Performance enhancement and size reduction of near-eye display

Cheng, Qijia

2020

Cheng, Q. (2020). Performance enhancement and size reduction of near-eye display.
Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/144165>

<https://doi.org/10.32657/10356/144165>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0
International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 19:05:10 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**PERFORMANCE ENHANCEMENT AND SIZE
REDUCTION OF NEAR-EYE DISPLAY**

CHENG QIJIA

SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING

2020

**PERFORMANCE ENHANCEMENT AND SIZE
REDUCTION OF NEAR-EYE DISPLAY**

CHENG QIJIA

SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2020

Statement of Originality

I certify that the work embodied in this thesis is the result of original research, is free of plagiarized materials, and has not been submitted for a higher degree to any other University or Institution.

16 Aug 2019

.....

Date

.....

Cheng Qijia

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

Aug.18, 2019

.....

Date

.....

Assoc. Prof. Zheng Yuanjin

Authorship Attribution Statement

This thesis contains material from 3 paper(s) published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

Chapter 3 is published as:

Qijia Cheng, Weitao Song, Feng Lin, Yue Liu, Yongtian Wang, and Yuanjin Zheng. "Resolution enhancement of near-eye displays by overlapping images." *Optics Communications* 458 (2020): 124723.

The contributions of the co-authors are as follows:

- Assoc. Prof. Zheng Yuanjin provided the initial project direction and edited the manuscript.
- Assoc. Prof. Lin Feng advised on the algorithm and edited manuscript.
- Dr. Song Weitao provided guidance on the experiment procedure, assisted the data analysis, and revised manuscript.
- I prepared the manuscript drafts.
- Dr. Liu Yue and Prof. Wang Yongtian revised the manuscript and provided suggestions.
- Dr. Song Weitao and I designed the experiment procedure. All the experiments were performed by me at the School of Electrical and Electronic Engineering, NTU. I analyzed the data.
- All system designs and setup were conducted by me.

Chapter 4 is published as:

Zhenfeng Zhuang, Qijia Cheng, Phil Surman, Yuanjin Zheng, and Xiao Wei Sun. "A compact and lightweight off-axis lightguide prism in near to eye display." *Optics Communications* 393 (2017): 143-151.

The contributions of the co-authors are as follows:

- Assoc. Prof. Zheng Yuanjin and Prof. Sun Xiao Wei provided the initial project direction and edited the manuscript.

- Dr. Zhuang Zhenfeng provided initial optical design.
- I performed the optimization and fabrication of the optical elements.
- Dr. Zhuang Zhenfeng and I wrote the manuscript and revised.
- All the experiments were performed by me at the School of Electrical and Electronic Engineering, NTU.
- Dr. Phil Surman revised the manuscript.

Chapter 5 is published as:

Weitao Song, Qijia Cheng, Phil Surman, Yue Liu, Yuanjin Zheng, Zhiping Lin, and Yongtian Wang. "Design of a light-field near-eye display using random pinholes." Optics Express 27, no. 17 (2019): 23763-23774.

The contributions of the co-authors are as follows:

- Assoc. Prof. Zheng Yuanjin and Assoc. Prof. Lin Zhiping provided the initial project direction and edited the manuscript.
- Dr. Song Weitao and I wrote the initial manuscript draft and revised the manuscripts.
- The manuscript was revised by Dr. Liu Yue and Prof. Wang Yongtian.
- Dr. Song Weitao and I designed the experiment procedure. I performed experiments at the School of Electrical and Electronic Engineering, NTU.
- System design, fabrication, and setup were conducted by me.
- Dr. Phil Surman revised the manuscript.

16 Aug 2019

Date

Cheng Qijia

Acknowledgements

I would like to take this chance to thank all the people that helped me during my Ph.D. program.

First of all, I would like to thank my supervisor, Professor Zheng Yuanjin, for giving me this opportunity to carry this state-of-the-art research on virtual reality and display technology. Thanks to his patience and forgiveness, I gradually picked up the required skills for my research project. I also thank my co-supervisor, Professor Lin Feng for guiding me through the project and answering my puzzles. His vast experience in computer science provides a different point-of-view.

I am thankful to Dr. Song Weitao for his assistance on the implementation of the research. His knowledge of optics and algorithm is proved to be helpful from time to time. I would like to thank Professor Sun Xiao Wei for his guidance in the first year of my Ph.D. From him, I learned lots of technopreneurship. I am also grateful to Professor Cuong Dang to let me work in his group in the last few months of my study and provide me future opportunities.

I would like to thank all the members at the Advanced Display Lab, NTU, including Dr. Zhang Xiangyu, Dr. Zhuang Zhenfeng, Dr. Phil Surman, Mr. Zhang Lei, Dr. Wang Shizheng, and Ms. Wang Hongjuan. Great thanks to Dr. Sun Mingyu for helping me on the procedure of submitting and preparing the thesis.

I am also grateful to my family for their constant support, without which my study cannot be accomplished.

Last but not least, I also thank Singapore National Research Foundation NRF-CRP11-2012-01 for sponsoring me through my Ph.D. program at Nanyang Technological University.

Table of Contents

Statement of Originality	III
Supervisor Declaration Statement.....	IV
Authorship Attribution Statement.....	V
Acknowledgements	VII
Table of Contents.....	IX
Summary	XII
List of Figures	XIV
List of Tables.....	XXII
Chapter 1 Introduction	1
1.1 Near-eye Display (NED).....	1
1.1.1 Virtual Reality	1
1.1.2 Near-eye Display	6
1.2 Motivation and Objectives	8
1.3 Major contributions of the Thesis.....	9
1.4 Organization of the Thesis	10
Chapter 2 Literature review.....	13
2.1 Resolution enhancement.....	13
2.1.1 Resolution limitation of near-eye display	13
2.1.2 Resolution enhancement methods.....	16
2.1.3. Display super-sampling and super-resolution	20
2.1.4. Display calibration.....	22
2.2 Near-eye display with free-form optical design	23
2.3 Light field near-eye display.....	24
Chapter 3 Resolution enhancement utilizing overlapping multiple display panels	28

3.1	Introduction	28
3.2	Method	30
3.2.1	System set-up	30
3.2.2	Image generation	37
3.2.3	Image formation from two overlapping images	41
3.2.4	Generic computational rendering algorithm for two-dimensional image and multiple display panels	45
3.2.5	Display calibration	47
3.3	Experiment setup	57
3.4	Results	62
3.5	Conclusion	75
Chapter 4 Compact off-axis near-eye display based on free-form optics		76
4.1	Introduction	76
4.2	Method	76
4.2.1	System description	76
4.2.2	Method of controlling clearances	77
4.2.3	Optimization of surface	81
4.3	Results	83
4.3.1	Specifications	83
4.3.2	Optical layout and performance	83
4.3.3	Analysis and performance evaluation	88
4.4	Conclusion	92
Chapter 5 Random pinhole light field near-eye display		93
5.1	Introduction	93
5.2	Method	95
5.2.1	System set-up	95
5.2.2	Rendering method	96
5.2.3	Viewing zone	98
5.3	Results	100
5.4	Conclusion	114
Chapter 6 Conclusion and Recommendations		116
6.1	Conclusion	116
6.2	Recommendations for further research	117

Author's Publications.....	119
Bibliography	121

Summary

Near-eye display (NED) paved its way in a variety of applications of modern virtual reality (VR). However, current NEDs have lots of limitations, which lead to compromised user experience and strongly limit their potential in real-world scenarios. Two of the most noticeable features that demanding improvements are resolution and size. On the one hand, applications of modern VR devices introduce more detailed textures and increased field-of-view (FoV), altogether create excessive demand for resolution. On the other hand, more mobile application scenarios of virtual reality and augmented reality (AR) bring the requirement of size reduction. The trend of developing high-resolution NED with small form-factor is with no doubt. During the last several decades, attempts were made to improve the resolution of NED. However, the utilization of eye-tracking and optical tiling adds a fairly amount of complexity to the NED system, which in turn makes the device even bulkier. Thus, a compact near-eye display device that can deliver excellent image quality is still in great demand by society. In this thesis, theory and methods are demonstrated to improve key performance factors in modern NED, including resolution, depth perception, and size.

In Chapter 3, to solve the main conflict between resolution and device size, a versatile resolution enhancement method for NED is implemented by overlapping multiple display panel images. Based on display super-resolution, this overlapping method effectively increases the perceived resolution without optical complexity. By investigating the algorithm theory of image generation and calibration, the perceived resolution can be enhanced and overcome the physical resolution of a single display panel. A NED prototype based on a semi-transparent mirror is designed and fabricated. The experiment demonstrated decent results of resolution enhancement. Compare with the conventional method of display tiling, my method offers an equivalent level of performance yet reduces system complexity, brings more flexibility to the design of NED.

Free-form optics enable innovative ways of implementing compact size NED. In Chapter 4, a design routine of off-axis near-eye display system is implemented. A compact near-eye display system described by x-y polynomial function is developed with a 4mm diameter exit pupil, a 21.5° diagonal full field-of-view (FoV) and 23 mm eye relief. Freeform surfaces are used to effectively reduce the size of NED while correcting the off-axis aberration. The designed system yields good optical performance yet very compact in terms of size, compared to traditional NED. A prototype is demonstrated with components available from the market. The prototype effectively indicated the feasibility of the design as a compact NED.

Light field receives great research attention in recent years as a feasible way to solve convergence-accommodation conflict in NED and stereoscopic display. Based on either microlens or pinhole, light field NED demonstrates an effective size reduction on the longitudinal direction of the device. In chapter 5, a light field NED with random pinhole design is demonstrated with unique viewing zone and optical simplicity.

Further research is recommended in chapter 6 to combine the individual technique as a hybrid device, where multiple light field display module can be used for super-resolution, with the design flexibility of free-form optics.

My research can act as guidance in the design of future NED to reduce device size, enhance resolution, and incorporate the depth sense of current virtual reality headset, thus further improve virtual reality user experience.

List of Figures

Figure 1.1 Peep show devices in 19th-century.

Figure 1.2 Virtual reality industry infrastructure.

Figure 1.3 Different kinds of near-eye displays.

Figure 2.1 Demonstration of resolution / field-of-view invariance. (a) High resolution, small field-of-view. (b) Large field-of-view, low resolution.

Figure 2.2 Demonstration of resolution / field-of-view invariance. (a) H is the size of the display, f is the effective focal length of the lens, and θ is the field-of-view. (b) H is the physical pixel size of the display, f is the effective focal length of the lens, and θ is the angular resolution in the image space.

Figure 2.3 Basic working principle of high-resolution area of interest. A low-resolution large FoV display is used as background. A high-resolution but small FoV display is superimposed on the background, acting as the “area of interest”.

Figure 2.4 Partial binocular overlap method to increase the FoV without sacrificing the resolution. (a) Normal binocular vision with full overlap. (b) Divergent partial binocular overlap. The viewing cone pointing outwards.

Figure 2.5 Dichoptic area of interest method. Left viewing channel is high-resolution but small FoV image. Right viewing channel is low-resolution but large FoV image. Two viewing channels merge together to form a high-resolution binocular image with large FoV.

Figure 2.6 Accommodation and convergence conflict. User’s eye focus on the display panel, gives accommodation distance. The binocular convergence distance in this case is further away. The difference between accommodation and convergence distance will cause discomfort and fatigue.

Figure 3.1 Schematic drawing of the overlapping near-eye display system with two display panels.

Figure 3.2 Positions of two display panels and the source image.

Figure 3.3 Low-resolution sub-frame images generate to display on (a) Display panel A, (b) Display Panel B.

Figure 3.4 (a) Simulated overlapping result of two display panels, (b) Zoom-in of source image at the native resolution of a single display panel. (c) Zoom-in of the result of overlapping.

Figure 3.5 Operation procedure of display overlapping. Start from a high-resolution image, sub-frame low-resolution images are generated and optimized to get the best results of overlapping. In the NED system, perceived image is formed by physically overlapping the sub-frame images.

Figure 3.6 (a) Analog one-dimensional signal, (b) Corresponding signal in greyscale representation, (c) Corresponding function of the signal in the frequency domain.

Figure 3.7 (a) Sampling signal, which is a comb function with period T . (b) Pixel format of the one-dimensional display. (c) sampling function in the frequency domain is also a comb function with period $f_s = \frac{1}{T}$. (d) Sampled signal $C(f)$ is a replication of the original spectrum in the frequency domain.

Figure 3.8 Kernel function of the display device.

Figure 3.9 (a) Reconstructed signal $R(f)$ from sampling, (b) Corresponding greyscale image of the reconstructed signal, (c) Signal reconstructed by kernel function in the frequency domain.

Figure 3.10 Explanation of overlapping mechanism for the special case when pixel offset value $l = \frac{T}{2}$. (a) The frequency spectrum of the continuous signal. (b) The frequency spectrum of the sampling signal, the period of the comb function is $2f_s$. (c) Frequency spectrum of the convolution signal. Yellow circle: no contamination from adjacent replica at a frequency f_s during overlapping.

Figure 3.11 Explanation of overlapping mechanism for general cases where pixel offset value $l \neq \frac{T}{2}$. (a) The frequency spectrum of the continuous one-dimensional signal. (b) The frequency spectrum of the sampling signal. (c) The frequency spectrum of the convolution signal. Yellow circle: contamination from adjacent replica at the frequency f_s during overlapping.

Figure 3.12 Overlapping simulation of two-dimensional images with two display panels. (a) one image with low resolution. (b) Overlapping result of two images at the same resolution with figure (a) with shift offset of one-fourth of pixel width. (c) The overlapping result with shift offset of half-pixel width. (d), (e) and (f) are the magnified images of the corresponding image (a), (b) and (c), respectively.

Figure 3.13 (a) Multi-frequency phase unwrapping strategy diagram. (b) Phase unwrapping results from intermediate phase λ_{12} and λ_{23} to the

final phase λ_{123} . The wavelength of the final phase should be the width of the display panel.

Figure 3.14 Sinusoidal stripe patterns used for phase unwrapping on horizontal direction, with wavelength λ_1 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.15 Sinusoidal stripe patterns used for phase unwrapping on horizontal direction, with wavelength λ_2 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.16 Sinusoidal stripe patterns used for phase unwrapping on horizontal direction, with wavelength λ_3 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.17 Sinusoidal stripe patterns used for phase unwrapping on vertical direction, with wavelength λ_1 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.18 Sinusoidal stripe patterns used for phase unwrapping on vertical direction, with wavelength λ_2 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.19 Sinusoidal stripe patterns used for phase unwrapping on vertical direction, with wavelength λ_3 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$. (e) Full black and (f) full white pattern to normalize the greyscale value.

Figure 3.20 Schematic diagram of a display calibration setup. By capturing the calibration pattern, a mapping can be established between camera image sensor pixel and display panel pixel.

Figure 3.21 Schematic design of dual-display NED.

Figure 3.22 The near-eye display prototype CAD design appearance.

Figure 3.23 The near-eye display prototype CAD design in a cross-section view. The red rectangle stands for the position of display panel 2. The yellow rectangle is the position of display panel 1. The light blue rectangle stands for the position of semi-reflective mirror. The deep blue oval represents the eyepiece. Green area is the driving circuit box. The object distance between the display panel and eyepiece is 95mm, result in an image distance at 1900mm.

Figure 3.24 NED prototype and experimental set-up. Digital compact camera Nikon Coolpix A was used for calibration and image capturing.

Figure 3.25 Complete captured calibration photo set consists of 52 photos of sinusoidal intensity patterns. (a) 4 step phase shift patterns with wavelength λ_1 on horizontal direction for display panel 1. (b) 4 step phase shift patterns with wavelength λ_2 on horizontal direction for display panel 1. (c) 4 step phase shift patterns with wavelength λ_3 on horizontal direction for display panel 1. (d) 4 step phase shift patterns with wavelength λ_1 on vertical direction for display panel 1. (e) 4 step phase shift patterns with wavelength λ_2 on vertical direction for display panel 1. (f) 4 step phase shift patterns with wavelength λ_3 on vertical direction for display panel 1. (g) Full black and full white for display panel 1. (h) 4 step phase shift patterns with wavelength λ_1 on horizontal direction for display panel 2. (i) 4 step phase shift patterns with wavelength λ_2 on horizontal direction for display panel 2. (j) 4 step phase shift patterns with wavelength λ_3 on horizontal direction for display panel 2. (k) 4 step phase shift patterns with wavelength λ_1 on vertical direction for display panel 2. (l) 4 step phase shift patterns with wavelength λ_2 on vertical direction for display panel 2. (m) 4 step phase shift patterns with wavelength λ_3 on vertical direction for display panel 2. (n) Full black and full white for display panel 2.

Figure 3.26 Display calibration result. Red color grid shows the position of display panel 1. Blue color grid shows the position of display panel 2. Relative positions of two display panels include translational and rotational offset.

Figure 3.27 High-resolution source image from a 16-mega-pixel DSLR camera with a high-quality lens. Image is cropped to the red area to get more details of the building.

Figure 3.28 Decomposed sub-frame low-resolution image at the native resolution 400x480 for each viewing channel of the display panel. (a) Image on rear display panel A. (b) Image on top display panel B.

Figure 3.29 Decomposed sub-frame low-resolution images shown on display panels with the native resolution 400x480 for one eye. (a) The image displayed on the rear display panel, transmit through the half-mirror combiner. (b) The image displayed on the top display panel, reflected by the half-mirror combiner.

Figure 3.30 (a) The overlapping of two low-resolution display panel images. (b) The down-sized image displayed by the rear display panel at the native display resolution 400x480.

Figure 3.31 (a) Zoom-in of the overlapping result of two low-resolution display panel images. (b) Zoom-in of the down-sized image displayed by the rear display panel at the native display resolution 400x480.

Figure 3.32 Video test results of overlapping NED. Decomposed low-resolution images for display panels: (a) (e) Image displayed on the rear display panel, transmit through the half-mirror combiner. (b) (f) The image displayed on the top display panel, reflected by the half-mirror combiner. (b) (g) The overlapping of two low-resolution display panel images. (d) (h) The down-sized image displayed by the rear display panel at the native display resolution 400x480.

Figure 3.33 Resolution chart test result of overlapping NED. (a) Image displayed on single display panel. (b) Zoom-in of red square area in figure (a). (c) Zoom-in of green square area in (d). (d) Image displayed by overlapping 2 display panels.

Figure 4.1 Cross-section of the proposed NED using two lightguide prisms combined with a doublet lens.

Figure 4.2 Determination of the intermediate image between two surfaces.

Figure 4.3 Determination of the incident angle of S1 and S3 to meet TIR conditions.

Figure 4.4 Imaging prism clearance.

