In this regard, various research groups have conducted perceptual studies to investigate the effect of VAC on user experience [47–49]. In the stereoscopic 3D, the objects within the Panum's fusional area are perceived as single images, and the objects falling outside the Panum's fusional area are perceived as diplopia (unfused double vision) [50]. According to those perceptual studies, people may feel discomfort if vergence and accommodation mismatch occurs more than 0.5 diopters in an unusual way. This tolerance range becomes wider at the peripheral vision [51] and is influenced by many factors [10]. Note that only the central vision is taken into consideration when defining the comfort range.

In addition to the focus cues discussed above, there are several additional challenges that must be addressed to improve the 3D theater experience, particularly in terms of the display system. First, the display system should have a high angular resolution so that the human visual system can make focused images of a virtual scene. Second, it is important to ensure that all audience members have a consistent viewing experience regardless of their seat location. Third, the display system should have a wide viewing zone to accommodate a large number of audience members. Although several studies have been introduced to improve 3D theater experiences, none of them have been able to address all three challenges, as they mainly aimed to realize a glasses-free environment. For instance, super multiview display [52] and integral imaging display [53] sacrifice spatial resolution to increase angular resolution. Holographic display [54] has a fixed viewing position, which is not appropriate in the theater environment where providing a wide viewing zone is a valuable factor.

2.1.3 Volumetric stereoscopic system

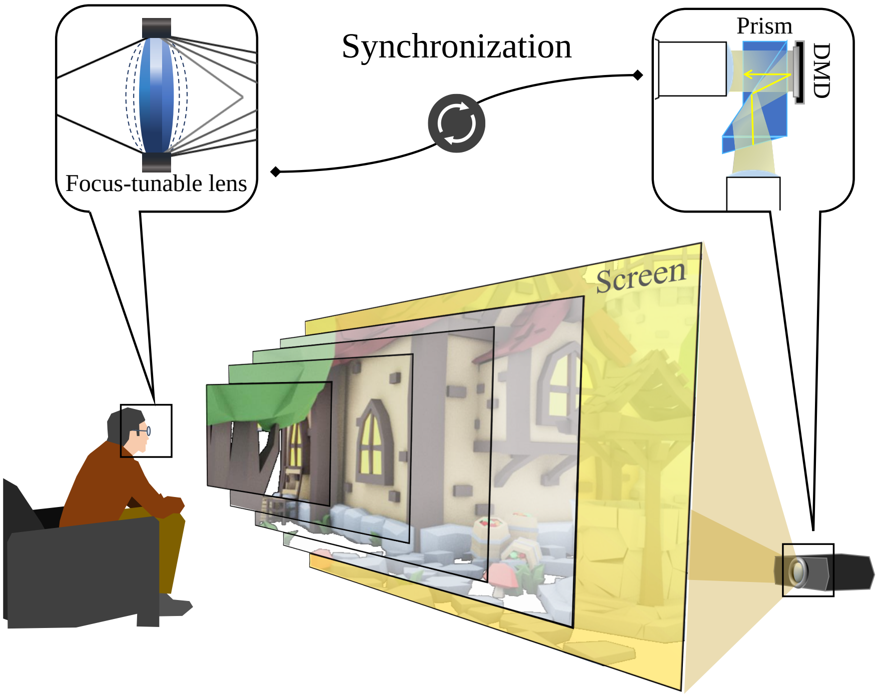

Figure 2.2 Volumetric stereoscopic system for use in large screen. The system consists of a tomographic projector that projects time-sequential images onto the screen and a focus-tunable lens located in front of the user's eye [1].

In Chapter 2, the author proposes a volumetric stereoscopic system that enables audiences to receive both focus cues and binocular disparity. The proposed system has several advantages, including the ability to provide a wide range of focus cues without sacrificing resolution, frame rate, or bit depth; ensuring a consistent viewing experience for all audiences; and providing a wide viewing zone that can accommodate many viewers. These notable features are achieved through a conceptual shift that adopts wearing stereoscopic glasses.

In the proposed system, every audience is supposed to wear a pair of focus-tunable optics, as illustrated in Fig. 2.2. These focus-tunable optics are synchronized with a tomographic projector reproducing multiple focal planes via temporal multiplexing. As the focus-tunable lens (FTL) rapidly changes the depth of the projection screen, the tomographic projector displays the corresponding images. At the same time, the binocular disparity could be achieved by applying either shutter glasses or polarization glasses. The tomographic projector consists of a digital micromirror device (DMD) and liquid crystal display (LCD) panel to refresh screen images at a fast frame rate. This combination of DMD and LCD panel allows for more efficient use of information with less loss in terms of bit depth or frame rate compared to previous methods [34, 35]. This system can be considered as a volumetric display with multiple planes positioned at various depths, allowing for a wider range of expression in 3D content. As a result, the proposed projection system enables users to view close objects in stereoscopic 3D without VAC and provides a more immersive 3D experience. In addition, it can alleviate the distortion problem caused by the difference in eye separation of each person in conventional stereoscopic 3D. Additionally, incorporating a DMD into the optical structure of a traditional projector is a feasible solution from an industrial perspective.

In this Chapter, the author first looks at the principle and background of the proposed system and demonstrates the tomographic projector's ability to create adequate focal planes. Second, the author thoroughly analyzes the geometric specifications of the proposed system in terms of viewing distance, viewing angle, and eye-box. Third, an optimization algorithm is introduced to expand the eye-box and present an accurate occlusion boundary without noticeable artifacts. The author concludes with discussions to enhance the experiences with the tomographic projection system.

2.2 Principle

2.2.1 Tomographic projector

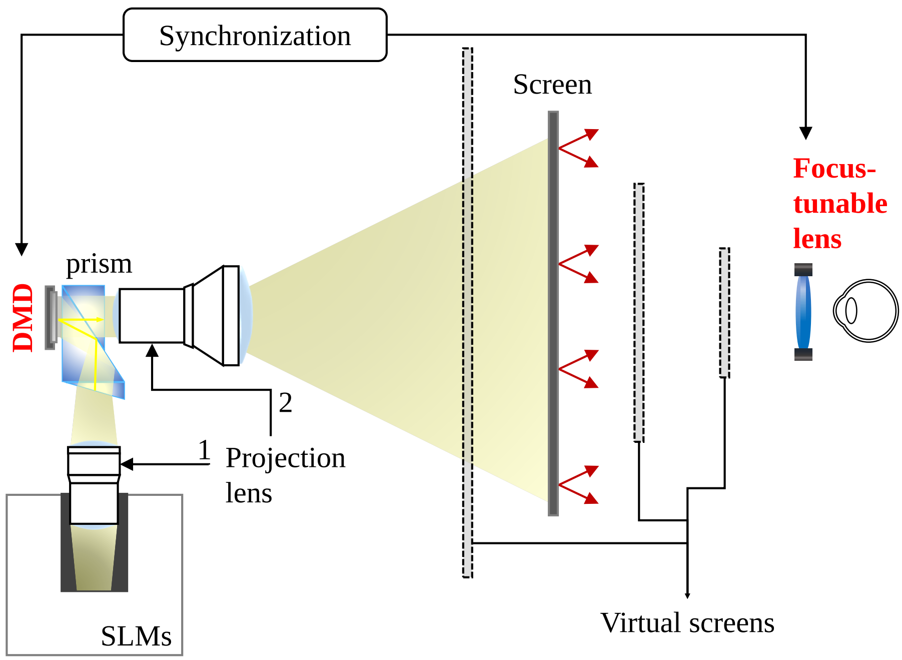

Figure 2.3 Illustration of the tomographic projection system. The system is divided into a tomographic projector, a screen, and focus-tunable lenses. Each audience is supposed to wear focus-tunable lenses synchronized with the tomographic projector.

Figure 2.3 shows a schematic diagram of the tomographic projector. The author adopts a FTL as an eyepiece that can sweep a wide dioptric range at the speed of 60 Hz. Through the eyepiece, the optical distance between the audience and the screen changes periodically from near to far and far to near. During each periodic cycle, the tomographic projector displays depth-sliced sequential images on the screen at the appropriate times. The human vision system recognizes the sequential images at different depths as a synthesized volumetric scene because the periodic cycle of FTL takes a short time. Namely, the tomographic projector is a temporal multiplexed volumetric display in which the user perceives multifocal plane images as the integration of time-sequential retinal images during the specific rate [55].

2.2.2 Depth-sliced sequential images

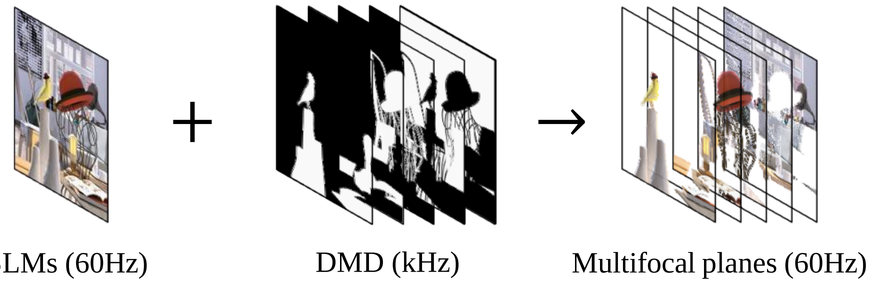

Figure 2.4 Schematic procedure of generating multifocal planes. By combining SLMs and DMD, the tomographic projector can refresh projection images at a fast frame rate resulting in a volumetric scene.

Depth-sliced sequential images can be generated by dividing the perspective image based on the depth of a volumetric scene. Appropriate focus cues are delivered when the tomographic projector is synced with the FTL to display each depth-sliced image at the desired depth. In other words, the tomographic projector should have a much higher refresh rate than 60 Hz to display a depth-sliced image at the desired moment. To reach such a fast refresh rate, the author combines the DMD and three amplitude SLMs. This combination enables much higher refresh rates while maintaining resolution and bit depth. The DMD, which functions as a localized and binarized filter for SLM pixels, converts SLM images into sequential images at a faster frame rate. Since DMD is possible to operate at 16 kHz, it has the capability to refresh projection images more than 260 times within 1/60 second [35]. Note that the resolution and bit depth of projection images are determined solely by the SLMs' specifications.

2.2.3 Artifacts at the occlusion boundary

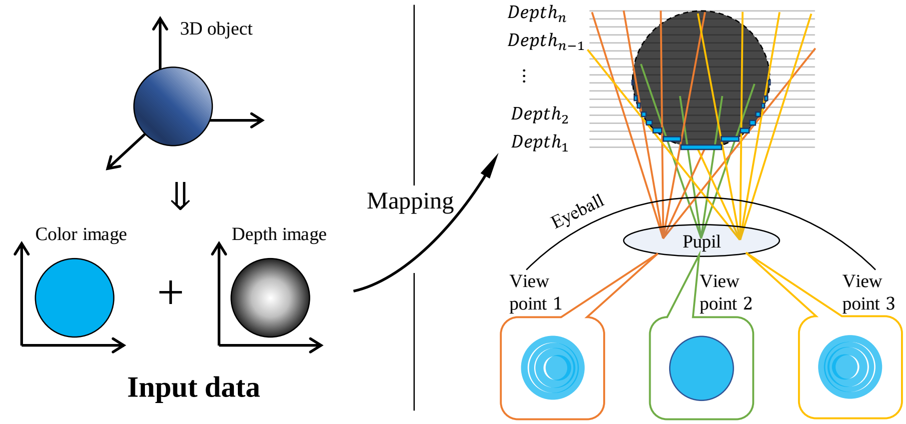

Figure 2.5 Depth-sliced decomposition of a volumetric scene when an intensity and depth map are provided. As illustrated on the right side, synthesized view images may contain artifacts such as separation or overlap according to the viewpoints within the eye-box [1].

By displaying each depth-sliced image at a corresponding distance through the FTL, the audience is expected to enjoy a comfortable viewing experience without the VAC. However, as described in Fig. 2.5, the difference between a synthesized scene and a real scene does not guarantee an immersive and realistic experience. In contrast to the natural environment, each depth-sliced image from the tomographic projector is unable to conceal light from a rear plane, as illustrated in Fig. 2.6. This limitation leads to artifacts in areas where the depth

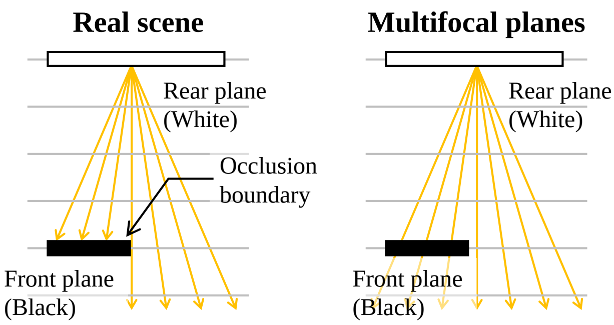

Figure 2.6 Illustration of occlusion boundary artifacts.

discontinuity is noticeable. To reduce these artifacts, alternative rendering approaches have been proposed in various studies with similar limitations. It has been demonstrated that it is possible to imitate occlusion of the light field by optimizing multifocal plane images [18]. Inspired by this idea, the artifacts could be reduced by optimizing both DMD image sequences and SLM image.

The optimization process for the DMD sequence is similar to the reconstruction process of discrete computed tomography (DCT) [56]. For reconstruction of tomographic images, DCT collects X-ray illumination intensity profiles of a volumetric tissue from several directions. The logarithm of an intensity profile is represented by a line integral of the attenuation coefficient. The volumetric profile of the attenuation coefficient can be reconstructed by back projection of the intensity profiles. As a result, DCT reconstructs low-bit depth tomographic images (e.g., binary images) that represent the volumetric information of the tissue. In the tomographic projector, the logarithm of an X-ray intensity profile corresponds to a perspective view image within the eye-box. The binary tomographic images are equivalent to the DMD image sequences. Thus, the author can optimize the DMD image sequences for optimal representation of volumetric scenes using a similar approach as tomographic reconstruction.

2.3 Implementation

2.3.1 Miniaturized tomographic projection system

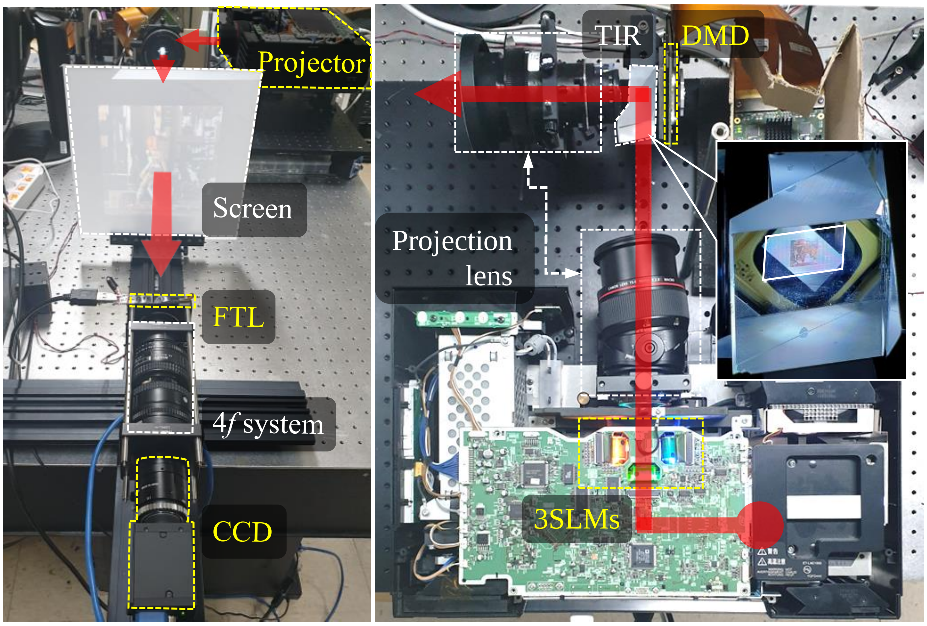

Figure 2.7 Photograph of the projection system (left) and the tomographic projector (right). The prototype is a 25 times miniaturized version of 3D theater [1].

To evaluate the feasibility of the tomographic projector, an experimental setup is built, as depicted in Fig. 2.7. The tomographic projection system includes three SLMs that generate a red/green/blue (RGB) image, a DMD with a total internal reflection (TIR) prism, projection lens units, and a FTL placed in front of the eye. More specifically, an image generated by the projector is relayed to the DMD via the first projection lens and TIR prism. The DMD then locally filters the image to refresh the projection images 60 times within a 1/60 second interval. After passing through the second projection lens, the filtered images are displayed on the screen. Each sequential screen image is shown to the user at the intended depth by the FTL that periodically adjusts its focal length in sync with the DMD. The FTL sweeps the depth of a virtual screen from 25 cm (4 D) to infinity (0 D) at a rate of 60 Hz. The depth range is divided into 60 focal planes with intervals of 0.07 D. As a result, the user perceives 60 focal plane images simultaneously.

Assuming eye relief, a distance from the eye to the FTL, is 14 mm, the system has a diagonal field of view (FOV) of 39 degrees, which is determined by the aperture of the FTL. The projector screen is 40 cm away from the FTL, and the image size is 20 cm. Note that the system can reproduce on a larger scale but is limited by the experimental space. It can be considered an environment where theater size is reduced by 25 times or more. The author believes the experimental results could be extended to the original scale environment without significant or unexpected artifacts. It is not expected that system resolution, occlusion artifacts, or focus cues will differ from the scaled-down results.

2.3.2 System design

The PT-AE1000E beam projector is disassembled and modified for three SLMs that produce an RGB image. A tilt-shift lens (Canon TS-E 80mm) is used to relay the image to the DMD. The micromirror of the DMD has a tilt angle of 12° between on- and off-state, so the illumination should be angled 24° (double the tilt angle) from the vertical. The tilt-shift lens allows the image planes to be slanted according to the Scheimpflug principle, which refers to the relationship between the image plane and a tilted imaging lens in geometrical optics [57]. The DLP9500 from Texas Instruments is used as the DMD with full high-definition (HD) resolution (1920×1080). As the illumination needs to enter the DMD at a 45-degree angle (perpendicular to the micromirror hinge-axis), the author only uses a resolution of 720×720, which is rotated 45 degrees as the

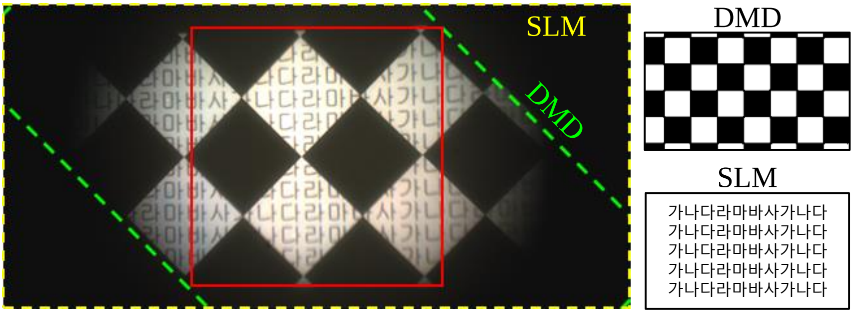

Figure 2.8 Photograph of the projection image. Both images from DMD and SLM are well-focused at the same time.

spatial modulation area as shown in Fig. 2.8. The resulting projection image has a resolution of 670×670 and a frame rate of 60 Hz.

The focus-tunable lens used in the experiment is the EL10-30-TC from Optotune. A negative offset lens is placed in front of the FTL, which allows the focus sweep range to extend to negative values. This offset lens makes FTL possible to express an object closer than the screen. With a focal length of -75 mm, the offset lens shifts the FTL's focal range to between -5 D and 6.7 D. Because the screen is 400 mm away in the experimental setup, the FTL diopter should sweep the range between -1.5 D and 2.5 D to represent 4 D. A $4f$ relay using lenses with an F-number of 1.4 (Nikon) is placed in front of the CCD camera. It is to ensure sufficient eye relief, which is limited by a C-mount lens.

For synchronization of FTL and DMD to display the images at an intended depth/an appropriate time, the data acquisition (DAQ) board from National Instrument is utilized. The DAQ board generates two reference clock signals, which are synchronized using LabView. As shown in Fig. 2.9(a), one is the triangle wave at 60 Hz varying the focal length of the FTL and the other one is for DMD to update the sequential backlight images. Both signals are generated with the same internal clock speed, so only synchronization at the starting point is required. The two signals have a sampling rate of 7200 Hz, resulting in 120 binary images being allocated for one period of the focus-tunable lens (1/60s). At this time, black images are inserted during the half cycle. This is done to reduce the crosstalk effect that may occur between adjacent frame images.

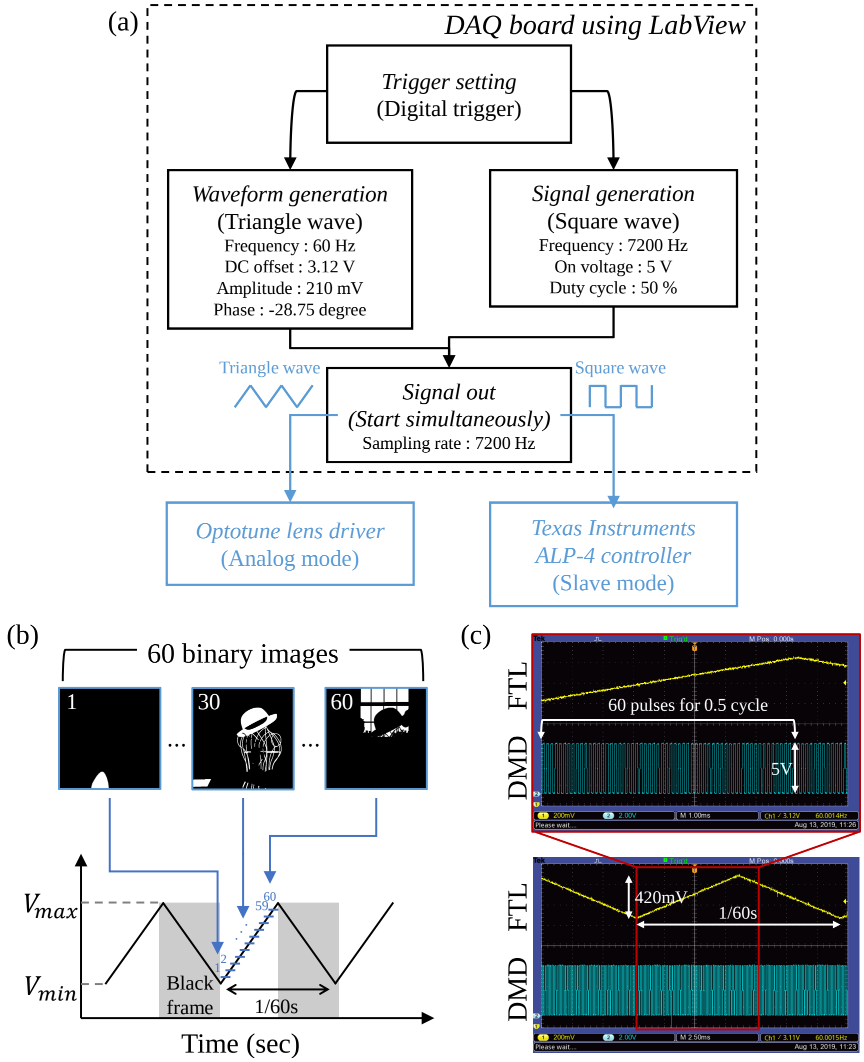

Figure 2.9 Illustration of (a) block diagram, (b) time-sequential process of synchronization, and (c) captured oscilloscope images [1].

2.4 Analysis

2.4.1 Determine viewing zone with uniform experience

One of the most important goals of 3D theater is to ensure uniform 3D experiences for all viewing positions. In stereoscopic 3D theaters, audiences perceive distorted depth cues due to variations in their viewing position, including distance from the screen and angle from the normal perspective view [58]. Similarly, the depth information provided by the tomographic projection system could also be distorted when the viewing position changes. Depth distortion of 3D contents may result in an unpleasant experience. The following section analyzes the relationship between the depth distortion and viewing position using a geometric analysis of the tomographic projector, as shown in Figs. 2.10 and 2.11.

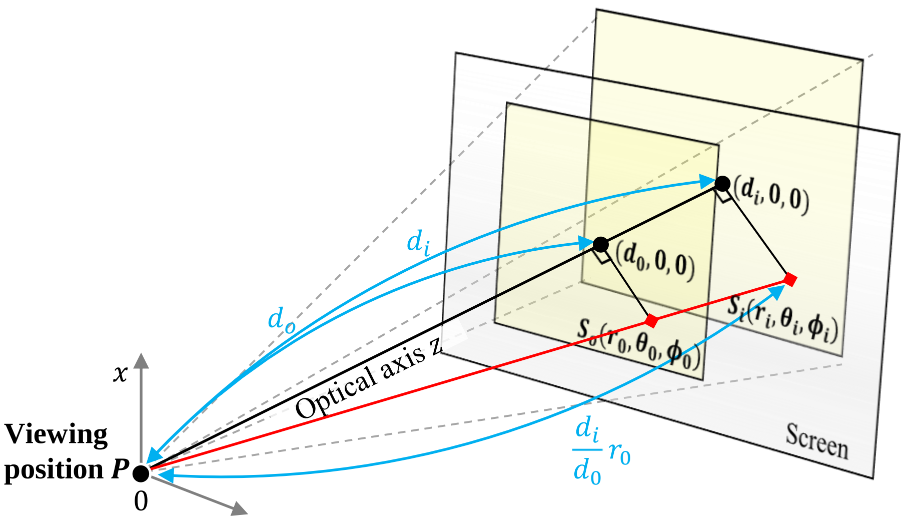

Figure 2.10 Illustration of the geometric analysis. In a spherical coordinate system, the virtual image magnified by a lens is formed in the same direction [1].

Geometric analysis in spherical coordinate

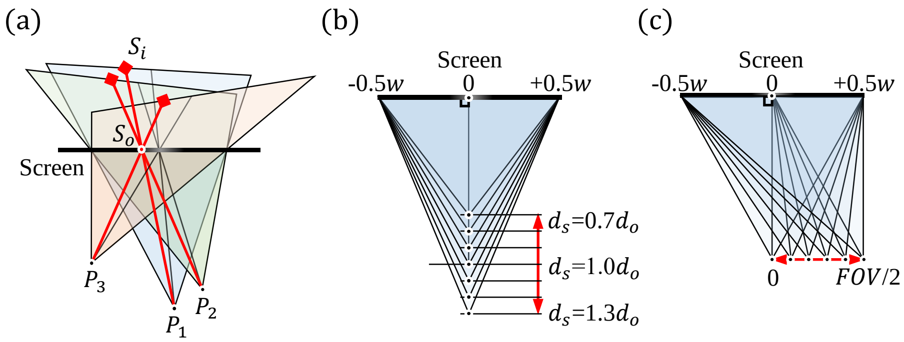

Figure 2.11 Illustration of the distortion cases. (a) The depth information of the virtual image would be different depending on the viewing position. The degree of geometric distortion varies according to the (b) viewing distance and (c) viewing angle [1].