Figure 4.5 Comparisons among different types of surface representation.

Figure 4.6 (a) Optical layout of the NED system; (b) MTF plot of the designed system with a 4 mm exit pupil.

Figure 4.7 (a) Definition of optical distortion; (b) Distortion grid of the designed NED optical system compared with real and paraxial rays.

Figure 4.8 (a) See-through mode of the NED system; (b) Distortion grid of the designed NED optical system caused by the imaging lightguide prism.

Figure 4.9 Design of the compensation prism with backward ray tracing. The distance between the projected image to the user is 1500 mm in this implementation. The human eye model is illustrated with single ideal lens with focal length 22mm. The ray is traced from image plane (i.e. retina) to the pupil,

then through the proposed system to the object plane. The see-through system has a field-of-view 24.8° .

Figure 4.10 Design layout of the compensation prism (a) in the yz-plane; (b) in the xz-plane.

Figure 4.11 Distortion grid of the designed compensation prism compared with real and paraxial rays.

Figure 4.12 The exit angle at the exit pupil versus incident angle of TIR (a) S1 surface; (b) S3 surface. The black, red, and blue lines indicate the exit angle at the center, top marginal, and bottom marginal of the exit pupil.

Figure 4.13 Surface contour of (a) S1, (b) S2, (c) S3, (d) S4 and (e) S5.

Figure 4.14 3D structure of the designed optical system. (a) Top view. (b) Front view. (c) Perspective view.

Figure 4.15 (a) Dimensions of lenses; (b) OLED display panel; (c) Near-eye system assembly.

Figure 4.16 Prototype appearance with the external driver circuit for the OLED panels.

Figure 4.17 (a) Sample image; (b) Experimental result.

Figure 5.1 Schematic explanation of the repeated viewing zones problem in light field NED. (a) The designed viewing position of light field NED with binocular vision. No depth confusion. (b) The practical situation with repeated viewing zone problem, caused by device alignment or eye position. Depth confusion occurs. (c) Generation of repeated viewing zones of periodic pinhole array light field display.

Figure 5.2 (a) Schematic diagram of the light field near-eye display with random pinholes. (b) Illustration of light field elemental images.

Figure 5.3 Rendering procedure of random pinhole light field NED.

Figure 5.4 Eye box demonstration of a light field pinhole near-eye display.

Figure 5.5 3D CAD design of the light field NED based on random pinholes.

Figure 5.6 The design pattern of (a) periodic pinhole array and (c) random pinhole. (b) and (d) are magnified images of (a) and (c).

Figure 5.7 Light field components fabricated and assembled in the NED system, including random pinhole film, acrylic board spacer, and an LCD device.

Figure 5.8 Photograph of the developed prototype and the camera (that simulates human eye).

Figure 5.9 Distribution of the viewing zones with different levels of signal-noise ratios (SNR) in the proposed NED system.

Figure 5.10 Signal-noise ratio (SNR) distribution with horizontal and vertical position at eye-relief distance (a) 40mm, (b) 50mm and (c) 60mm.

Figure 5.11 Rendered light field pattern for sample image “Lena”. (a) Elemental image for periodic pinhole array. (b) Enlarged image pattern for periodic pinhole array. (c) Elemental image for the random pinhole. (d) Enlarged image pattern for the random pinhole.

Figure 5.12 Perceived periodic pinhole array light field image by the camera with the NED prototype. (a) Correct light field image at the designed viewing zone. (b) Image during the transition between viewing zones. (c) Light field image at the repeated viewing zone with wrong depth information.

Figure 5.13 Perceived random pinhole light field image by the camera with the NED prototype. (a) Correct light field image at the designed viewing zone. (b) Decreased SNR and image quality when moving out of the viewing zone. (c) Light field image out of the viewing zone, the image is blurry.

Figure 5.14 Rendered light field pattern for sample image “Mandrill”. (a) Image pattern for periodic pinhole array. (b) Enlarged image pattern for periodic pinhole array. (c) Image pattern for the random pinhole. (d) Enlarged image pattern for the random pinhole.

Figure 5.15 Perceived periodic pinhole array light field image by the camera with the NED prototype. (a) Correct light field image at the designed viewing zone. (b) Image during the transition between viewing zones. (c) Light field image at the repeated viewing zone with wrong depth information.

Figure 5.16 Perceived random pinhole light field image by the camera with the NED prototype. (a) Light field image at the designed viewing zone. (b) Decreased SNR and image quality when moving out of the viewing zone. (c) Light field image out of the viewing zone, the image is blurry.

Figure 5.17 (a) Rendered light field image pattern for periodic pinhole array. (b) Perceived correct periodic pinhole light field image at the designed viewing zone.

Figure 5.18 Perceived periodic pinhole array light field image by the camera with the NED prototype. (a) Image during the transition between viewing zones. (b) Light field image at the repeated viewing zone with wrong depth information.

Figure 5.19 (a) Rendered light field image pattern for the random pinhole. (b) Perceived correct periodic pinhole light field image at the designed viewing zone.

Figure 5.20 Perceived random pinhole light field image by the camera with the NED prototype. (a) Decreased SNR and image quality when moving out of the viewing zone. (b) Light field image out of the viewing zone, the image is blurry.

List of Tables

Table 1.1 Performance comparison of different virtual reality system.

Table 3.1 Specification of the liquid crystal display (LCD) display panel for overlapping NED prototype construction.

Table 4.1 Specifications of the OLED panel and optical parameters of the NED system.

Chapter 1 Introduction

1.1 Near-eye Display (NED)

1.1.1 Virtual Reality

Virtual reality (VR) generally refers to the technology that uses virtual reality related special hardware, together with computer software, to generate and simulate a believable and immersive scenario of real-world or an alternate arbitrary world. In daily life, though the word “virtual reality” is the most popular one to be used, we see different terminologies, including *Virtual Reality*, *Virtual Environments*, *Synthetic Experience*, *Artificial Worlds*, *Artificial Reality*, etc., which all have same or nearly equivalent meaning [1]. However, sometimes the term *Virtual Environments* is used instead of *Virtual Reality* to emphasize that the simulated content is created from imagination instead of a remodeling of the real world.

The origin of virtual reality can be traced back to the imagination from ancient times, where stories were formed. Then technology stepped into imagination when the virtual stories were described by books, radio, and visible films. With the help of latest computer technology, an interactive virtual environment can be generated with computer graphics. With the concept of virtual reality, the imagined environment can be generated inside out to show human a virtual reality world.

The history of virtual reality can date back to the peep show in the early modern period (17th-century) when peep show was also called peep box or raree show [2]. In the application of peep show, objects are generated by a series exhibition of pictures. The objects are viewed by the user through a small hole or a magnifying glass. A peep show is typically carried in a wooden box where the viewer can look inside it. Different motion picture devices were invented in 19th-century, including Mutoscope invented by W.K.L. Dickson and Herman Casler [3] and Kinetoscope invented by famous Thomas Edison [4]. Both Mutoscope

and Kinetoscope represent very successful commercialization of early virtual reality devices.

The 19th-century Chinese peep shows are famous for its long-lasting impact, known as Xi Yang Jing (西洋镜) or La Yang Pian (拉洋片). Usually performed with puppet, this kind of peep show is still available in some cities in China.

As can be observed in the old photos, ancient peep shows already demonstrated the basic working mechanism of virtual reality, which in turn act as the fundamentals of modern virtual reality devices.

Figure 1.1 Peep show devices in 19th-century [5, 6, 7].

The modern history of virtual reality device is recognized to start from 1962 when a simulator called Sensorama was invented [8, 9]. Created by Morton Heilig, Sensorama is a simulator that augments sound, wind, and vibrations. Although the machine is not interactive, it is the first device to create such a multi-sensory environment. In 1963, Ivan Sutherland published his doctoral thesis, naming “SKETCHPAD: stereo HMD, position tracking and graphics engine”, which is considered as the earliest virtual reality head-mounted display (HMD), with proper head tracking, marks the beginning of modern virtual reality [10]. After that, in 1965 and 1966, he proposed a further solution to interactive virtual reality with a concept of graphics, feedback, sound, smell, etc. In the 1970s, the first force feedback system was realized. Known as “GROPE”. The first virtual environment was created by Myron Krueger team in 1975. In the 1980s, U.S.

NASA played a very important role in the development of virtual reality. Multiple systems, including “Virtual Visual Environmental Display” (VIVED) and “Virtual Interface Environment Workstation” (VIEW) were developed to meet the needs of simulation and training. In 1985 and 1988, DataGlove and EyePhone HMD were developed respectively as the first commercialized virtual reality product. In the 1990s, virtual reality grew rapidly with Virtual Reality Modeling Language (VRML) invented in 1995. During the 1990s, virtual reality induced market grew from ~\$50 million in 1993 to \$1.4 billion in 2000, which is about 28 times in 7 years⁸.

Up to now, the application of virtual reality ranging from engineering, architecture, military, to entertainment and education, provides infinite possibilities to human. The industry of virtual reality had strong growth in the last 10 years. It is predicted that the market size of virtual reality will reach an unimaginable 100 billion by the end of 2021 [11]. As big players stepping in, including Google, Facebook, and China’s BAT (Baidu, Alibaba, and Tencent), the future growth of virtual reality is undoubted. The industry infrastructure of virtual reality is shown in Figure 1.2.

Figure 1.2 Virtual reality industry infrastructure [12].

Although it can be very difficult to use a single standard to categorize all virtual reality systems, we can still categorize them based on the presence of real-world in the virtual environment. Virtual reality systems can be simply divided into three big categories, immersive virtual reality, non-immersive virtual reality, and semi-immersive virtual reality.

Non-immersive virtual reality, as the name suggests, is the system provides least immersive feeling to the user. Typically, the non-immersive system refers to a desktop virtual reality system where mostly desktop screen monitor is used to show the virtual environment. Non-immersive virtual reality does not require

a sophisticated interactive device, thus typically has low build-up cost. However, the experience of non-immersive virtual reality system can be easily outperformed by immersive virtual reality system, where the user is isolated from a real-world environment. A semi-immersive virtual reality is one kind of large-scale device, sometimes called “fish tank”, which indicates the user is inside the system. Semi-immersive system typically uses a big cinematic screen or multiple projection screens, sometimes simply IMAX screen with 3D shutter glasses. With such wide field-of-view and a binocular visual effect, the semi-immersive system holds its own application, including aircraft simulation and industrial design. Table 1.1 shows a general comparison between different kinds of virtual reality systems.

FEATURES	NON- IMMERSIVE VR	SEMI- IMMERSIVE VR	FULL- IMMERSIVE VR
RESOLUTION	High	High	Medium
SCALE	Low	Medium-High	High
SENSE OF AWARENESS	Low	Medium	High
FIELD OF REGARD	Low	Medium	High
SENSE OF IMMERSION	None-Low	Medium-High	High
MOTION LAG	Low	Low	Medium-High

Table 1.1 Performance comparison of different virtual reality system [13].

In virtual reality, the immersive feeling is delivered by all possible ways to generate a believable simulation of a real or alternative world. The ultimate goal of virtual reality is to bypass or cheat our senses so that we cannot tell whether the environment is generated by the human. However, even based on state-of-the-art technology, this goal is nearly impossible to be achieved, because some of the human feelings are hard to simulate [14], but we can still create an immersive virtual reality feeling by delivering the senses in possible ways, including [15]:

- Eye vision: the vision received by eyes created by the display, preferably an HMD.
- Ear hearing: the audio generated by external audio devices, ranging from typically ~20Hz to 20kHz, with an identifiable stereo sound stage.
- Somatic senses: including touch sensations, the status of body tissue and organs, body position, pain, and vibrations.
- Equilibrium: including the sense from the vestibular system, sense of gravity, head rotation, and acceleration.

Among the four main aspects of virtual reality perception, eye vision is intuitively chosen as the most important virtual reality element for commercialized VR. Improper display of visual content will cause severe sensory conflict to the brain, thus in turn result in discomfort and sickness, including mental discomfort, visual discomfort, and somatic discomfort. Visual factors that can cause virtual reality sickness, including poor image quality, poor resolution, display time lag, tracking error, stereo mismatch, narrow field of view, wrong interpupillary distance, etc. Besides, device discomfort is another reason for virtual reality sickness, for instance, heavy device weight, bulky device size, and uncomfortable mounting.

In this research, we mainly focus on the visual aspect of virtual reality, which is referred to as head-mounted display or near-eye display.

1.1.2 Near-eye Display

Near-eye display (NED), also referred as a head-mounted display (HMD) or head-worn display (HWD), is the display device worn by the user on the head with a close distance to the eye for content perception and interaction. Modern immersive virtual reality system is typically implemented with NED, where a fully immersive visual effect can be offered with stereo binocular vision. With the emerging of faster wireless connection technology and mobile computing, the application of NED becoming more and more common, where more visualized user interface is demanded.

The types of NED can be determined based on whether the display content is only computer-generated content or a combination of computer-generated content and the view of real-world. Most NEDs only display virtual content; however, in some applications, the computer-generated image is superimposed with a real-world scenario, usually with an optical see-through mirror or with digital method. This is referred to as augmented reality (AR) or mixed reality (MR).

Figure 1.3 Different kinds of near-eye displays [16, 17].

NED is miniaturized to contain 1 or 2 pieces of display panels inside to provide high-quality vision. Different image sources can be provided to NED, including most common liquid crystal display (LCD), organic light-emitting diode (OLED), micro-electro-mechanical systems (MEMS) and liquid crystal on silicon (LCoS). As an essential part of the display system, the market of the display panel has been growing rapidly, especially OLED technology, driven by the growing energy-saving demand. Besides, OLED easily edges over other competing technologies on contrast ratio, flexibility, and simple construction. Because of the advantages listed, the application of OLED display panels is growing among virtual reality devices and consumer electronics. The design of the NED can be an interdisciplinary task, guided by the application of each individual case. The aspects involved during the design process include optical engineering, optical coatings, optical materials, electronics, physiology, etc [18].

Eyeglass based near-eye display are better accepted because of mobility and better aesthetics. However, though bulkier NED like Oculus Rift is not as acceptable aesthetically as glass-based NED, they typically offer better immersion and performance features including resolution and field-of-view, thus

becoming popular in applications that do not feature mobility as the primary limitation. In other words, although people prefer thin and small optics from a NED, the performance of compact NED system still cannot match the expectation in terms of resolution and field-of-view. Meanwhile, with the development of the latest optics design technique, including freeform optics, innovative compact near-eye display improves with time, provides more choice to the user for a more portable and comfortable virtual reality experience.

1.2 Motivation and Objectives

In this thesis, theory and methods are demonstrated to improve key performance factors in the modern near-eye display (NED), including size and resolution, to provide a hint to the design of future VR and AR devices. Targets and objectives achieved can be listed as follows:

1. Develop an image overlapping super-resolution method to enhance the resolution in the near-eye display. Implement the method with the computational program and achieve image decomposition from high-resolution image input.
2. Implement a high accuracy calibration method according to the property of NED and achieve pixel-level alignment.
3. Design a dual display panel NED prototype to demonstrate the effectiveness of image overlapping. Carried a theoretical derivation to discuss the generation procedure of sub-frame image and the calibration principle to super-impose multiple images. Compare the result to the traditional one display NED setup and validate the result for both image and video applications.
4. Develop a freeform surface based off-axis NED to achieve compact size and good optical performance. Investigate and x-y polynomial surfaces with other different freeform optics surface types.
5. Fabricate the designed optical elements with commercialized fabrication technique. Design and fabricate an optical NED prototype. Evaluate compact NED performance for further enhancement.

6. Investigate the design method and property of random pinhole, including the rendering method, determination of eye box size, and viewing zone distribution.
7. Design and fabricate a random-pinhole-based NED prototype to demonstrate the effectiveness of the proposed method. Analyze the behavior of the random pinhole. Compare the perceived light field image quality from random pinhole over periodic pinhole array.

1.3 Major contributions of the Thesis

In Chapter 3, a novel resolution enhancement method for NED is proposed by overlapping the image of multiple display panels. The theory and sampling model of NED image overlapping are developed to illustrate the principle of resolution enhancement. To generate sub-frame image for resolution enhancement of multiple display panels, a computational algorithm is developed, taking consideration of the intrinsic alignment errors such as shift and rotation. Next, to achieve image overlapping and super-resolution, a special display calibration method is introduced to align the sub-pixel of the NED display panels, which is crucial for resolution enhancement. A simple yet effective optical solution for NED image overlapping is proposed by using a semi-reflective mirror, which is cost effective and has very low image distortion. To validate the theory and computational model, an overlapping NED prototype is designed and fabricated with common commercial liquid-crystal display (LCD) panels and semi-transparent mirror at reasonable cost of 400 Singapore dollars. To demonstrate the resolution enhancement result, an experiment is performed using the NED prototype and a mirrorless camera. This experiment shows that the NED resolution is effectively improved by using the proposed method, which overcomes the resolution limit of a single display panel. Furthermore, without complex lens and tracking mechanism, the proposed method reduces the system complexity and design difficulty. Compare to optical tiling, overlapping NED can easily switch between single display panel mode and multiple display panel mode, where the image is continuous without seams in between display panels. This provides more possibility and flexibility in the design of NED for high-resolution priority applications.

In Chapter 4, we proposed a improved design of compact off-axis near-eye display system described by x-y polynomial surface. The design features OLED micro display and multiple times of reflection, resulting in very compact size and better human head shape clearance. A compensation prism is designed to enable augmented reality. Here, we proposed a designing method of off-axis NED, including clearance control, free-form surface optimization, and characterization. The developed prism is manufactured with commercial diamond turning technique. A prototype is fabricated and assembled with an OLED micro display, resulting in an usable NED with 4-mm diameter exit pupil, 21.5° diagonal full field-of-view (FoV) and 23 mm eye relief. Test results show that the proposed near-eye display has good optical performance yet compact in size. Overall, the proposed design is a good option in the application of state-of-the-art wearable virtual reality.

In Chapter 5, a novel random pinhole based light field NED with a unique viewing zone is proposed to eliminate the repeated viewing zone problem of traditional periodic pinhole array, thus improve the depth perception. The modeling and principle of the novel random pinhole light field display is described, including system setup, image rendering method, eye box size, viewing zone distribution and resolution. Based on the proposed theory, we found a random pinhole setup with big eye box and pupil size for human eye. Next, a random pinhole-based NED prototype is fabricated with cheap inkjet-printed pinhole film and LCD panel, which is simple and compact. To validate the behavior of random pinhole light field, an experiment is designed that characterizes the viewing zone of the light field display. The result shows the uniqueness of the random pinhole-based light field, which can effectively eliminate the repeated viewing zone of traditional periodic pinhole array and provide more consistent performance of the depth perception.