In the tomographic projector, the focal plane depth for each image is determined by the focal length of the FTL, as described by the thin lens formula below.

$$1/d_o + 1/(-d_i) = 1/f, \tag{2.1}$$

where d_o is the distance to the screen and f is the focal length of the FTL, the distance $d_i > 0$ of the virtual image can be calculated.

$$M = d_i/d_o. \tag{2.2}$$

The magnification ratio M of the virtual image is proportional to the floating distance d_i of the image. When there is a pixel of S_o as shown in Fig. 2.10, it will be magnified by the lens and placed at S_i. In the spherical coordinate, S_i can be expressed as the following principle of similar triangles.

$$r_i = Mr_o, \theta_i = \theta_o, \phi_i = \phi_o \tag{2.3}$$

where r, θ, and ϕ are spherical coordinates with the origin at the center of the FTL. As shown in Eq. 2.3, θ and ϕ remain constant regardless of the magnification ratio. This means that the pixel does not undergo any angular movement as the focal length of the FTL changes. Note that it is valid regardless of the FTL's optical axis direction or central position.

Viewing distance

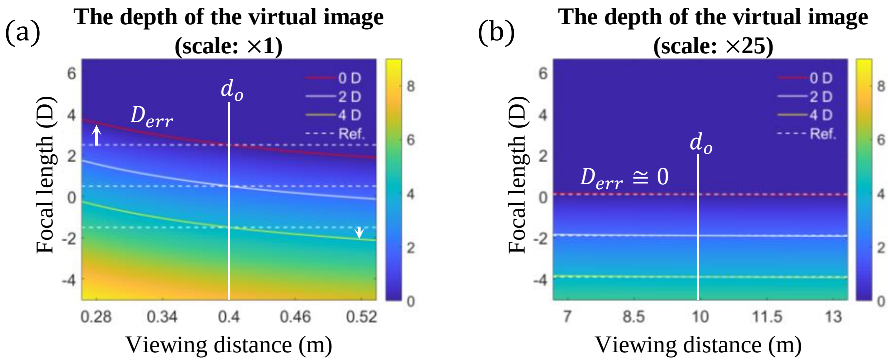

Figure 2.12 The simulation results of focus cue error when the viewing distance is changed by $\pm 33\%$. (a) In a laboratory setting, focus cues are shifted when the viewing distance is adjusted from the reference at d_o. Compared to the white dashed line that indicates desired performance, generated focus cues may contain dioptric errors. (b) In 25 times scale-up environment, however, focus cues follow the guideline regardless of viewing distance [1].

According to Eq. 2.1, the virtual screen depth is determined by both the focal length of the FTL and the viewing distance d_o. Here, if the viewing distance deviates from the reference value, focus cues will be distorted. Figure 2.12 illustrates how changes in viewing distance can accumulate focus cue errors. The color bar indicates the focal depth of the screen. The red and yellow lines represent conditions where the focal depths are fixed at 0 D and 4 D, respectively.

The error in focus cues caused by a change in viewing position is given by

$$D_{err} = |1/d_s - 1/d_o|, \tag{2.4}$$

where d_o is the reference viewing distance to the screen and d_s is the changed viewing distance. It is confirmed that the depth distortion occurs in the experimental environment, which had the specifications miniaturized by a factor of 25.

However, in practical terms, the depth distortion caused by variations in viewing distance is negligible for the following reasons. First, the variation of viewing distance becomes much less influential in the actual theater environ-

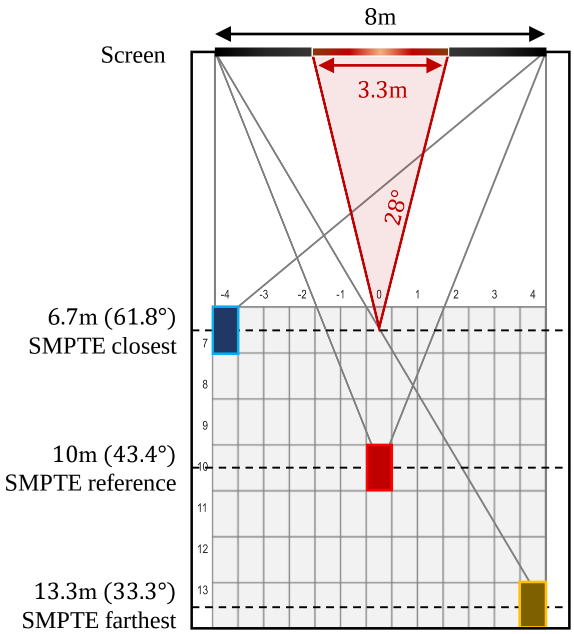

Figure 2.13 Simulated viewing zone of an actual theater environment where the screen is 8 meters wide. The viewing zone is determined according to SMPTE standard EG 18-1994 which is the guideline of the horizontal viewing angle for movie theaters, and the number of seats is 7 (row)$\times$13 (column) [1].

ment where the scale of the system is increased (25 times larger), as shown in Fig. 2.13. This is because the focus cue error is inversely proportional to the reference viewing distance, as indicated by Eq. 2.4. Secondly, severe depth variations can be corrected by adjusting the sweep range of the FTL or the focal length of the offset lens. Since the optical power of FTL increases with current linearly, the change of the sweep range can be controlled by shifting the offset value of the input signal (white line in Fig. 2.12). Thirdly, the focus cue error is usually much smaller than the binocular depth distortion that occurs in a conventional stereoscopic 3D system.

Viewing angle

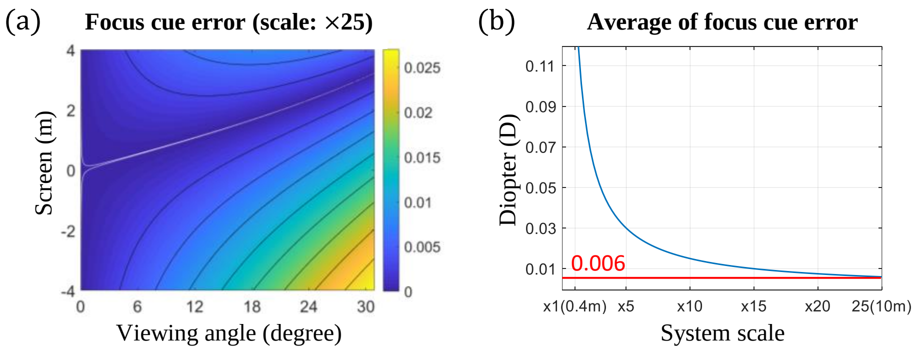

Figure 2.14 (a) Focus cue error analysis when the viewing angle shifts up to 31° in 25 times scale-up environment. The color bar refers to the focal depth of the screen. Similar to viewing distance analysis, (b) the degree of focus cue error caused by viewing angle shift is also inversely proportional to the reference viewing distance [1].

Most theaters are unable to provide central seating for all audiences. Therefore, some audiences should watch the screen from different slanted vertical or horizontal angles. However, variation in viewing angle may lead to depth distortion in the tomographic projection system. This is because the viewing angle causes the optical axis of the lens to no longer be perpendicular to the screen, resulting in depth distortion. In other words, the virtual screen image becomes slanted based on the viewing angle. This phenomenon leads to focus cue errors that can be expressed as follows.

$$D_{err} = |1/(d_s \sec\theta - x \sin\theta) - 1/d_s|, \tag{2.5}$$

where θ is the angle from the normal perspective view and x is the displacement away from the center on the screen. Note that the author assumes that the optical axis of the FTL is directed toward the center of the screen.

The simulation results in Fig. 2.14 show focus cue error caused by viewing angle variation. The focus cue error caused by viewing angle variation is calculated in the horizontal direction of the screen while the viewing angle is changed up to 30.9 degrees, which is the half FOV of the closest seat based on SMPTE standards. In the experimental setting, a depth distortion of up to 0.5 D is observed. Unlike the focus cue error caused by variations in viewing distance, it is difficult to individually correct the focus cue error caused by viewing angle variations. However, the distortion is decreased to a negligible amount in 25 times scale-up environment as shown in Fig. 2.14(b). This can also be explained by the fact that the focus cue error is inversely proportional to the viewing distance. Therefore, it can be concluded that the tomographic projection system can provide uniform 3D experiences in terms of focus cues. Under the practical environment where the screen is at a distance of meters, the proposed system has a wide viewing zone.

Vergence-accommodation conflict

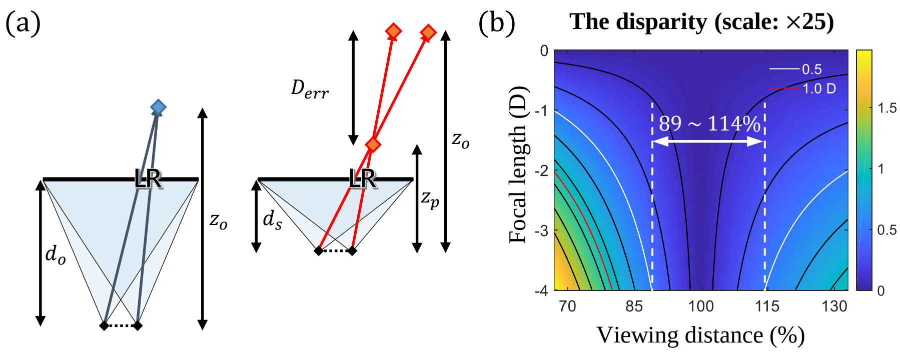

Figure 2.15 Illustration of (a) convergence distance error and (b) the degree of VAC according to viewing positions. As illustrated in the figure, the convergence angle is dependent on the viewing position. This convergence distance error gives rise to unexpected VAC problems [1].

In previous sections, the author has verified the capability of the proposed system to support accurate focus cues regardless of the viewing position in theaters. However, another distortion of binocular cues should be considered according to the viewing position [12] as it can give rise to unexpected VAC. When the viewing distance changes from d_o to d_s, the convergence distance of z_o changes to z_p, as shown in Fig. 2.15(a). When z_o is infinity (0 D), z_s is also infinity regardless of the viewing distance, meaning that the entire depth range is inversely proportional to the distance of the screen.

$$z_p = z_o d_s / d_o \tag{2.6}$$

While the tomographic projector provides uniform focus cues regardless of the viewer's position, the convergence distance is unstable. This may cause discomfort to the viewers due to VAC, which is given by

$$D_{err} = |1/z_p - 1/z_o|. \tag{2.7}$$

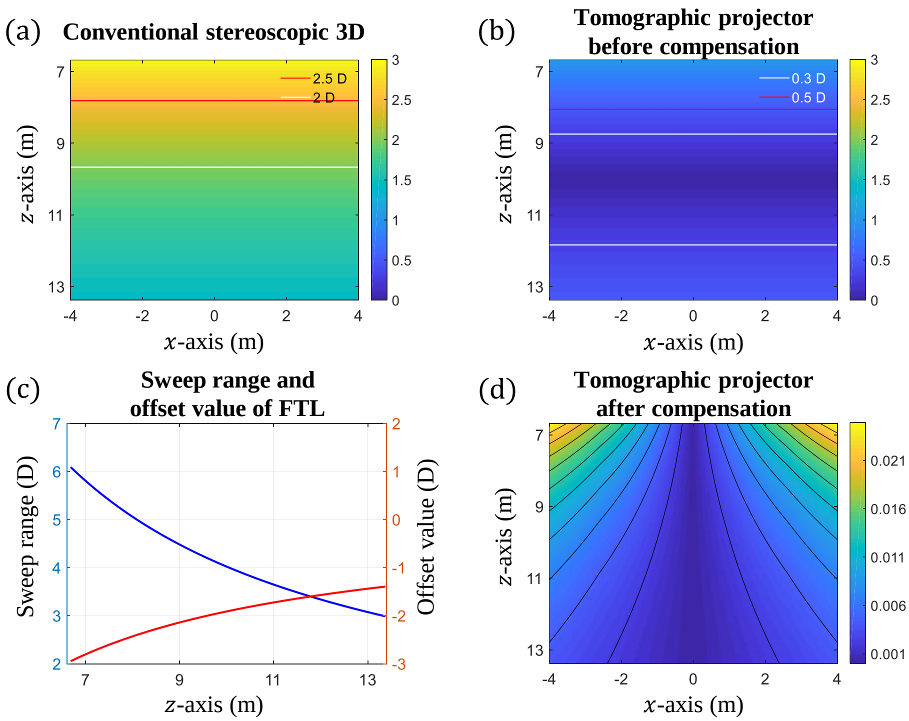

Figure 2.16 The VAC analysis for (a) conventional stereoscopic 3D system and the (b) tomographic projection system in the theater environment of Fig. 2.13. The color bar indicates an average value of D_{err} in the dioptric unit over the entire depth range. (c) The VAC can be further alleviated if FTL is employed for the correction of the convergence error. It is done by changing the sweep range (blue line) and the offset value (red line) corresponding to each viewing distance. (d) The compensation result is represented. After the depth distortion is compensated, no significant VAC is observed regardless of viewing position [1].

According to previous research, people may feel discomfort with focus cue errors above 0.5 $\sim$ 1 D [47]. Typically, the screen is a few meters away, so a focus diopter is less than one. Therefore, to avoid discomfort to audiences, the convergence distance of the contents should be farther than 1 m. As shown in

Fig 2.16(a), if the convergence distance is provided from 0 D up to 4.0 D, the disparity is larger than 3.0 D, causing a severe conflict between accommodation and convergence at all seats.

On the other hand, the proposed system has some comfortable viewing zone, as shown by the red line in Fig. 2.16(b). In this specification, 71 out of 91 seats offer a comfortable viewing zone where the average D_{err} is less than 0.5 D, allowing audiences to experience a volumetric scene. Furthermore, by individually adjusting focus cues at each position, audiences can enjoy comfortable viewing experiences without significant VAC in all seats, as shown in Fig. 2.16(d). In summary, the tomographic projection system can effectively alleviate the VAC problem according to the viewing distance compared to conventional stereoscopic 3D.

2.4.2 Eye-box for individual audience

In the tomographic projection system, it is also important to analyze the tolerance of individual viewing experience because a viewer wears a pair of focus-tunable lenses to see virtual images. The pair of focus-tunable lenses allow the viewer to see sufficient volumetric imagery within the eye-box or exit pupil. The exit pupil is the area where scattered light from the projection screen is delivered. To view the projection screen without experiencing the vignetting effect, the viewer's eye should be located within the exit pupil. The eye-box is referred to as a region where the viewer can observe accurate volumetric imagery. If the viewer's eye is located outside the eye-box, they may recognize the artifacts in the volumetric scenes, as previously shown in Fig. 2.5. Therefore, the tomographic projector should secure a sufficiently large exit pupil and eye-box to have tolerance for pupil movement.

The tomographic projector reconstructs several focal plane images that are synthesized to formulate volumetric imagery. The alignment of focal plane images is essential to reconstruct accurate volumetric scenes. However, the alignment of focal plane images changes based on the position of the pupil. If the pupil is not at the desired point, the viewer may recognize the separation or overlap of focal plane images. As a result, the eye-box of the tomographic projection system is supposed to be restricted to a fixed point. This also leads to artifacts in the representation of occlusion boundaries [18]. However, the eye-box could be enlarged by applying a computational optimization that solves a binary least squares problem. Inspired by previous research that alleviates occlusion boundary artifacts in multifocal displays, the author proposes an algorithm that solves the least squares problem for the tomographic projector.

Optimization algorithm

The binary least squares problem has similarity with least squares problems for multifocal displays [18, 59, 60]. In multifocal displays, the author finds optimal focal plane images that reconstruct accurate retinal focal stacks or pupil view images. The tomographic projector uses a similar approach to optimize the reconstructed focal plane images. However, it is necessary to consider the distinct features of the tomographic projector. One characteristic of the tomographic projector is the correlation between its focal plane images. The tomographic projector's focal plane images are determined by the multiplication of a single 24-bit image on the SLM and 1-bit images on the DMD. In other words, all focal plane images share the information given by the identical 24-bit image on the SLM. Second, it is important to consider that the DMD only supports 1-bit images, which is referred to as a binary constraint of the least squares problem. The binary constraint makes the least squares problem non-deterministic polynomial-time hard (NP-hard).

Here, the author deals with the relaxation of the NP-hard problem by applying an alternating least-squares (ALS) strategy. The SLM image and the DMD image sequence are updated independently. Each iteration consists of two least squares problems to update the SLM and DMD images. For instance, the DMD image sequences are assumed constant when the SLM image is updated, and vice versa. To make least squares problems easy to solve, The binary constraint is ignored for the DMD images at each iteration. Without the binary constraint, the two least squares problems can be solved by using SART [61]. The image sequence is initialized with a given red/green/blue/depth (RGB-D) image, and 100 iterations are performed. Every 30 iterations, the DMD images are transformed into a binary image sequence that shows minimum errors in both energy and variance. The energy is the sum of the DMD image sequence, and the variance is the absolute difference caused by the updates. For example, if there is a DMD pixel supposed to be 0.25 for the first three images and 0.3 for the last image, the corresponding binary image pixel becomes 1 for the last image to balance the energy.

Tolerance for interpupillary distance variation.

Eye separation (interpupillary distance, IPD) is an important factor to consider when producing 3D content. Generally, it has a value of around 63 mm, but there is a deviation for each person. If a person's IPD differs from the value used when creating 3D content, it can cause depth distortion while viewing. Especially if a person's IPD is smaller than the reference value, it can cause the eyes to diverge outward, resulting in double vision (diplopia) without convergence. This ocular divergence can occur in people with a smaller IPD, such as children and women, which causes visual fatigue [62].

If the tolerance of the eye-box is increased by applying optimization, it is

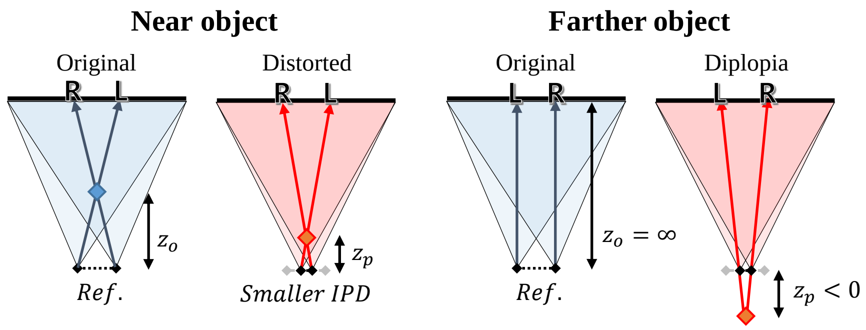

Figure 2.17 Convergence distortion caused by IPD variation. When IPD is smaller than the reference, the user experiences exaggerated 3D than the ground truth. An object in front of the screen gets closer, and another object beyond the screen gets farther. The binocular disparity of the screen larger than IPD causes diplopia for the farther object [1].

possible to compensate for various IPD per user. In other words, the proposed system could address the IPD mismatch caused by human variation in conventional stereoscopic 3D. This can be accomplished by setting the space between the focus-tunable lenses to the reference value. In this case, the focus cues are optically adjusted based on how much a person's IPD deviates from the reference, allowing for compensation of the distortion within the eye-box of the system.

2.5 Results

2.5.1 Multifocal plane generation

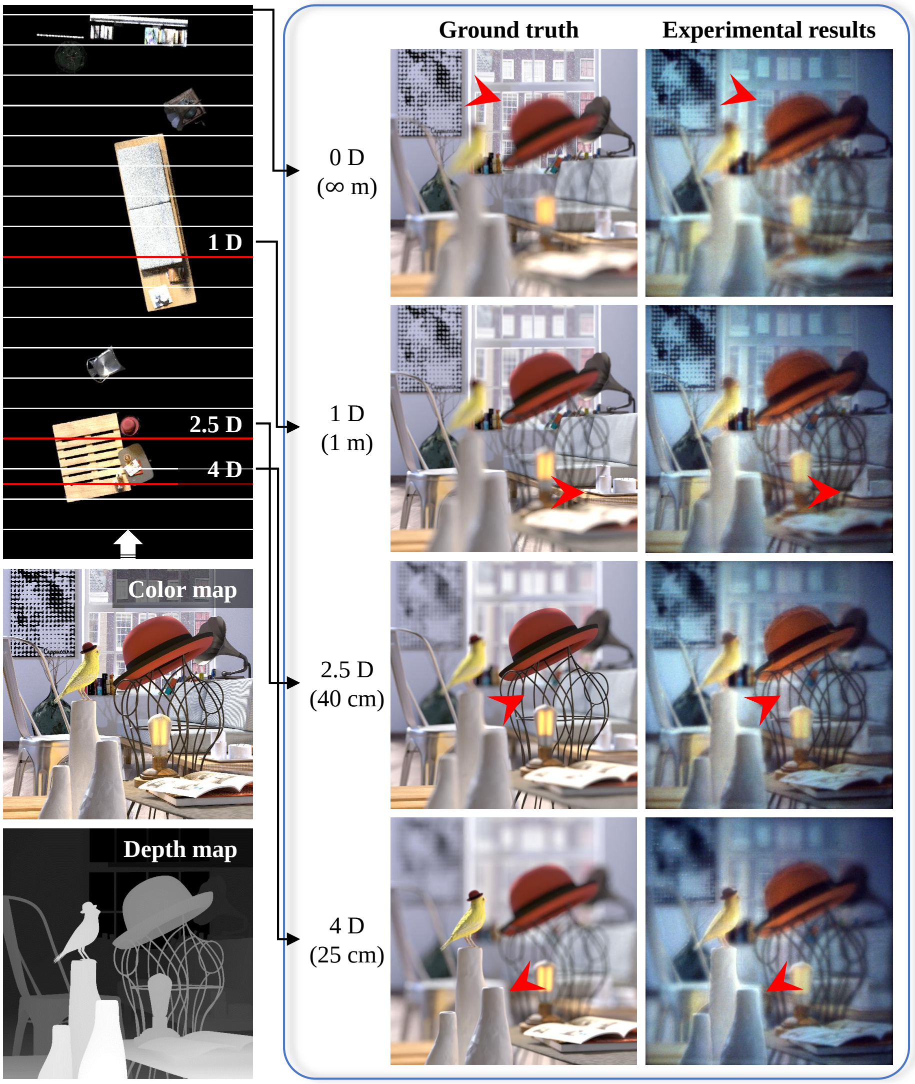

Figure 2.18 Experimental results to demonstrate the feasibility of tomographic projector. The target volumetric scenes are reconstructed successfully with quasi-continuous focus cues.

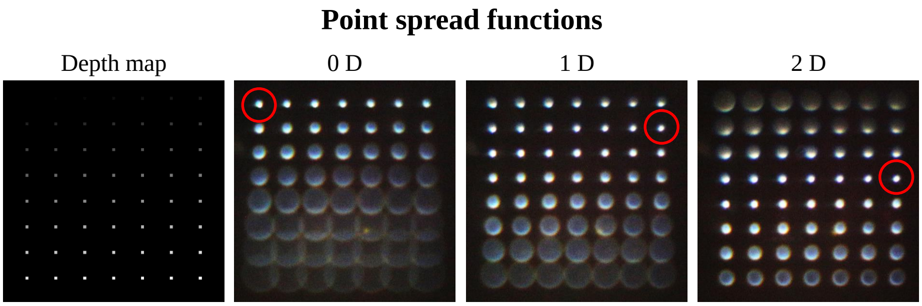

Figure 2.19 Experimental results of point spread functions of the prototype.

The results of the tomographic projector prototype are shown in Fig. 2.18. The perspective view image and its corresponding depth map are depicted with a top view of the target scene on the left of the figure. A volumetric scene is along the depth range between 0 D and 4 D. To demonstrate the ability of tomographic projectors, reconstructed images are captured by changing the CCD camera's focal depth. As shown in the results, the depth information of 3D contents is well conveyed while maintaining high resolution and contrast. In addition, as shown in Fig. 2.19, point spread functions are captured in different focal planes at the rate of 60 Hz. In this figure, the depth of the point gradually increases from the lower right (25 cm) to the upper left (infinity). The red arrows and circles indicate areas that are in focus. The changes in the point-spread function confirm that multifocal planes are well formed.

2.5.2 Parallax and appropriate focal blur

To provide accurate parallax and enlarge the eye-box size, it is necessary to reconstruct the four-dimensional light field introduced to the pupil plane. Accordingly, 7×7 perspective view images spaced 1 mm apart on the pupil plane are used as a target for the optimization [31]. The eye-box size of interest is 6 mm. The SLM and DMD images are updated in order to minimize the er-

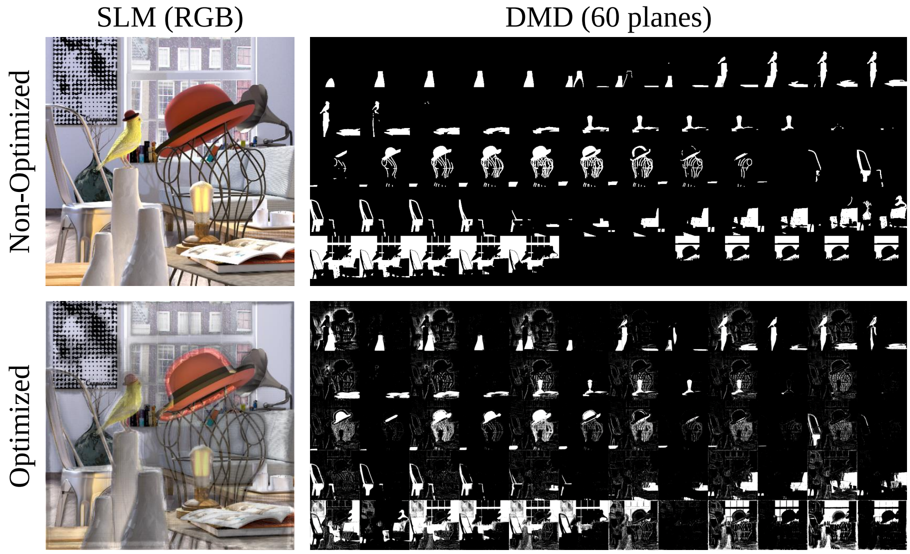

Figure 2.20 Illustration of SLM image and DMD image sequence with and without optimization.

rors between the ground truth and the reconstructed pupil view images. Figure 2.20 illustrates the optimized SLM image and DMD image sequences. When these images are applied for the tomographic projection, separation or overlap of focal plane images is alleviated, as shown in Fig. 2.21. The feasibility of the optimization is verified by changing the viewpoint and focal length of the CCD camera. As indicated by the red arrows, the tolerance for pupil movement and the eye-box size are increased. Additionally, the occlusion boundary artifact is also reduced because the pupil view images are reconstructed with enough accuracy. The results confirm that the presented optimization algorithm allows for improvements in the tomographic projection in terms of fidelity and tolerance.