1.4 Organization of the Thesis

Chapter 1 of the thesis introduces the development of virtual reality, which is the biggest driving force of near-eye display research in recent years. The definition and development of the near-eye display are also given.

In Chapter 2, literature reviews are conducted on previous researches on several aspects of near-eye display researches, including the invariance between resolution and field-of-view, resolution enhancement methods, display overlapping, free-form optics near-eye display, and light-field near-eye display. Based on existing researches, my ideas of improvement are proposed.

The target of the thesis is to increase the resolution, enlarge the field-of-view, improve accommodation & convergence conflict properties, improve depth acquisition, and reduce the NED size. To achieve these targets, individual solutions are proposed in Chapter 3, 4 and 5 to address different aspects of the task. In Chapter 3, we proposed a resolution enhancement method via super-resolution. In Chapter 4, we demonstrated a free-form off-axis design which result in a very compact NED. In Chapter 5, a light-field-based NED is implemented with eliminated repeated viewing zones, which effectively improves depth acquisition.

In Chapter 3, to solve the main conflict between resolution and device size, a versatile resolution enhancement method for NED is implemented by overlapping multiple display panel images. Based on display super-resolution, this overlapping method effectively increases the perceived resolution without optical complexity. By investigating the algorithm theory of image generation and calibration, the perceived resolution can be enhanced and overcome the physical resolution of a single display panel. A NED prototype based on a semi-transparent mirror is designed and fabricated. Experiments demonstrated decent results of resolution enhancement. Compare with the conventional method of display tiling, my method offers an equivalent level of performance yet reduces system complexity, brings more flexibility in the design of NED.

In Chapter 4, a design routine of off-axis near-eye display system is implemented. Freeform surfaces are used to effectively reduce the size of NED

while correcting the off-axis aberration. The designed system yields good optical performance yet very compact in terms of size, compared to traditional NED. A prototype is demonstrated with components available from the market. The prototype effectively indicated the feasibility of the design as a compact NED.

In chapter 5, a light field NED with random pinhole design is demonstrated with unique viewing zone and optical simplicity to solve convergence-accommodation conflict in NED. Based on either microlens or pinhole, light field NED demonstrates an effective size reduction on the longitudinal direction of the device.

Further research is recommended in chapter 6 to combine the individual technique as a hybrid device, where multiple light field display module can be used for super-resolution, with the design flexibility of free-form optics.

Chapter 2 Literature review

2.1 Resolution enhancement

2.1.1 Resolution limitation of near-eye display

State-of-the-art applications of virtual reality range from engineering, architecture, military to entertainment and education, enables infinite possibilities to human. As the core component in the implementation of virtual reality system, modern near-eye display (NED) requires both field-of-view and resolution towards a satisfactory experience for the most demanding immersive virtual reality, e.g., simulation and video games. From the optics point of view, the two main factors taken into consideration are field-of-view and resolution, which determine most applications of the NED.

In real-world, the application of NED can be decided mostly by field-of-view. Each human eye can see 140 degrees horizontally and 110 degrees vertically. If two eyes are put in a pair, the horizontal field-of-view can reach 195 degrees [19]. A low field-of-view (<40 degrees) design is usually used in compact symbol or text-based application where the requirement of field-of-view is not as critical as immersive applications where a wide field-of-view is demanded, e.g., pilot simulation and video games. In mobile virtual reality applications, where the NED moves together with the user's head, large field-of-view is critical. Typically, 60 degrees vertical and 75 degrees horizontal field-of-view is required¹⁸.

Resolution is another important aspect when talking about NED. In fact, it is hard to get the eye to focus on a blur or low-quality image. The maximum resolution at the center of the retina can be as small as 1 arcsec. Typically, 1 arcmin is at least required for application with long wear time. There are some applications where resolution is critical, including surgery and inspection.

Figure 2.1 Demonstration of resolution / field-of-view invariance. (a) High resolution, small field-of-view. (b) Large field-of-view, low resolution.

There is a fundamental trade-off in NED between field-of-view and resolution. As the pixels are spread evenly across the field-of-view, once field-of-view is enlarged, the density of pixel per unit angle must decrease accordingly. As shown in Figure 2.1, two simulated images are generated from the same source, one with small field-of-view, another one with a large field-of-view. As we can observe from the figure, the image with a larger field-of-view has a much rougher edge.

The relationship of field-of-view and angular resolution can be described by the expression:

$$H = f \tan \theta \quad (2.1)$$

where H is the physical pixel size of the display, f is the effective focal length of the lens, and θ is the angular resolution in the image space [20]. In this case, high-resolution means we want θ to be smaller. Or alternatively, in this equation, H is the size of the display, and θ is the field-of-view. Thus, for a single display panel, the angular resolution and field-of-view is decided by the focal length f . Here H acts as an invariance for a particular display device. The field-of-view and resolution of the display must compromise for each other. The demonstration of the resolution / field-of-view invariance is shown in Figure 2.2.

Figure 2.2 Demonstration of resolution / field-of-view invariance. (a) H is the size of the display, f is the effective focal length of the lens, and θ is the field-of-view. (b) H is the physical pixel size of the display, f is the effective focal length of the lens, and θ is the angular resolution in the image space.

As discussed above, taking 1 arcmin as the minimum angular resolution requirement for NED, the display resolution can be calculated by:

$$R = \frac{FOV}{1arcmin} \quad (2.2)$$

where FOV is the abbreviation for field-of-view. If we take the field-of-view to be 45° vertical by 60° horizontal, the total rows and columns needed are 2700 and 3600, respectively. If we take the most popular field-of-view in commercial video game near-eye display, including Oculus Rift and HTC Vive, which show about 110° of horizontal field-of-view, the total number of lines of pixels needed is 6600. In comparison, the resolution of the display panel used by Oculus Rift and HTC Vive are only 2160 x 1200 for both eyes. In these systems, left and right viewing channels for each eye are isolated to avoid crosstalk. Thus, the horizontal display resolution for a single eye is only limited to 1080, while vertical resolution is at 1200. These numbers are only 1/6 of the resolution expected.

Thus, for cutting-the-edge applications including flight simulation, video game, and most of the other immersive applications, both field-of-view and resolution

are demanded. The user needs to see a wide range while small items must be identifiable for better simulation experience and reduced fatigue. However, the physical resolution of the display is limited by the fabrication technology and manufacturing cost. With the latest commercial display panels, 2K horizontal resolution is a sweet spot which is taken by many commercial NED including Oculus Rift and HTC Vive, while 4K display panel is a commercial limit for displays about 5 inches in size. Fit in an extremely large display panel inside NED is also very hard considering the mobility required. As a result, in some applications, for instance, medical surgery, field-of-view must be reduced to get better angular resolution, which in turn bring big compromise to the viewing experience. Thus, to achieve better resolution and field-of-view, other methods need to be taken.

2.1.2 Resolution enhancement methods

Solutions were developed, including the high-resolution area of interest, partial binocular overlap, optical tiling, and dichoptic image [21, 22, 23, 24].

The high-resolution area of interest method was developed to increase effective resolution without sacrificing the field-of-view [25, 26, 27, 28]. As shown in Figure 2.3, a large field-of-view image with low resolution is used as a background. A small field-of-view high-resolution image is superimposed onto the background, according to the eye-tracking signal, known as the area of interest.

Figure 2.3 Basic working principle of high-resolution area of interest. A low-resolution large FoV display is used as background. A high-resolution but small FoV display is superimposed on the background, acting as the “area of interest”.

With the high-resolution image moving together with the user’s eye vision, a display image with high resolution at the center can be yielded, with large field-of-view at the border as peripheral. However, this system requires a sophisticated eye-tracking mechanism for the area of interest to move with the user’s eye vision.

The partial binocular overlap method is a good way to enlarge the field-of-view, which is, turn the optical relay inward or outward to make the viewing cones overlap, which is called convergent and divergent configuration, respectively [29]. Figure 2.4 shows the comparison of normal full binocular overlap with partial overlap.

Figure 2.4 Partial binocular overlap method to increase the FoV without sacrificing the resolution. (a) Normal binocular vision with full overlap. (b) Divergent partial binocular overlap. The viewing cone pointing outwards.

Binocular overlap can increase the horizontal field-of-view without decreasing the resolution. Divergent packaging makes it very easy to pack the dual-display around the user's head. The main challenge of binocular overlap is the distortion correction as distortion will cause the non-linear mapping of the image. The overlapped area needs to be carefully controlled to avoid eye fatigue.

Dichoptic area of interest is similar to the high-resolution area of interest method. Instead of tracking the user's eye, in this method, a large field-of-view image is presented to one of the user's eyes while a high-resolution small field-of-view image is presented to the other eye [30]. As shown in Figure 2.5, left viewing channel is high-resolution but small FoV image, while right viewing channel is low-resolution but large FoV image.

Figure 2.5 Dichoptic area of interest method. Left viewing channel is high-resolution but small FoV image. Right viewing channel is low-resolution but large FoV image. Two viewing channels merge together to form a high-resolution binocular image with large FoV.

Instead of using tracking, this high-resolution area is typically fixed at the center of the viewing area. Thus, when the left and right eye image are merged as a binocular image, an area with high resolution can be obtained at the center. The solution used in night vision.

Optical tiling is another solution. In optical tiling, multiple high-resolution screens with small field-of-view are tiled optically to form a large field-of-view [31, 32, 33], just like what people did with projection screens at early years. The overall field-of-view is equal to the addition of each small modules. Optical tiling is primarily used in flight simulation, which has the most critical demand for large field-of-view and high resolution. As the most scalable method to increase FoV or resolution, the tiling method is presented in some other display systems, such as video wall [34], projecting screen [35], autostereoscopic display [36], light-field display [37], and holographic display [38].

In the conventional implementation of a large FoV near-eye display, multiple eyepieces are use with a common rotating center to form an array. The eyepiece

types investigated in previous studies include Fresnel lens [39], doublet, triplet [40], and aspherical [41]. Recent development of tiled NED includes rotational symmetrical eyepieces [42] and the free-form optics solution [43, 44, 45]. These novel systems managed to solve some problems of traditional tiling NED and make user's eye on the optical axis. The most challenging part in display tiling can be the alignment of each screen. A seamless alignment can be very important to give the user a continuous view of the virtual scene. The potential of tiling NED is limited by the manufacturing cost and alignment difficulty during assembly.

As discussed above, the existing methods to achieve high resolution together with large field-of-view all have their different advantage thus fit in different applications. However, they all have different disadvantages, either introducing optical complexity to the system or harm the viewing experience. Thus, aside from improving the physical resolution of the display panel, other solutions should be developed with latest digital science and computer graphics. Display super-resolution is a possible way to increase the resolution but haven't been applied to NED.

2.1.3. Display super-sampling and super-resolution

Primarily used in signal processing and computer graphics, super-sampling is a method to generate alias-free images. In computer graphics, polygons and surface textures need to be rendered carefully to avoid the presence of alias, which is caused by the low sampling rate of display pixels. With super-sampling, multiple low-resolution data are computationally combined, to produce high-quality data sample, i.e., pixel information [46]. From a signal processing point of view, the information above the Nyquist frequency of a given image generated during the rendering process is the reason for bothering alias [47]. Conventionally, anti-aliasing is performed in the way such as full-scene anti-aliasing (FSAA), where the high-resolution image is rendered from 2 times to even 16 times the original resolution, then filtered and down-sampled to the demanded resolution. However, this operation requires much more computational power, which can be a limitation in the application of computer

graphics. In the operation of super-sampling, different samples are taken at a different location with respect to a standard point, in the application of image processing, the pixel point. Alternatively, the sampling location can be chosen jittered or stochastic to get a high-quality rendering of alias-free images [48, 49]. The samples are computationally combined with filtering to produce a smooth image [50, 51]. Compare to FSAA, where full image is rendered at an enhanced resolution; in super-sampling, multiple images are rendered at the same resolution of the display with sub-pixel offset.

The method of super-sampling can be treated as the inverse dual of super-resolution method in the image capturing and computer vision, where a high-resolution image can be captured as a combination of two comparably lower resolution images [52]. In 2004, the concept of a jittered camera was developed based on a sub-pixel level shift to improve the recorded video quality [53]. This method in image capturing is widely used in many applications, which can overcome the physical resolution limit of the camera itself and get a high-resolution image [54].

Display super-resolution is to take advantage of multiple display samples and physically combine them to get high-resolution images when it is more practical to apply multiple low-resolution devices instead of a single high-resolution display, to achieve a similar result [55]. Originally, the idea of display super-resolution started from projecting displays. Presented by Christopher Jaynes and Divya Ramakrishnan in 2003, multiple projectors are applied with an overlapping display area to significant improvements of resolution [56]. Sub-pixel shifts and calibration of multiple sub-frames are the key yielding resolution exceeding individual projectors. In 2005, “wobulation” was presented as an effective solution to increase the resolution of a single digital projection system, where subframe images are shifted by sub-pixel level by a wobbled opto-mechanical image shifter [57]. The perceived image is superimposed by two sub-frames with an offset in succession, leading to high-resolution display result as in comparison with the traditional projector. Previous research demonstrated that calibration is needed in multi-projector research of super-resolution [58]. Previous research already paved the way of multiple alternative algorithms for

effective implementation of display super-resolution. In this thesis, display super-resolution is implemented using display calibration as an important step.

2.1.4. Display calibration

When resolution increases in the system, small offset change become significant. System calibration is required under this situation. Traditional single camera-based phase measuring profilometry was developed. The measured object is projected with a calibration image while using the camera to shoot the surface of the measured object to detect the deformation of the calibration image. Then the deformation of the image is processed, calculate the phase information representing the position of the object surface. Finally, according to the phase information and the calibrated parameters, overlapping image formation can be carried [59, 60].

The digital camera is a computer vision system to obtain image information. Information and parameter in world coordinates can be calculated from the image coordinate system, or from the three-dimensional space information or two-dimensional image coordinates. To do this, correspondence between the image point position and the space point need to be established, that is, the position of the camera in the world coordinate system and orientation (external parameters), and the camera's own geometry and optical parameters (internal parameters). The parameters should be determined by the camera's imaging model. Experiment and calculation are required to determine the above-mentioned internal and external parameters. The process of determining these parameters is called calibration [61].

With the development of computer vision and photogrammetry, research on camera calibration technology is becoming popular. The calibration sign also changes from three-dimensional to two-dimensional and one-dimensional. The camera calibration techniques are becoming mature with a lot of available works of literature and results. The traditional camera calibration method described in the previous section requires that a calibration be placed in the scene. More importantly, in some cases, due to the camera's focal length to be constantly

adjusted, the location will be based on the surrounding environment. Therefore, it is necessary to calibrate the internal and external parameters of the camera after each adjustment. To overcome these shortcomings, camera self-calibration technique was developed rapidly and became an important research direction in the field of computer vision [62, 63, 64, 65]. In general, since self-calibration technology can be calibrated directly from the image sequence within the camera parameters without the calibration sign, better flexibility can be achieved. However, the traditional method can still achieve better accuracy, which is critical in the measurement of pixels. The choice of calibration method should be tested and considered based on the requirement of the display super-resolution requirement.

2.2 Near-eye display with free-form optical design

Towards the application of military, engineering, entertainment, and education, where a compact device is more demanded instead of traditional HMD, the near-eye display is developed to provide portability and mobility. Optically see-through designs are especially favored. Various early designs and products were presented by companies including Zeiss [66, 67], Sony [68, 69, 70, 71], Google [72, 73], and other international companies [74, 75, 76] to adapt the needs of consumer and professional markets, where users are extremely in favor of devices with compact design, lightweight, portable wearing, and non-intrusive layout. To reach this target and achieve an “eye-glasses-type” of NEDs, methods were employed including freeform optics [77], scanning optics [78], diffractive optics [79, 80, 81] and waveguide optics [82, 83, 84].

Off-axis optical system is defined as the optical system where the optical axis and the mechanical axis are not overlapping with each other. The catadioptric method was deployed to design and off-axis see-through NED [85], result in a helmet shape device which is too bulky for commercial use. Taking advantage of total internal reflection, waveguide-based NED was designed [86, 87]. However, due to inevitable stray light and chromatic aberration, the field-of-view is limited, which is a constraint when it comes to some application scenarios. A

full-color holographic waveguide-based NED was proposed by Mukawa et al. using in- and out-coupling holographic elements [88].

Freeform optics utilizes the surfaces which are not translational or rotational symmetric. Freeform optics offers good miniaturization and simplicity of components in many modern illumination and imaging systems. Freeform surface design for NED was firstly presented by Cakmakci, where a diffractive optical element (DOE) surface was used together with a freeform mirror. Following this work, a design with 3mm exit pupil and 24 degrees field-of-view was presented [89, 90, 91]. A double mirror freeform design was proposed by Bauer et al. The side effect of double mirror results in extra difficulty in elements alignment [92]. Systems consist of multiple aspherical mirrors, and relay lenses were proposed where alignment remained as a problem [93, 94, 95]. Remarkable designs were proposed where distortions were noticeable in off-axis design [96]. A prism-based NED was presented by Cheng et al. with the x-y polynomial surface [97]. The further design was developed with a tiled bottom surface [98]. A design is further produced by Cakmakci with four freeform surfaces and an off-axis design [99].

2.3 Light field near-eye display

Advancement of application in virtual reality and augmented reality evolves near-eye display in performance and form factors, makes it more suitable for wearable and mobile computing [100].

Traditional near-eye display relies on binocular parallax to provide user depth cues. As an effective way to generate depth information, binocular parallax offers promising depth effect. However, in traditional NED, with normal display panel setup, accommodation / convergence conflict (A/C conflict) occurs. Figure 2.6 demonstrates the accommodation and convergence conflict. In this figure, user's eye focuses on the display panel, which shows the image content. The accommodation distance is given by the distance between the eye and the display panel. The convergence distance is given by the binocular parallax of the image content of each viewing channel. In this case shown in Figure 2.6, the convergence distance is much bigger than the accommodation distance.

Figure 2.6 Accommodation and convergence conflict. User's eye focus on the display panel, gives accommodation distance. The binocular convergence distance in this case is further away. The difference between accommodation and convergence distance will cause discomfort and fatigue.

Traditional NED only has one image plane, which is the plane of the flat panel display. When binocular parallax generates depth cues, the user perceives the depth at a different plane with the image plane, which leads to visual discomfort and fatigue [101]. In fact, A/C conflict becomes more severe when combined with motion sickness, which is a major obstacle towards better user experience.

To solve A/C conflict, multiple focal planes must be generated. Previous researches proposed several solutions, including temporal-multiplexed [102, 103] and spatial-multiplexed [104, 105] method. However, this kind of implementation has to introduce additional display or optical elements, including display panels, a semi-transparent mirror, a spatial light modulator, etc., which in turn drastically increases the complexity and cost of the NED system.

Another solution of A/C conflict is Maxwellian view display [106, 107, 108]. In the Maxwellian view display, images are projected on the user's retina directly through the pupil. Aside from the complexity, Maxwellian view display suffers

from providing a large exit pupil size to cover the whole eye socket, resulting in a very sensitive pupil position and interpupillary distance, as well as vignetting.

Light field display is a promising solution to provide focusing capability at different depth. In a light field display, directional rays are generated to mimic the depth of focus. We saw various methods developed by researchers to generate the light field, including multi-layer computational light field display [109, 110, 111], which uses pixels on liquid crystal display (LCD) as a light modulator, to control the directional rays and form light field. However, as a light modulation device, LCD cannot control individual ray, thus require a large amount of computation to get the light field correctly optimized, which in turn limited the application.

Microstructure array so far remains one of the simplest ways to generate the light field. Microstructure based near-eye display draws great attention in recent development trends of the near-eye display [112, 113, 114, 115, 116, 117]. Microstructure array is widely used in stereoscopic display and integral imaging. Pinhole device are used in the implementation of super-multi view (SMV) display [118, 119, 120]. Typical implementations include microlens array and pinhole array. In both cases, the device can angularly sample the ray of light and form light field. Different variety of microstructure devices have been developed to suit the various applications of VR and AR, including optically see-through ones [121]. The rendering procedure, resolution, and the design of the eye box have been investigated [122].

To fully reproduced the light field, enough rays should be formed in the eye box. However, in real practice, the performance of the light field display still has some limitations. One of the challenges is the presence of repeated viewing zones. In the design of light field display, the pupils and viewing zones are expected to remain fixed at one position, however, in a real-world scenario, the pupil always drifts from the designed position, either caused by the movement of eyeball or equipment fitment. The alignment of the microstructure and the display panel may cause the viewing zone itself to drift drastically as well. Both scenarios lead to the problem of repeated viewing zones, which refers to the wrong perceived

image when the pupil drifts out of the designed viewing eye box [123, 124]. The repeated viewing zone problem also presents in the autostereoscopic display researches. The perceived image from neighbor viewing zone usually differs from the design position, thus causing the problem of depth acquisition.

One solution to tackle the repeated viewing zone is to make the exit pupil larger. However, based on the latest design principle, the pupil size is usually limited to less than 10mm [125, 126, 127], due to variation of human pupil size (2-8mm) and interpupillary distances. Another solution is introducing a visual barrier [128] or specially illuminated backlight system [129], similar to the method used in the autostereoscopic display system. However, the NED system's compact nature means it's very difficult to introduce additional components, unlike the autostereoscopic display system. A similar conclusion applies to the eye or gazes tracking solution in NED.

Inspired by the non-uniform barrier solution in the autostereoscopic display system [130], here we would like to propose a random-pinhole-based solution for light field NED. Instead of using periodic microstructure as a spatial angular light modulation device, a random microstructure can be used. Instead of generating repeated viewing zones, the random microstructure device can only produce one unique viewing zone in space. Thus, there's only one unique position where the user can perceive a clear image from the light field NED, which eliminates the chance of getting the wrong image with confusing depth cues.

The method and implementation of random microstructure based light field NED are presented in Chapter 5.

Chapter 3 Resolution enhancement utilizing overlapping multiple display panels

In this chapter, we demonstrate a super-resolution near-eye display (NED) capable of displaying a high-resolution image. A dual display panel NED design was introduced to enable super-resolution in a near-eye display. Algorithms were developed to achieve super-resolution. Key steps, including image generation and display calibration, are discussed. A NED prototype was constructed. Experimental assessments show enhanced resolution and overall image quality with less alias and blur. The proposed method is an effective and low-cost solution to improve resolution, thus has potential application in virtual reality application where resolution is highly demanded.

3.1 Introduction

The applications of near-eye display (NED) are growing rapidly, thanks to the development of virtual reality (VR), augmented reality (AR), wearable devices, and mobile computing [131, 132, 133, 134].

Various NED systems were designed since then, to meet the demands from consuming market [135, 136]. However, it remains challenging to implement a NED design with to achieve large field-of-view (FoV) and high angular resolution at the same time due to resolution/FoV invariance explained in Chapter 2, where the angular resolution is determined by a given FoV and number of effective pixels of the display panel. As a result, compromises need to be made between FoV and angular resolution perceived by the user, if no additional pixels added to the system, no matter what optical solution paths are taken. The method proposed by previous researchers, including the high-resolution area of interest method, partial binocular overlap, dichoptic area, and optical tiling, can all be treated as proper solutions to bring effective additional pixels to the NED system. In this chapter, we are going to discuss how to bring more effective pixels by image overlapping and display super-resolution.

As discussed in Chapter 2, image overlapping was initially developed to increase the resolution of a projecting display system, which was demonstrated effective to break the resolution/FoV invariance, thus widely applied. Near-eye display has already been proved to be capable of multiple display panels, where multiple display panels are usually tiled together. In this scenario, each display panel is in charge of a single viewing zone channel. The overall FoV is equivalent to the addition of a separate display channel.

Meanwhile, we see NED devices with multiple display panel layers, where liquid-crystal-display (LCD) panel layers are stacked to form a display system with multiple focal planes [137, 138, 139]. This kind of systems is built mainly based on the purpose of solving the accommodation and convergence problem and offer more viewing comfort for users. High dynamic range is another good result provided by overlapping. A dynamic range of 14 orders was demonstrated with LCoS display panel [140].

The feasibility of overlapping display in NED can be transferred from similar implementation from projector systems. However, the difference appears when it comes to the alignment of the display panel. To achieve sub-pixel level alignment and super-resolution, both systems require calibration. In a common scenario of the multi-projector system, each projector is very likely to be placed separately and independently, which makes it tricky when it comes to the tolerance of super-resolution. Meanwhile, in a NED system, the position of each display panel is pre-determined during assembly, thus more stable and feasible in real practice. Conventional augmented reality (AR) implementations using optical combiner or semi-transparent mirror can be used in the design of overlapping NED. Conventional combiner designs are easy to achieve but usually result in bulky devices. To reduce the NED size, novel overlapping method can be achieved with various state-of-the-art augmented reality (AR) solutions, including waveguide, freeform optics, and relay optics, etc. [132,133]

Based on the discussion above, we would like to propose a NED resolution enhancement technique based on overlapping display images. The proposed method can be implemented by various designs to achieve display overlapping.

To simplify the technique and demonstrate the effectiveness of display overlapping, here the NED system was implemented by a conventional reflective semi-transparent mirror combiner. A prototype was fabricated, which showed a clear enhancement of resolution. Calibration technique was introduced and discussed, which is crucial for the resolution enhancement, i.e., makes the additional pixels effective to the user. Because of the conventional optics setup, the resulted NED is still bulky in size, but it proves overlapping to be an effective alternative in the resolution enhancement of NED. The overlapping method does not introduce noticeable artifacts compared to existing methods, e.g., tiling seam or mismatch, which is very helpful when it comes to user experience. It brings more design possibility and flexibility.

3.2 Method

3.2.1 System set-up

To super-impose two or multiple images into one, waveguide, freeform optics or relay optics can be used as effective ways to achieve the overlapping. Here, to demonstrate the method, the simplest but elegant way was taken, which is a half-mirror combiner. A half-mirror is not only easy to fabricate, but also brings less aberration to the system, which is very important during the calibration process. **Figure 3.1** shows the schematic drawing of the proposed system. Two display panels are used in this design, one on the rear, shown as Panel A in the figure, directly across the eye; another on the top side of the system, shown as Panel B. As human eyes have nearest focus distance of 6.5cm [141], it is necessary to introduce an eyepiece for near distance viewing, same as traditional NED. Display panel A and B are placed perpendicular to each other in this case. A half-mirror is inserted between display panel A and the eyepiece with an angle of 45 degrees. The positions of panel A panel B and half-mirror are carefully designed so that the two display panels are symmetric across the half-mirror. The illuminating rays from the pixels on both display panels are reflected and transmitted through the half-mirror to the user's eyes, respectively. As a result, the image from display panel A is still visible to the user. Meanwhile, the reflected virtual image of display panel B is expected to overlap with the real

image of display panel A on the rear, to achieve display overlapping. The transmission/reflection ratio of the half-mirror should be 50/50, in this case, to achieve two images with identical brightness.

Figure 3.1 Schematic drawing of the overlapping near-eye display system with two display panels.

However, in real practice, it is impossible to make the images of display panel A and B fully overlap. In other words, misalignments are everywhere during the assembly due to manufacturing tolerance. On the one hand, alignment error may cause a challenge during image overlapping and super-resolution, which makes calibration more critical. On the other hand, sub-pixel level alignment offset is one of the necessities for resolution enhancement. Here the key is to control the misalignment and take advantage of it with the calibration technique.

Figure 3.2 Positions of two display panels and the source image.

We can set up a two-dimensional Cartesian coordinate on the same plane of the rear display panel A. The x-y axis can be defined based on the viewing direction of NED user. Here we assume the virtual image from display panel A is in the same plane of display panel B, as in a NED system, the effect of out-of-plane tilt error is minor, compared to in-plane alignment error, e.g., shift and rotation, which are our biggest concern in this case. Figure 3.2 shows the positions of two display panels and the source image. In the Cartesian coordinate system defined, the image of each display panel is subject to rotation and shift error, caused by manufacturing tolerance of each component.

We define the rotation of image A and image B to be R_1 and R_2 , respectively.

The rotation behavior of the two sub-frame images can be described as:

$$R_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix} \quad (i = 1,2) \quad (3.1)$$

where angle θ is the angle of rotation of each separate image from the different display panel.

The translation vector of each frame is defined as:

$$T_i = \begin{bmatrix} Tx_i \\ Ty_i \end{bmatrix} \quad (i = 1,2) \quad (3.2)$$

where Tx and Ty are the translation of the images in the coordinate.

The perceived image by the NED user is an overlap combination of the two sub-frame low-resolution images. Thus, to display a proper combined image, the content on each display panel should be calculated. Specifically, to achieve resolution enhancement, i.e. super-resolution, the image generation of each display panel should be discussed. It should be noted here the scales of rotation and translation in Figure 3.2 are exaggerated to demonstrate the effect. In real practice, the alignment error can be invisible to the user. However, the alignment errors are usually on the same scale or larger than the size of a single pixel.

Figure 3.3 Low-resolution sub-frame images generate to display on (a) Display panel A, (b) Display Panel B.

Figure 3.3 shows the image content to be display on display panel A and B for resolution enhancement. It can be observed that, because display panel B is subject to an exaggerated rotation and translation, the image content has to counter that transformation, result in a content image with clockwise rotation. The combined image from two display panels is simulated and shown in Figure 3.4 (a).

(a)

(b)

(c)

Figure 3.4 (a) Simulated overlapping result of two display panels, (b) Zoom-in of source image at the native resolution of a single display panel. (c) Zoom-in of the result of overlapping.

To make an easy comparison, the source image was downsampled to the native resolution of a single display panel. To make the result easy to observe, magnified images of the downsampled image and the overlapping result are shown in Figure 3.4 (b) and (c). The overlapping result shows clear enhancement in the perceived resolution.

Super-resolution utilizes multiple low-resolution sub-frame images to superimpose high-resolution perceived image. The idea of super-resolution can be found in various applications, including inkjet printing, image capturing, and

projector display. The overall procedures of display super-resolution are image decomposition (i.e., sub-frame generation) and image superimposition.

The image decomposition process is also called the rendering process or sampling process. In the decomposition procedure, sub-frame low-resolution images are calculated from the original high-resolution image source. As the inverse process of camera super-resolution problem, the goal of sub-frame generation of display super-resolution is to get the best possible low-resolution sub-frames with which the original image can be well reproduced.

For NED system, the display panel components are subject to shift, rotation, or tilting. Similar to the super-resolution in other optical systems, when it comes to sub-frame generation, the alignment of display panels in NED system is an important concern. The pixel size of the display inside NED system is much tinier than the pixel size can be found in a larger display system. For a 5.5 inch display panel used in NED, the long edge is 12.2cm. If the resolution is 2560×1440 , then the pixel size is around 0.05mm. To achieve good pixel alignment, the mechanical fabrication and assembly accuracy need to be controlled under 0.05mm or even less, which is challenging. Depend on the system design, the possible alignment design of the display panels in a NED system can be divided into two categories: active or passive. For active alignment design, we assume the display panel can be accurately positioned to achieve the expected alignment. The alignment of the display panels can also be tuned to get different performance levels. On the other hand, if the alignment of the display panels cannot be accurately controlled to a sub-pixel level, it can be considered as passive alignment design. With passive alignment, although the positioning cannot be tuned, we want to extract most super-resolution performance out of the package. Thus, regardless of the alignment method, the behavior of super-resolution at different alignment conditions should be studied to provide more consistent resolution enhancement.

An ideal model of critical uniform sampling or recurrent nonuniform sampling may not be the best model. As a generalized model, pixel sampling tends to become non-uniform. We consider a general image model with n sub-frames

where the super-resolution process consists of reconstruction (up-sampling), shift and pre-filtering:

$$X' = R \sum_{k=1}^n \Delta_k S^T x_k \quad (3.3)$$

where X' is the superimposed image by n sub-frames x_k , shifted on sub-pixel level by Δ_k , reconstructed by R and sampled by transposed sampling operator matrix S^T .

Thus, to get the best sub-frame with the best alias cancellation, we need to minimize the difference between the original image and reconstructed image, which is given by:

$$\varepsilon = \|X - X'\| \quad (3.4)$$

where X stands for the original source image. Thus, the best x_k for super-resolution is obtained when the error ε is minimized.

The image superimposition process is also called the image formation process. The image formation process is more straightforward in this case, compare to the image decomposition process. In a NED system, an optical solution must be used to let images superimpose.

Figure 3.5 shows the procedure of super-resolution operation described above with simulation results. A 100 pixels square source image is used to generate 2 image sub-frames at a resolution of 50 pixels. Here a half-pixel diagonal offset is assumed. The sub-frame images look rough with big grains and alias. The resolution of the simulated superimposed image is well enhanced. The details are well improved, as can be observed at the hat decoration. However, the perfect half-pixel offset is almost impossible in practical applications. Thus, calibration plays a very important role in super-resolution.

Thus, in the following sections, the detailed methods of image generation, overlapping and calibration will be discussed.

Figure 3.5 Operation procedure of display overlapping. Start from a high-resolution image, sub-frame low-resolution images are generated and optimized to get the best results of overlapping. In the NED system, perceived image is formed by physically overlapping the sub-frame images.

3.2.2 Image generation

In previous researches of projector system super-resolution, various methods and principles were developed and validated, which are good references for our trial of overlapping super-resolution in NED.

To understand the behavior of pixel signals, we start from the most basic example of a one-dimensional signal train. In real practice, a one-dimensional signal means a single line of pixels. Figure 3.6 shows a one-dimensional continuous swept-frequency cosine signal in continuous format and a corresponding greyscale format. The swept-frequency cosine function can be expressed as $x(t) = \cos(\theta(t))$, where $\theta(t)$ is the instantaneous phase of the

chirp signal. The spread function of the signal in the frequency domain is shown in (c), denoted as $P(f)$.

We will use this signal to explain the image sampling process and the reconstruction process on a single display panel.

Figure 3.6 (a) Analog one-dimensional signal, (b) Corresponding signal in greyscale representation, (c) Corresponding function of the signal in the frequency domain.

In a display panel device, all images are delivered by the illumination of different pixels. Thus, the signal must be sampled and pixelized to be displayed. Here we define the size of the pixel to be T , which stands for the distance between neighbor pixels on the display panel. Based on this assumption, Figure 3.7 shows the sampling signal, where (a) is the plot of the sampling signal, which is a comb function with period T , (b) is the pixel elements of the one-dimensional display device and (c) is the sampling function in the frequency domain.

Figure 3.7 (a) Sampling signal, which is a comb function with period T . (b) Pixel format of the one-dimensional display. (c) sampling function in the frequency domain is also a comb function with period $f_s = \frac{1}{T}$. (d) Sampled signal $C(f)$ is a replication of the original spectrum in the frequency domain.

With Fourier transformation, we know that the sampling function in the frequency domain is also a comb function with period $f_s = \frac{1}{T}$. The period of the delta function in the frequency domain is exactly the frequency of the sampling signal in spatial domain. The full representation of the frequency function is:

$$S(f) = \sum_{n=-\infty}^{+\infty} \delta(f - nf_s) = \sum_{n=-\infty}^{+\infty} \delta\left(f - \frac{n}{T}\right) \quad (3.5)$$

According to Nyquist signal sampling, the signal must be bandlimited to $\frac{f_s}{2}$, because of the information-carrying ability of the sampling signal f_s .

To get the greyscale signal for the one-dimensional display, the original one-dimensional signal is sampled. The sampling process can be described as a

convolutional process between the continuous original signal and the sampling comb signal. As a result, the sampled signal $C(f)$ is a replication of the original spectrum in the frequency domain. Thus, the sampling result can be denoted as:

$$C(f) = P(f) \otimes S(f) = P\left(f - \left\lfloor \frac{f + \frac{f_s}{2}}{f_s} \right\rfloor f_s\right) \quad (3.6)$$

where $\lfloor \cdot \rfloor$ is a floor function and $\otimes$ is convolution operation, to make replication of the original function.

After sampling, the signal is reconstructed to reflect the real behavior of a display device. Here we define a kernel function of the display device, which is a spread function of every single pixel, shown in Figure 3.8. In real practice, this pixel spread function can sometimes be very dependent on the fabrication process and display device design, where the illumination behavior is very diverse.

Figure 3.8 Kernel function of the display device.

Here for the simplest way, we define the kernel function $K(f)$ as a bandlimited function to $\frac{f_s}{2}$, same reason as before. Then, the original frequency spectrum can be reconstructed as:

$$R(f) = K(f) \otimes C(f) \quad (3.7)$$

where $R(f)$ is the reconstructed signal. As the frequency information within $\frac{f_s}{2}$ of function $C(f)$ is already known, the frequency information of reconstructed signal $R(f)$ can be solved. The information of reconstructed signal $R(f)$ is shown in Figure 3.9. We can see that the shape of the original signal is reconstructed with alias behavior, limited by the nature of pixels.

Figure 3.9 (a) Reconstructed signal $R(f)$ from sampling, (b) Corresponding greyscale image of the reconstructed signal, (c) Signal reconstructed by kernel function in the frequency domain.