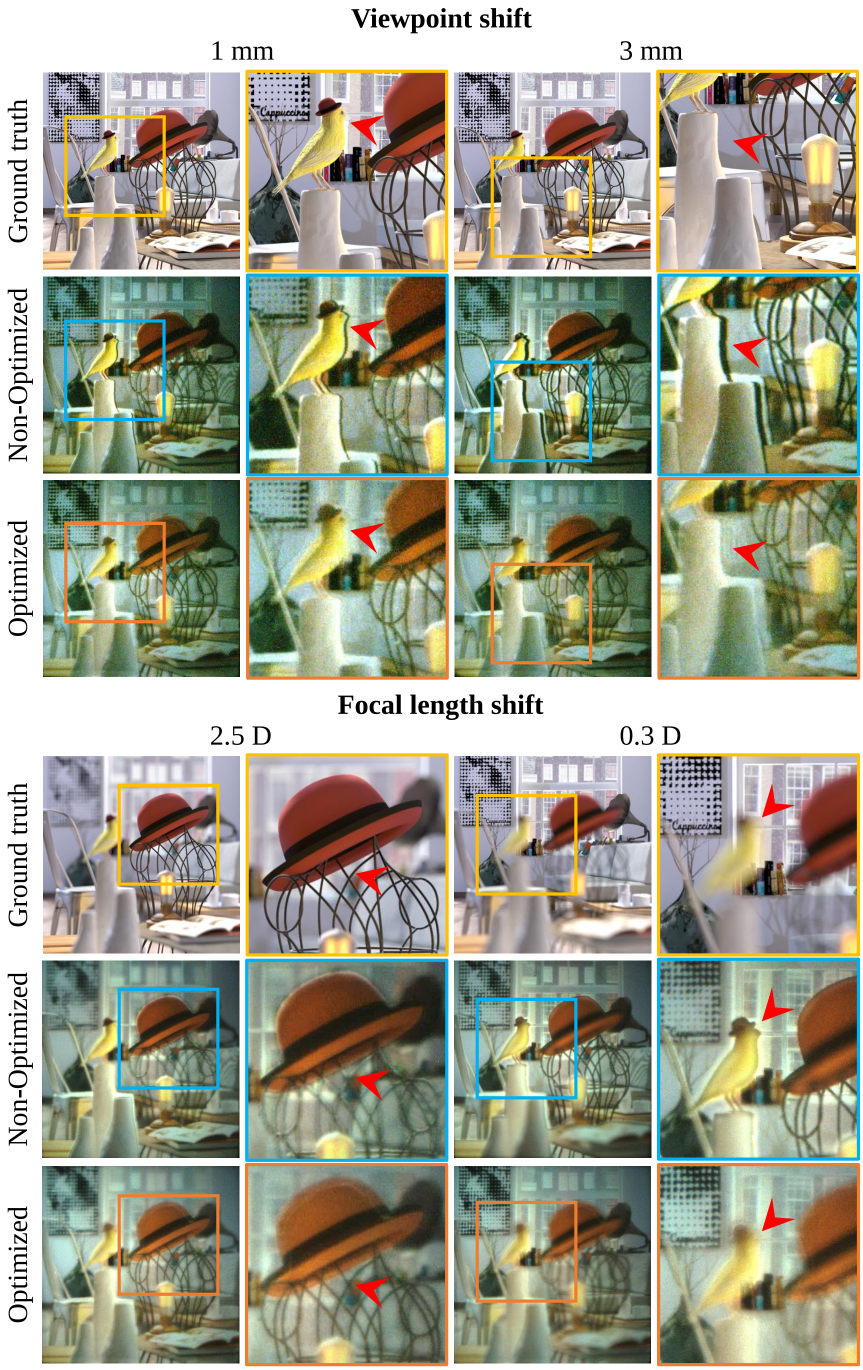

Figure 2.21 Experimental result to demonstrate the validity of the optimization.

2.6 Discussion

2.6.1 Compensation of convergence distortion

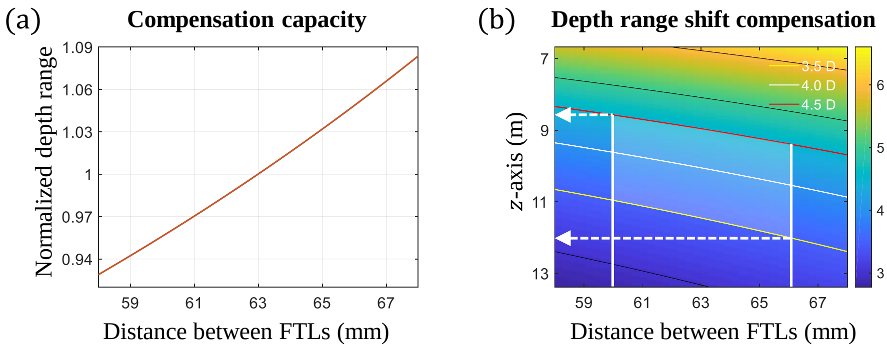

Figure 2.22 Simulation results of the compensation by changing the interspace between FTLs. Graph (a) shows the value of the normalized depth range that users perceive. The reference IPD is considered as 63 mm. In that case, the boundary compensating for convergence distortion is represented in graph (b). The scale bar indicates the depth range [1].

As described in Fig. 2.15, stereoscopic 3D theaters provide different convergence angles according to the viewing position. For a uniform 3D experience, it is necessary to minimize convergence distortion. Inspired by the fact that IPD influences the convergence distance, the author presents a method for compensating for the convergence distortion caused by viewing distance. If the interspace between FTLs deviates from the reference IPD, another convergence distortion occurs. These two types of distortion can be canceled in the appropriate conditions. For example, an audience who experiences more exaggerated 3D due to sitting close to the screen can wear FTLs with a smaller separation distance to compensate for it. Using the compensation method, the viewing zone can be extended where uniform convergence is provided, as shown in Fig. 2.22(b).

2.6.2 Advanced system

Degradation in resolution

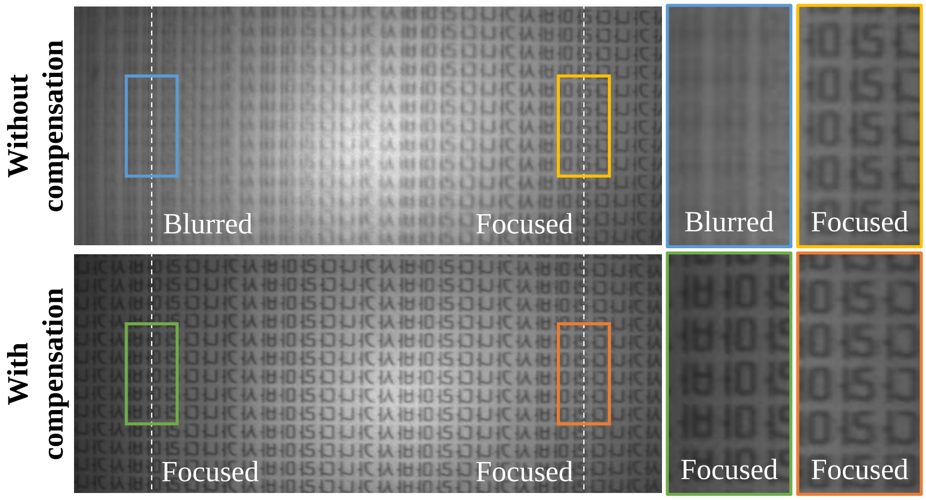

Figure 2.23 Problem of oblique incidence. According to the Scheimpflug principle, the tilt-shift lens is adopted to correct the oblique incidence.

The proposed projection system provides focus cues by combining three SLMs for full-color 8-bit depth images and DMD in charge of local filtering. The most significant advantage of this system is that it does not have critical trade-offs in providing focus cues to users at different positions. However, resolution degradation occurs due to the mismatch of image plane when relaying SLMs to the DMD. Since DMD consists of an array of tilted mirrors, there should be an oblique incidence using TIR. For the oblique incidence, the relay lens is tilted according to the Scheimpflug principle in order to make the focal plane identical to the DMD plane. To do this, a tilt-shift lens (Canon TS-E 80mm) is utilized. Figure 2.23 shows the effect of the compensation. Before tilting the projection lens, the right part of the image is focused while the left part is blurred. After applying the tilting, the image plane is well-aligned on the DMD plane so that the image is in-focus throughout all parts of the plane. However, tilting the relay lens could also result in off-axis aberration.

Advanced system design

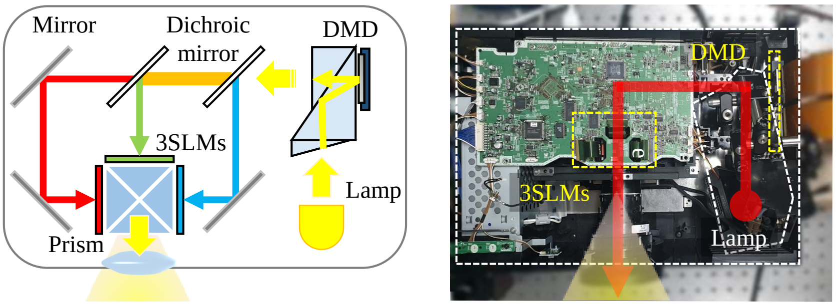

Figure 2.24 Schematic diagram and implemented a prototype of a tomographic projector that employs DMD as a backlight.

It can be improved by changing the systematic strategy. The proposed suggestion is to insert DMD inside the projector and spatially modulate the backlight, as shown in Fig. 2.24. Specifically, the light coming from a projector's lamp converges to DMD in even brightness through an integrator. This spatially modulated backlight is relayed to the three SLMs and passes through a projection lens. Implementing a prototype using this method confirms that the resolution of images is comparable to that of commercial projectors, as shown in Fig. 2.25(a). Note that the white image is displayed on the DMD with a fixed focal length only to check the spatial resolution of SLMs. In that case, the resolution of DMD's active area is 432×768 but can be improved by changing the optical system in the projector. However, there is an issue where the DMD image appears flipped in the red color due to the $4f$ lenses used to compensate for the

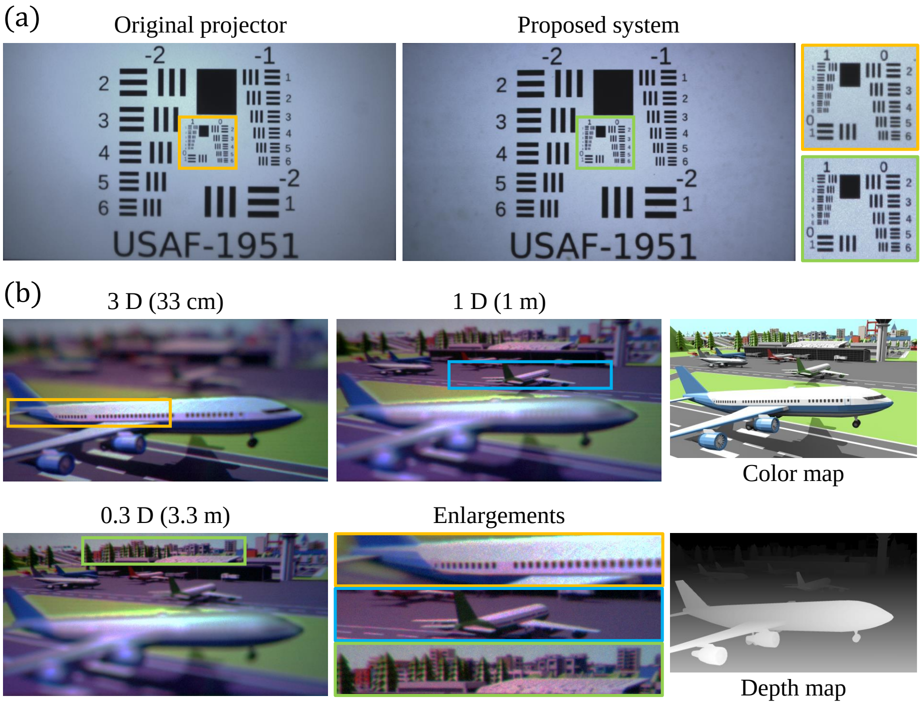

Figure 2.25 Experimental results of the (a) resolution target (USAF1951) for comparison with the original projector and (b) volumetric scene with its enlargements [1].

differences in the optical path among the colors. The author believes this problem would be easily solved by adjusting the optical path between each color. Figure 2.25(b) shows the results of combining RGB images taken separately. Appropriate focus cues are provided with the original resolution of the projector.

The author also believes that rearranging SLMs and DMD could better cope with brightness loss when enlarged to real environment size. The proposed system has brightness loss caused by the limited duty cycle of DMD projection. The duty cycle of the system is defined as the ratio of on-time to off-time for the half cycle. The optimal duty cycle, as determined through optimization methods, has a brightness of approximately 0.3 [63]. This means that the backlight unit should be three times brighter than a conventional one, potentially limiting the screen size. However, if the backlight first meets DMD, which is likely more thermally tolerant, and then passes SLMs, it would be feasible to increase backlight brightness enough.

2.6.3 Improvement in the focus-tunable lens

There are several methods to change the focal length [64]. The FTL used in the experiment adopts a shape-changing method based on a combination of optical fluids and a polymer membrane. It is a precision product with an operating speed of 60 Hz with a suitable form factor. However, the diagonal FOV is calculated as 39 degrees when the aperture is 10 mm and the eye relief is 14 mm. In this case, it can only cover the farthest range as specified by the SMPTE standard. Therefore, the aperture of FTL can be the bottleneck in the theater environment where a wide range of visibility is essential for an immersive experience. Among commercial products, there is FTL with an aperture of 16 mm from Optotune. It has a wide diagonal FOV of 59 degrees but can cause a flicker because its operating speed is 50 Hz. However, with technological advancements, the author believes that the FTL will soon reach a larger aperture that covers the entire FOV.

In addition, for practical use, the mechanical problems occurring during the operation of FTL need to be improved. For example, during the synchronization, FTL has two major errors caused by motor delay and arbitrary vibration. In the prototype, the motor delay is compensated in the calibration step by finding the appropriate phase delay of the FTL signal (-28.75 degrees $\approx$ 2.8 ms). On the other hand, arbitrary vibration could not be compensated as the proposed system does not have a real-time focal length tracker. While the tomographic projector demonstrated reliable focal plane reconstruction, artificial vibrations may lead to a dioptric error. It is believed that the error amount is within an acceptable range considering the VAC range, but it should be improved for a comfortable viewing experience in the commercialization step.

2.6.4 Magnification and field of view

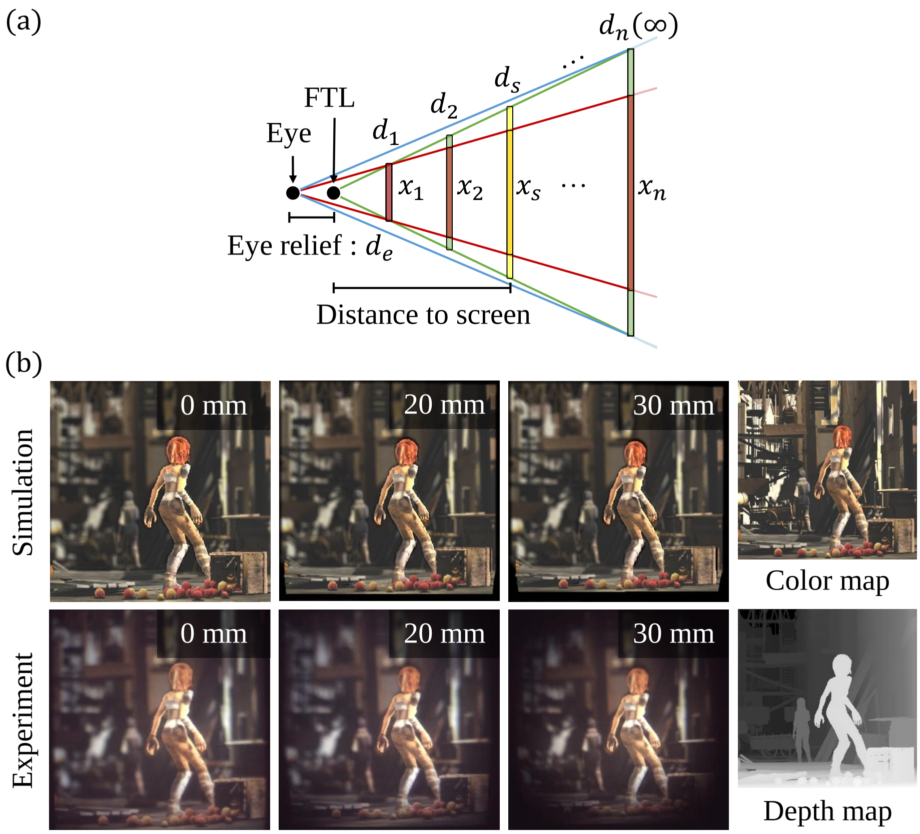

Figure 2.26 (a) Illustration of FOV variation according to the depth of focal plane and eye relief. (b) Corresponding simulation and experimental results are demonstrated when the eye relief changes from 0 to 30 mm [1].

In the tomographic projector, users wear the FTL in the form of eyewear.

Eyewear types require eye relief, the distance between the lens and eye more than 10 mm. Due to the eye relief, the magnification of each depth plane becomes different. As shown in Fig. 2.26(a), the FOV of each focal plane diminishes as the distance gets closer. It can be expressed as

$$FOV_n = 2\tan^{-1}\left(\frac{x_n}{2(d_e + d_n)}\right), \tag{2.8}$$

where $x_n = (d_n/d_s)x_s$. d_e is the distance of eye relief, d_n is the distance to n-th focal plane, d_s is the distance to the projection screen, x_s is the projection screen size, and x_n is the size of virtual screen image. The upper-bound of FOV is given by

$$FOV_{max} \cong 2\tan^{-1}\frac{x_s}{2d_s}, \tag{2.9}$$

where the distance from the screen is infinite or the eye relief is zero. The maximum difference in screen size is calculated below according to Eq. 2.8.

$$\Delta Size_{max} = \left(\frac{\%}{100}\right) = \frac{\tan FOV_{max}}{\tan FOV_1} - 1 = \frac{d_e}{d_1} \tag{2.10}$$

According to the above equation, the defect gets notable when the difference in depths gets wider. Additionally, the defect gets more noticeable as the distance from the center gets larger. In the proposed system, the combination of DMD and SLMs increases the focal plane number without scarifying bit depth or frame rate. However, this configuration has a drawback: the correlation of each focal plane image. Therefore, it is hard to digitally compensate for the magnification variations across each depth plane. However, magnification is not a significant problem in a typical viewing environment. Since the optical axis and gaze direction of a user are identical, less error is observed in the fovea. Recent research demonstrates that the artifacts caused by magnification variation can also be reduced through the use of a modified optimization algorithm [65]. Moreover, taking retinal blur into consideration, the focal blur effect could be strong enough to make the error hardly noticeable where the depth is discontinuous, as shown in Fig. 2.26(b). The 3D content is sourced from the work of [66].

2.6.5 Visual perception

As briefly mentioned in Chapter 1, humans interpret the surrounding environment through the human visual system. This ability is called visual perception, and it does not only include seeing an object or content but also understanding all 3D space. Eye and head motion, for example, allow us to take in much more information and provide a continuous stream of navigating through our surroundings. Conventional displays (e.g., a television and monitor) have simply provided information for some purpose. In contrast, to turn seeing into an immersive experience (e.g., a 3D theater and VR), real-world-like senses should be fulfilled. As part of this purpose, the author has presented the system reproducing the 3D scene with appropriate focus cues. However, it is also important to consider visual perception factors in addition to hardware performance since the viewer performs the space awareness of the 3D environment.

There are multiple visual parameters that affect the visual perception of 3D space. The first one is the visual field, which is similar to FOV, but it represents the domain on the retina without moving the gaze. Generally, each human eye has a visual field of about 135 degrees in height and 160 degrees in width [67]. Therefore, a display system that can fill the user's visual field as wide as possible can improve the viewing experience, such as immersion [68]. As a way to cover the entire visual field in a projection-based system, a domed environment has been proposed [69]. This approach utilizes a single hemispheric screen onto which images are projected to bring visual immersion and can be seamlessly applied to the projection-based proposed system. As mentioned in the previous section, the system's FOV is limited by the small aperture of the currently available FTL, but the author believes the proposed system can cover a wider visual field using advanced FTL in the future.

The second is visual acuity. The clarity of vision decreases gradually from the center of the visual field (fovea) to the periphery [70]. Therefore, the viewer mainly obtains visual information from the foveal region, which is less than two degrees of the visual field. Naturally, providing a high-resolution image to the fovea while stimulating all visual fields for an immersive experience requires an increase of effective pixels. Even though the resolution for the parafoveal and peripheral regions is significantly lower, a foveated display scheme using head tracking cannot be applied because the viewers should be able to recognize the same image [71]. With the current technology, a multi-projection system using multiple projectors could be effectively utilized.

The third is the gaze movement required to dynamically update the scene in the visual field since the vision is only sharp in the central area. Generally, a stereoscopic display using a screen cannot provide an appropriate motion parallax from head displacement, although the motion parallax is essential to enhance viewing experiences such as visual comfort and a sense of presence [72]. However, this becomes less important in a movie theater environment of the proposed method where the viewing position is fixed. Instead, correct images should be ensured within the eye-box that is wide enough to allow freedom of pupil movement. As the FTL is added to 3D glasses, the drawback is that off-axis aberrations could arise when the gaze moves toward the edge of the FTL, which people may find uncomfortable. So, an optimization methodology considering gaze movement should be developed for future work to avoid user discomfort.

The last is sensory cues, such as auditory, tactile, and vestibular, which contribute to visual perception in some situations. Especially the vestibular, which is a sensory system for the purpose of normal movement and equilibrium, is considered important in stereoscopic displays because it can potentially induce motion sickness [73]. This is because the signals from the eyes and body do not match in the virtual 3D environment, where viewers are expected to stare at the display without movement. This sensory mismatch causes visual-vestibular conflict. Furthermore, it is known that the higher the motion velocity in the scene, the stronger the cybersickness [74]. Therefore, methods that provide realistic motion and multi-sensory cues through physical devices have been developed and are successfully commercialized [75]. The proposed system can also apply the same methodology for an immersive experience.

The bottom line is that various visual perception factors should also be considered along with the system performance to evaluate the quality of viewing experience. In this dissertation, however, a more detailed evaluation is not pursued, as it will require a more extensive perceptual study and is beyond the scope of this dissertation.

2.7 Conclusion

In this chapter, the author presents a stereoscopic system to give more immersive and comfortable experiences on a large screen. The main contributions of this work are summarized in Fig. 2.27. The tomographic projector combines SLMs for color and contrast and DMD for fast spatial filtering, which are synchronized with the FTL. A theater using a tomographic projector enables audiences to perceive multiple focal planes that were previously unavailable. The author implements a miniaturized theater environment of theater and confirms that quasi-continuous focus cues are reconstructed. The author also thoroughly analyzes the viewing zone where uniform 3D experience is delivered regardless of seat position. According to the analysis, the tomographic projector does not significantly distort focus cues in practical environments where the screen is several meters away from viewers. Additionally, it is verified that the proposed system could address some of the limitations of conventional stereoscopic 3D, including vergence-accommodation conflict and various IPD issues. To enhance the viewing experience, the author proposes an optimization algorithm to reduce occlusion boundary artifacts. The artifact mitigation is demonstrated through simulations and experiments. Finally, the author conducts an in-depth discussion about challenges that should be considered for practical use.

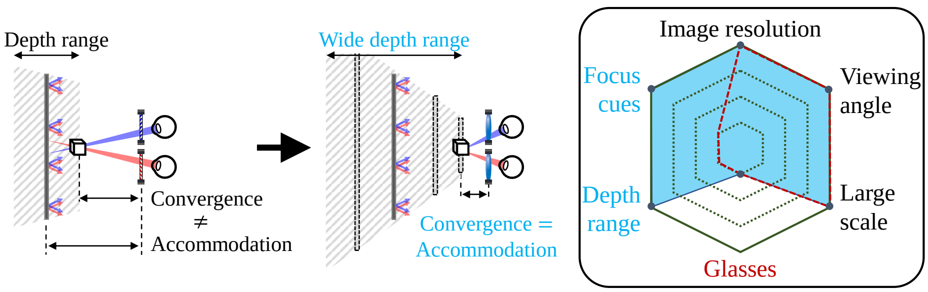

Figure 2.27 Summary of Chapter 2.

Chapter 3. Effective resolution enhancement of light field display via optical transmission

3.1 Introduction

The previous chapter introduced the effective method for realizing a volumetric scene in stereoscopic displays through the use of appropriate focus cues. However, the stereoscopic display has an inherent limitation of the necessity of wearing glasses, which will cause another inconvenience. Therefore, making viewing conditions convenient without the need for special glasses has been an important research topic in 3D displays.

3.1.1 Trade-off in light field displays

The interest in glasses-free 3D displays over the past decades has led to the development of light field displays [16, 24, 25, 31, 76–78]. Light field displays have the ability to modulate both the direction and intensity of light, enabling the reconstruction of 3D objects in free space without glasses. From the integral

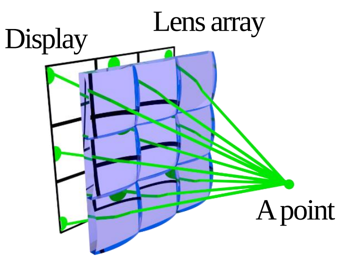

Figure 3.1 Light field display requires very dense ray bundles even to represent a single point.

imaging (InIm) [79] to the compressive light field [80], various light field methods have been proposed to achieve angular resolution through the use of barrier, lenses, and stacking of panels. However, all of these approaches are inherently limited by the information capacity of a flat panel display, which is defined as the number of pixels. In the light field display, a bundle of converging-diverging rays should be used to represent a single point, as shown in Fig. 3.1. Since the ray and the pixel are generally one-to-one matched, it reduces available information by its number. Therefore, a critical issue of the light field display is the need for enormous amounts of information to reach enough visual quality, full parallax, and focus cues. This may be the prime reason why light field displays have been difficult to put into practical use.

3.1.2 Comparison of light field displays

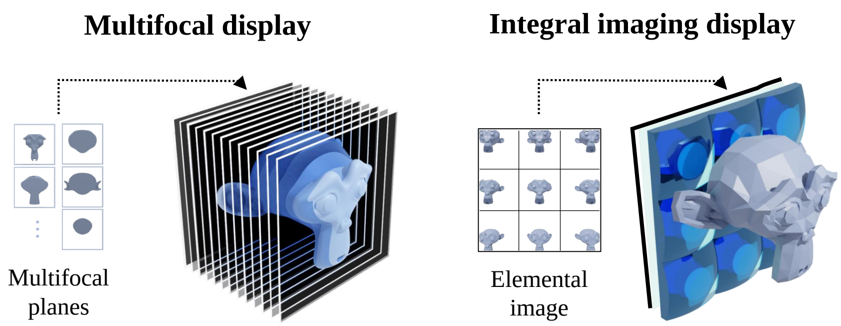

Figure 3.2 Illustration of two types of light field displays for comparison.

Efforts have been made to relieve the trade-off by using multiple displays or applying temporal multiplexing techniques. In this context, multifocal displays were proposed as a way to increase the information capacity and extend 3D volume [76, 80]. A multifocal display is a system that uses spatial or temporal multiplexing to stack multiple 2D planes in the depth direction. It is also known

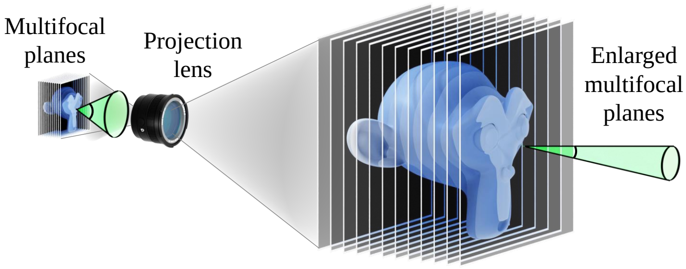

Figure 3.3 Illustration of the reduction in viewing angle when projecting multifocal planes on a large scale.

as a multi-layer display, depth-fused display, or tomographic display. This approach has the potential to reduce the cost of spatial or angular resolutions by computationally optimizing the light field [31, 77]. This optimization technique effectively compresses the spatial-angular information by considering the correlation among the light field rays. Benefiting from the recent improvements in the multifocal displays [1, 34, 35, 81], a large number of focal planes can be produced using a time-multiplexing technique. Furthermore, as described in Chapter 2, it is possible to generate continuous and natural parallax by adopting the computational optimization method [63, 82, 83]. However, there is a problem that such a time-multiplexed system cannot be directly implemented on a large scale, as shown in Fig. 3.3. This is because the viewing angle of the multifocal planes decreases as they are expanded over a larger area. Even applying a scattering screen would result in the loss of all angular information. A space-multiplexed system stacking multiple physical panels also has difficulties that require a bulky space and a large number of display panels for large scale.