3.2.3 Image formation from two overlapping images

In the last section, we discussed the procedure of image generation for a one-dimensional signal. In this section, the procedure is completed by followed overlapping image formation. For a one-dimensional signal, shifting is the only transformation available. Here we define the amount of shift to be l . For repeated pixel patterns, when shift displacement is larger than the pixel pitch, the behavior of the display is repeated with the last shift cycle. Thus, here we define $l \leq T$, where T is the pixel pitch defined in the previous section. The sampling behavior of two overlapping one-dimensional images can be described by two periodic comb function. Here we still use the swept-frequency cosine signal denoted as $P(f)$ used in the previous section.

We start from a special occasion when the offset $l = \frac{T}{2}$, which means the shift is exactly half a pixel. Under this special scenario, the sampling signal is evenly distributed, with a sampling frequency at the doubled original frequency. The

amplitude of the overlapping signal will be increased to double the intensity, considering the illuminance of the two display panels are added together. Simply overlapping two pixels together, the intensity turns to double the original one. Thus, the frequency spectrum of the overlapped result can be expressed as the following function:

$$S_o(f) = 2 \sum_{n=-\infty}^{+\infty} \delta(f - 2nf_s) \quad (3.8)$$

Based on our discussion in the last section, the convolution result can be expressed as:

$$C_o(f) = P(f) \otimes S_o(f) = P\left(f - \left\lfloor \frac{f + f_s}{2f_s} \right\rfloor 2f_s\right) \quad (3.9)$$

where $\lfloor \cdot \rfloor$ is floor operation function, which is used to deal with the frequency spectrum replication. The signal then can be reconstructed as:

$$R_o(f) = K_o(f) \otimes C_o(f) \quad (3.10)$$

where $K_o(f)$ is the kernel function of the reconstructed signal. In this case, for the overlapping image, the kernel function $K_o(f)$ remains same with original signal.

As the frequency information within f_s of function $C_o(f)$ is already known, the frequency information of the reconstructed signal $R_o(f)$ can be solved. Thus, using a kernel function with a width f_s , the reconstructed overlapping signal has twice the frequency compared to the original one display signal.

The graphical demonstrations of the operations are shown in Figure 3.10, where figure (a) shows the frequency spectrum of the continuous signal, figure (b) shows the frequency spectrum of the sampling signal with a period of the comb function at $2f_s$, figure (c) shows the frequency spectrum of the convolution signal $C_o(f)$.

Figure 3.10 Explanation of overlapping mechanism for the special case when pixel offset value $l = \frac{T}{2}$. (a) The frequency spectrum of the continuous signal. (b) The frequency spectrum of the sampling signal, the period of the comb function is $2f_s$. (c) Frequency spectrum of the convolution signal. Yellow circle: no contamination from adjacent replica at a frequency f_s during overlapping.

For a general scenario where pixel offset value $l \neq \frac{T}{2}$, the sampling signal of overlapping in the frequency domain can be expressed as:

$$S_o(f) = 2S(f) \cos^2 \pi f l = 2 \sum_{n=-\infty}^{+\infty} \delta(f - n f_s) \cos^2 \pi f l \quad (3.11)$$

The graphical demonstration for general cases where pixel offset value $l \neq \frac{T}{2}$ is shown in Figure 3.11. Here, figure (a) shows the frequency spectrum of the continuous signal. Figure (b) shows the frequency spectrum of the sampling signal in equation 3.11.

When it comes to the convolution process, it should be noted that the frequency spectrum of the overlapped signal is contaminated at the frequency f_s , as a result of the non-zero value of the sampling signal in Figure 3.11 (b).

The convolution signal for general overlapping can be expressed as:

$$C_o(f) = 2P \left(f - \left\lfloor \frac{f}{f_s} \right\rfloor f_s \right) \cos^2 \pi \left\lfloor \frac{f}{f_s} \right\rfloor l + 2P \left(f - \left\lceil \frac{f}{f_s} \right\rceil f_s \right) \cos^2 \pi \left\lceil \frac{f}{f_s} \right\rceil l \quad (3.12)$$

The graphical plot of the frequency spectrum of the convolution signal is shown in Figure 3.11 (c). As we can see in the yellow circle, at the frequency f_s , there is contamination of the spectrum.

Figure 3.11 Explanation of overlapping mechanism for general cases where pixel offset value $l \neq \frac{T}{2}$. (a) The frequency spectrum of the continuous one-dimensional signal. (b) The frequency spectrum of the sampling signal. (c) The frequency spectrum of the convolution signal. Yellow circle: contamination from adjacent replica at the frequency f_s during overlapping.

In this general case, we can conclude that although the sampling of overlapping can be improved, but not as good as the special case when $l = \frac{T}{2}$, where a doubled sampling can be achieved, i.e., doubled resolution. However, in the proposed design of NED, controlling the pixel offset on purpose can be very hard. Thus, we consider a computational optimization method for the case of two-dimensional image overlapping.

3.2.4 Generic computational rendering algorithm for two-dimensional image and multiple display panels

In real practice, to achieve a designed offset in a NED system is sometimes not feasible, due to alignment error and manufacturing tolerance. To solve this problem, we would like to propose an optimization method for general two-dimensional image overlapping.

The feasibility of more than two display panels has been demonstrated by previous researches [142] [143]. Recall the two-dimensional Cartesian coordinate system we defined in section 3.2.1. The positions of the display panels can be described by rotation matrices $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_k$ and translation vectors $\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_n$. To get the accurate position information of the multiple display panels, calibration methods are used, which will be discussed in detail in the next section. Here we discuss the image generation method first.

To get the sub-frame low-resolution images for overlapping, we must have the input information, which is a high-resolution source image. Ideally, the resolution of the input image should be as high as possible for better mapping of pixels during image generation. Basically, an image resolution twice as each individual display panel is preferred to get a resolution enhancement. The input source image is used as a target image for optimization.

During the image generation process, the target image is put at the world coordinate with the edge of the source image parallel to the coordinate axes of the world coordinate. We denote the resolution of the source image to be M horizontally and N vertically. The pixel size of the source image is defined as P^w_x horizontally and P^w_y vertically. Thus, using the rotation matrices and translation vector, the pixel position of the overlapping image can be mapped to the source image with the relationship below:

$$\begin{bmatrix} x^k \\ y^k \end{bmatrix} = \mathbf{R}^k \begin{bmatrix} \left(i - \frac{M}{2}\right) p_x^w \\ \left(j - \frac{N}{2}\right) p_y^w \end{bmatrix} + \mathbf{T}^k \quad (k = 1, 2, \dots) \quad (3.13)$$

$$\begin{cases} i^k = \text{round}\left(\frac{x^k}{p_x^k} + \frac{M^k}{2}\right) \\ j^k = \text{round}\left(\frac{y^k}{p_y^k} + \frac{N^k}{2}\right) \end{cases} \quad (k = 1, 2, \dots) \quad (3.14)$$

Here, (i, j) is the pixel coordinate of the pixel value on the source image, (i^k, j^k) the corresponding pixel coordinate on the k -th sub-frame display panel. P_x^1 and P_y^1 are the pixel size of display panel A, while P_x^2 and P_y^2 are the pixel size of the display panel B. The pixel size of each individual display panel can be different, this optimization method can offer more flexibility in the real system. (x^k, y^k) are the world coordinates which are used as a media to perform the pixel mapping.

As we discussed before, the error between the greyscale value of the overlapping result and source image should be minimized. This relationship can be expressed as:

$$\begin{aligned} \underset{i=1, j=1}{\operatorname{argmin}} \sum_{i \leq M, j \leq N}^2 & \left\| I_{i,j}^w - I_{i^1, j^1}^1 - I_{i^2, j^2}^2 - \dots - I_{i^k, j^k}^k \right\| \\ \text{subject to } & 0 \leq I_{i^1, j^1}^1, \dots, I_{i^k, j^k}^k \leq 1 \end{aligned} \quad (3.15)$$

In this equation, the variable $I_{i,j}^w$ stands for the greyscale value of the source image at the location (i, j) , while I_{i^1, j^1}^1 to I_{i^k, j^k}^k are the greyscale values for display panel 1 to display panel k at location (i^1, j^1) to (i^k, j^k) , respectively.

With the proposed optimization method, a simulation result of two-dimensional image overlapping is shown in Figure 3.12. Figure (a) shows one display image at a comparably low resolution, limited by the number of pixels of the display sampling. Figure (b) shows the overlapping result of two images at the same resolution with figure (a) with a shift offset of one-fourth of pixel width. Figure (c) is the optimum overlapping result with shift offset of half-pixel width. Figure (d) (e) and (f) are the magnified images of the corresponding image (a), (b) and (c),

respectively. As we can see from the figure, compared to the one display low-resolution image, overlapping brings a fairly amount of improvements to the overall resolution. Even a quarter offset overlapping image reproduces more details, compare to the original one display image. However, the best result comes when the offset is exactly half a pixel, validates the theory analysis proposed.

Figure 3.12 Overlapping simulation of two-dimensional images with two display panels. (a) one image with low resolution. (b) Overlapping result of two images at the same resolution with figure (a) with shift offset of one-fourth of pixel width. (c) The overlapping result with shift offset of half-pixel width. (d), (e) and (f) are the magnified images of the corresponding image (a), (b) and (c), respectively.

3.2.5 Display calibration

In practical applications, the display may not be perfectly aligned. Display panels may subject to shift, rotate and even out-of-plane tilt. It is necessary to calibrate the display panel so that the low-resolution sub-frame pixels can be correctly mapped to the high-resolution image space to display the right high-resolution

image. In NED, the pixel size is small, which means the calibration method must be accurate to the sub-pixel level.

Multi-frequency phase unwrapping method takes advantage of phase shift interferometry to re-construct the structural information of the targeted object [144]. Various fringe projection profilometry (FPP) methods are discussed, which can be applied to display calibration. Comparing to the decoding of 3D surface features in fringe projection profilometry, in display calibration, there is no requirement of a dedicated projector light source. The display panels can be used to present high quality calibration patterns. Instead of calculating distorted phase information on the surface, pixel positions and alignments are concerned. High accuracy can be achieved down to pixel level thus become a proper way to calibrate display panel position in our case.

During the calibration process, multiple phase shift interferometry patterns are used where the intensity is given by [145]:

$$I_n(x, y) = I_0(x, y) + I_1(x, y) \cos [\varphi(x, y) + \frac{n-1}{2} \pi] \quad (3.16)$$

where I_0 is the base intensity level, I_1 is intensity modulation, $\varphi(x, y)$ is the phase we want to solve. n is the steps of phase shift, which is a positive integer. Here we take total shift steps to be 4 to solve 3 unknowns. Thus, the phase can be solved as:

$$\varphi(x, y) = \arctan \left(\frac{I_4 - I_2}{I_1 - I_3} \right) \quad (3.17)$$

The interferometry pattern consists of multiple stripes, which means the solved phase information is periodic. To make the phase information unique at every point, multiple frequency patterns can be temporally applied to form heterodyne wavelength:

$$\lambda_{12} = \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2} \quad (3.18)$$

where λ_{12} is the wavelength of the low-frequency phase function as demanded, λ_1 and λ_2 are the wavelength of the original high frequency of the patterns. Multi-frequency heterodyne can be repeatedly performed to obtain a low

frequency, so that phase information is spatially unique while maintaining the accuracy of high-frequency interferometry pattern. Thus, to further enhance the calibration accuracy, 3 stripe patterns with different frequencies are used where heterodyning can be performed 3 times to produce a continuous phase wrapping, as shown below:

Figure 3.13 (a) Multi-frequency phase unwrapping strategy diagram. (b) Phase unwrapping results from intermediate phase λ_{12} and λ_{23} to the final phase λ_{123} . The wavelength of the final phase should be the width of the display panel.

Here, λ_1 , λ_2 and λ_3 are the wavelength high-frequency base calibration stripe patterns to be physically applied to the target object, which ensures

good accuracy. Then, λ_1 and λ_2 are heterodyned to be λ_{12} while λ_2 and λ_3 are heterodyned to be λ_{23} . λ_{12} and λ_{23} are further heterodyned into λ_{123} . The value of the base frequency patterns should be selected so that the wavelength of the final phase should be the width of the display panel. which means a unique and continuous phase across the whole frame from 0 to 2π .

For the display panel, instead of projecting external structural light, we can simply show the calibration pattern on it and capture. Sinusoidal intensity calibration patterns were generated in groups of 4. Within each group, the stripes are shifted by $\frac{\pi}{2}$ in 4 steps, while 3 groups of patterns with different frequencies are used to get heterodyning. Phase unwrapping is performed independently on horizontal and vertical directions. Thus, 26 calibration patterns are used for one display panel, including one solid black and one solid white. Figure 3.14-19 shows the patterns we use as λ_1 , λ_2 and λ_3 for multi-frequency heterodyne phase unwrapping on both horizontal direction and vertical direction. Each stripe pattern has 4 phase shift variations. Figure 3.19 shows an additional pair of full black and full white pattern for brightness intensity normalization.

Figure 3.14 Sinusoidal stripe patterns used for phase unwrapping on horizontal direction, with wavelength λ_1 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.15 Sinusoidal stripe patterns used for phase unwrapping on horizontal direction, with wavelength λ_2 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.16 Sinusoidal stripe patterns used for phase unwrapping on horizontal direction, with wavelength λ_3 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.17 Sinusoidal stripe patterns used for phase unwrapping on vertical direction, with wavelength λ_1 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.18 Sinusoidal stripe patterns used for phase unwrapping on vertical direction, with wavelength λ_2 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$.

Figure 3.19 Sinusoidal stripe patterns used for phase unwrapping on vertical direction, with wavelength λ_3 . (a), (b), (c) and (d) have phase step difference of $\frac{\pi}{2}$. (e) Full black and (f) full white pattern to normalize the greyscale value.

The phase information $\varphi(x, y)$ contains the accurate position information of the display panel. The phase unwrapping solves the phase of the object, which in our case the display panel. The phase sets up a mapping between the camera pixel and the world coordinate of the display pixel. For a given display pixel location, a camera pixel location can be uniquely mapped, as shown in Figure 3.20. Here pinhole camera model is applied to simplify the calibration calculation. We suppose the world coordinate of the selected pixel point on the display panel is (x_w, y_w, z_w) , the camera coordinate of the pixel point is (x_c, y_c, z_c) , the center point of the camera image sensor is (u_0, v_0) , the projection point on the image sensor is (u, v) .

Figure 3.20 Schematic diagram of a display calibration setup. By capturing the calibration pattern, a mapping can be established between camera image sensor pixel and display panel pixel.

The mapping from camera coordinate to the image sensor is given by:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} p_x & 0 & u_0 \\ 0 & p_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} \quad (3.19)$$

where p_x and p_y is the pixel density (pixels/mm) of the image sensor.

The transformation between world coordinate and camera coordinate can be expressed as

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (3.20)$$

where $\mathbf{R}$ and $\mathbf{T}$ are 3×3 rotation matrix and 3×1 translation vector, respectively. [146]

Using equation (3.19) and (3.20) we can derive the relationship between the world coordinate and the projection point on the image sensor:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f \cdot p_x & 0 & u_0 & 0 \\ 0 & f \cdot p_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (3.21)$$

where f is the focal length of the camera. Now we want to establish the relationship between the absolute phase and the world coordinate. To simplify the process, we assume the display panels are perpendicular with the optical axis of the camera, where off the plane tilting is ignored. Thus, z_c becomes a constant while z_w becomes 0. The world coordinate is defined as the display panel coordinate, to get the rotation matrix and translation vector of the display panel. Then we can derive the relationship of world coordinates with phase:

$$\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} L_x \cdot p_x & 0 & 0 \\ 0 & L_y \cdot p_y & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\varphi_x}{2\pi} \\ \frac{\varphi_y}{2\pi} \\ 0 \end{bmatrix} \quad (3.22)$$

where L_x and L_y are the resolution of the display panel on horizontal and vertical directions. φ_x and φ_y are the final phase of the selected point.

In the experiment, multiple points are selected across the display panel to make the computation more accurate. The choice of the points is not limited to any special location. Arbitrary points can be selected on the display panel. One of the possible solutions is to calculate via four corner points of the display panel. The corner point C_1 to C_4 can be determined by finding the point where the x or y phase equal to 2π or 0, respectively. The center point can be derived when both phases on x and y directions equal to π . More points will make the computation results converge better. For each point with specific phase value, equation 3.19-3.22 can be use to solve the rotation matrix and translation vector.

Based on the discussion above, the rendering method of general two-dimensional overlapping are proposed as follows:

Step 1:

Display the calibration patterns on the display panels and perform display calibration procedure to get the position information of each display panels.

Various calibration methods can be used. Here we used a multi-frequency heterodyne phase unwrapping method.

Step 2:

With the calibration result, derive the corresponding rotation matrices and translation vectors for both display panel. Use the information to obtain accurate mapping between the source image and sub-frame image.

Step 3:

Use the pixel mapping obtained in Step 2 to perform computational optimization. Output optimized results.

Step 4:

Display the output image on the display panel and evaluate results.

3.3 Experiment setup

According to the principle of display super-sampling and super-resolution, at least two pieces of display need to be superimposed for the operation of super-resolution. Based on the theoretical discussion above, a NED prototype was designed with two pieces of 5.0-inch liquid crystal display (LCD) display panels available from the market with reasonable price at 80 Singapore dollars (SGD) each set together with driving circuit. The display panels are separately placed perpendicular with each other. The image from display panel on top of the NED is reflected by 90 degrees by a semi-transparent mirror with reflection enhancement coating to be superimposed with another display. The mirror is placed diagonally in between the two display panels with 45 degrees of the lean angle so that the reflected image can be exactly aligned with another display. Figure 3.21 shows the schematic demonstration of the design.

Figure 3.21 Schematic design of the dual-display NED.

The display panel is chosen from a list of commercialized liquid crystal display panels. To ensure the best pixel shape and size match during the test, two identical display panels are chosen. To better demonstrate the effectiveness of super-resolution, the proposed display panel is measured 5 inches diagonally while the resolution is 800 pixels horizontally times 480 pixels vertically with RGB subpixels aligned vertically. The display comes with a driver circuit board, therefore, can be connected to a PC with High Definition Multimedia Interface (HDMI) connector. Details of the display panel are shown in Table 3.1:

SPECS	DATA
MODEL	HSD050IDW1-A
VENDOR	Hannstar Display Corp.
MODULE DIMENSION	114.8mm x 73.2mm x 1.43mm
ACTIVE AREA DIMENSION	108.0mm (H) x 64.8mm (V) (5.0 inch diagonally, 15:9)
RESOLUTION	800 x 3 (H, RGB) x 480 (V)
DISPLAY MODE	Normally white
PIXEL PITCH	0.135mm
INTERFACE	HDMI/VGA

Table 3.1 Specification of the liquid crystal display (LCD) display panel for overlapping NED prototype construction.