In contrast, the projection-based light field system has the advantage of being scalable at will [16, 78]. Additionally, using multiple projectors can easily increase the information capacity [24, 25]. Integral imaging is the most straight-

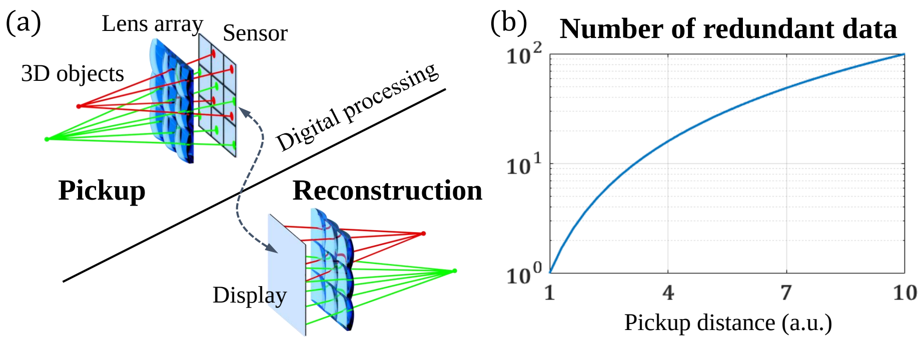

Figure 3.4 (a) Basic two processes of the InIm system: pickup and reconstruction. (b) The number of picked up points increases according to distance [2].

forward and suitable method for the projection-based light field. In the InIm, angular information is recorded and reproduced in the form of spatial information called an elemental image (EI), as shown in Fig. 3.4(a). However, the amount of information is limited by the spatial resolution of the pickup sensor or display device. Moreover, since ray bundles of the same intensity should be displayed as an element image to express angular information of a 3D scene, redundant information occurs, as shown in Fig. 3.4(b). Thereby, the InIm shows significantly lower image quality than the original resolution of the display panel. The author thinks that there exists a capability to improve the performance of InIm projection by removing repeated and inefficiently used information.

In summary, integral imaging can be implemented on a large-scale system, but it technically limits the resolution of 3D images. On the other hand, multifocal displays are better at image quality but limited in enlarging images. As a way to overcome the limitations of each system, the author proposes a new optical configuration that combines the advantages of multifocal displays and integral imaging.

3.1.3 Combination of light field techniques

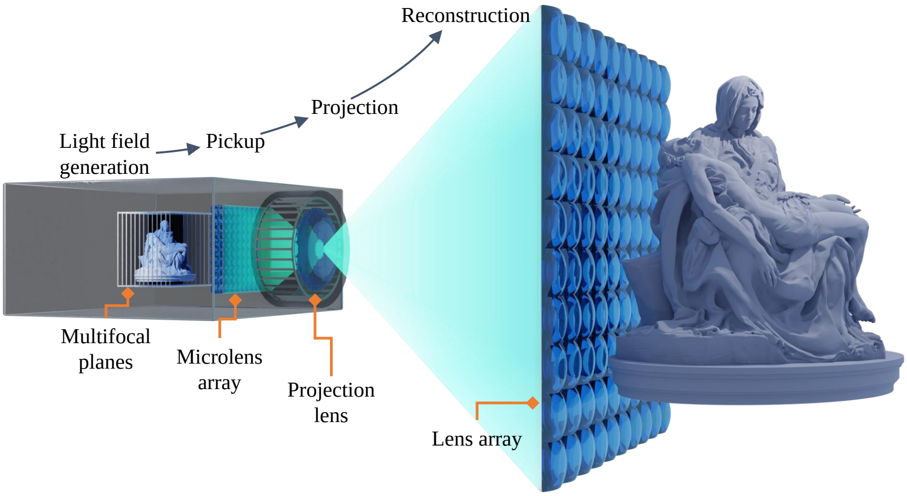

Figure 3.5 Illustration of the proposed system. It consists of 4 steps, optically combining multifocal display with integral imaging.

In Chapter 3, the author introduces a projection-type light field display system. A novel optical design that combines multifocal display with InIm technology enables an autostereoscopic 3D display with effective light modulation. Here, the tomographic approach generates a high-resolution volumetric scene, and InIm allows for the reconstruction of the volumetric scene on a large screen through projection. As shown in Fig. 3.5, since all the processes are realized optically without digital processing, the proposed system can overcome the performance limitations related to the number of pixels in conventional InIm displays. The author builds a prototype display and demonstrates that the proposed optical design has the potential to achieve massive resolution with full parallax in a single device.

3.2 Principle

3.2.1 Optical transform of information

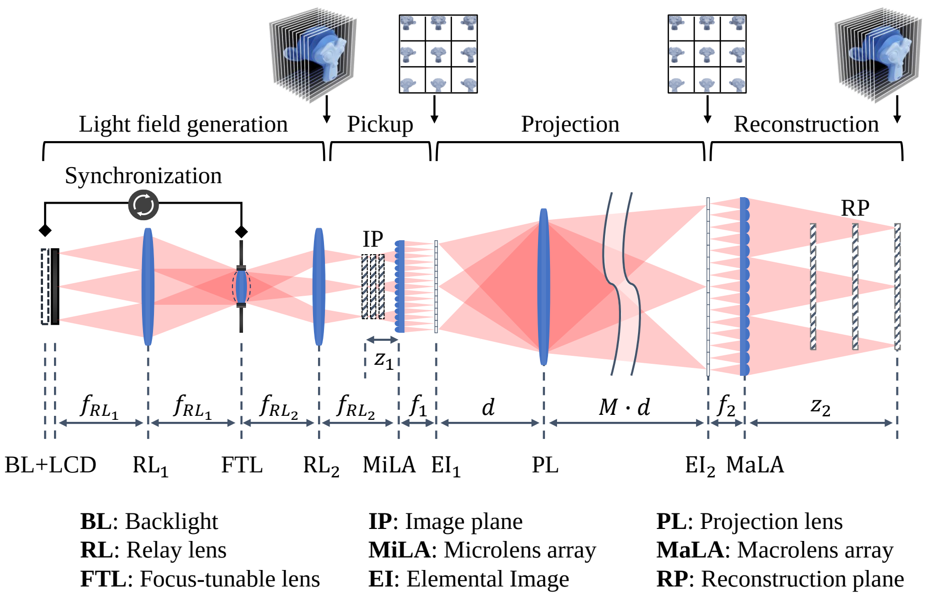

Figure 3.6 Simplified concept diagram of the optical structure. Note that the z_1 can be a negative value, which means pickup as a virtual image [2].

Figure 3.6 illustrates the principle of light field optical transmission, which involves four steps: light field generation, pickup, projection, and reconstruction. The main idea is to connect all the processes optically using a projection system. With this configuration, the automatically mapped EI plays a key role in avoiding the inefficient use of information, as represented by the trade-off relationship in the InIm. Furthermore, unlike the previous InIm systems, the proposed design prevents hardware-related information reduction at the capture and display stage and allows the EI to be super-sampled without discrete division.

For the first step, the tomographic display is adopted to generate a light field. As used in Chapter 2, this method can produce a volumetric scene over a wide depth range by creating dozens of planes placed at different depths [1, 63, 82, 83]. The multifocal planes (MFPs) are generated from an RGB-D image by synchronizing a binary backlight with a FTL. Here, the FTL is utilized as an aperture stop for the MFPs. While the FTL controls the floating position z_1, each focal plane has the same divergence angle and size with the telecentric relay [21]. Then, in the next step, the synthesized light field from the MFPs is picked up by a microlens array (MiLA). In the EI_1, the overlap between each lenslet can be alleviated by matching the F-number ($F\#$) of MFPs to that of MiLA. Thirdly, through the projection lens, the EI_1 is enlarged by the magnification factor M on the EI_2 plane. The author places the screen here. However, similar to previous projection InIm techniques, it is possible to use either a screen or direct projection method here. The differences between the two methods have been thoroughly described using parameters such as fill factor and depth range [84, 85]. As the final step, the light field after the screen is reconstructed as it passes through the macrolens array (MaLA) in the reverse process of the pickup. The proposed workflow is depicted in Fig. 3.7.

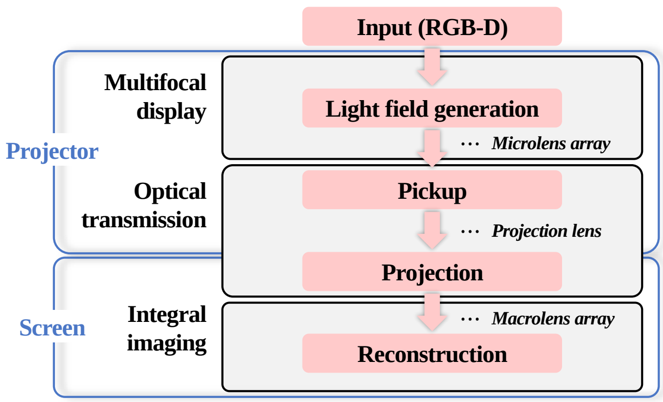

Figure 3.7 Schematic workflow diagram.

3.2.2 Large-scale projection

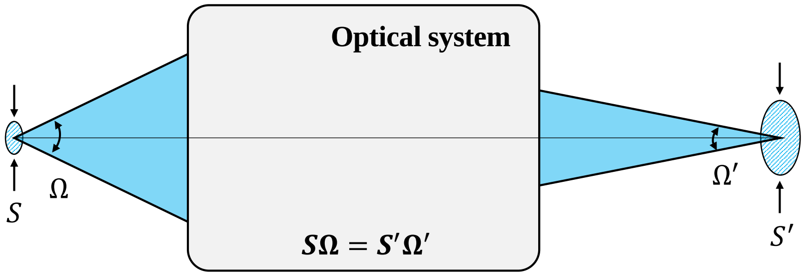

Figure 3.8 Conservation of étendue

As described in Fig. 3.8, the throughput (also called étendue), which is defined as the product of the area S and solid angle Ω, is conserved in an optical system. Therefore, as depicted in Fig. 3.3, bringing the multifocal planes to the big screen is difficult because the wider the area during magnification, the narrower the angle of each display pixel [86]. However, the author uses InIm techniques to convey angular information while projecting the volumetric scene. These techniques allow angular information of the volumetric scene to be converted into spatial information during the pickup process. Accordingly, even if the divergence angle for each pixel is reduced when projected, it is restored while the EI is returned back to the angular information in the reconstruction process. In other words, the proposed approach not only avoids the information loss in the InIm but also effectively solves the enlargement problem of the multifocal displays.

3.3 Analysis

3.3.1 Light field analysis

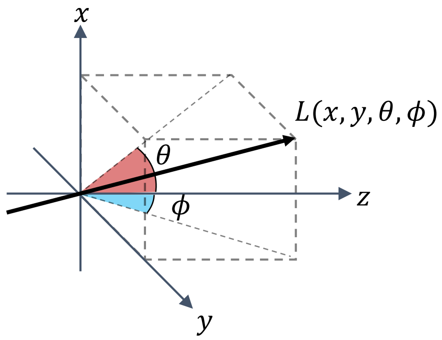

Figure 3.9 Parameterization of a ray by position (x, y) and direction (θ, ϕ)

Light field displays are based on ray optics to represent 3D objects. Therefore, ray-tracing analysis is the most straightforward way to explain the proposed method. Unlike a flat panel display, the light field has directional information to describe a vector of light in space. As depicted in Fig. 3.9, each ray located at a depth z has a two-dimensional direction in addition to a two-dimensional position. By calculating all light fields generated by the system, the analysis of the optical system can be performed.

As an example, Fig. 3.10 shows a 2D light field simulation for a single plane. In this analysis, the author only counts a single dimension x and analyzes the light field as an ordered pair of position and angle. This is because there is no difference in the axial direction due to the design's symmetry of the lens. The horizontal axis represents the x-direction position of the light field, and the vertical axis represents the tangent value of its angle. Note that $(x, \tan\theta)$ is used instead of (x, θ) in the non-paraxial model. The light field emanating from a pixel has the same color, and $\tan\theta$ is limited to arbitrary values $(-1, 1)$. Here,

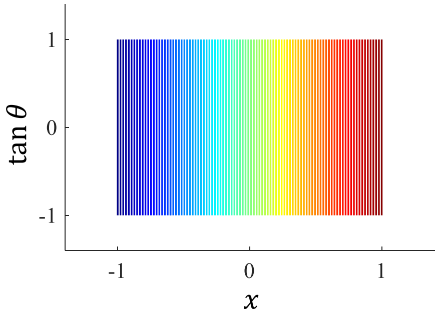

Figure 3.10 Two-dimensional light field simulation.

the change in the light field can be explained in two cases: when it propagates in free space and when it passes through a lens.

When a ray propagates by a distance of z, the light fields on the $(x, \tan\theta)$ plane are horizontally twisted and transformed into a parallelogram as shown in Fig. 3.11(a). For the simulation, the position x_2 and direction θ_2 after the free-space propagation is calculated as follows.

$$x_2 = x_1 + z\tan\theta_1, \tag{3.1}$$

$$\tan\theta_2 = \tan\theta_1. \tag{3.2}$$

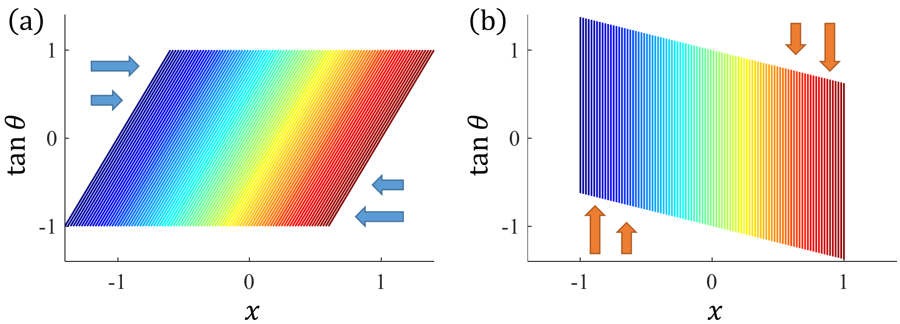

Figure 3.11 Light field simulation after (a) free-space propagation and (b) passing through a convex lens.

Similarly, when a ray passes through an ideal lens, the transformed light field $(x_2, \tan\theta_2)$ is calculated as follows.

$$x_2 = x_1, \tag{3.3}$$

$$\tan\theta_2 = -x_1/f + \tan\theta_1, \tag{3.4}$$

where the f is the focal length of the lens. As opposed to free-space propagation, there is a twist in the vertical direction as $\tan\theta_2$ decreases proportionally to x. By using only the two equations described above, all the light fields of the proposed system can be analyzed.

Figures 3.12(a)-3.12(f) illustrate the light field analysis of the proposed system for a focal plane. First, the light field for a single plane has a rectangular shape with a spatial length of L_x and a tangent value of $1/F\#_1$, where $F\#_1$ is the $F\#$ of MiLA. Note that the limited divergence angle of the image plane prevents information overlap between the lenslet. After it propagates to the MiLA, the light field is transformed into a parallelogram, as shown in Fig. 3.12(b). Since it propagates a distance of z_1, the width widens to $L_x + z_1/F\#_1$ according to Eq. 3.1. Then, the light field is divided spatially by the interval D which is the lenslet pitch of MiLA. As it passes through the MiLA, the maximum tangent value doubles according to Eq. 3.4. Here, the MiLA optically arranges the pixels in the EI_1 plane according to the spatial position of each lenslet. Through this pickup process, angular information of the light field is converted into spatial information as the parallax between the lenslets.

As shown in Fig. 3.12(d), the width of the light field increases by a factor of M after projection and magnification. The divergence angle, which decreases during the magnification, is expanded again by diffusing at the screen plane. Since the reconstruction is realized in the focus mode [22], the screen is placed at the focal length of the MaLA. Therefore, the light field after the MaLA has a

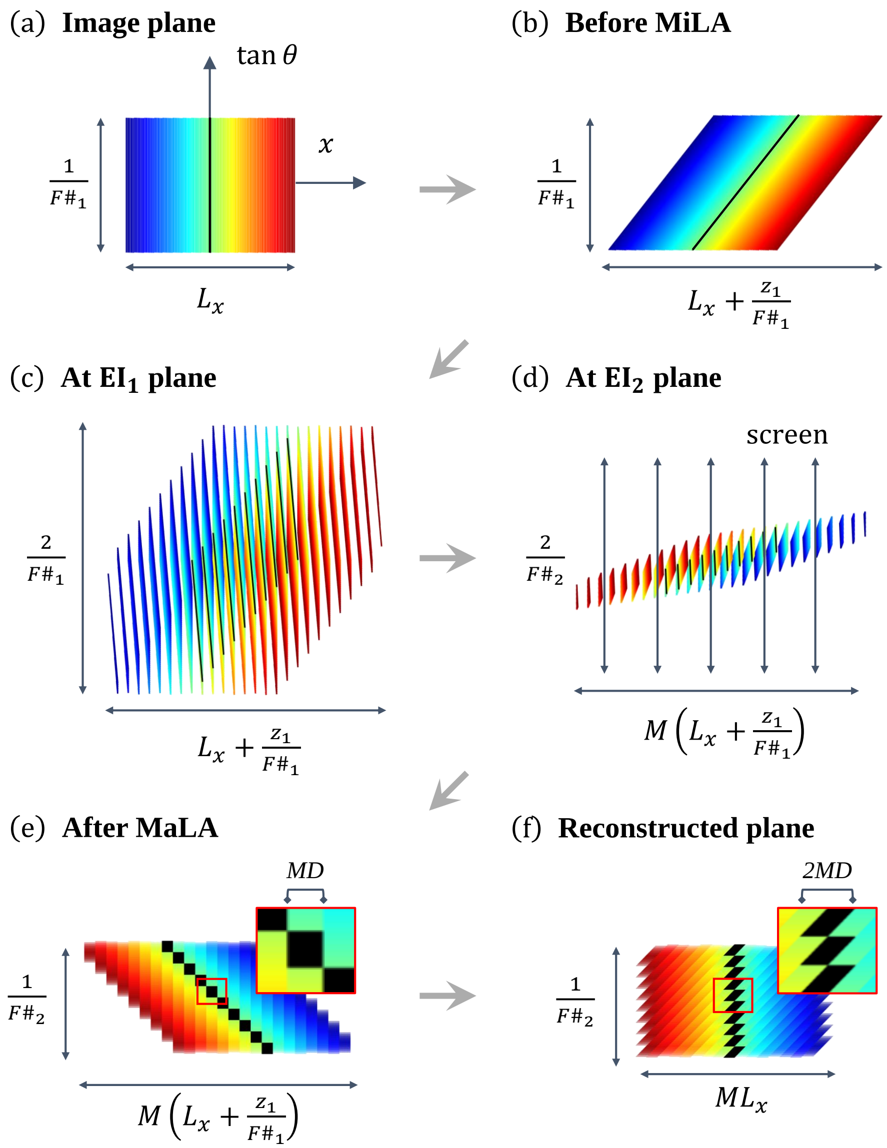

Figure 3.12 (a)-(f) Ray-tracing results of the proposed method. The center pixel is marked in black to trace the shape of the light field.

rectangular shape, as shown in the inset of Fig. 3.12(e). Furthermore, the tangent value of the light field reproduced by MaLA is $1/F\#_2$, defining the viewing angle [22]. Through this reconstruction process, the EI_2 is returned back to the angular information. Finally, as shown in Fig. 3.12(f), the MaLA reproduces the focal plane where each display pixel appears as a sampled form after propagating a distance of z_2. Here, the distance z_2 of the reconstructed image plane is derived as $z_1 MF\#_2/F\#_1$ using Eq. 3.1. Because the image depth is proportional to $F\#_2/F\#_1$ and the viewing angle is proportional to $1/F\#_2$, these two parameters can be customized by adjusting the $F\#$ between the pickup and the reconstruction.

3.3.2 Resolution of the optical system

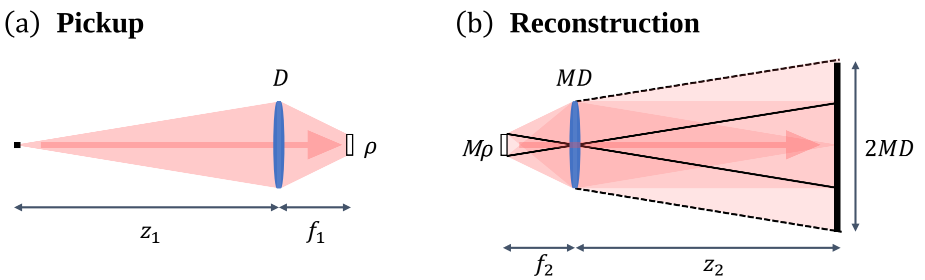

Figure 3.13 Illustrations of the changes in spot size during (g) pickup and (h) reconstruction process [2].

To evaluate the system's performance, the size of a point at the reconstruction plane is calculated based on ray optics. As shown in Fig. 3.13(a), the blur spot size ρ at the EI_1 plane is calculated as Df_1/z_1 while picking up a point at z_1 distance. After projecting and scattering, the blur spot reconstructs the point at z_2 distance, as shown in Fig. 3.13(b). The lateral size of the reconstructed point is $MD + M\rho z_2/f_2$. By substituting ρ and z_2, the reconstructed point size is

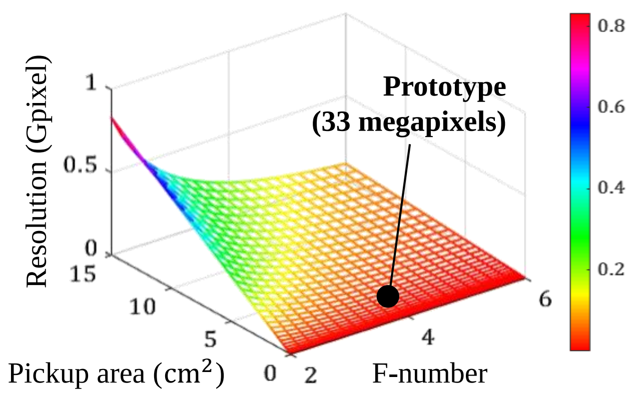

Figure 3.14 Maximum resolution of the system corresponding to the Rayleigh criterion when a wavelength is 550 nm. Here, the prototype has a pickup area of 2.4 cm^2 with $F/4$ [2].

calculated as a constant value of $2MD$, regardless of the pickup distance z_1.

However, the blur spot size cannot be smaller than the diffraction limit for the MiLA during the pickup process. Considering the diffraction effect, the minimum spot ρ_m formed by the plane wave is $2.44\lambda F\#_1$. In accordance with the Rayleigh criterion, the maximum spatial resolution of the system can be considered as $4A/{\rho_m}^2$, where A is the pickup area. By increasing the numerical aperture of the MiLA and the pickup area, the proposed optical design enables large-scale volumetric display equivalent to the InIm of megapixels or higher, as shown in Fig. 3.14. For example, using a typical projection lens with a pickup area of 24 mm$\times$36 mm and $F/1.2$, it is possible to achieve a volumetric display with a resolution of up to 1.3 gigapixels. For the experiment, due to a lack of suitable off-the-shelf MaLA, the pickup area of the prototype is set to 15.4 mm$\times$15.4 mm.

3.4 Implementation

3.4.1 Light field generation

The light field generation system is implemented using a DMD (DLP9500) from Texas Instruments as the binary backlight. The DMD has a full HD resolution and 16 kHz operation speed with 20.7 mm$\times$11.7 mm size. Due to the DMD mirror characteristic, which rotates at an angle of 45 degrees, a resolution of 763$\times$763 is used as the modulation area. The backlight image projected from the DMD is relayed to the transparent LCD with a magnification of two times. The LCD model used is Sharp LS029B3SX, and the effective resolution is 465$\times$465 (0.22 megapixels).

For the FTL, the Optotune EL10-30-TC with a 10 mm aperture is selected. As the MiLA of RPC photonics is used at a F-number of 4 and a pitch of 100 µm, the focal length of RL_2 is set to 40 mm to match the $F\#$ between the MFPs and MiLA. Here, the focal length of the FTL, f_{FTL}, can sweep between -5 D and 6.7 D with a negative offset lens of -75 mm. A floating distance z_1 can

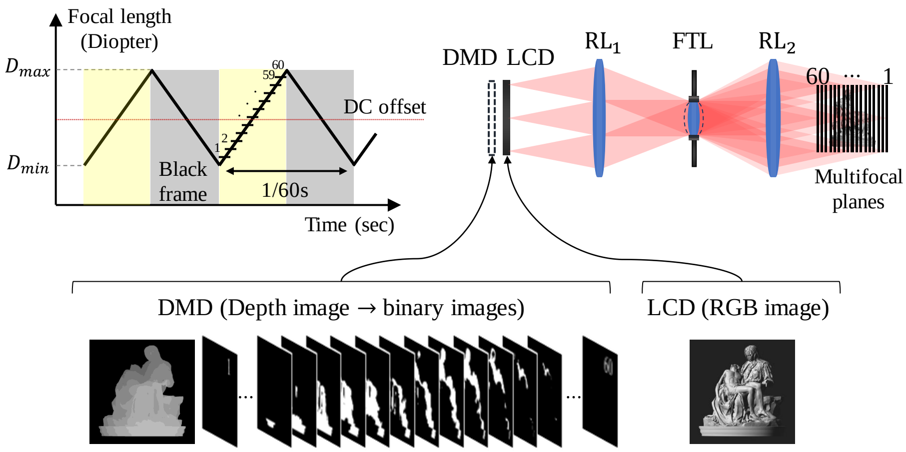

Figure 3.15 Illustration of the time-sequential process.

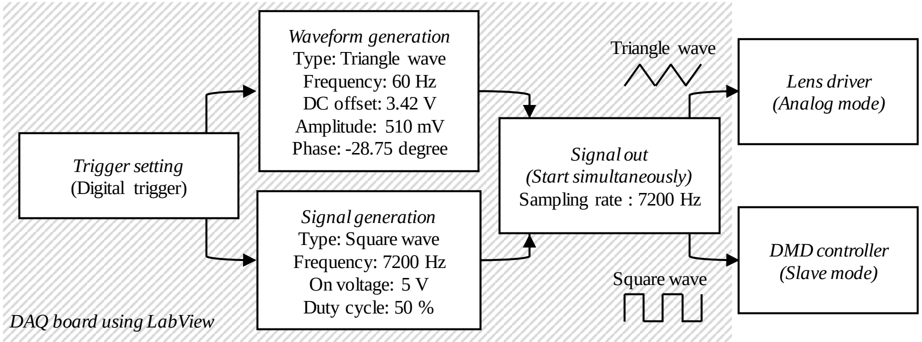

Figure 3.16 Block diagram of the synchronization using DAQ board.

then be derived as ${f_{RL_2}}^2/f_{FTL}$ as the FTL is placed at the Fourier plane [87]. Therefore, the maximum depth range of the volumetric scene is 18.7 mm, of which the author uses 16 mm in the experiment.

In the light field generation step, the DMD and FTL are synchronized through the DAQ from National Instrument, as shown in Fig. 3.16. The signals that control the two devices are generated together with the same internal clock speed. The sampling rate is set to 7200 Hz and one period of FTL is 1/60 s considering the flicker. Here, 120 binary images can be allocated. However, as illustrated in Fig. 3.15, to eliminate the crosstalk between adjacent frames, a black frame is inserted for half cycle [82]. In the experiments, 60 planes are generated for the volume of 13.4 mm$\times$13.4 mm$\times$16 mm with the real-time operation. In the experiment, each depth appears six times for one period with a duty cycle of 0.1.

3.4.2 System design

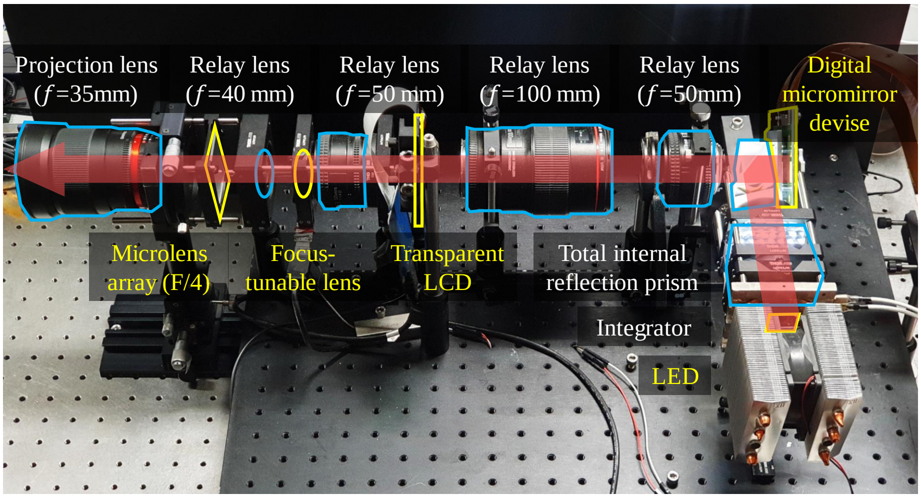

Figure 3.17 Photograph of the prototype system [2].

Figure 3.17 shows the entire experimental setup. In detail, the light emitted from LED is relayed to the DMD through the TIR prism. Here, the author utilizes an integrator that is comprised of two lens arrays with a collimation lens. It makes uniform illumination and easily modifies the divergence angle on the DMD. First, the binary backlight of the DMD is enlarged by $\times 2$ relay optics composed of two camera lenses (Nikkor AF 50 mm $F/1.4$ and Canon EF 100 mm $F/2.8$). Then, the 24-bit image of the LCD is picked up by the MiLA after passing through relay lenses (Nikkor AF 50 mm $F/1.4$ and achromatic lens of 40 mm focal length). The number of MiLA is 154×154. Because of the experimental space limitation, the magnification of M is set to 16, and the projection length is 60 cm. However, the system scale can be expanded without restrictions. A holographic diffuser with expanding angle of 30 degrees is utilized for the screen. The MaLA has a pitch of 1.6 mm with $F/5$. With this configuration, the system reconstructs the volumetric scene of 21.4 cm$\times$21.4 cm$\times$32 cm.

3.5 Results

3.5.1 Imaging result

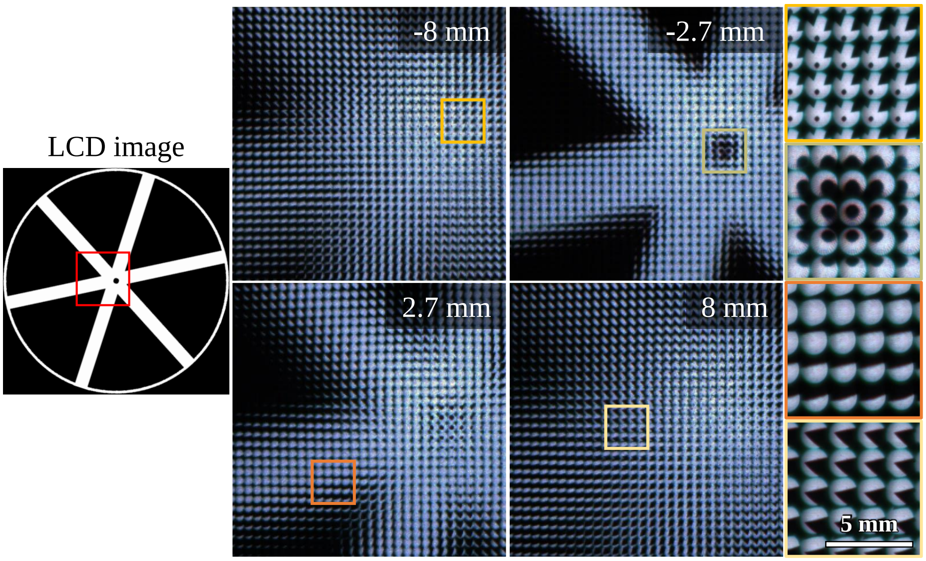

Figure 3.18 Captured elemental images at the screen plane. White wheel image is sampled differently according to the pickup distance z_1 [2].

Figure 3.18 shows the cropped EI_2 for varying pickup distance z_1 from -8 mm to 8 mm. As the $F\#$ is matched, the EI is confined within a boundary of each lenslet. The number N of lenslets required to represent a pixel is calculated as $|z_1/f_1|$. Even though the scene created by the tomographic display contains the information matched with the RGB-D image in a one-to-one ratio, the optical pickup process allows the generation of light fields mapped in a one-to-N ratio like conventional InIm. As such, the proposed design can effectively perform the InIm projection with N times higher resolution.

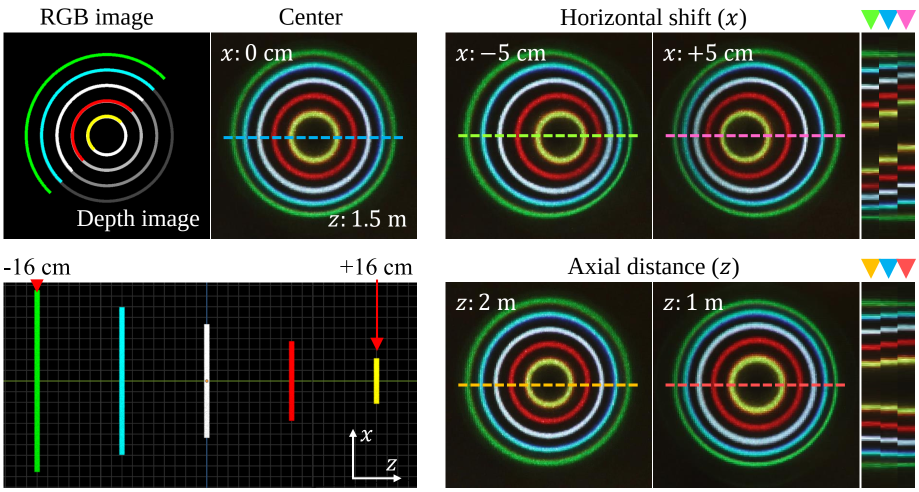

Figure 3.19 Parallax according to the horizontal and axial distance. Five circles are constructed with a given RGB-D image. To emphasize the parallax, the images sliced horizontally are shown on the right [2].

The author constructs the multifocal planes for verifying the support of full-parallax and appropriate focus cues. The results are captured by a CMOS camera (Canon EOS 5D Mark III) using a 50 mm focal length camera lens with $F/1.4$ at a distance of 1.5 m. As shown in Fig. 3.19, the proposed method can support continuous parallax not only in the horizontal direction but also in the axial direction within a viewing zone.

Figure 3.20 shows the integral imaging system's viewing zone where the viewing angles of all the lenslets are overlapped. Here, the viewing angle is calculated by $2\tan^{-1}(1/2F\#_2)$ and viewing zone is obtained by $z_v/F\#_2 - L$, where z_v is the viewing distance, and L is the length of the lens array. In the prototype, the viewing angle is 11.4 degrees, and the viewing zone is 15.4 cm at a distance of 2 m. Note that the FOV depends on the viewing distance, and the maximum FOV is equal to the viewing angle.

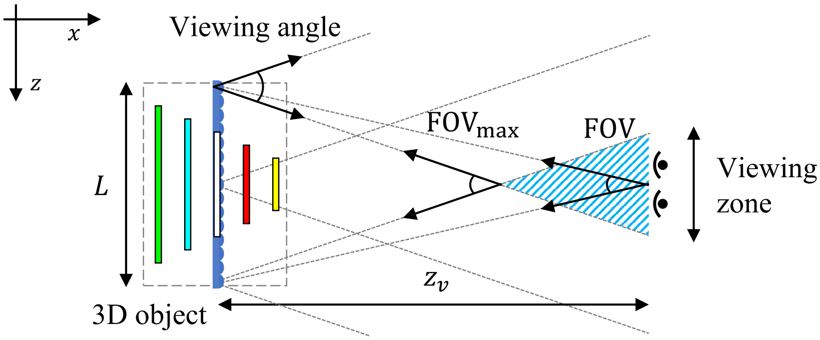

Figure 3.20 Viewing zone of integral imaging system

Since these parameters are proportional to $1/F\#_2$, they can just be expanded by replacing the lens array with a lower $F\#$. For example, when using a lens array of $F/3$, the viewing angle and area will increase to 18.9 degrees and 42 cm at the same viewing distance. However, the volumetric depth decreases from 32 cm to 19.2 cm, as described in the light field analysis. So, it is beneficial to form the MFPs as wide as possible in the light field generation. Besides, it is possible to expand further by applying previously developed technologies such as multiple projections and gaze tracking methods [25, 88].

Figure 3.21 shows the experimental results of volumetric scenes. By changing the focal length of the camera lens, it is confirmed that the depth information of the volumetric scene is well reconstructed. Since the EI is super-sampled and relayed directly, it represents an anti-aliased image in the focus plane.

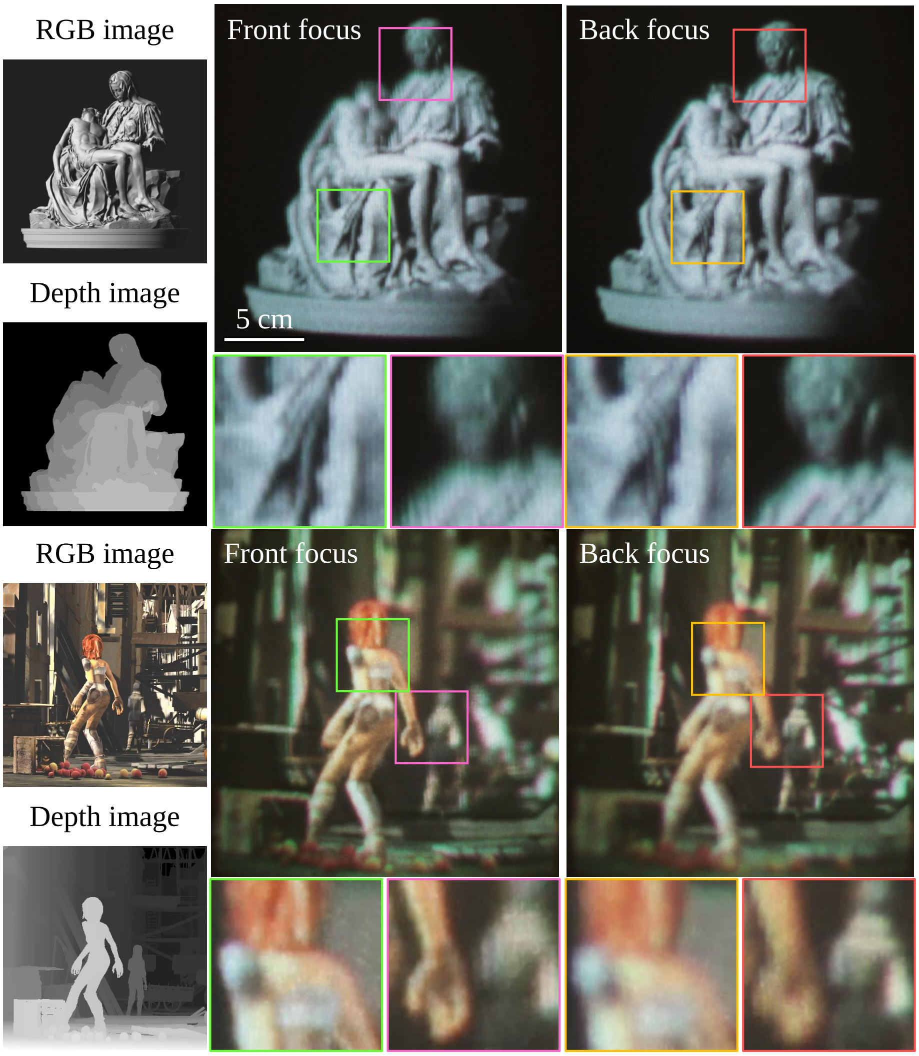

Figure 3.21 Volumetric scenes of $Pieta$ and $Market$ captured with front and back focus. The magnified images demonstrate the validity of 3D reconstruction [2].

3.5.2 Resolution assessment

To verify the capability to implement high-resolution InIm, the author examines the resolution by measuring the modulation transfer function (MTF).

Elemental image

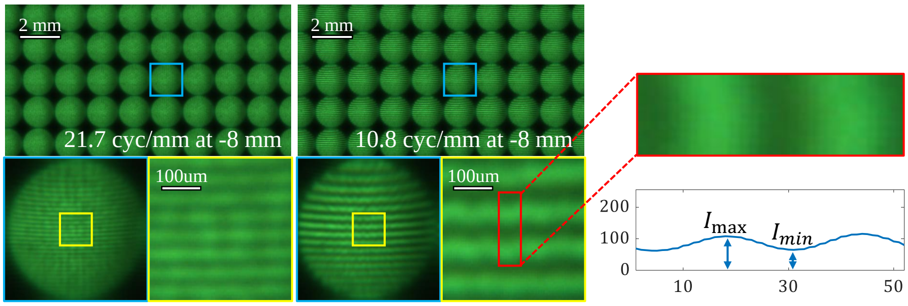

Figure 3.22 Captured EIs when for the two cases of 21.7 cycles/mm and 10.8 cycles/mm [2].

First, the author has measured the MTF to analyze the resolution of the EI_2. Here, the CCD camera (GS3-U3-89S6C-C FLIR) without a lens is placed at the EI_2 plane. A method of displaying binary gratings is utilized in the two cases of ± 8 mm. Then, the contrast of the captured grating with a certain spatial frequency is calculated.

For instance, Fig. 3.22 shows the results for two frequencies when the pickup distance is -8 mm. Two cycles near the center are only considered due to the distortion at the boundary that the FTL causes. For the area of interest, the contrast is calculated as

$$\text{Contrast} = (I_{max} - I_{min})/(I_{max} + I_{min}), \tag{3.5}$$

where the I_{max} is the maximum value, and the I_{min} is the minimum value. This process is repeated 14 times by changing the frequency of the binary grating.

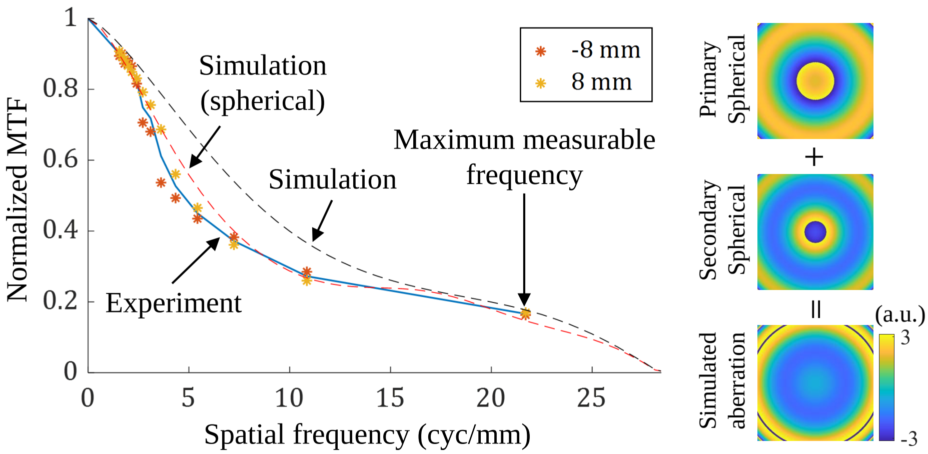

Figure 3.23 Experimental MTF results for the EI planes when z_1 is ± 8 mm. The simulated aberrations are depicted on the right side.

As shown in Fig. 3.23, compared to the simulation result, it fits quite well except for a minor mismatch due to the lenslet aberration. For the simulation, the author considers a single point (an impulse) on a single focal plane at a certain depth and calculates the impulse response function h. Since each lenslet has a square aperture, the field upon crossing the virtual aperture plane is multiplied by the pupil function of the lenslet and the lens quadratic phase. Using the angular spectrum method, the resultant field then propagates in free space to a screen plane. After the impulse response function h is in hand, the MTF can be obtained by the following equation.

$$\mathrm{MTF} = \left|\mathrm{FT}[|h|^2]\right|, \tag{3.6}$$

where $\mathrm{FT}[\cdot]$ and $|\cdot|$ denote a 2D Fourier transform and an absolute, respectively.

Unlike the ideal lens in the simulation, the lens array has various aberrations due to lens thickness, surface impurities, and irregularities in manufacturing processes. This difference between the experiment and the simulation is analyzed numerically using the Zernike polynomial to quantify the aberration of the lens [89]. For the numerical analysis of the MTF degradation, only the primary and secondary spherical aberration (Z_4^0, Z_6^0) are estimated because of the rotational symmetry of the lens. The MTF curve with the appropriate Zernike coefficients is more consistent with the experimental result, which is plotted as the dashed red line in Fig. 3.23.

When the MTF is 17 %, the prototype can generate the EI with a resolution of up to 5347×5347 (28.6 megapixels). Even using the relatively lower resolution LCD (0.22 megapixels) and DMD (0.58 megapixels), 36 times higher resolution is obtained. Consequently, the proposed method can realize the projection-type InIm, which features ultrahigh resolution that previously could not be achieved with a display panel.

Reconstruction image

For the reconstruction image, a Siemens star target is utilized to evaluate the contrast and imaging performance qualitatively, as shown in Fig. 3.24. Each spoke's angle represents the reconstruction distance and the radius from the center corresponds to the spatial frequency. Since each spoke is located at a different depth, out-of-focused spokes are gradually blurred as getting away from a focused spoke which is marked with a red arrow. In addition, the line trace along the arrow experimentally shows the contrast change. From these results, it is clear that the depth information from the MFPs is well transmitted via the EI.

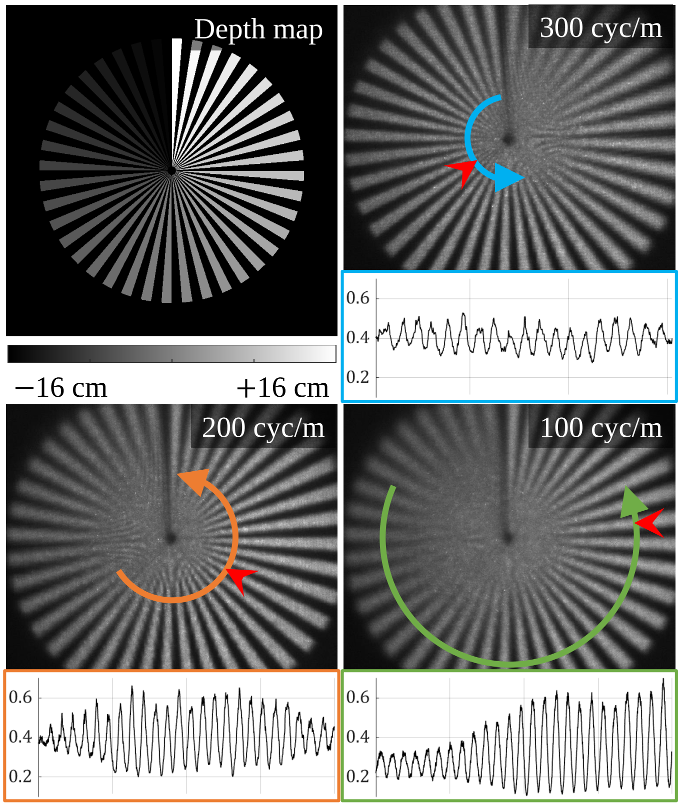

Figure 3.24 Qualitative results of Siemens star target for the green channel [2].

For the quantitative evaluation of the resolution, the MTF of the reconstruction plane is measured in the same way used to analyze the EI. The results are taken with a camera lens of $F/1.4$ and a focal length of 50 mm. As shown in Fig 3.25, the contrast is averaged over the central square area due to unwanted peaks occurring on the diffuser.

The experimental MTF curve for the reconstructed image is illustrated in Fig. 3.26. The results demonstrate that the proposed method can regenerate the light field on a large scale from high-resolution EI. However, the measured MTF is less than the wave simulation result, which is thought to be caused by aberrations in the projection lens and the MaLA. By performing the numerical anal-

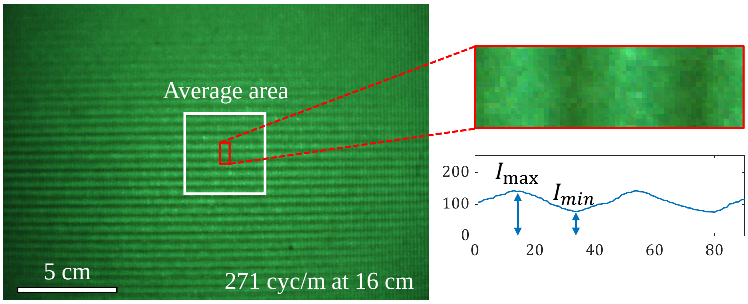

Figure 3.25 Captured image when the reconstruction distance is 16 cm, and the frequency is 271 cycles/m [2].

ysis on the reconstructed image, it can be estimated that spherical and off-axis aberrations mainly affect system imaging. Here, off-axis aberrations are caused because non-central elemental images are also involved in the superposition in the processing of reconstructing the focal plane. These aberrations, which potentially degrade the image resolution, could be alleviated by improving the manufacturing and design of the lens array [90].

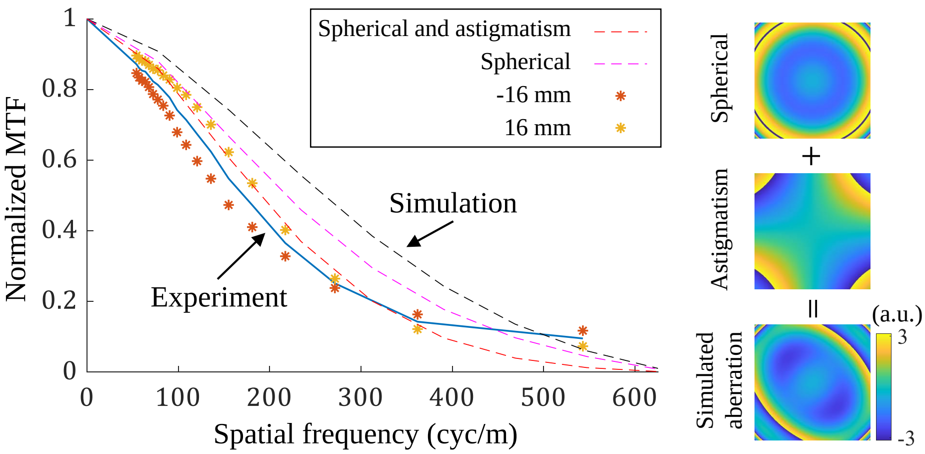

Figure 3.26 Experimental MTF results for the reconstruction planes. The simulated aberrations are depicted on the right side.

3.6 Discussion

3.6.1 Light field optimization

As introduced in Chapter 2, the multifocal planes can deal with an occlusion effect by solving an optimization problem. In a similar way to the previous method, the author optimizes the LCD image and DMD image sequence, as shown in Fig. 3.27. The image sequence is initialized with a given RGB-D im-

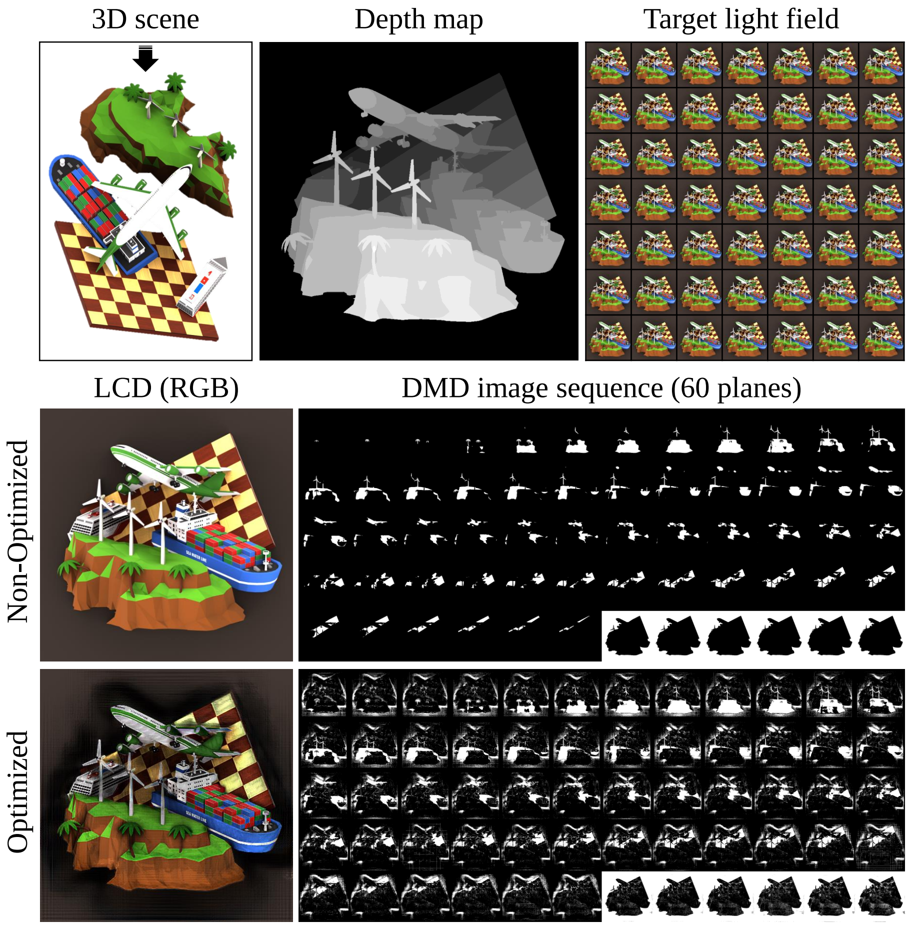

Figure 3.27 Illustration of LCD and DMD image sequence with target light field.

age, and the number of iterations is set to 300. For every 30 iterations, the DMD images are binarized and updated while minimizing errors. For the target light field, 7 by 7 images rendered in Blender are used.

Figure 3.28 shows that the optimization could reduce the artifact caused by the occlusion. Since the author generates an additive light field that shares an LCD image, there may be limitations in handling an occlusion effect fully. However, the proposed optical design could be seamlessly integrated with existing light field displays. For better performance, other strategies for generating independent focal planes or high-speed LCD can be applied [34, 35, 81].