A semi-transparent mirror is designed to super-impose images generated by two display panels. An ideal coating of the semi-transparent mirror is expected to split reflection over transmission into 50%:50% ratio across all wavelengths, with which two image sub-frames can be at identical brightness and colors. The semi-transparent mirror was split into two identical pieces to accommodate the left and right viewing zone as well as reducing the cost of manufacturing.

A NED case was designed to hold the optical setup for super-resolution, based on the principle shown in Figure 3.1. The external shell of the NED was designed with proper dimension to fit in both the display panel and semi-transparent mirror. While the display panel on top was fixed in the main NED body, the rear panel can be rotated and shifted upon assembly, allow more adjustment of screen alignment. The left and right viewing zone were separated to avoid unnecessary crosstalk. A front cover was designed with eyepieces. The interpupillary distance, in this case, was set to be the human average interpupillary distance of 64mm. The eyepiece lens was expected to have a focal length of ~100mm, to accommodate the larger distance between the eyepieces to the rear display panel, considering the 45 degrees mirror structure and the display panel on top. The NED was modeled in 3D CAD software Solidworks. With this design, two

sub-frame low-resolution image can be physically combined for the operation of calibration and super-resolution. Figure 3.22 and 3.23 shows the appearance of the NED prototype design, which consist of 3 main parts: from front to rear, a front panel that holds eyepieces, the main body that holds two pieces of the display panel and semi-transparent half-mirror, and a rear attachment box to hold the driving circuit.

Figure 3.22 The near-eye display prototype CAD design appearance.

Figure 3.23 shows the cross-section view of the prototype design. As we can observe, the eyepieces are attached to the front panel on the left. The 45 degrees mirror is installed at the middle, with the display panels on top and rear.

Figure 3.23 The near-eye display prototype CAD design in a cross-section view. The red rectangle stands for the position of display panel 2. The yellow rectangle is the position of display panel 1. The light blue rectangle stands for the position of semi-reflective mirror. The deep blue oval represents the eyepiece. Green area is the driving circuit box. The object distance between the display panel and eyepiece is 95mm, result in an image distance at 1900mm.

To accommodate the display panel on the top portion of the NED and the half-mirror, the distance between the front eyepiece and the display panel on rear is designed to be 95mm. An eyepiece with a focal length of 100mm. According to simple lens equation: $\frac{1}{f} = \frac{1}{o} + \frac{1}{i}$, where f , o and i are focal length, object distance and image distance, respectively. With f equal to 100mm, o equal to 95mm, we can calculate i equal to -1900mm, which indicates a virtual image at distance 1.9 meters.

The prototype was fabricated by 3D printing process and painted into gloss black to reduce unnecessary reflection. The two display panels were connected to a computer via HDMI Cable. The optical components were assembled into the system.

During the experimental process, a digital compact camera Nikon Coolpix A was used for calibration and result image capturing. This camera has a large image

sensor (APS-C) format without optical low-pass-filter (LPF), which is helpful to perform the evaluation on the pixel level.

The video capturing task during the experiment was performed by Panasonic GH4 mirrorless camera for better video quality.

3.4 Results

Figure 3.24 shows the experimental system, including the designed NED system and a digital camera for calibration and image capturing.

Figure 3.24 NED prototype and experimental set-up. Digital compact camera Nikon Coolpix A was used for calibration and image capturing.

Based on the previous discussion, to render the low-resolution image for each display panel, the first step should be calibrating the NED system to get position information of two display panels. A calibration image set consists of 26 sinusoidal intensity patterns on both horizontal and vertical directions were displayed on each panel, to perform multi-frequency heterodyne phase unwrapping. The displayed image patterns were captured by the camera and processed to calibrate the display panels prior to super-resolution. It should be noted that during the process of acquiring 52 images, components of the experimental system, including cameras and NED, should avoid possible

displacement or impact to the system. **Figure 3.25** shows the captured calibration patterns on each display.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

(k)

Figure 3.25 Complete captured calibration photo set consists of 52 photos of sinusoidal intensity patterns. (a) 4 step phase shift patterns with wavelength λ_1 on horizontal direction for display panel 1. (b) 4 step phase shift patterns with wavelength λ_2 on horizontal direction for display panel 1. (c) 4 step phase shift patterns with wavelength λ_3 on horizontal direction for display panel 1. (d) 4 step phase shift patterns with wavelength λ_1 on vertical direction for display panel 1. (e) 4 step phase shift patterns with wavelength λ_2 on vertical direction for display panel 1. (f) 4 step phase shift patterns with wavelength λ_3 on vertical direction

for display panel 1. (g) Full black and full white for display panel 1. (h) 4 step phase shift patterns with wavelength λ_1 on horizontal direction for display panel 2. (i) 4 step phase shift patterns with wavelength λ_2 on horizontal direction for display panel 2. (j) 4 step phase shift patterns with wavelength λ_3 on horizontal direction for display panel 2. (k) 4 step phase shift patterns with wavelength λ_1 on vertical direction for display panel 2. (l) 4 step phase shift patterns with wavelength λ_2 on vertical direction for display panel 2. (m) 4 step phase shift patterns with wavelength λ_3 on vertical direction for display panel 2. (n) Full black and full white for display panel 2.

With the calibration images captured by the camera, the information of the display panel can be extracted. The rotation matrix and translation vector can be calculated. For this case of calibration pattern,

$$\mathbf{R}_1 = \begin{bmatrix} 0.9993 & 0.0174 \\ -0.0273 & 1.0121 \end{bmatrix}$$

$$\mathbf{R}_2 = \begin{bmatrix} 1.0002 & -0.0172 \\ 0.027 & 0.9875 \end{bmatrix}$$

$$\mathbf{T}_1 = \begin{bmatrix} -0.0386 \\ 0.2232 \end{bmatrix}$$

$$\mathbf{T}_2 = \begin{bmatrix} 0.0424 \\ -0.2194 \end{bmatrix}$$

The experimental calibration results are shown in Figure 3.26. Two rectangular mesh grids are drawn to show the position of two display images. Here, the red color grid indicates the image position of the rear display panel, while the blue color one stands for the reflected top panel image. As we can tell from the figure, the two display panels are subject to shift and rotation.

Figure 3.26 Display calibration result. Red color grid shows the position of display panel 1. Blue color grid shows the position of display panel 2. Relative positions of two display panels include translational and rotational offset.

Figure 3.27 High-resolution source image from a 16-mega-pixel DSLR camera with a high-quality lens. Image is cropped to the red area to get more details of the building.

The High-resolution target image is shown in Figure 3.27. This image is from a 16-mega-pixel DSLR camera with a high-quality lens. At a redundant resolution of 4928x3280, the image is further cropped to the red area to get more details of the building for evaluation. Super-resolution images were generated based on the calibration results above. Because the NED assembly is a handcrafted prototype, during the entire experiment, the system was prevented from any major modification or external impact to make the calibration data effective. With the optimization method proposed in section 3.2.4, images to be displayed on each display panel are rendered based on the calibration results. Figure 3.28 shows the rendered images at the native resolution of the display panel, 400x480 for one viewing channel.

Figure 3.28 Decomposed sub-frame low-resolution image at the native resolution 400x480 for each viewing channel of the display panel. (a) Image on rear display panel A. (b) Image on top display panel B.

As we can see, the decomposed images have some uneven brightness distribution. This is caused by the global optimization, where local smoothness is ignored to get global optimum result. These two images are displayed at the native resolution on the display panel, shown in Figure 3.29. Figure (a) is the image displayed on the rear display panel, transmit through the half-mirror. Figure (b) is the image displayed on the top display panel, reflected by the half-

mirror combiner. These two images are captured by illuminating one display panel at a time while keeping the other one fully black.

Figure 3.29 Decomposed sub-frame low-resolution images shown on display panels with the native resolution 400x480 for one eye. (a) The image displayed on the rear display panel, transmit through the half-mirror combiner. (b) The image displayed on the top display panel, reflected by the half-mirror combiner.

The overlapping result is demonstrated in Figure 3.30 and Figure 3.31. Figure 3.30 (a) shows the overlapping of two low-resolution display panel images, while 3.30 (b) shows the down-sized image displayed by the rear display panel at the native display resolution 400x480 for comparison. The down-size process is achieved with bicubic interpolation, which is one of the most common algorithms in image scaling. In the real usage of the NED, the field-of-view coverage of the display is very broad. Thus, an enlarged view of the overlapping image is shown in Figure 3.31 (a). Figure 3.31 (b) shows the zoom-in of the down-sized image displayed by the rear display panel at the native display resolution 400x480.

Figure 3.30 (a) The overlapping of two low-resolution display panel images. (b) The down-sized image displayed by the rear display panel at the native display resolution 400x480.

Figure 3.31 (a) Zoom-in of the overlapping result of two low-resolution display panel images. (b) Zoom-in of the down-sized image displayed by the rear display panel at the native display resolution 400x480.

As we can see, the overlapping result does not contain the discontinuous area in each decomposed image. The overlapping image achieves a good-looking

image very close to the target image. From the comparison of the overlapping results and down-sized control image, we can see an obvious enhancement in the perceived resolution. The pattern details of the window are reproduced on the overlapping image. The lines are continuous and smooth, with very little alias. In comparison, the native resized image has lots of alias on the white line, makes it discontinuous and unrecognizable. The small shadow detail on the overlapping image is well recognized. While in the controlled sample, the shadow cannot be recognized. The pixel edge on the overlapping image is also eliminated, resulting in a smoother image with less grain. In the zoom-in image, we can clearly observe the presence of a higher sampling rate.

A video test was performed to confirm the proposed method in a video application. A video was created by making pan and zoom-in motion on the high-resolution source image. The video was decomposed into separate frames and feed into the two-dimensional image rendering algorithm to generate one pair of low-resolution images for overlapping each frame. The resulted images were converted back into a video to playback at 15 frames-per-second (FPS). The video playback was captured by a video camera Panasonic GH4 at 3840x2160 resolution / 30 FPS. Figure 3.32 shows one frame from the video captured, which demonstrates the experiment results of the video capability of the NED prototype.

Figure 3.32 Video test results of overlapping NED. Decomposed low-resolution images for display panels: (a) (e) Image displayed on the rear display panel, transmit through the half-mirror combiner. (b) (f) The image displayed on the top display panel, reflected by the half-mirror combiner. (b) (g) The overlapping of two low-resolution display panel images. (d) (h) The down-sized image displayed by the rear display panel at the native display resolution 400x480.

Figure 3.32 (a)-(h) illustrate the results of video test on the proposed NED prototype. Single frame image was selected from the captured video and zoomed accordingly to make the details distinguishable. Low-resolution sub-frame images are shown in (a), (c) (e) and (f), which are rendered from a high-resolution video source. Overlapping video results are shown in (c) and (g). Figure (d) and (h) show the capture of the down-sampled video displayed by the rear display panel at the native display resolution 400x480. Very similar to the results of a static image compare to single display panel images, the overlapping images are rich in details with less alias, which become obvious judging from the shape of the line details. As a result of resolution enhancement, the device is capable of reproducing complex image pattern, while the image from a single panel is rough and blur.

A resolution charts was tested to show the enhancement of resolution using overlapping method. Figure 3.33 shows the testing results. Figure 3.33 (a) is a captured photo of perceived image from a single display panel. The original high-resolution image is resized with traditional method (bicubic) to fit the low-resolution single display panel. Figure 3.33 (d) is the perceived image generated by overlapping 2 display panels. To compare the resolution difference, each figure is zoomed-in, shown as Figure 3.33 (b) and (c). As we can observe, the line separation and clarity in the overlapping image is enhanced. The 5 lines pattern in Figure 3.33 (b) cannot be distinguished. The resolution and contrast of the line patterned cannot be reproduced by a single display panel. In comparison, the image generated by overlapping two display panels shows good line separation and contrast. This shows clear evidence of the resolution enhancement.

Figure 3.33 Resolution chart test result of overlapping NED. (a) Image displayed on single display panel. (b) Zoom-in of red square area in figure (a). (c) Zoom-in of green square area in (d). (d) Image displayed by overlapping 2 display panels.

In conclusion, the proposed method demonstrates its effectiveness in both static image and motion video. The experiment shows that the image overlapping method can bring a big improvement to the overall image quality, especially for images with an excessive amount of details. The proposed method is demonstrated with two display panels in this thesis. However, the method can

be applied to multiple display panel applications with proper optical and computational implementations. Possible overlapping method includes free-form optics and waveguide optics. The key here is how to implement a design that can fulfil multiple display channels without causing additional viewing distortions. The recommendation for future research is further discussed in Chapter 6.

3.5 Conclusion

A super-resolution method was developed together with calibration to enhance the resolution in the near-eye display. A dual display panel NED prototype was designed to demonstrate the effectiveness of image overlapping. By adding an extra display in the NED system, the perceived resolution is improved by a big margin compared to traditional one display NED setup, validated by the test for both image and video. The proposed overlapping method is more versatile, compared to existing resolution enhancement methods. It introduces no extra artifacts and can be combined with free-form optics or other overlapping implementation. This research can act as a guide to achieve high-resolution NED with low additional cost, thus has great potential applications in VR/AR with demanding resolution requirements.

Chapter 3 is published as:

Qijia Cheng, Weitao Song, Feng Lin, Yue Liu, Yongtian Wang, and Yuanjin Zheng. "Resolution enhancement of near-eye displays by overlapping images." *Optics Communications* 458 (2020): 124723.

Chapter 4 Compact off-axis near-eye display based on free-form optics

4.1 Introduction

In this chapter, a near-eye display design is proposed with off-axis lightguide. An imaging lightguide prism and a compensation prism is presented in the system. The imaging lightguide prism magnifies the image from the display source. It is formed with four freeform surfaces described by x-y polynomial functions. The light ray is then corrected by the compensation prism to yield an image with low distortion. As a result, a near-eye system is developed with a 4-mm diameter exit pupil, a 21.5° diagonal full field-of-view (FoV) and 23 mm eye relief. A near-eye display prototype is fabricated with monocular vision. Test results indicated that the proposed near-eye display has good optical performance yet compact in size. Overall, the proposed design is a good option in the application of state-of-the-art wearable virtual reality.

4.2 Method

4.2.1 System description

Figure 4.1 Cross-section of the proposed NED using two lightguide prisms combined with a doublet lens.

The basic architecture of the near-eye display system is shown in Figure 4.1. The system consists of two lightguide prisms and a doublet lens. The exit pupil is defined by the eye-pupil of the user. A coordinate system is established with the center of exit pupil as the origin. The z-axis is the direction of viewing while the y-axis along the horizontal left direction. The image from the display source is magnified to the exit pupil. The system consists of multiple refracting and reflecting surfaces. The light beam is refracted 4 times and reflected 3 times before it reaches the exit pupil.

To achieve a better shape of the NED, the display panel is designed on the side of the user's head. This design meanwhile prevents the display from blocking the user's front vision, thus enables the possibility of augmented reality. An intermediate image is formed by the doublet lens inside the prism to correct the aberrations from optical tiling and off-axis. A compensation prism is deployed with 50/50 beam splitter at surface S2.

4.2.2 Method of controlling clearances

Total internal reflection (TIR) condition is the main reason for the complexity when designing the system. Formation of an intermediate image inside the lightguide prism is the key. Besides, intercepting rays need to be optimized during optimization. All surfaces are tilted and decentered respect to the optical axis. Constraints according the physical and anatomy of the human user are required to ensure a reasonable design without any conflict with human head shape. Additional constraints are added during software optimization to meet this requirement.

Figure 4.2 Determination of the intermediate image between two surfaces.

The position of the intermediate image is calculated based on the following principle. Figure 4.2 only shows rays with zero-degree field-of-view. Ray A is the marginal ray on top of the exit pupil. k is the slope. b is the intercept. Thus, k and b are given by:

$$\begin{cases} k_u = \frac{\cos(\alpha)}{\cos(\beta)} \\ b_u = y_u - k_u \times z_u \end{cases} \quad (4.1)$$

where $\cos(\alpha)$ and $\cos(\beta)$ are the cosines of ray A on y-direction and z-direction, respectively. They can be expressed as:

$$\begin{cases} \cos(\alpha) = \frac{\mathbf{v} \cdot \mathbf{e}_y}{|\mathbf{v}|} \\ \cos(\beta) = \frac{\mathbf{v} \cdot \mathbf{e}_z}{|\mathbf{v}|} \end{cases} \quad (4.2)$$

Here, $\mathbf{v}$ is the direction vector of ray A. $\mathbf{e}_y$ and $\mathbf{e}_z$ are the unit vectors on y and z-axis. Thus, for the bottom ray at the exit pupil, namely ray B, the same equation can be given. Here we define the coordinates of the intermediate image to be (z_0, y_0) . Rays A and B must have z-intercept to be:

$$z_0 = \frac{(b_b - b_u)}{(k_u - k_b)} \quad (4.3)$$

where k_b and b_b are the slope and the y-intercept of ray B, respectively. The coordinates of the intermediate image must satisfy the following conditions, according to the special limitation of the prism:

$$\begin{cases} 21 \text{ mm} \leq z_0 \leq 32 \text{ mm} \\ 6 \text{ mm} \leq y_0 \leq 18 \text{ mm} \end{cases} \quad (4.4)$$

Figure 4.3 Determination of the incident angle of S1 and S3 to meet TIR conditions.

As shown in Figure 4.3, light ray enters the lightguide prism via surface S4. Then the light is reflected by surface S3, then S1 and S2 respectively. Surfaces S2 and S3 are tiled together. A coating is applied to surface S2 to achieve a very low loss of reflection of more than 95%.

On surface S1, the upper marginal ray at the exit pupil has the minimum incident angle. Our goal is to keep this ray fully reflected while maintaining a compact structure as well. Thus, the maximum incident angle should be limited. A large incident angle will result in a bulky design, while a small incident angle will cause light leakage. Therefore, inequality can be derived:

$$\arcsin\left(\frac{1}{n}\right) < \theta_{u1}, \theta_{u2} < \arcsin\left(\frac{1}{n}\right) + 5^\circ \quad (4.5)$$

where θ_{u1} and θ_{u2} are the incident angle of the upper marginal ray at exit pupil on surfaces S1 and S3, respectively. n is the refractive index of the material used to build the imaging lightguide prism.

Figure 4.4 Imaging prism clearance.