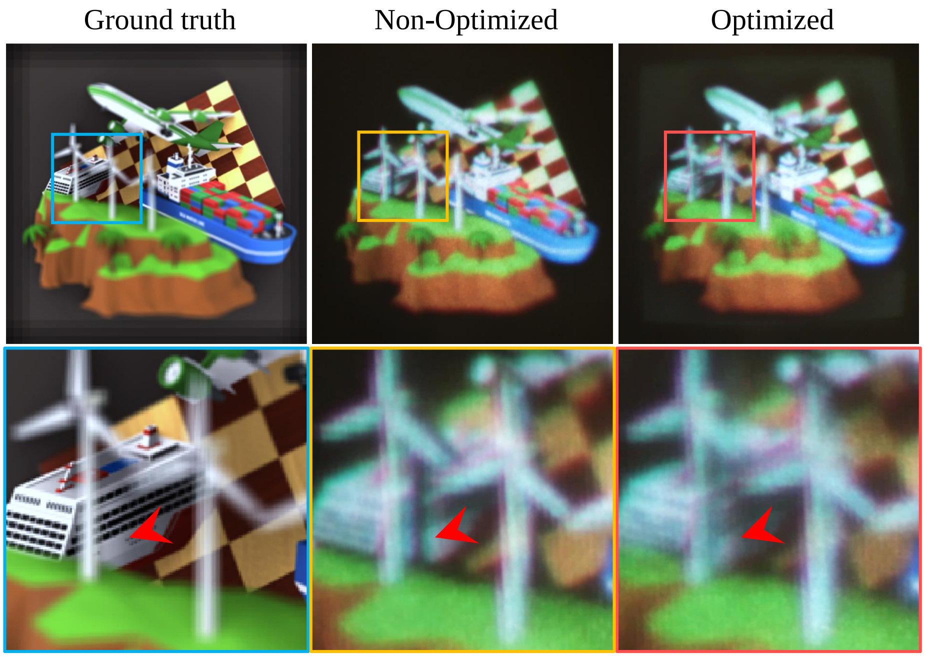

Figure 3.28 Captured images for comparison with and without the optimization. As indicated by the red arrow in the enlarged images, the artifact at the occlusion boundary can be mitigated [2].

3.7 Conclusion

In this chapter, the author presents a new optical configuration that effectively brings spatial-angular information to the big screen, as shown in Fig. 3.29. The large amount of information demanded has been challenging to deal with in a glasses-free 3D display. The proposed method optically connects the multifocal and InIm displays using a projection system. It is a perspective beyond the fundamental trade-off relation rather than merely combining existing studies. This approach effectively generates the spatial-angular information of the light field and reconstructs a volumetric image on a large screen. With the novel design, the experimental results demonstrate that optical pixel mapping can realize the EI close to 28.6 megapixels at 17 % MTF and mitigate the information loss associated with the repetitive images in the InIm. Consequently, the resolution is improved by 36 times, and it is verified that a full-parallax volumetric scene can be implemented on a large scale through projection. The proposed optical design could be widely integrated with existing light field displays, and high visual quality can be achieved based on previous projection-type InIm techniques. The author hopes that the perspective of optically manipulating spatial-angular information will inspire further developments for large-scale 3D displays.

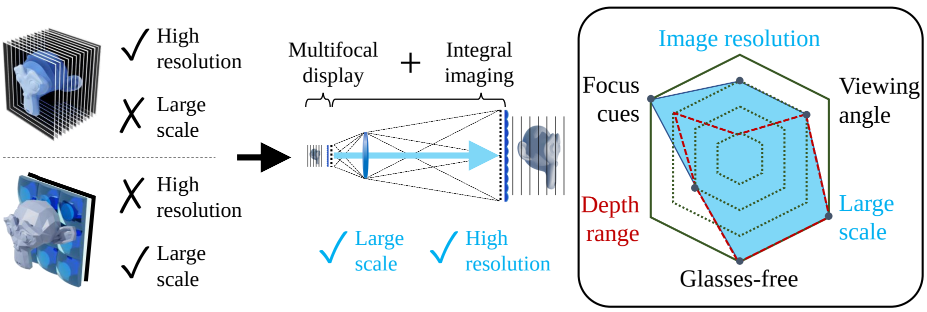

Figure 3.29 Summary of Chapter 3.

Chapter 4. Bandwidth enhancement of holographic display using multi-illumination strategy

4.1 Introduction

The previous chapter introduced a method that can bypass the spatial-angular trade-off and dramatically increase the effective light field resolution. This method could be the most feasible approach for viewing high-resolution 3D images on a large screen without glasses. However, as the light field is reconstructed through the lens optics, there is a fixed focal plane at a certain depth. This results in pixel resolution decreasing as far from the central focal plane, as shown in Fig 4.1. So the system's 3D depth is fundamentally limited to a range where the focus error equals the image pixel size [22, 37]. Such a limited depth range of light field displays can be another issue that hinders the immersive 3D experience.

In contrast, the holographic display is based on wave optics to generate 3D objects in space, allowing focal planes to be positioned freely at depth. This property of controlling the wavefront of light gives a wide depth range than light field displays. Furthermore, the hologram can express a natural parallax

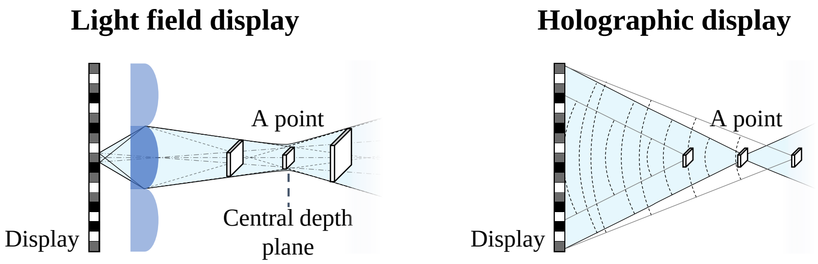

Figure 4.1 Illustration of different ways of representing a point in space.

without the occlusion effect discussed in the previous chapters. Thus, the holographic display is considered the ultimate 3D display with the potential to meet advanced performance. Since this advent of holographic displays, efforts have been steadily made to replace the existing flat panel display [91–95]. However, holographic displays have an inherent trade-off relationship between the viewing angle and the display size, described by the space-bandwidth product (SBP).

4.1.1 Space-bandwidth product

Unlike the flat panel display using an incoherent light source, a hologram is reproduced by a coherent beam and SLM. The periodic pattern of the SLM generates high-order diffraction, which is the repeated pattern of the zero-order signal [38]. Figure 4.2 illustrates the spatial frequency domain within the ± 1st diffraction order. Here, when generating a signal that exceeds the zero-order region of the SLM, the spatial aliasing problem occurs. Therefore, the effective bandwidth of the SLM is confined within the Nyquist frequency so as not to violate the sampling theorem.

In this regard, the spatial frequency bandwidth of a hologram is between

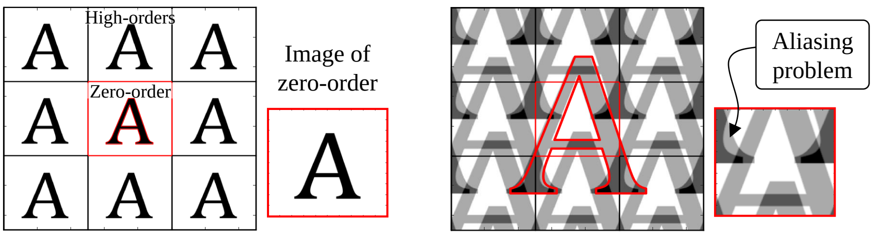

Figure 4.2 Illustration of the aliasing problem in the spatial frequency domain. The effective region of the SLM is confined to the zero-order, as indicated by the red square line.

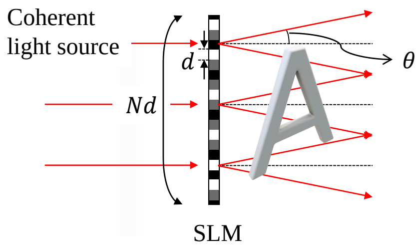

Figure 4.3 The product of the display size and the viewing angle is proportional to the number of pixels.

$-1/2d$ and $1/2d$, where d is the SLM's pixel pitch [96]. As shown in Fig 4.3, since the display size is Nd, the product of the display size and spatial frequency bandwidth, the space-bandwidth product, is equal to the number of the SLM's pixels N. In other words, this SBP determines the overall performance, such as image size and viewing angle in holographic displays. However, even with the 4K-8K resolution of the most advanced SLM, there is a fundamental limitation of having a much smaller viewing angle and size than flat panel displays. Accordingly, with the innovation in the manufacturing industry to increase the number of SLM pixels, a lot of research has been actively done to overcome the bandwidth limitations of the holographic display.

The multiplexing technique in time or space can effectively extend the limited bandwidth of the SLM, so it has been widely utilized in holographic displays. However, the spatial multiplexing method using multiple SLMs is bulky and expensive [39]. And the case of the temporal multiplexing method requires relatively complex optics and consumes the system's frame rate as the number of multiplexing [91, 92]. Recently, several studies have shown that a scattering medium can be utilized to expand the viewing angle of the holographic display [93, 94]. However, they require highly precise alignment between the scattering medium and the SLM, and the image quality could be degraded by unexpected speckle noise due to scattering. This speckle noise caused by using a coherent light source could also be a critical issue of the holographic display that decreases the contrast and is potentially harmful to the eyes.

4.1.2 Compression of information beyond hardware limits

In this chapter, aiming at the ultimate 3D display with a fully advanced experience, the author introduces a practical method to deal with the challenges of holographic display. Here, multi-angle illumination using multiple laser diodes (LDs) is adopted to expand the limited diffraction angle of the SLM. The multi-angle illumination strategy has been applied in several studies as a practical approach to extending the bandwidth of holographic displays [91, 97]. However, since each light source shares the same SLM pattern, a temporal multiplexing method or additional eye tracking should be required. In this dissertation, to solve the problem of signal repetitions, the author additionally utilizes a random binary mask (BM). This approach filters the duplicated signals through different mask patterns, giving a degree of freedom to express different information. Since much information is generated simultaneously from a single phase pattern, it can be considered that the information is compressed. Consequently, the proposed method can expand the bandwidth of the holographic display by alleviating the bottleneck of hardware limitations. The author demonstrates via simulations and experiments that the method effectively increases the bandwidth with sufficient image quality. Furthermore, the speckle noise can be reduced by the advantage of incoherent summation in the reconstruction plane.

4.2 Principle

4.2.1 Multi-illumination strategy

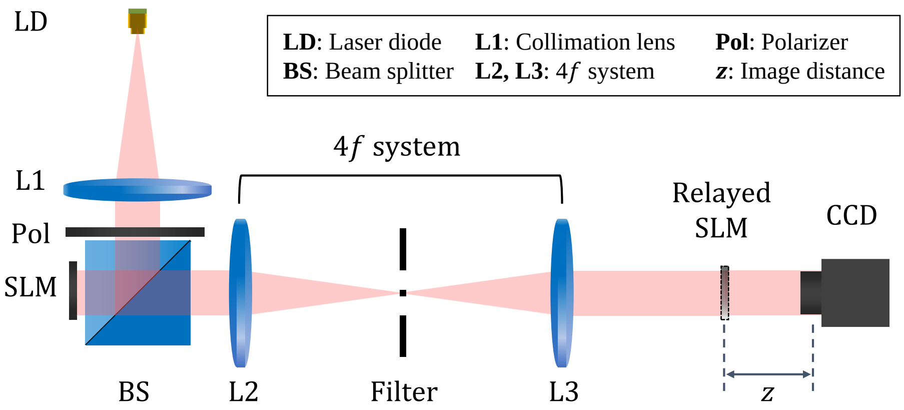

Figure 4.4 Schematic diagram of the conventional holographic display.

Figure 4.4 shows a conventional holographic display with filtering optics. Here, the high-order diffraction caused by the periodic structure of the SLM appears as a duplicate of the original signal and is generally removed through optical filtering as it causes aliasing between the signals. However, in the proposed method, the intensity of the high diffraction order is further strengthened by the multi-illumination strategy to expand the bandwidth.

Each direction (i, j) of illumination is set to match with that of high diffraction orders as follows.

$$\theta_{ij} = \sin^{-1}[m_{ij}\lambda/d], \tag{4.1}$$

where λ is the wavelength, d is the pixel pitch of the SLM, m_{ij} denotes the $(i, j)^{\text{th}}$ diffraction order, and i, j take positive and negative integer values.

As shown in Fig. 4.5, by illuminating multiple laser diodes from different angles, the spatial frequency range of the hologram becomes wider with the

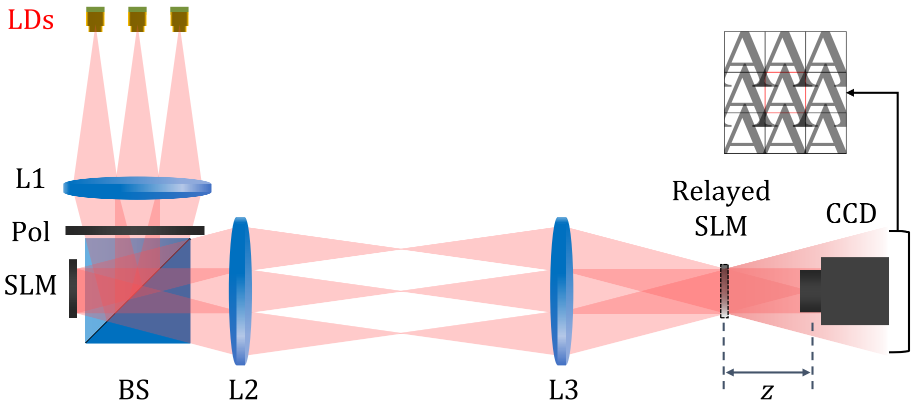

Figure 4.5 Schematic diagram of the holographic display with multi-illumination strategy. The multiple illuminations can increase the bandwidth of the holographic display, but duplicated signal causes spatial aliasing.

extended energy distribution, while the incoherent summation of LDs reduces speckle noise. However, the wavefront from each LD transfers identical information of the SLM that differs only in the carrier frequency, resulting in indistinguishable crosstalk between the signals.

4.2.2 Method of handling the aliasing problem

Figure 4.6 shows a schematic diagram of the proposed holographic display. The main idea is to optically break the equivalence of the information by utilizing a random BM at the frequency domain. In other words, each signal is filtered by the BM of different positions, which breaks the correlation of the signal. Finally, by making the duplicated information meaningful with the BM, it is possible to seamlessly extend the entire bandwidth of the holographic display over the number of illuminations. At the same time, the information is also efficiently compressed into an SLM phase pattern.

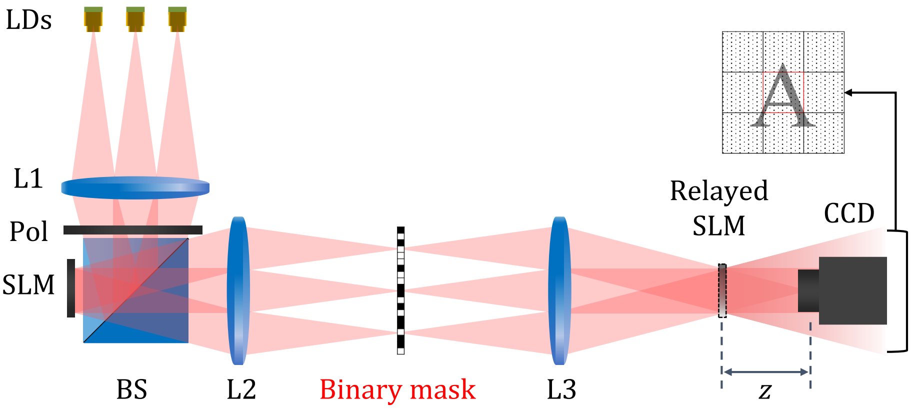

Figure 4.6 Simplified concept diagram of the optical structure. Multiple collimated beams from LDs are modulated in the phase-only SLM simultaneously. Then in the spatial frequency domain, the wavefronts are separately filtered with a BM to reduce their information similarity. After a $4f$ system, the relayed SLM has extended bandwidth in proportion to the number of illuminations [3].

4.2.3 Algorithm for optimization

Briefly, a multi-illumination is realized by adopting the LD array. Then, the BM in the spatial frequency domain acts as the filter and gives the degrees of freedom to modify each information of LDs individually. Here, the BM of arbitrary shape and incoherent illuminations are considered in generating the computer graphic hologram (CGH).

Figure 4.7 illustrates the proposed algorithm to synthesize the CGH. The phase of the SLM is optimized by using the stochastic gradient descent (SGD) approach. First, the high-order diffraction is considered [95], described as follows.

$$U = \mathrm{comb}[df_x, df_y] \otimes \mathrm{FT}[e^{i\phi}], \tag{4.2}$$

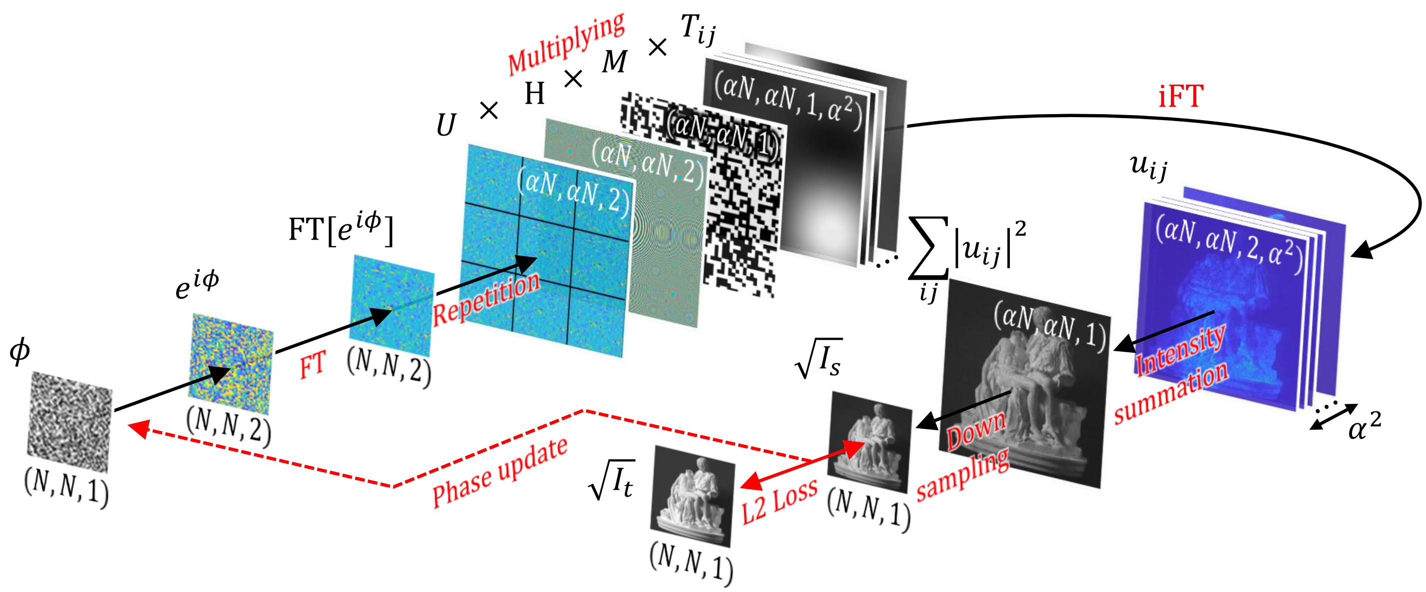

Figure 4.7 Schematic diagram of the proposed algorithm. The values in the bracket indicate the range of the spatial and frequency domain. Note that the number of 2 in the third dimension indicates complex values. The wavefront $e^{i\phi}$ from the phase-only SLM is tiled in the frequency domain to account for the high-order diffraction. The spectrum U is then multiplied by the binary mask M, transfer function H, and sinc amplitude T_{ij}. Each energy distribution of the $(i,j)^{th}$ LD is calculated individually in the fourth dimension. After the inverse Fourier transform, the reconstructed images $|u_{ij}|^2$ are summed on an intensity basis. Then, the loss is calculated to update the phase pattern taking the amplitude of the reconstructed result I_s [3].

where U is the tiled spectrum, $\mathrm{comb}[\cdot,\cdot]$ denotes the Dirac comb function, the symbol $\otimes$ is convolution operator, $e^{i\phi}$ is the modulated wavefront at the SLM, and f_x,f_y are the spatial frequencies.

The tiled spectrum U is then multiplied by the binary mask $M \in \{1,0\}^{N_{xy}}$ and transfer function H of the angular spectrum method after padding in the frequency domain [98]. With the propagation distance z, H is defined as

$$H = \begin{cases} e^{i\frac{2\pi}{\lambda}z\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}}, & \text{if } \sqrt{{f_x}^2+{f_y}^2} < \frac{1}{\lambda} \\ 0, & \text{otherwise} \end{cases}. \tag{4.3}$$

The author accounts for the carrier frequencies of the LDs and the fill factor of the SLM in the frequency domain by multiplying sinc amplitude. When the SLM is illuminated by the $(i,j)^{\text{th}}$ LD, the 2D sinc amplitude T_{ij} is given by

$$T_{ij} = \mathrm{sinc}[af_x + i/d, af_y + j/d], \tag{4.4}$$

where a is the width of the active area of the pixel. Note that each carrier frequency is set to be matched with high-order direction, as shown in Fig 4.8, but not mandatory if taken into account in the optimization. The hologram can be optimized even if the carrier frequency and wavelength are set differently. So a full-color image could be produced using a time-sequential operation without any hardware change.

From Eqs. 4.2-4.4, the complex amplitude u_{ij} for $(i,j)^{\text{th}}$ LD at the distance z is represented as

$$u_{ij} = \mathrm{iFT}[U \circ H \circ M \circ T_{ij}], \tag{4.5}$$

where $\mathrm{iFT}[\cdot]$ denotes an inverse 2D Fourier transform operator and $\circ$ is elemental-wise multiplication. Finally, the optimal phase pattern is obtained based on the SGD approach by solving the following problem:

$$\underset{\phi}{\mathrm{minimize}}\, \mathcal{L}(s \cdot \sqrt{I_s}, \sqrt{I_t}), \tag{4.6}$$

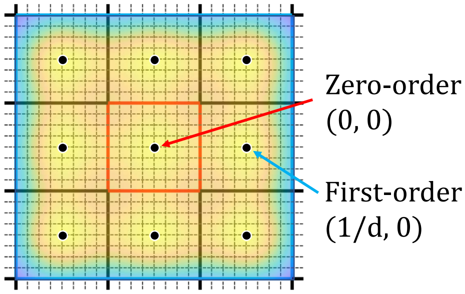

Figure 4.8 Illustration of the spatial frequency domain (f_x, f_y). The multi-illumination angles are adjusted toward each central point of the high diffraction orders. Accordingly, the intensity is widely distributed over the extended bandwidth.

where $\mathcal{L}$ is the cost function representing squared L_2 loss, s is the scale parameter to adjust the energy, and I_t is the target intensity. I_s is the sum of all the intensity from each directional illumination defined as

$$I_s = \sum_{i,j=-(\alpha-1)/2}^{(\alpha-1)/2} |u_{ij}|^2, \tag{4.7}$$

where α is the number of LDs for x- or y-coordinates. The $\sqrt{I_s}$ is downsampled before calculating the loss to match the resolution between the SLM and target. This process makes the optimization problem less difficult and ensures better image quality.

4.3 Implementation

4.3.1 System design

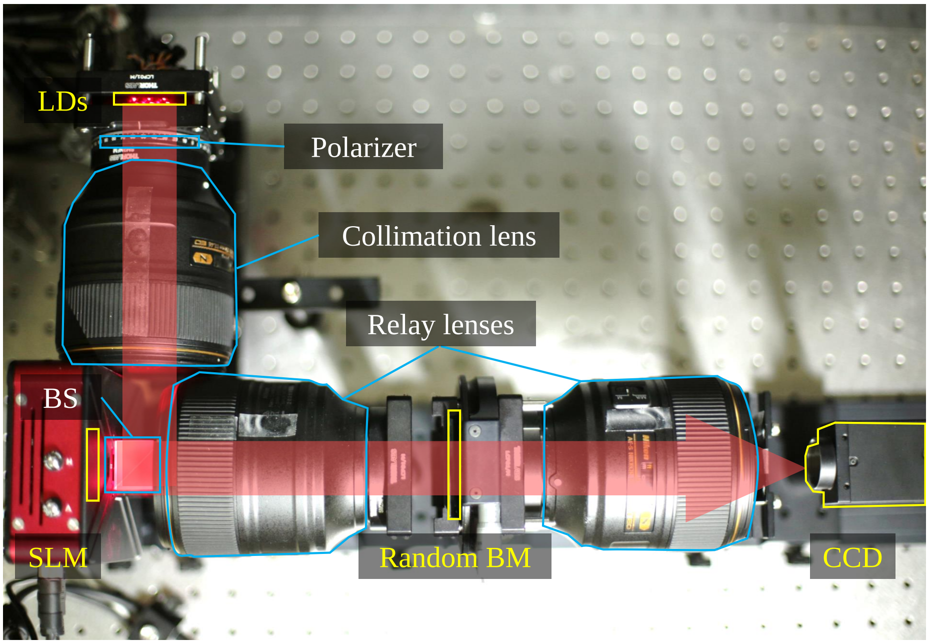

Figure 4.9 Photograph of the prototype [3].

Figure 4.9 shows the experiment setup. The author uses Thorlabs Exulus HD2 for the SLM and CPS635F laser module with 635 nm wavelength. A 60 mm cage system from Thorlabs and a two-axis linear stage are utilized for the mask alignment. Three Nikon 105 mm camera lenses are used for the collimation and $4f$ relay optics in the experiment.

As shown in Fig 4.10, 3×3 LD array is assembled with a plastic housing made with a 3D printer. The Blender software is used for 3D modeling of the LD housing. There is a minor misalignment error in the fabricated LD housing due to the relatively large nozzle size of the 3D printer (0.4 mm, Ultimaker 3).

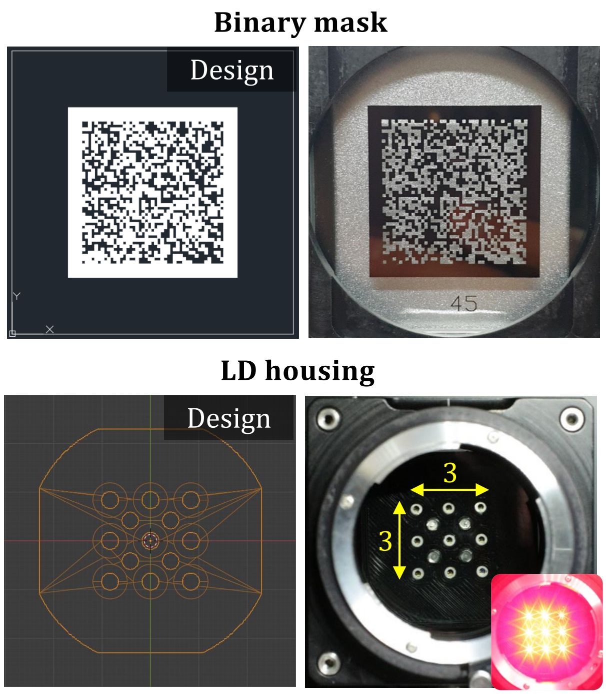

Figure 4.10 Designed binary mask and LD housing.

This issue could be solved through precise engineering in the future. In contrast, the variance of the commercial laser diode is rather small experimentally.

The BM is fabricated on glass (50 mm$\times$50 mm square, 2.3 mm thick) coated with patterned chromium. The AutoCAD program is used to design the mask pattern $m(x, y)$. Here, The spatial frequency domain (f_x, f_y) is converted to spatial domain (x, y) as follows.

$$m(x, y) = f_L \tan[\sin^{-1}[\lambda M(f_x, f_y)]], \tag{4.8}$$

where f_L is the focal length of the relay lens and λ is the wavelength. From Eq. 4.8, the pixel pitch of the mask is about 0.56 mm when the f_L is 105 mm, which provides high robustness in alignment compared to previous approaches using scattering media.

4.3.2 System alignment issue

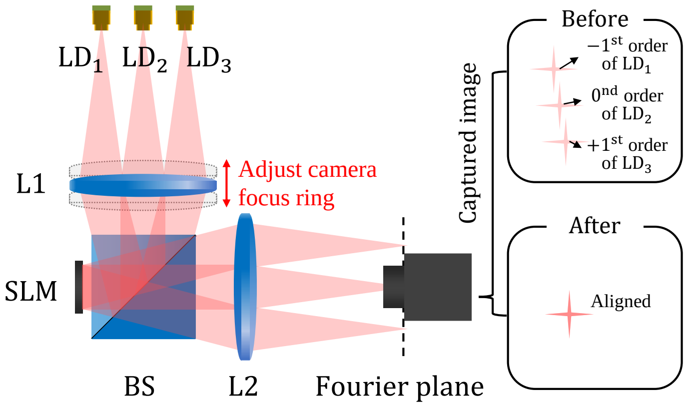

Figure 4.11 Alignment of multiple laser diodes.

For hardware implementation, it is essential to accurately match the focal length of the camera lens and the gap distance between the LDs, as in the simulation environment. Therefore, since the focal length of the camera lens is slightly changed according to the position of the focus ring (about between 104 mm and 106 mm), as illustrated in Fig 4.11, The focus ring of the camera lens is adjusted to match the target focal length of 105 mm till each diffraction order converges to a point on the Fourier plane.

4.4 Analysis

The feasibility of the proposed method is verified by measuring the reconstructed image quality using the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) index in the simulation. To confirm the essential characteristic of the BM, The author compares the proposed method against single LD and multiple LDs cases without the mask. For the simulation, 3×3 LDs with a wavelength of 635 nm and virtual SLM of 1000×1000 resolution with 8 µm pixel pitch are utilized. The SLM fill factor (a^2/d^2) is set to 0.92.

4.4.1 Determine binary mask pattern

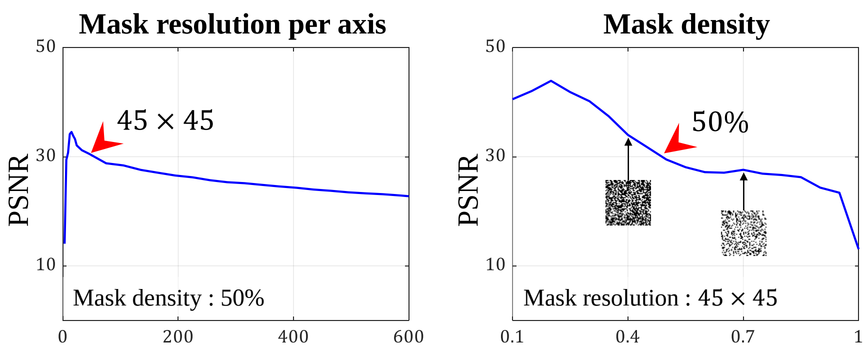

Figure 4.12 Quality analysis according to the mask resolution and density [3].

First, the author investigates the properties of the BM pattern. For the simulation, the learning rate is set to 1.00, and the number of iterations is 1000. A 'Dog' image is used at a propagation distance of 50 mm [99]. The phase value ϕ is randomly initialized in the range of $[-\pi, \pi)$ for optimization. Figure 4.12 shows that the image quality of the system is not significantly affected by the resolution of the BM. Here, very low-resolution BM may cause a noticeable discontinuity in the viewing angle. On the other hand, high-resolution BM requires precise alignment and decreases the optimization performance with a sparse spectrum. Besides, the BM with a lower density shows higher PSNR values as it increases the capacity to modify the information of each source. However, it penalizes energy efficiency. Therefore, concerning the practical aspects, the BM with 50% density and 45×45 resolution is selected.

4.4.2 Comparison of image quality

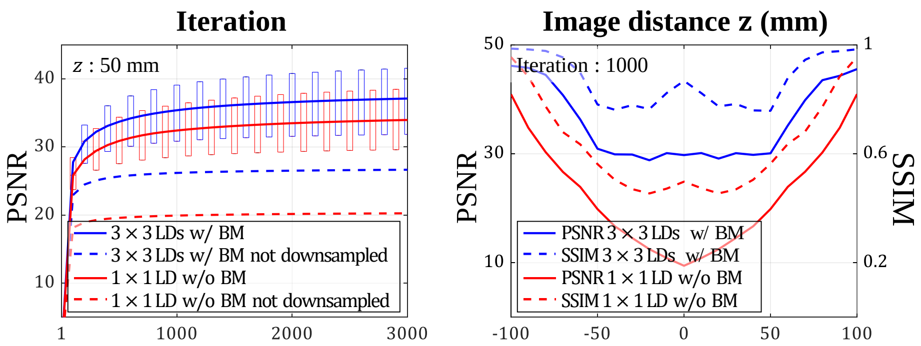

Figure 4.13 Quantitative evaluation according to the iteration number and the distance from the SLM. Box plot denotes the median, 25^{th}, and 75^{th} percentiles [3].

Then, the reliability is investigated using 50 images in the DIV2K test dataset [99]. As shown in Fig. 4.13, the proposed method achieves higher image quality than the cases not applied with the BM. Even in the case without downsampling, it gives better performance. Although the performance can vary depending on the image used, the proposed method not only yields a PSNR improvement but also is more robust near the SLM, where the information overlap is aggravated. Furthermore, the proposed approach can cover the wide depth range as the noise from the duplicated information is highly mitigated compared to the case without the BM. The simulation results verify that the proposed method could support high image quality with the multi-illumination strategy.

Figure 4.14 directly shows the validity of the random BM. The simulation is conducted to compare aliasing artifacts with conventional systems for 25mm and 50mm distances. In the case of not applying the BM, the duplicated images of the zero-order signal are placed in the shifted location according to each carrier frequency of LDs. Moreover, the aliasing gets worse as the propagation distance gets smaller [100]. In contrast, the proposed method effectively suppresses the noise originating from duplicated information. After each wavefront of the LD is blocked differently in the BM, the correlation sharing the single SLM is broken. As a result, the target image can be reproduced individually from the multiple illuminations.

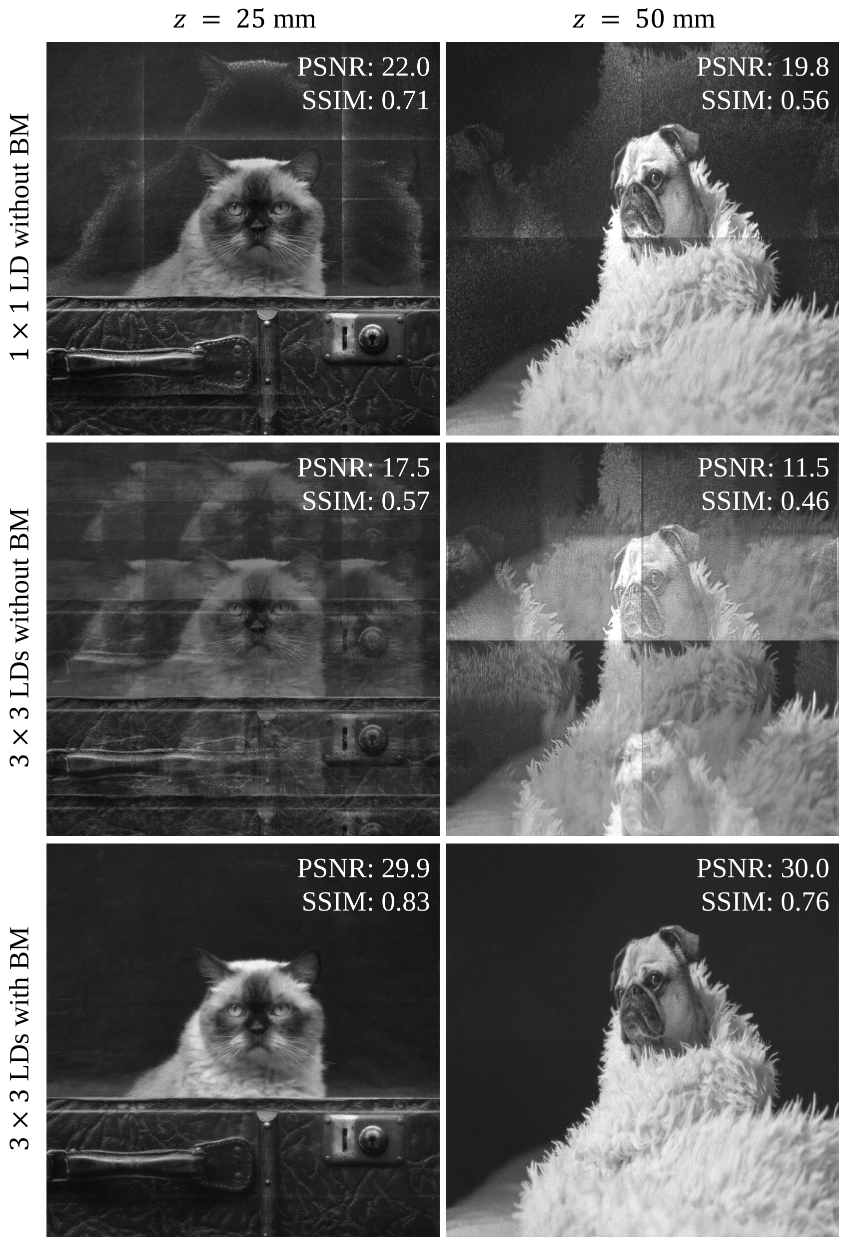

Figure 4.14 Comparison of reconstructed images with and without the BM. The images with a dark background are presented to observe the noise better. In the case of not using the BM, 'Dog' gets a lower PSNR as the background is darker than 'Cat' [3].

4.4.3 Holographic stereogram

The proposed algorithm can be applied to a method of holographic stereograms. This method composes sub-holograms for each viewing angle based on multi-view images and tiles them in the spatial frequency domain. To verify that the information shared by each LD is segmented through the BM, nine sub-holograms (3×3) covering different images are generated in the simulation. In this case, the Fourier transform is calculated for each spatial frequency area that is independent spatially without propagation. Then, the phase pattern is optimized using multiple targets.

Figure 4.15 represents the reconstructed results from each $(i, j)^{th}$ frequency area. Although these areas share the same phase pattern, the proposed method can reconstruct the independent images simultaneously for each carrier frequency. As a result, a wider viewing angle can be achieved in holographic displays with extended bandwidth.

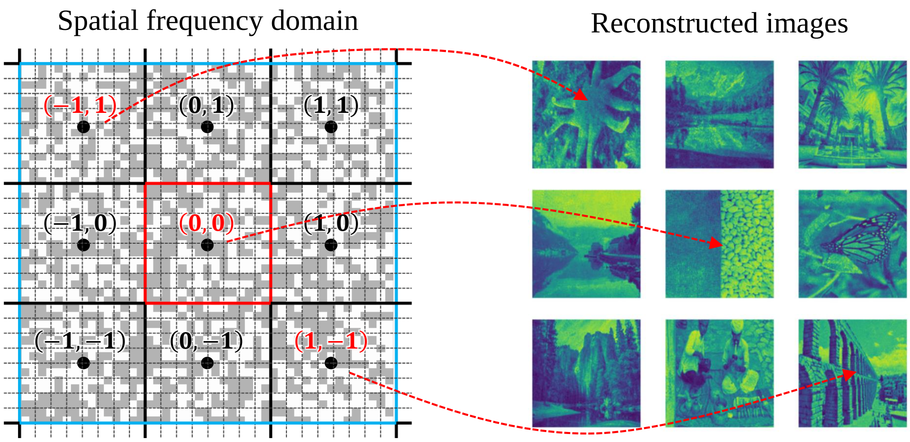

Figure 4.15 Reconstructed images for each sub-hologram (or for each direction of (i, j)). The information originating from the same SLM pattern is individually modulated after passing through the BM and represents the target images.

4.5 Results

4.5.1 Imaging result

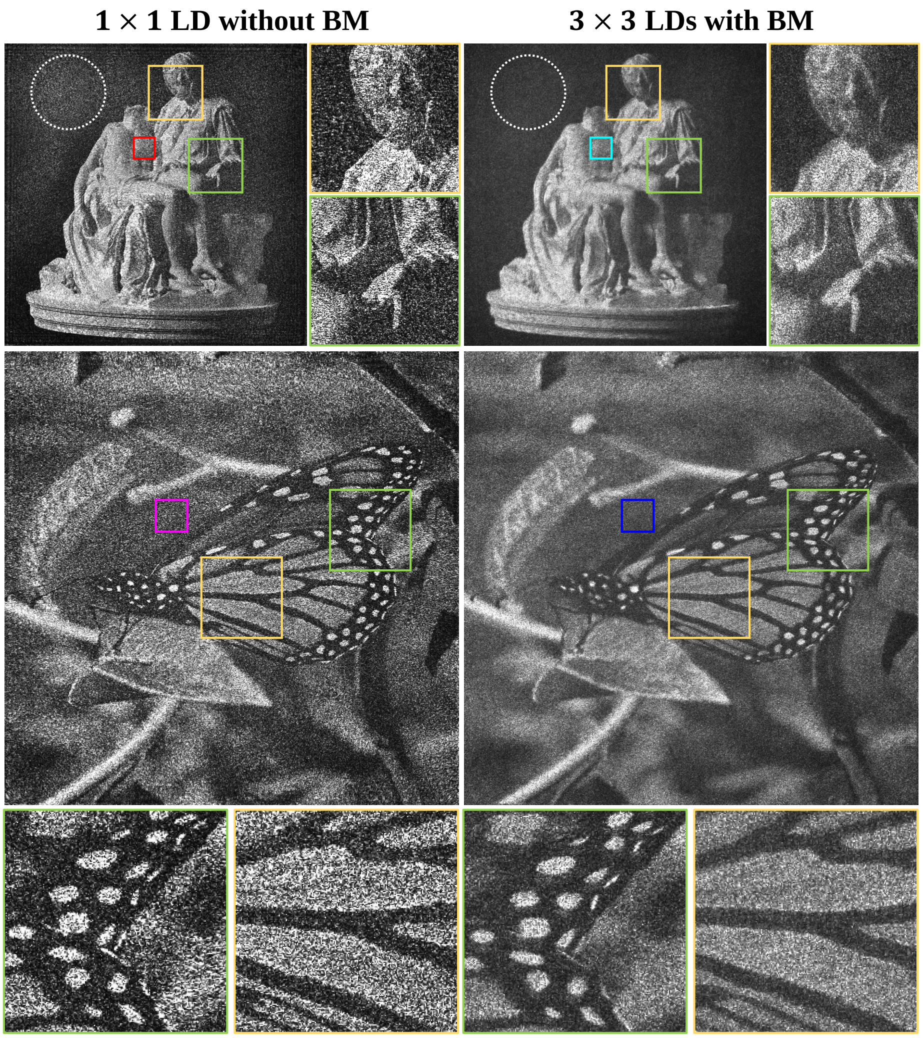

Figure 4.16 Experimental results of the prototype with a propagation distance of 50 mm. The noise by duplicated information is well suppressed. The results are captured using the CCD camera without a lens [3].

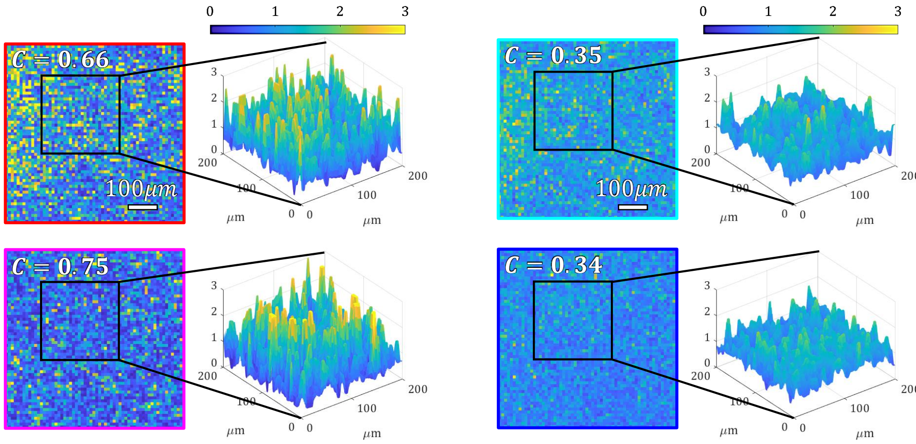

Figure 4.17 Comparison results to demonstrate the feasibility of speckle noise reduction. The speckle contrast C is calculated by dividing the standard deviation by the mean intensity [3].

In the experiment, the author compares the proposed method with the previous research [100] to confirm the benefit of noise signal suppression. As shown in Fig. 4.16, the aliasing artifacts from the duplicated signal are clearly suppressed even on a black background (white circle). The captured results also show the novel feature of decreasing unwanted speckle noise. Overall, the proposed method produces a face more smoothly compared to the standard case using a single LD source without the mask. Figure 4.17 presents enlarged images cropped from the result of Fig. 4.16. The speckle reduction effect is proportional to the overlap ratio of the regions illuminated by the LDs in the reconstruction plane. Since each central point of the high diffraction orders is placed at the edge of the image for a propagation distance of 50 mm, the speckle contrast reduces to about half as four independent speckle patterns are superimposed [91, 92]. Consequently, this approach has the distinct advantage of not sacrificing the system's frame rate.

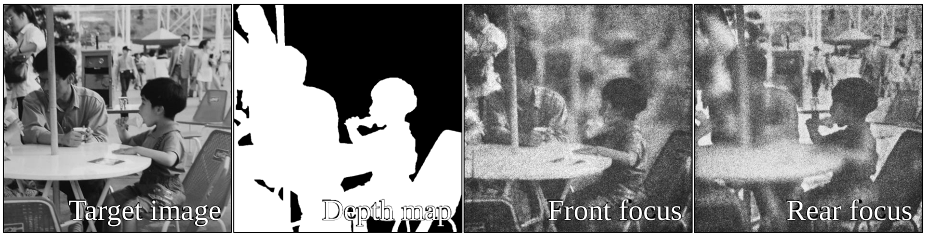

Figure 4.18 Experimental results of a multi-depth scene. The distances are 50 mm and 53 mm, respectively [3].

Furthermore, the 3D hologram can be expressed in the same way as in previous studies [101, 102]. The phase pattern is optimized with an amplitude loss function for two target planes to obtain the results of Fig. 4.18. The images are well reproduced experimentally at each depth.

As demonstrated in the simulation results, the proposed method can widen the viewing angle with extended bandwidth. To verify this experimentally, a hologram is generated by dividing the extended bandwidth into three sub-holograms. Each sub-hologram sequentially represents different images from ‘A’ to ‘C’ as the viewing angle changes in the horizontal direction. Figure 4.19 demonstrates that the aliasing artifacts are effectively alleviated, and the viewing angle is tripled while exceeding the single LD case, $\pm 2.3°$.

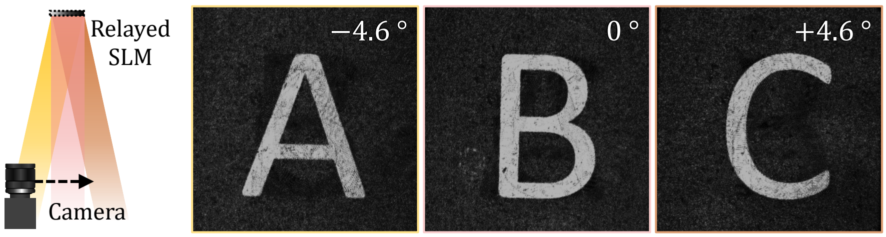

Figure 4.19 Experimental results in changing the observation angles. The results are captured using a CCD camera with a 100 mm lens [3].

4.6 Discussion

4.6.1 Limitations

The proposed method achieves the viewing angle expansion through the multiple illumination strategy. The author believes the proposed method can further extend the laser diodes as much as desired if hardware factors are thoroughly considered in the simulation. However, the performance of a liquid crystal-based SLM varies according to the angle of incidence, resulting in poor reconstruction and limiting the maximum space bandwidth. Therefore, as the angle of incidence increases, it is necessary to measure the changing factors such as fill factor, energy distribution, light efficiency, and phase modulation according to the incident angles, respectively. In particular, as in the paper [103], depolarization light and dynamic range change occurring at the high incident angle would reduce the image quality. However, if the technology of optical compensation film (e.g., triacetyl cellulose, TAC) used in conventional LCD is applied to the SLM, the off-axis phase modulation error could be alleviated. The practical bottleneck of the proposed system is the limited numerical aperture and vignetting effect of the relay lens. Therefore, implementing the system without a relay lens would be the next step of this method in the future.

In this chapter, the core of the proposed method is enabling multiple tilted beams incident on the SLM to achieve a wider viewing angle rather than making the system larger. Here, since there is the trade-off between the viewing angle and display size in holographic displays, the proposed method to address the limited SBP of available SLMs also contributes significantly in terms of system size. Although still not allowing a large enough size, this approach can be leveraged for further development and realization of large-scale holographic displays as hardware advances.

4.7 Conclusion

In this chapter, the author introduces the new optical configuration utilizing the multi-illumination and BM to address the limited bandwidth, which has been a hurdle to the practical use of the holographic display. This approach effectively circumvents the trade-off between viewing angle and display size. The contribution of this work can be summarized as shown in Fig. 4.20. The experimental results demonstrate that the multi-illumination strategy achieves sufficient image quality with the extended bandwidth and effectively reduces the speckle noise. The proposed method has practical significance as it is less sensitive to errors than the methods utilizing a scattering device which requires a high-precision component and micro-scale alignment optics to improve the bandwidth [93,94]. Even though it takes more time to synthesize the hologram with the expanded bandwidth, previous work has shown that training propagation models make real-time hologram synthesis possible [104]. Furthermore, the proposed method can be integrated with previous works optimizing light sources to further reduce the speckle noise and replace the random BM with a grayscale mask to improve light throughput and image quality. The author believes the proposed concept of multi-illumination with the mask would generally be applicable to other systems such as near-eye or table-top holographic displays.

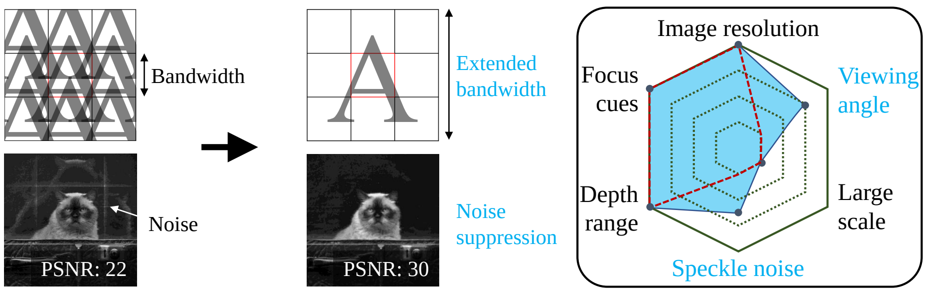

Figure 4.20 Summary of Chapter 4.

Chapter 5. Conclusion

Over the past century, information displays have been developed for a more realistic and immersive experience. With the improvement of technology, the current ultrahigh-resolution display provides images that are difficult to distinguish the pixels with the human eye, and a high level of immersion can be experienced from the large screen that fills the theater. However, whether these state-of-the-art displays can fully represent real-world experiences remains a question mark. What is clear is that, although not easy, realistic 3D representation is the direction in which display technology is headed. Along with this, in this dissertation, the author focuses on the realization of more realistic and immersive 3D displays.

Many studies and efforts have developed various methodologies to represent 3D images on a 2D display. In Chapter 1, the author introduces previous research and highlights the limitations of each candidate technique for true 3D display. Since extending 2D to 3D is challenging, each 3D technique has its own trade-offs. The author explores ways to solve the existing problems of promising 3D displays from the perspectives of large-scale realization and user experience. In the following chapters, three newly proposed solutions are introduced to overcome the various limitations for a more immersive experience.

In Chapter 2, the author introduces a solution that can create a more immersive experience by providing monocular focal cues in the most popular 3D environment, such as theater. Current stereoscopic display systems without appropriate focus cues have a problem reducing the expressible depth range of 3D objects to ensure viewer comfort. The author proposes a new stereoscopic display system that combines a multifocal display, one of the light field methods for appropriate focus cues. Experimentally the author shows that the proposed system can support a wide depth range from 25 cm to optical infinity with sufficient tolerance while preserving high resolution and contrast. This novel design will provide the technical background to place 3D objects right in front of the users' eyes for the stereoscopic display field. Furthermore, in the simulation of applying the proposed system to a large space, it is confirmed that the viewing experience can potentially be more uniform and comfortable, regardless of the viewing position. Consequently, the author believes that the proposed system could provide a new era of immersive viewing experiences in 3D theater. However, the stereoscopic display has an intrinsic problem requiring glasses. In contrast, autostereoscopic displays allow for a comfortable viewing environment as the users do not require to wear any glasses. However, in the absence of glasses that can efficiently provide binocular parallax, the system must produce different images for different viewing angles. It results in a huge amount of information to be processed.

In Chapter 3, the author introduces a solution to address the information requirements for autostereoscopic 3D displays. In this method, the author combines two different light field methods of multifocal display and integral imaging using an optical projection approach. Since the proposed optical design inherits the merits of two methods, it has the representative advantage of generating a large amount of spatial-angular information on a large scale. As a result, the author not only achieves ultrahigh-definition light field but also effectively solves the enlargement problem of multifocal display and the information loss problem of integral imaging. The performance of the proposed system is demonstrated and verified qualitatively and quantitatively. The prototype can synthesize the light field equivalent to 28.6 megapixels, which is 36 times higher resolution than the original. The author believes that the proposed system has the potential to support high-quality 3D images without glasses, so it can be effectively utilized where immersive 3D experience is competitive in everyday life, such as digital signage, media art, and head-up displays. However, the expressible depth range of the 3D objects is fundamentally limited near the focal plane of the lens array. Considering this, the holographic display could be the final solution with its advanced properties, such as wide depth range and natural parallax.

In Chapter 4, the author presents a solution to address the limited bandwidth, which is a critical limitation for large-scale implementation of holographic displays. This method combines two different components of multiple laser diodes and random BM. Here, while the LDs expand the limited diffraction angle of the SLM, the BM is placed at the frequency domain to give the degrees of freedom to modulate each wavefront of LDs differently. The author verifies that random BM solves the problem of signal repetitions caused by sharing the same phase pattern of the SLM through the experiments. This approach is a novel perspective that can effectively compress even more information beyond the fundamental trade-off relation. As a result, a hologram with nine times more bandwidth than the zero-order signal is achieved. Further, with the merit of mutual incoherent light sources, it has the added advantage of reducing the speckle noise of the hologram while widening the viewing angle. Although the extended bandwidth is still unsatisfactory for practical use, better hardware will be developed, and the author believes that the proposed method could eventually contribute to the ultimate 3D display realization.

For future work, the author proposes several meaningful and interesting topics related to the proposed solutions. First, in the stereoscopic displays of Chapter 2, another approach to generate multifocal planes can be applied. In the tomographic projection method, the combination of the DMD and transparent

LCD can effectively generate dozens of multifocal planes, but a complex optical system is required, and a lot of light is also lost in the process. In this regard, micro- and mini-LED displays are a suitable alternative in all aspects, such as operation speed, light efficiency, and image quality. It is even more encouraging that the movie theater with a LED display has already been commercialized.