Figure 4.4 shows the definition of the clearance constraints of the system. P is the rear vertex of the main prism, Q is the front vertex of the doublet lens, N is the rear vertex of the doublet lens, D is the front incident point of the central ray, and M is the center point of the OLED display panel. Eye relief L_{OD} in the system is defined as the distance between the eye and the surface S1. Eye relief also provides focus adjustment for the user. A minimum eye relief value is defined as 15 mm. However, eye relief cannot be infinitely large. A large eye relief would give insufficient clearance to the system [147]. The clearance is defined as the following, where L_{AM} is the clearance from the exit pupil to the center of the display panel:

$$\begin{cases} 25 \text{ mm} \leq L_{AM} \leq 45 \text{ mm} \\ 15 \text{ mm} \leq L_{OD} \leq 25 \text{ mm} \end{cases} \quad (4.6)$$

As for the doublet lens, proper decenter and/or tilt is a possible way to enhance the image quality. Mechanical clearances are added:

$$\begin{cases} L_{PQ} > 5 \text{ mm} \\ L_{MN} < 14 \text{ mm} \\ -70^\circ < \alpha < -20^\circ \end{cases} \quad (4.7)$$

where L_{PQ} is the distance between P and Q, L_{MN} is the distance between M and N, α is the angle of tilt of the doublet lens with respect to the imaging lightguide prism.

4.2.3 Optimization of surface

In the design of the lightguide prism, the aberrations caused by off-axis optics should be noted. Generally, a freeform surface has more degree of freedom. The design routine of freeform optics also varies from traditional ones. A mathematical expression is expected to define the surface, which can rapidly converge during the optimization. Different expressions are evaluated in the design process, such as spherical surface (SPS), aspherical surface (ASS), odd polynomial surface (OPS) and x-y polynomial surface (XYP).

The aspherical surface (ASS) can be defined as:

$$z(r) = \frac{c * r^2}{1 + \sqrt{1 - (1 + k) * c^2 * r^2}} + a_4 * r^4 + a_6 * r^6 + \dots \quad (4.8)$$

where $z(r)$ is sag; c is the curvature of the surface; $r = \sqrt{x^2 + y^2}$ is the height above the optical axis. $a_4, a_6, \dots$ are constants called aspheric coefficients. k is the conic constant, which determines the surface types. When the aspherical coefficients and the conic constant all equal to zero, the surface becomes a spherical surface. In traditional aspherical surfaces, odd orders are avoided because of manufacturing and design difficulties.

The merit functions of different types of surfaces are shown in Figure 4.5. According to the figure, OPS and XYP have superior performance than the SPS and ASS. Both OPS and XYP can achieve lower merit function compare to other types.

Figure 4.5 Comparisons among different types of surface representation.

As shown by its name, OPS is the surface that contains odd order of polynomials, while in traditional aspherical surface, only even orders are used. OPS is used to initially express the freeform surface to correct the asymmetrical. The expression of OPS is denoted as:

$$z(r) = \frac{c * r^2}{1 + \sqrt{1 - (1 + k) * c^2 * r^2}} + a_1 * r + a_2 * r^2 + \dots + a_{30} * r^{30} \quad (4.9)$$

The definition of the variables same as stated in the definition of aspherical surface. $a_1, a_2, \dots, a_{30}$ are the coefficients of the odd polynomial equation used to correct aberrations.

Aberration existing in our system caused by tilt and decenter is very hard to correct with normal aspherical design. As a rotational-symmetric surface, OPS seems to be inappropriate when it comes to off-axis aberrations. Thus, XYP remains as the choice. The XYP is expressed as follows:

$$z_{(x,y)} = \frac{cr^2}{1 + \sqrt{1 - (1 + k)c^2r^2}} + \sum_{j=2}^{66} C_j x^m y^n \quad (4.10)$$

The index j is given by:

$$j = \frac{(m + n)^2 + m + 3n}{2} + 1 \quad (4.11)$$

where z, k, c , and r remain the same definition as the OPS expression. C_j ($j=2, 3, \dots, 66$) are the coefficients of $x^m y^n$ terms used to correct aberrations. m

and n are integers to indicate the order of the polynomial. Higher order coefficient C_j provides higher accuracy for the described surface. The range of C_j is determined by the simulation toolbox to be 2 to 66.

4.3 Results

4.3.1 Specifications

The display panel is a key component in the near-eye display system. A 0.38-inch OLED display is selected to achieve compact volume. Its resolution is 873 × 500 with 10 μm pixel size. Contrast ratio of a display device is defined as the ratio between the highest and lowest luminance the display can achieve. The contrast ratio of the selected OLED display panel is as high as 10,000:1, which is helpful for better image quality [148]. The brightness of the display panel can reach 500 Cd/m², thus helpful when the light is reflected multiple times. Table 4.1 shows the overall specification of the system. The optical system is designed to achieve a field-of-view of 21.5° with an exit pupil diameter of 4 mm. The effective focal length (EFL) of the system can be determined by measuring the distance between the rear focal point and the rear principle point of the system, which is about 25 mm in this design. Eye relief is set to be 23 mm to accommodate possible eyeglasses worn by the viewer.

Full-color OLED		NED system	
Items	Specifications	Items	Specifications
Screen size	0.38 in. (Diagonal)	Configuration	Off-axis configuration
Active area	8.73(W) × 5.00(H) mm	Spectrum	Photopic
Pixel size	10 μm × 10 μm	Exit pupil diameter	4 mm
Resolution	873 × 500	Effective focus length	25 mm
Power consumption	25 mW	Eye relief	23 mm
--	--	Field of view	17.2° × 13°

Table 4.1 Specifications of the OLED panel and optical parameters of the NED system.

4.3.2 Optical layout and performance

System simulation and optimization are carried using OpticStudio (ZEMAX). The modulation transfer function (MTF) of the system at spatial frequency f is defined as

$$Modulation(f) = \frac{I_{max}(f) - I_{min}(f)}{I_{max}(f) + I_{min}(f)} \quad (4.12)$$

where $I_{max}(f)$ and $I_{min}(f)$ represent the maximum and minimum intensities of the reproduced perfect solid lines patterns at spatial frequency f . The perfect solid line pattern used in the test has a sharp edge transition from full white to all black, vice versa. The MTF is a good measurement of the system's ability to transfer image contrast at different resolutions. The MTF of the proposed near-eye display is shown in Figure 4.6 (b). At low pattern frequency, the system can reproduce the spatial solid line pattern very well with a modulation contrast close to 1. As the frequency of the line pattern increases to a fairly high frequency, the modulation starts to drop, which means the system cannot fully reproduce the sharp edge transition of the perfect pattern. The resulted image will lack contrast and resolution. The absolute Nyquist frequency of the display panel is calculated as 50 cycles/mm. To compensate the pixel construction and plain area in between sub-pixels, a correction Kell factor 0.7 is applied [149], result in an effective Nyquist frequency at 35 cycles/mm, determined by the pixel size and resolution of the display panel. The tangential and sagittal MTF values for all field are greater than 40% at 30 cycles/mm, which is sufficient for the application of the visual system [147].

Figure 4.6 (a) Optical layout of the NED system; (b) MTF plot of the designed system with a 4 mm exit pupil.

Distortion is an important factor to be considered for a near-eye display system. Analysis needs to be taken to reduce the distortion of the system. Optical distortion is defined by:

$$D = \frac{y' - y'_o}{y'_o} \times 100\% \quad (4.13)$$

where y' is the distance between the center and corner of the observed image while y'_o is the counterpart of the original image, as shown in Figure 4.7 (a). The proposed system has a maximum distortion rate of 1.8%, according to the simulation, thus typically unnoticeable by human eyes [150].

Figure 4.7 (a) Definition of optical distortion; (b) Distortion grid of the designed NED optical system compared with real and paraxial rays.

Figure 4.8 (a) See-through mode of the NED system; (b) Distortion grid of the designed NED optical system caused by the imaging lightguide prism.

A 50/50 beam splitter coating is applied to surface S2. User can observe the image from the display panel and the image of the real world at the same time through the imaging prism. However, the real-world image will subject to a shift compared to the naked-eye view, if no compensation prism is applied, also indicated by the simulation result from the Figure 4.8. A maximum shift of -7.4% is suggested. A compensation prism is required to correct the shift and reduce the distortion from the imaging prism. A backward ray tracing is applied in the design of compensation prism, as shown in Figure 4.9. Instead of tracing the ray from the object to the exit pupil, ray is traced from the user's eye to the object. The focal length of human eyes is set as 22 mm [151]. An ideal lens with

22 mm focal length is thus simulated at the position of the exit to mimic human eye. The distance between the projected image and the front surface of the compensation prism is 1500 mm in this implementation. The field-of-view that allows the user to see-through is around 24.8° , which can cover the whole view of the artificial content.

Figure 4.9 Design of the compensation prism with backward ray tracing. The distance between the projected image to the user is 1500 mm in this implementation. The human eye model is illustrated with single ideal lens with focal length 22mm. The ray is traced from image plane (i.e. retina) to the pupil, then through the proposed system to the object plane. The see-through system has a field-of-view 24.8° .

The composition model of the imaging prism and compensation prism is shown in Figure 4.10 (a) and (b) below. S5 is introduced as a surface with yz-plane symmetry. The distortion is reduced from -7.4% to only 1.3% in the grid chart, which indicates the effectiveness of the compensation lens.

Figure 4.10 Design layout of the compensation prism (a) in the yz-plane; (b) in the xz-plane.

Figure 4.11 Distortion grid of the designed compensation prism compared with real and paraxial rays.

4.3.3 Analysis and performance evaluation

Total internal reflection (TIR) constraints are applied to the rays. Ray tracing method is applied to understand the incident angle of rays on the freeform surface. The Figure 4.12 shows the relationship between the exit angle at the exit pupil position and the incident beam angle of the freeform surface. The center, top and bottom rays of the exit pupil are evaluated here. The exit angle is set to range from -8.7° to 8.7° . A decrease of all the incident angles at S1 and

S3 are observed. The minimum angle is observed at top marginal ray, which is 39.5° at S1 and 39.8° at S3. For a simple reference, the material used in making the prism is Polycarbonate, which has a refractive index of $n=1.585$, suggesting a critical angle of 39.1° . Thus, the calculated result indicates that all the internal angles are larger than the critical angle which satisfy the TIR condition.

Figure 4.12 The exit angle at the exit pupil versus incident angle of TIR (a) S1 surface; (b) S3 surface. The black, red, and blue lines indicate the exit angle at the center, top marginal, and bottom marginal of the exit pupil.

The surface contour of the x-y polynomial surface is shown in Figure 4.13. The surface profile changes continuously, indicated by the smooth color change in the figure. The surface is symmetrical along the x-axis.

Figure 4.13 Surface contour of (a) S1, (b) S2, (c) S3, (d) S4 and (e) S5.

The Figure 4.14 shows the 3D model of the optical components.

Figure 4.14 3D structure of the designed optical system. (a) Top view. (b) Front view. (c) Perspective view.

Figure 4.15 (a) Dimensions of lenses; (b) OLED display panel; (c) Near-eye system assembly.

The lightguide prism and the doublet lens was fabricated to verify the proposed system, as shown in Figure 4.15. Figure (a) shows the dimension of the lightguide lens fabricated by ultra-precision diamond turning in Singapore. The doublet lens is fabricated in China. Figure (b) shows the 0.38-inch OLED panel. Unfortunately, limited by the budget, we didn't manage to fabricate the

compensation prism. Without the compensation prism, the system subject to - 7.4% of distortion. The augmented reality capability of the near-eye display prototype cannot be fully tested in the experiment. The external shell was 3D with ABS material. The prototype is assembled shown in figure (c). The prototype is in a wearable form. It is thin and light for wearable applications. The experimental setup is shown in the Figure 4.16.

Figure 4.16 Prototype appearance with the external driver circuit for the OLED panels.

As illustrated in Figure 4.17 (a) below, a sample image is used in the test. Figure 4.17 (b) shows an image captured by a camera. A smart phone camera is used in the test to mimic the pupil size of human eye. The phone camera has a lens with 5mm focal length and F number $F/1.8$. The aperture size can be calculated to be 2.8mm, which falls into human pupil size range [152]. Virtual imagery of Rubik's cubes can be viewed by the user superimposed on real-world object. Minor stray light exists on the edge of the display, which can be effectively eliminated with a proper light reduction. Because of the absence of the compensation prism, the see-through performance of the system is limited. Augmented reality mode can only be achieved with a proper setup including compensation prism.

Figure 4.17 (a) Sample image; (b) Experimental result.

4.4 Conclusion

In this chapter, a near-eye display design is proposed with off-axis lightguide. An imaging lightguide prism and a compensation prism is presented in the system. The imaging lightguide prism magnifies the image from the display source. It is formed with four freeform surfaces described by x-y polynomial functions. The light ray is then corrected by the compensation prism to yield an image with low distortion. As a result, a near-eye system is developed with a 4-mm diameter exit pupil, a 21.5° diagonal full field-of-view (FoV) and 23 mm eye relief. A near-eye display prototype is fabricated with monocular vision. Test results indicated that the proposed near-eye display has good optical performance yet compact in size. Overall, the proposed design is a good option in the application of state-of-the-art wearable virtual reality.

Chapter 4 is published as:

Zhenfeng Zhuang, Qijia Cheng, Phil Surman, Yuanjin Zheng, and Xiao Wei Sun. "A compact and lightweight off-axis lightguide prism in near to eye display." *Optics Communications* 393 (2017): 143-151.

Chapter 5 Random pinhole light field near-eye display

5.1 Introduction

Advancement of application in virtual reality and augmented reality evolves near-eye display in performance and form factors, makes it more suitable for wearable and mobile computing.

As discussed in Chapter 2, Traditional near-eye display relies on binocular parallax to provide user depth cues, which may cause accommodation and convergence conflict (A/C conflict) issue, thus, in turn, leads to visual discomfort and fatigue. A/C conflict is a major obstacle towards better user experience.

Light field display is a promising solution to A/C conflict. Microstructure array so far remains one of the simplest and feasible ways to generate light field, including using microlens array and pinhole array. Practically, the position drift of eye pupil, as well as the alignment issue of the microstructure array and the display panel, may result in the viewing zone position to vary. Both scenarios lead to the problem of repeated viewing zones, generated by the periodic nature of the microstructure array. The perceived image from neighbor viewing zone usually differs from the design position, thus causing the problem of depth acquisition. The initial target of introducing light field NED is to eliminate the visual sickness caused by A/C conflict. However, if repeated viewing zone occurs, visual discomfort is not solved but becomes even worse. A feasible solution to prevent repeated viewing zone from occurring is demanded.

Figure 5.1 Schematic explanation of the repeated viewing zones problem in light field NED. (a) The designed viewing position of light field NED with binocular vision. No depth confusion. (b) The practical situation with repeated viewing zone problem, caused by device alignment or eye position. Depth confusion occurs. (c) Generation of repeated viewing zones of periodic pinhole array light field display.

Figure 5.1 shows the schematic mechanism of the repeated viewing zones problem in light field NED. Figure 5.1 (a) shows the designed viewing position of light field NED, where depth cues are correct. Figure 5.1 (b) shows the practical situation with repeated viewing zone problem, caused by device alignment or eye position. When the user's eye falls into the neighbor viewing zone, depth confusion occurs. In this case, the user may have some difficulty to converge. Figure 5.1 (c) demonstrates the mechanism of repeated viewing

zones generated by periodic pinhole array. Because of the periodic nature of the pinhole array, ray can pass through the adjacent pinhole of the expected primary pinhole to form repeated viewing zone next to the designed primary viewing zone along the viewing plane. Here, 2 repeated viewing zones next to the primary viewing zone are drawn in the figure for demonstration. In practice, the viewing zones are continuous at different viewing angle across the whole viewing range.

As we discussed, increasing exit pupil size can be very challenging in the NED design, while adding eye-tracking or visual blocking device into a NED may greatly increase the system complexity. To solve the repeated viewing zone, we must go back to the periodic nature of the light field NED, which is the periodic arrangement of the microstructure array.

Inspired by the non-uniform barrier solution in the autostereoscopic display system [153, 154, 155], here we would like to propose a random-microstructure-based solution for light field NED. Instead of using periodic microstructure as an angular light generation device, a random microstructure can be used, which can only produce one unique viewing zone in space. Thus, there is only one unique position where the user can perceive a clear image from the light field NED, which eliminates the chance of getting the wrong image with confusing depth cues. In the following section, method and system setup are discussed.

5.2 Method

5.2.1 System set-up

Figure 5.2 (a) shows the basic schematic setup of a random-pinhole-based light field display. The system consists of three layers. From front to back, they are random pinhole mask, liquid crystal display (LCD) and backlight element. On the LCD we have the elemental image presented. As we can see, the modification from a traditional periodic pinhole array is kept at a minimum level. The only change made is the pinhole opening pattern on the film mask. It should

be noted that in a practical setup, to maintain the distance between the pinhole film and the LCD, a transparent spacer should be added.

5.2.2 Rendering method

generated with a basic two-dimensional random number generator. However, an additional restraint should be added, which is the minimum distance between adjacent pinhole. Without the minimum distance control, the pinholes are very likely to overlap. The minimum distance between adjacent pinholes also determines the eye box size of the NED, which is discussed in section 5.2.3.

Similar to the rendering procedure of integral imaging, the pinhole-based light field can be rendered with an inverse process of light field camera, which can be implemented with a ray-tracing algorithm. The element image displayed on the LCD panel can be treated as cameras sensors with perspective projection. The axes of the projection are in line with the connection from the eye box center to the pinhole center. For periodic pinhole, the element image will be periodic as well. But in the proposed random pinhole method, the back image is not periodic. The most important process of rendering is how to find the corresponding pinhole for all the pixels on the display panel.

Figure 5.3 Rendering procedure of random pinhole light field NED.

As shown in Figure 5.3, here we define point O to be the center of the exit pupil on the eye relief plane, which stands for the designed eye position of the user.

It is not compulsory for O to be at the center axis of the display panel. Instead, the position of the pupil can be defined based on the need. Here we define the pixel on the display panel to be P_{ij} , which is at the i^{th} column and j^{th} row of the display. The line segment OP_{ij} intersects with the pinhole film plane at the point H_{ij} . The position of H_{ij} can be easily determined by using similar triangles. Once H_{ij} is determined, the distance between H_{ij} and all the pinholes in the film can be determined. Thus, the nearest pinhole can be chosen. $H_{near,ij}$ is defined as the nearest pinhole from H_{ij} , which has the shortest distance to H_{ij} . Thus, the pinhole at $H_{near,ij}$ is the corresponding pinhole of the given pixel P_{ij} . Let i and j running through all the pixels on display, the corresponding pinhole for each pixel can be confirmed.