Second, in Chapter 3, a more practical and advanced light field projection system can be developed by applying the previous integral imaging projection technique. In the reconstruction process, Current systems require heavy optics such as lens arrays and scattering screens. A thinner and lighter screen can be realized if a well-researched concave mirror array is applied here. In addition, the effective resolution can be expanded by adopting a multi-projection system.

Finally, in Chapter 4, research to make the system more compact while extending the bandwidth of the holographic display can be studied. The extension of SBP is an important research topic for near-eye displays such as holographic VR. However, in order to apply the proposed method, the bulky size of the $4f$ relay optics could be a problem. In this case, using a pair of lens arrays instead of relay lenses is expected to seamlessly extend the bandwidth of the holographic display with a compact form factor. Also, finding an optimal mask pattern that is robust to alignment and enables high performance can be a meaningful topic.

The author believes that displays that provide realistic, immersive, and indistinguishable experiences will eventually be developed. In this context, the author hopes that this dissertation will contribute to the further development of the 3D display field, which is facing the most difficult challenge of realizing a volumetric 3D object from a 2D display. In particular, the author hopes that the proposed methods will inspire the realization of more advanced large-scale 3D systems and allow the viewing experience to be expanded without boundaries.

Bibliography

[1] Y. Jo, S. Lee, D. Yoo, S. Choi, D. Kim, and B. Lee, "Tomographic projector: Large scale volumetric display with uniform viewing experiences," *ACM Transactions on Graphics*, vol. 38, no. 6, p. 215, 2019.

[2] Y. Jo, K. Bang, D. Yoo, B. Lee, and B. Lee, "Ultrahigh-definition volumetric light field projection," *Optics Letters*, vol. 46, no. 17, pp. 4212–4215, 2021.

[3] Y. Jo, D. Yoo, D. Lee, M. Kim, and B. Lee, "Multi-illumination 3D holographic display using a binary mask," *Optics Letters*, vol. 47, no. 10, pp. 2482–2485, 2022.

[4] M. J. Tovée, *An introduction to the visual system.* Cambridge University Press, 1996.

[5] N. A. Dodgson, "Variation and extrema of human interpupillary distance," in *Stereoscopic Displays and Virtual Reality Systems XI*, vol. 5291, pp. 36–46, International Society for Optics and Photonics, 2004.

[6] J. P. Wann, S. Rushton, and M. Mon-Williams, "Natural problems for stereoscopic depth perception in virtual environments," *Vision Research*, vol. 35, no. 19, pp. 2731–2736, 1995.

[7] B. Lee, "Three-dimensional displays, past and present," *Physics Today*, vol. 66, no. 4, p. 36, 2013.

[8] S. Reichelt, R. Häussler, G. Fütterer, and N. Leister, "Depth cues in human visual perception and their realization in 3D displays," in *Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV*, vol. 7690, pp. 92–103, International Society for Optics and Photonics, 2010.

[9] F. Speranza, W. J. Tam, R. Renaud, and N. Hur, "Effect of disparity and motion on visual comfort of stereoscopic images," in *Stereoscopic Displays and Virtual Reality Systems XIII*, vol. 6055, pp. 94–103, International Society for Optics and Photonics, 2006.

[10] M. T. Lambooij, W. A. IJsselsteijn, and I. Heynderickx, "Visual discomfort in stereoscopic displays: A review," in *Stereoscopic Displays and Virtual Reality Systems XIV*, vol. 6490, p. 64900I, International Society for Optics and Photonics, 2007.

[11] S.-N. Yang, T. Schlieski, B. Selmins, S. C. Cooper, R. A. Doherty, P. J. Corriveau, and J. E. Sheedy, "Stereoscopic viewing and reported perceived immersion and symptoms," *Optometry and Vision Science*, vol. 89, no. 7, pp. 1068–1080, 2012.

[12] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks, "The zone of comfort: Predicting visual discomfort with stereo displays," *Journal of Vision*, vol. 11, no. 8, pp. 11–11, 2011.

[13] H. Choi, J. Kim, S.-W. Cho, Y. Kim, J. B. Park, and B. Lee, "Three-dimensional-two-dimensional mixed display system using integral imaging with an active pinhole array on a liquid crystal panel," *Applied Optics*, vol. 47, no. 13, pp. 2207–2214, 2008.

[14] T. Peterka, R. L. Kooima, D. J. Sandin, A. Johnson, J. Leigh, and T. A. DeFanti, "Advances in the dynallax solid-state dynamic parallax barrier autostereoscopic visualization display system," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 3, pp. 487–499, 2008.

[15] Y. Kim, J. Kim, J.-M. Kang, J.-H. Jung, H. Choi, and B. Lee, "Point light source integral imaging with improved resolution and viewing angle by the use of electrically movable pinhole array," *Optics Express*, vol. 15, no. 26, pp. 18253–18267, 2007.

[16] N. Efrat, P. Didyk, M. Foshey, W. Matusik, and A. Levin, "Cinema 3D: Large scale automultiscopic display," *ACM Transactions on Graphics*, vol. 35, no. 4, p. 59, 2016.

[17] D. Lanman, M. Hirsch, Y. Kim, and R. Raskar, "Content-adaptive parallax barriers: Optimizing dual-layer 3D displays using low-rank light field factorization," *ACM Transactions on Graphics*, vol. 29, no. 6, p. 163, 2010.

[18] R. Narain, R. A. Albert, A. Bulbul, G. J. Ward, M. S. Banks, and J. F. O'Brien, "Optimal presentation of imagery with focus cues on multi-plane displays," *ACM Transactions on Graphics*, vol. 34, no. 4, p. 59, 2015.

[19] J. An, K. Won, Y. Kim, J.-Y. Hong, H. Kim, Y. Kim, H. Song, C. Choi, Y. Kim, J. Seo, *et al.*, "Slim-panel holographic video display," *Nature Communications*, vol. 11, p. 5568, 2020.

[20] P.-A. Blanche, "Holography, and the future of 3D display," *Light: Advanced Manufacturing*, vol. 2, no. 4, pp. 446–459, 2021.

[21] M. Martínez-Corral and B. Javidi, "Fundamentals of 3D imaging and displays: A tutorial on integral imaging, light-field, and plenoptic systems," *Advances in Optics and Photonics*, vol. 10, no. 3, pp. 512–566, 2018.

[22] B. Lee, J.-H. Park, and S.-W. Min, "Three-dimensional display and information processing based on integral imaging," in *Digital Holography and Three-Dimensional Display*, pp. 333–378, Springer, 2006.

[23] Y. Kim, K. Hong, and B. Lee, "Recent researches based on integral imaging display method," *3D Research*, vol. 1, no. 1, pp. 17–27, 2010.

[24] M. A. Alam, G. Baasantseren, M.-U. Erdenebat, N. Kim, and J.-H. Park, "Resolution enhancement of integral-imaging three-dimensional display using directional elemental image projection," *Journal of the Society for Information Display*, vol. 20, no. 4, pp. 221–227, 2012.

[25] H. Watanabe, N. Okaichi, H. Sasaki, and M. Kawakita, "Pixel-density and viewing-angle enhanced integral 3D display with parallel projection of multiple uhd elemental images," *Optics Express*, vol. 28, no. 17, pp. 24731–24746, 2020.

[26] Y. Kim, J.-H. Park, H. Choi, J. Kim, S.-W. Cho, and B. Lee, "Depth-enhanced three-dimensional integral imaging by use of multilayered display devices," *Applied Optics*, vol. 45, no. 18, pp. 4334–4343, 2006.

[27] J.-H. Park, K. Hong, and B. Lee, "Recent progress in three-dimensional information processing based on integral imaging," *Applied Optics*, vol. 48, no. 34, pp. H77–H94, 2009.

[28] T. Zhan, J. Xiong, J. Zou, and S.-T. Wu, "Multifocal displays: Review and prospect," *PhotoniX*, vol. 1, no. 1, pp. 1–31, 2020.

[29] D. Lanman and D. Luebke, "Near-eye light field displays," *ACM Transactions on Graphics*, vol. 32, no. 6, p. 220, 2013.

[30] A. Maimone and H. Fuchs, "Computational augmented reality eyeglasses," in *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 29–38, IEEE, 2013.

[31] F.-C. Huang, K. Chen, and G. Wetzstein, "The light field stereoscope: Immersive computer graphics via factored near-eye light field displays with focus cues," *ACM Transactions on Graphics*, vol. 34, no. 4, p. 60, 2015.

[32] K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks, "A stereo display prototype with multiple focal distances," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 804–813, 2004.

[33] D. Kim, S. Lee, S. Moon, J. Cho, Y. Jo, and B. Lee, "Hybrid multi-layer displays providing accommodation cues," *Optics Express*, vol. 26, no. 13, pp. 17170–17184, 2018.

[34] J.-H. R. Chang, B. V. Kumar, and A. C. Sankaranarayanan, "Towards multifocal displays with dense focal stacks," *ACM Transactions on Graphics*, vol. 37, no. 6, p. 198, 2018.

[35] K. Rathinavel, H. Wang, A. Blate, and H. Fuchs, "An extended depth-at-field volumetric near-eye augmented reality display," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 11, pp. 2857–2866, 2018.

[36] C. Chang, K. Bang, G. Wetzstein, B. Lee, and L. Gao, "Toward the next-generation VR/AR optics: A review of holographic near-eye displays

from a human-centric perspective," *Optica*, vol. 7, no. 11, pp. 1563–1578, 2020.

[37] M. Yamaguchi, "Light-field and holographic three-dimensional displays," *JOSA A*, vol. 33, no. 12, pp. 2348–2364, 2016.

[38] H. Zhang, J. Xie, J. Liu, and Y. Wang, "Elimination of a zero-order beam induced by a pixelated spatial light modulator for holographic projection," *Applied Optics*, vol. 48, no. 30, pp. 5834–5841, 2009.

[39] H. Sasaki, K. Yamamoto, K. Wakunami, Y. Ichihashi, R. Oi, and T. Senoh, "Large size three-dimensional video by electronic holography using multiple spatial light modulators," *Scientific Reports*, vol. 4, no. 1, p. 6177, 2014.

[40] A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," *ACM Transactions on Graphics*, vol. 36, no. 4, p. 85, 2017.

[41] C. Jang, K. Bang, G. Li, and B. Lee, "Holographic near-eye display with expanded eye-box," *ACM Transactions on Graphics*, vol. 37, no. 6, p. 195, 2018.

[42] B. Mendiburu, *3D movie making: Stereoscopic digital cinema from script to screen*. Focal press, 2012.

[43] Y. Nojiri, H. Yamanoue, A. Hanazato, M. Emoto, and F. Okano, "Visual comfort/discomfort and visual fatigue caused by stereoscopic HDTV viewing," in *Stereoscopic Displays and Virtual Reality Systems XI*, vol. 5291, pp. 303–314, International Society for Optics and Photonics, 2004.

[44] F. L. Kooi and A. Toet, "Visual comfort of binocular and 3D displays," *Displays*, vol. 25, no. 2-3, pp. 99–108, 2004.

[45] K. N. Ogle, *Researches in binocular vision*. WB Saunders, 1950.

[46] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of Vision*, vol. 8, no. 3, pp. 33–33, 2008.

[47] A. S. Percival, "The relation of convergence to accommodation and its practical bearing," *Ophthal. Rev.*, vol. 11, pp. 313–328, 1892.

[48] K. J. MacKenzie, D. M. Hoffman, and S. J. Watt, "Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control," *Journal of Vision*, vol. 10, no. 8, pp. 22–22, 2010.

[49] M. Mauderer, S. Conte, M. A. Nacenta, and D. Vishwanath, "Depth perception with gaze-contingent depth of field," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 217–226, ACM, 2014.

[50] S. E. Palmer, *Vision science: Photons to phenomenology*. MIT press, 1999.

[51] B. Wang and K. J. Ciuffreda, "Depth-of-focus of the human eye in the near retinal periphery," *Vision Research*, vol. 44, no. 11, pp. 1115–1125, 2004.

[52] Y. Takaki and N. Nago, "Multi-projection of lenticular displays to construct a 256-view super multi-view display," *Optics Express*, vol. 18, no. 9, pp. 8824–8835, 2010.

[53] S.-g. Park, J. Yeom, Y. Jeong, N. Chen, J.-Y. Hong, and B. Lee, "Recent issues on integral imaging and its applications," *Journal of Information Display*, vol. 15, no. 1, pp. 37–46, 2014.

[54] K. Wakunami, P.-Y. Hsieh, R. Oi, T. Senoh, H. Sasaki, Y. Ichihashi, M. Okui, Y.-P. Huang, and K. Yamamoto, "Projection-type see-through holographic three-dimensional display," *Nature Communications*, vol. 7, p. 12954, 2016.

[55] M. Kalloniatis and C. Luu, *Temporal resolution*. University of Utah Health Sciences Center, 2007.

[56] G. T. Herman and A. Kuba, *Discrete tomography: Foundations, algorithms, and applications*. Springer Science & Business Media, 2012.

[57] L. Larmore, *Introduction to photographic principles*. Dover Publications New York, 2d ed. ed., 1965.

[58] Z. Gao, A. Hwang, G. Zhai, and E. Peli, "Correcting geometric distortions in stereoscopic 3D imaging," *PloS one*, vol. 13, no. 10, p. e0205032, 2018.

[59] O. Mercier, Y. Sulai, K. Mackenzie, M. Zannoli, J. Hillis, D. Nowrouzezahrai, and D. Lanman, "Fast gaze-contingent optimal decompositions for multifocal displays," *ACM Transactions on Graphics*, vol. 36, no. 6, p. 237, 2017.

[60] L. Xiao, A. Kaplanyan, A. Fix, M. Chapman, and D. Lanman, "Deepfocus: Learned image synthesis for computational displays," *ACM Transactions on Graphics*, vol. 37, no. 6, 2018.

[61] A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique: A superior implementation of the art algorithm," *Ultrasonic imaging*, vol. 6, no. 1, pp. 81–94, 1984.

[62] K. Ukai and P. A. Howarth, "Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations," *Displays*, vol. 29, no. 2, pp. 106–116, 2008.

[63] S. Choi, S. Lee, Y. Jo, D. Yoo, D. Kim, and B. Lee, "Optimal binary representation via non-convex optimization on tomographic displays," *Optics Express*, vol. 27, no. 17, pp. 24362–24381, 2019.

[64] R. Stevens, D. Rhodes, A. Hasnain, and P.-Y. Laffont, "Varifocal technologies providing prescription and vac mitigation in HMDs using alvarez lenses," in *Digital Optics for Immersive Displays*, vol. 10676, p. 106760J, International Society for Optics and Photonics, 2018.

[65] S. Lee, S. Lee, D. Kim, and B. Lee, "Distortion corrected tomographic near-eye displays using light field optimization," *Optics Express*, vol. 29, no. 17, pp. 27573–27586, 2021.

[66] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black, "A naturalistic open source movie for optical flow evaluation," in *European Conference on Computer Vision*, pp. 611–625, Springer, 2012.

[67] R. H. Spector, "Visual fields," 2011.

[68] J. Hou, Y. Nam, W. Peng, and K. M. Lee, "Effects of screen size, viewing angle, and players' immersion tendencies on game experience," *Computers in Human Behavior*, vol. 28, no. 2, pp. 617–623, 2012.

[69] E. Lantz, "Future directions in visual display systems," *ACM SIGGRAPH Computer Graphics*, vol. 31, no. 2, pp. 38–42, 1997.

[70] D. J. Marmor and M. F. Marmor, "Simulating vision with and without macular disease," *Archives of ophthalmology*, vol. 128, no. 1, pp. 117–125, 2010.

[71] F. B. Neves and J. Page, "Head and/or eye tracked optically blended display system," 1987. US Patent 4,634,384.

[72] S. Kongsilp and M. N. Dailey, "Motion parallax from head movement enhances stereoscopic displays by improving presence and decreasing visual fatigue," *Displays*, vol. 49, pp. 72–79, 2017.

[73] H. Akiduki, S. Nishiike, H. Watanabe, K. Matsuoka, T. Kubo, and N. Takeda, "Visual-vestibular conflict induced by virtual reality in humans," *Neuroscience Letters*, vol. 340, no. 3, pp. 197–200, 2003.

[74] R. H. So, A. Ho, and W. Lo, "A metric to quantify virtual scene movement for the study of cybersickness: Definition, implementation, and verification," *Presence*, vol. 10, no. 2, pp. 193–215, 2001.

[75] K. Hirota, S. Ebisawa, T. Amemiya, and Y. Ikei, "A system for creating the content for a multi-sensory theater," in *International Conference on Virtual and Mixed Reality*, pp. 151–157, Springer, 2011.

[76] K. Osmanis, G. Valters, R. Zabels, U. Gertners, I. Osmanis, L. Kalnins, U. Kandere, and A. Ozols, "Advanced multiplanar volumetric 3D display," in *Emerging Liquid Crystal Technologies XIII*, vol. 10555, p. 1055510, International Society for Optics and Photonics, 2018.

[77] S. Lee, C. Jang, S. Moon, J. Cho, and B. Lee, "Additive light field displays: Realization of augmented reality with holographic optical elements," *ACM Transactions on Graphics*, vol. 35, no. 4, p. 60, 2016.

[78] M. Hirsch, G. Wetzstein, and R. Raskar, "A compressive light field projection system," *ACM Transactions on Graphics*, vol. 33, no. 4, p. 58, 2014.

[79] G. Lippmann, "Epreuves reversibles photographies integrals," *Comptes-Rendus Academie des Sciences*, vol. 146, pp. 446–451, 1908.

[80] G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar, "Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays," *ACM Transactions on Graphics*, vol. 30, no. 4, p. 95, 2011.

[81] K. Suzuki, Y. Fukano, and H. Oku, "1000-volume/s high-speed volumetric display for high-speed HMD," *Optics Express*, vol. 28, no. 20, pp. 29455–29468, 2020.

[82] S. Lee, Y. Jo, D. Yoo, J. Cho, D. Lee, and B. Lee, "Tomographic near-eye displays," *Nature Communications*, vol. 10, no. 1, p. 2497, 2019.

[83] D. Yoo, S. Lee, Y. Jo, J. Cho, S. Choi, and B. Lee, "Volumetric head-mounted display with locally adaptive focal blocks," *IEEE Transactions on Visualization and Computer Graphics*, 2020.

[84] S.-G. Park, B.-S. Song, and S.-W. Min, "Analysis of image visibility in projection-type integral imaging system without diffuser," *Journal of the Optical Society of Korea*, vol. 14, no. 2, pp. 121–126, 2010.

[85] Y. Lou, J. Hu, F. Wu, and A. Chen, "Bandwidth-enhanced depth priority integral imaging using a band-limited diffusing illumination technique," *Applied Optics*, vol. 58, no. 5, pp. A234–A241, 2019.

[86] M. S. Brennesholtz and E. H. Stupp, *Projection displays*. John Wiley & Sons, 2008.

[87] C. Zuo, Q. Chen, W. Qu, and A. Asundi, "High-speed transport-of-intensity phase microscopy with an electrically tunable lens," *Optics Express*, vol. 21, no. 20, pp. 24060–24075, 2013.

[88] G. Park, J.-H. Jung, K. Hong, Y. Kim, Y.-H. Kim, S.-W. Min, and B. Lee, "Multi-viewer tracking integral imaging system and its viewing zone analysis," *Optics Express*, vol. 17, no. 20, pp. 17895–17908, 2009.

[89] L. N. Thibos, R. A. Applegate, J. T. Schwiegerling, and R. Webb, "Standards for reporting the optical aberrations of eyes," in *Vision science and its applications*, p. SuC1, Optical Society of America, 2000.

[90] D. Wu, Q.-D. Chen, L.-G. Niu, J. Jiao, H. Xia, J.-F. Song, and H.-B. Sun, "100% fill-factor aspheric microlens arrays (amla) with sub-20-nm precision," *IEEE Photonics Technology Letters*, vol. 21, no. 20, pp. 1535–1537, 2009.

[91] B. Lee, D. Yoo, J. Jeong, S. Lee, D. Lee, and B. Lee, "Wide-angle speckleless dmd holographic display using structured illumination with temporal multiplexing," *Optics Letters*, vol. 45, no. 8, pp. 2148–2151, 2020.

[92] S. Lee, D. Kim, S.-W. Nam, and B. Lee, "Speckle reduced holographic displays using tomographic synthesis," *Optics Letters*, vol. 45, no. 17, pp. 4686–4689, 2020.

[93] H. Yu, K. Lee, J. Park, and Y. Park, "Ultrahigh-definition dynamic 3D holographic display by active control of volume speckle fields," *Nature Photonics*, vol. 11, no. 3, pp. 186–192, 2017.

[94] S.-H. Baek, E. Tseng, A. Maimone, N. Matsuda, G. Kuo, Q. Fu, W. Heidrich, D. Lanman, and F. Heide, "Neural étendue expander for ultra-wide-angle high-fidelity holographic display," *arXiv preprint arXiv:2109.08123*, 2021.

[95] K. Bang, C. Jang, and B. Lee, "Compact noise-filtering volume gratings for holographic displays," *Optics Letters*, vol. 44, no. 9, pp. 2133–2136, 2019.

[96] J. W. Goodman, *Introduction to Fourier Optics*. McGraw-Hill, 1968.

[97] D. Lee, K. Bang, S.-W. Nam, B. Lee, D. Kim, and B. Lee, "Expanding energy envelope in holographic display via mutually coherent multi-directional illumination," *Scientific Reports*, vol. 12, no. 1, pp. 1–13, 2022.

[98] K. Matsushima and T. Shimobaba, "Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields," *Optics Express*, vol. 17, no. 22, pp. 19662–19673, 2009.

[99] E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 126–135, 2017.

[100] M. Gopakumar, J. Kim, S. Choi, Y. Peng, and G. Wetzstein, "Unfiltered holography: Optimizing high diffraction orders without optical fil-

tering for compact holographic displays," *Optics Letters*, vol. 46, no. 23, pp. 5822–5825, 2021.

[101] J. Lee, J. Jeong, J. Cho, D. Yoo, B. Lee, and B. Lee, "Deep neural network for multi-depth hologram generation and its training strategy," *Optics Express*, vol. 28, no. 18, pp. 27137–27154, 2020.

[102] C. Chen, B. Lee, N.-N. Li, M. Chae, D. Wang, Q.-H. Wang, and B. Lee, "Multi-depth hologram generation using stochastic gradient descent algorithm with complex loss function," *Optics Express*, vol. 29, no. 10, pp. 15089–15103, 2021.

[103] A. Lizana, N. Martín, M. Estapé, E. Fernández, I. Moreno, A. Márquez, C. Iemmi, J. Campos, and M. Yzuel, "Influence of the incident angle in the performance of liquid crystal on silicon displays," *Optics Express*, vol. 17, no. 10, pp. 8491–8505, 2009.

[104] L. Shi, B. Li, C. Kim, P. Kellnhofer, and W. Matusik, "Towards real-time photorealistic 3d holography with deep neural networks," *Nature*, vol. 591, no. 7849, pp. 234–239, 2021.

Appendix

Portions of the work discussed in this dissertation were also presented in the following publications:

[Chapter 2] Y. Jo, S. Lee, D. Yoo, S. Choi, D. Kim, and B. Lee, "Tomographic projector: Large scale volumetric display with uniform viewing experiences," *ACM Transactions on Graphics*, vol. 38, no. 6, article 215, 2019.

[Chapter 3] Y. Jo, K. Bang, D. Yoo, B. Lee, and B. Lee, "Ultrahigh-definition volumetric light field projection," *Optics Letters*, vol. 46, no. 17 pp. 4212-4215, 2021.

[Chapter 4] Y. Jo, D. Yoo, D. Lee, M. Kim, and B. Lee, "Multi-illumination 3D holographic display using a binary mask," *Optics Letters*, vol. 47, no. 10, pp. 2482-2485, 2022.

초 록

디스플레이가 개발된 이후로 대중들은 더욱 실제와 같은 영상을 원해왔다. 기술의 발달로 고화질의 대화면 디스플레이가 가능해진 지금, 보다 몰입감 있는 경험을 향해 3차원 디스플레이 분야가 발전해나가고 있다. 그러나, 기존 2차원 화면에서 3차원 영상을 표현하는 일은 전보다 더 많은 정보를 필요로 한다. 이는 현재 디스플레이의 가용한 정보 수준을 훨씬 넘어서기 때문에 그동안 여러 형태의 시스템적인 방법론들이 해결을 위해 개발되어 왔다. 그런데도 3차원 디스플레이 기술들은 저마다 본질적인 한계를 갖게 되었고, 결과적으로 시각적 피로, 줄어든 해상도, 좁은 시청 각도와 같은 시청 경험을 저해하는 요인을 만들었다. 결국 완전한 몰입감의 진정한 3차원 디스플레이를 위해서는 하드웨어의 발전뿐만 아니라, 제한된 정보량에 기인한 한계점을 해소할 수 있는 광학적으로 새로운 형태의 방법론이 개발되어야 한다.

본 학위논문에서 저자는 기존 개발된 방법론을 적극적으로 활용하여 새로운 관점에서 이들을 조합하는 것으로 장점은 더하면서 단점은 보완하는 연구를 진행한다. 그 과정에서 기존의 한계를 극복하고 몰입 경험을 향상할 수 있는 세 가지의 새로운 광학 시스템을 소개한다. 첫 번째로, 안경형 디스플레이 시스템에서 다초점면 기술을 도입하여 알맞은 단안 초점 단서를 제공하는 토모그래픽 프로젝션 시스템을 제안한다. 이를 통해 발생 가능한 시각적 피로를 완화하고, 3차원 물체의 깊이 범위를 바로 눈앞까지 확장하면서 영화관과 같은 공간에서 더욱 몰입감 있는 시청 경험을 만들어낼 수 있다. 두 번째로, 안경 없는 간편한 시청 환경에서도 높은 성능의 3차원 영상을 만들어낼 수 있는 라이트 필드 프로젝션 시스템을 제안한다. 이때 집적 영상과 다초점면 디스플레이 간의 다른 정보를 광학적으로 연장하여 이전에 없던 초고해상도의 라이트 필드를 달성한다. 세 번째로, 저자는 자유롭게 3차원 물체 표현이 가능한 홀로그래픽 디스플레이에서 한계점인 대역폭을 확장하였다. 여기서 다중

광원 방법을 도입하여 시청 각도를 넓히면서 스페클 노이즈를 경감시켰으며, 동시에 이진 마스크로 더 많은 정보를 표현할 수 있는 자유도를 부여하였다. 결과적으로 최적화를 통해 정보를 9배 압축하여, 효율적으로 홀로그램을 구현할 수 있는 홀로그래픽 시스템을 제시한다.

눈으로 구분할 수 없는 초고해상도 화면부터, 건물을 수놓은 초대형 디스플레이까지 상상 속 기술들이 이미 현실로 다가왔다. 메타버스로 가상 세계가 점차 현실에 가까워지는 지금, 2차원을 넘어 실제 같은 몰입감을 줄 수 있는 완전한 3차원 디스플레이 또한 언젠가 반드시 개발될 것이라고 믿는다. 그 과정에 있어서 이 학위논문연구가 실용적인 새로운 관점들을 제시하고 더 나아가 3차원 디스플레이의 발전에 기여하기를 희망한다.

주요어: 3차원 디스플레이, 체적 디스플레이, 안경형 디스플레이, 라이트 필드 디스플레이, 홀로그래픽 디스플레이, 대화면 디스플레이, 몰입형 경험

학번: 2017-27135