When images are displayed on the LCD panel, the light emissions of pixels are not directional, which covers the front half of the space. However, when the pinhole film is applied to the display, it acts as a beam selection device. The only beam passes through the pinhole can be transmitted through, while other beams are blocked by the opaque film. In Figure 5.3, three beams from the pixel P_{ij} are shown that can transmit through the pinholes. In this condition, the ray that can reach the pupil is $P_{ij}H_{near,ij}$. By this way, for a given eye box, the image pattern can be rendered with ray tracing. In the proposed system, the positions of the pinholes and LCD panel are fixed. Thus, the coordinate of the effective rays can be calculated prior to the rendering process, which effectively reduces the amount of computation required and makes it possible to generate in real-time.

5.2.3 Viewing zone

For periodic pinhole light field display, the eye box size is determined by the pitch of the pinhole. For random pinhole, there is no fixed pitch. Instead, the pinhole distribution can be characterized by statistics. The mechanism between the eye box size and pinhole distribution is demonstrated in Figure 5.4. Here, a two-dimensional calculation is carried while the other dimension conforms to the same rule. Thus, instead of using the pinhole notation H_{ij} , we use H_i to denote

the pinhole distribution from the cross-section view from the side of the pinhole. Further, we denote three successive pinholes to be H_{i-1} , H_i and H_{i+1} .

The image to be displayed on the display panel is rendered with the eye box parameter fixed. Here, the eye box is set at a distance l_r with the pinhole film, which is also referred to as eye relief. g is defined as the gap between the pinhole film and display panel. As mentioned, there is no periodic pitch for the randomly distributed pinholes. Thus, the distance between neighbor pinholes $H_{i-1}H_i$ and H_iH_{i+1} are random. Here we define E_i as the size of the elemental image formed by pinhole H_i . Based on the rendering process discussed in the last section, E_i is given by:

$$E_i = \frac{l_r}{2l_r + 2g} (H_{i-1}H_i + H_iH_{i+1}) \quad (5.1)$$

Figure 5.4 Eye box demonstration of a light field pinhole near-eye display.

Figure 5.4 is a demonstration of the eye box determination of random pinhole NED. The eye box of pinhole H_i is the set of position where the user can perceive the ray from pinhole H_i . In Figure 5.4 there are 5 pinholes, denote as H_1 to H_5 . We denote the eye box size at the observer plane to be D_i . The corresponding elemental image area to be E_1 to E_5 . Then we have 5 eye boxes generated by each pinhole, denoted as D_1 to D_5 . We define the eye box of the

light field NED system to be D . Thus, D is the intersection of all the eye boxes from each pinhole. According to the theorem of similar triangles, D is given by:

$$D = \min_i D_i = \min_i \left(\frac{l_r E_i}{g} \right) = \frac{l_r^2}{l_r g + g^2} \min_i H_i \quad (5.2)$$

where $\min_i H_i$ is the minimum distance between neighbor pinholes H_i . This conclusion can be generalized to a three-dimensional light field NED, where the eye box size at a certain direction is proportional to the minimum pinhole distance at the corresponding direction. Based on the discussion above, we need to carefully control the minimum pinhole distribution distance to get the expected size of the eye box.

In this case, we expect an eye box size of 8mm, which means $\min_i D_i = 8mm$. The acrylic spacer between the pinhole and the display panel has a refractive index of 1.49. The thickness of the spacer is 5mm. Thus, the equivalent gap g compare to air is 3.356mm. Based on the $l_r = 50mm$ eye relief design, the minimum distance of the random pinhole can be calculated by equation 5.2: $\min_i H_i = D \cdot \frac{l_r g + g^2}{l_r^2} = 8mm \cdot \frac{50mm \cdot 3.356mm + (3.356mm)^2}{50mm^2} = 0.57mm$. The pixel size of the display is 0.0475mm. Thus, the minimum pinhole distance to get 8mm eye box is equivalent to 12 pixels. Here we set the minimum distance to be equivalent to 15 pixels (0.7125mm) for pinhole pattern generation to get an eye box at around 10mm. The pinhole film patterns are shown in Figure 5.6. Figure (a) and (b) are the patterns of a periodic pinhole array with a pitch of 0.7125mm. Figure (c) and (d) are random pinhole patterns with a calculated minimum distance of 0.7125mm. The pinhole opening size in both cases was set at 0.15mm.

5.3 Results

A light field NED prototype was developed to validate the proposed random pinhole method. The NED prototype was modeled in 3D CAD software Solidworks, shown in Figure 5.5. The interpupillary distance of the NED was set

at 64mm for the average human. The eye relief was designed to be 50mm. The left and right viewing channels are separated by a thin plastic wall to avoid unexpected crosstalk. During the rendering process, the left and right eye viewing channel are rendered separately as well. The designed prototype is fabricated by 3D printing and painted into gloss black.

Figure 5.5 3D CAD design of the light field NED based on random pinholes.

Figure 5.6 The design pattern of (a) periodic pinhole array and (c) random pinhole. (b) and (d) are magnified images of (a) and (c).

The multilayer random pinhole light field display is assembled into the system. Here we chose a commercialized 5-inch LCD panel for mobile and VR application with high-resolution 2560×1440 , available for purchase from Japan Display Inc. (JDI). The display panel has a pixel size 0.0475mm with RGB 8bit color and a refresh rate of 60Hz. The LCD module contains an LED backlight and one pair of orthogonal polarizers at the front and back of the liquid crystal layer. An opaque plastic thin film with random pinhole openings was chosen as the light modulator. To reduce the fabrication cost, the random pinhole pattern

was fabricated by a laserjet printer. For a comparison test, a periodic pinhole pattern was fabricated with exactly the same method. A transparent acrylic spacer is fabricated to be the same size as the display. The thickness of the spacer is set to be 5mm. The refractive index of the acrylic material is 1.49, which gives an equivalent optical path 3.356mm. To accommodate the viewing channel of two eyes, the display panel and random pinhole film are divided into two parts horizontally, which mean for each eye pupil 1280×1440 possible rays can be generated to form the light field. Random pinhole patterns and displayed image patterns can be calculated independently for each viewing channel. The NED prototype was designed to make the pinhole film interchangeable so that during the experiment, different pinhole film can be tested.

Figure 5.7 shows the light field display components used in the experiment, corresponding to the design in Figure 5.2. On top of the stack is the pinhole layer. The pinhole layer can be either random pinhole or periodic pinhole array. The intermediate layer is a transparent spacer layer made of acrylic. The LCD display assembly is at the back of the stack. The LCD assembly includes backlight and polarizer.

Figure 5.7 Light field components fabricated and assembled in the NED system, including random pinhole film, acrylic board spacer, and an LCD device.

During this experiment, a digital camera Nikon Coolpix A was employed to simulate human pupil. The NED prototype was connected to the computer with

an HDMI interface for data transfer and a USB cable for power supply. The experiment setup is shown in Figure 5.8.

Figure 5.8 Photograph of the developed prototype and the camera (that simulates human eye).

Because of the random distribution of pinholes, the viewing zone distribution is random as well. Here a simulation method is developed to investigate the viewing zone shape. From an arbitrary point in space, a ray can be formed by connecting the point with a pinhole. The ray is classified as a signal ray if the ray also passes a corresponding pixel to the pinhole on the display panel. Otherwise, it is a noise ray. The viewing zone can be simulated by sampling the points in the space. Viewing zone is the set of all the points in space which yield signal rays with all the pinholes. Figure 5.9 demonstrates the simulation result of the viewing zone with multiple levels of signal-to-noise ratio (SNR). The viewing zone indicator in the blue, green, and red colors stand for the SNR at 100%, 80%, and 60%, respectively. The red area with SNR 100% is the best viewing position for the user, where almost no noise ray is observed. When the viewing position drifts away from the red zone, the SNR starts to drop so that the user will perceive more noise rays and blurry image.

Figure 5.9 Distribution of the viewing zones with different levels of signal-noise ratios (SNR) in the proposed NED system.

Three planes are shown in Figure 5.9, which has eye relief distances 40, 50, and 60mm. The distribution of the viewing zone distribution across these three planes is analyzed in detail in Figure 5.10. Here (a), (b) and (c) shows the viewing zone at distances 40, 50, and 60mm. As the figures show, the eye box size with 100% SNR at 50mm is around 10mm at the designed eye relief. As long as the human pupil is within this area, high-quality light field image can be perceived. However, if the human pupil moves out of the high SNR area, signal rays will be replaced by noise rays, and the user cannot perceive the light field image. Instead, the image quality will drop drastically and becomes blurry. It should be noted the viewing zone is unique, while there is no repeated viewing zone in the space.

Figure 5.10 Signal-noise ratio (SNR) distribution with horizontal and vertical position at eye-relief distance (a) 40mm, (b) 50mm and (c) 60mm.

After the viewing zone simulation, the NED system was tested by experiment. Figure 5.11-13 show the experiment results of the system with a classic test image “Lena”. Figure 5.11 shows the rendered light field pattern for sample image “Lena”. Figure (a) and (b) show the image pattern for periodic pinhole array, while figure (c) and (d) show the image pattern for the random pinhole. As we can see, the elemental image of periodic pinhole array shows periodicity, while the elemental image of the random pinhole shows no periodicity. For the random pinhole, the size of each image element is not the same either, determined by the rendering method.

Figure 5.11 Rendered light field pattern for sample image “Lena”. (a) Elemental image for periodic pinhole array. (b) Enlarged image pattern for periodic pinhole array. (c) Elemental image for the random pinhole. (d) Enlarged image pattern for the random pinhole.

Figure 5.12 shows the perceived periodic pinhole array light field image by the camera with the NED prototype. Figure (a) shows the original correct light field image at the designed viewing zone. However, in real practice, the user very likely cannot perceive the correct image due to pupil position variation. When the pupil moves between viewing zones, the perceived image behaves like

figure (b), which is a transition of images from two viewing zones. If the pupil moves further to the neighboring viewing zone, the user perceives a clear image shown in figure (c) which is very similar with the original image from the look, but with wrong depth information. The repeated zone image is so similar to the correct image, so the user cannot differentiate in most cases.

Figure 5.12 Perceived periodic pinhole array light field image by the camera with the NED prototype. (a) Correct light field image at the designed viewing zone. (b) Image during the transition between viewing zones. (c) Light field image at the repeated viewing zone with wrong depth information.

Figure 5.13 shows the perceived random pinhole light field image. Similar to the periodic light field image, we start from the correct light field image at the designed viewing zone in figure (a). Then the pupil position drifts to the edge of the viewing zone where the image quality starts to drop due to decreased SNR. After that, the behavior of the viewing zone starts to differ from the periodic one. As we can see in figure (c), when the pupil is out of the viewing zone, instead of showing another clear image, the image of the random pinhole is very blurry. This way, random pinhole prevents repeated viewing zone from happening, with its unique viewing zone distribution across space.

Figure 5.13 Perceived random pinhole light field image by the camera with the NED prototype. (a) Correct light field image at the designed viewing zone. (b) Decreased SNR and image quality when moving out of the viewing zone. (c) Light field image out of the viewing zone, the image is blurry.

Figure 5.14 to 5.16 show another sample image “Mandrill” with the same experimental procedure as the sample image “Lena” in Figure 5.11-13. The behavior is further validated that there is no repeated viewing zone for the random pinhole.

Figure 5.14 Rendered light field pattern for sample image “Mandrill”. (a) Image pattern for periodic pinhole array. (b) Enlarged image pattern for periodic pinhole array. (c) Image pattern for the random pinhole. (d) Enlarged image pattern for the random pinhole.

Figure 5.15 Perceived periodic pinhole array light field image by the camera with the NED prototype. (a) Correct light field image at the designed viewing zone. (b) Image during the transition between viewing zones. (c) Light field image at the repeated viewing zone with wrong depth information.

Figure 5.16 Perceived random pinhole light field image by the camera with the NED prototype. (a) Light field image at the designed viewing zone. (b) Decreased SNR and image quality when moving out of the viewing zone. (c) Light field image out of the viewing zone, the image is blurry.

A resolution test was carried with a standard resolution chart image. Figure 5.17 and 5.18 show the resolution test from a periodic pinhole array. Figure 5.19 and 5.20 show the resolution test from a random pinhole. As we can see, the

resolution of the two systems is on par. The resolution from the period pinole array is slightly better, due to more pinhole quantity. Based on our generation method, the minimum random pinhole distance is 0.7125mm. Thus, lots of pinholes have larger distance than this minimum number. However, for the periodic array, all the distance between pinholes are kept at this value, which means the periodic pinholes are packed tighter than the random ones, result in a slightly higher resolution.

Figure 5.17 (a) Rendered light field image pattern for periodic pinhole array. (b) Perceived correct periodic pinhole light field image at the designed viewing zone.

Figure 5.18 Perceived periodic pinhole array light field image by the camera with the NED prototype. (a) Image during the transition between viewing zones. (b) Light field image at the repeated viewing zone with wrong depth information.

Figure 5.19 (a) Rendered light field image pattern for the random pinhole. (b) Perceived correct periodic pinhole light field image at the designed viewing zone.

Figure 5.20 Perceived random pinhole light field image by the camera with the NED prototype. (a) Decreased SNR and image quality when moving out of the viewing zone. (b) Light field image out of the viewing zone, the image is blurry.

In conclusion, the experiments show clearly the difference between the random pinhole and the periodic pinhole when it comes to the viewing zone. The random pinhole does not present a repeated viewing zone. Instead, a unique viewing zone in space is presented, controlled by the minimum spacing of the random pinholes. This property effectively prevents wrong information from passing to the user.

5.4 Conclusion

In this research, the random pinhole is demonstrated as a light field generation device in the near-eye display. By using random pinhole, the repeated viewing zone problem of light field display is solved. The design method of random pinhole film is discussed in detail, including the rendering method, determination of eye box size, and viewing zone distribution. A random-pinhole-based NED prototype was designed and fabricated to demonstrate the effectiveness of the proposed method, which clearly shows the advantage of the random pinhole. The perceived light field image from random pinhole has equal image quality over periodic pinhole array but with unique viewing zone.

Chapter 5 is published as:

Weitao Song, Qijia Cheng, Phil Surman, Yue Liu, Yuanjin Zheng, Zhiping Lin, and Yongtian Wang. "Design of a light-field near-eye display using random pinholes." *Optics Express* 27, no. 17 (2019): 23763-23774.

Chapter 6 Conclusion and Recommendations

6.1 Conclusion

In this thesis, theory and methods are demonstrated to improve key performance factors in modern NED, including size, resolution, and depth perception.

To solve the main conflict between resolution and device size, a versatile resolution enhancement method for NED is implemented by overlapping multiple display panel images. Based on display super-resolution, this overlapping method effectively increases the perceived resolution without optical complexity. By investigating the algorithm theory of image generation and calibration, the perceived resolution can be enhanced and overcome the physical resolution of a single display panel. A NED prototype based on a semi-transparent mirror is designed and fabricated. The experiment demonstrated decent results of resolution enhancement. Compare with the conventional method of display tiling, my method offers an equivalent level of performance yet reduces system complexity, brings more flexibility in the design of NED.

Free-form optics enable innovative ways of implementing compact size NED. A design routine of off-axis near-eye display system is proposed with off-axis lightguide. Four freeform surfaces described by x-y polynomial functions are used to effectively reduce the size of NED while correcting the off-axis aberration. An imaging lightguide prism and a compensation prism is presented in the system. The system achieved 4-mm diameter exit pupil, a 21.5° diagonal full field-of-view (FoV) and 23 mm eye relief. The designed system yields good optical performance yet very compact in terms of size, compared to traditional NED. A prototype is demonstrated with components available from the market. Test results indicated that the proposed near-eye display has good optical performance yet compact in size. The prototype effectively indicated the feasibility of the design as a compact NED, which can be a good option in the application of wearable virtual reality.

Light field can be treated as a feasible way to solve convergence-accommodation conflict in NED and stereoscopic display. A light field NED with random pinhole design is demonstrated with unique viewing zone and optical simplicity. In this research, the random pinhole is demonstrated as a light field generation device in the near-eye display. By using random pinhole, the repeated viewing zone problem of light field display is solved. The design method of random pinhole film is discussed in detail, including the rendering method, determination of eye box size, and viewing zone distribution. A random-pinhole-based NED prototype was designed and fabricated to demonstrate the effectiveness of the proposed method, which clearly shows the advantage of the random pinhole. The perceived light field image from random pinhole has equal image quality over periodic pinhole array but with unique viewing zone.

Further research is recommended in section 6.2 to combine the individual technique in Chapter 3, 4 and 5 as a hybrid device, where multiple light field display module can be used for super-resolution, with the design flexibility of free-form optics.

My research can act as guidance in the design of future NED to reduce device size, enhance resolution, and incorporate the depth sense, thus further improve virtual reality user experience.

6.2 Recommendations for further research

Further researches are recommended in several aspects.

The display overlapping method can be further developed to incorporate higher precision calibration method, including better accuracy and involve more misalignment error caused by optical aberration. During the calibration process, the out of the plane tilting of the display panels are ignored to simplify the model. Camera distortion and intrinsic parameters are ignored, which may compromise the uniformity of the overlapping result.

Active alignment method can be investigated, including active actuators. Possible mechanism can be designed to control the alignment with the assistance of calibration. The computational algorithm can be further developed for real-time image generation. Further image overlapping optical design can be developed to make the device more compact, including waveguide implementations. The computational method has already been discussed in chapter 3 to support more two pieces of display panels. An overlapping NED with more than two display panel components can be developed to further improved the perceived resolution.

The freeform off-axis NED of chapter 4 can be further improved with the tunable liquid lens, to solve the accommodation convergence conflict problem. The random pinhole light field NED can be further improved in display performance with better hardware. For example, lithography fabricated pinhole mask or random microlens.

The three proposed method can be combined, which is a straightforward plan. The image overlapping method can be implemented with an off-axis freeform design, which can result in a compact device with great resolution. The image overlapping method can be used in the light field display, which can greatly improve the resolution problem demanded by the light field display.

Author's Publications

Qijia Cheng, Weitao Song, Feng Lin, Yue Liu, Yongtian Wang, and Yuanjin Zheng. "Resolution enhancement of near-eye displays by overlapping images." *Optics Communications* 458 (2020): 124723.

Weitao Song, **Qijia Cheng**, Yue Liu, Yuanjin Zheng and Yongtian Wang. "Design of a light field near-eye display using random pinholes." *Optics Express* 27, no. 17 (2019): 23763-23774.

Weitao Song, **Qijia Cheng**, Xuming Liu, Yandong Wang, Yetao Huang, Yuanjin Zheng, Yue Liu and Yongtian Wang, "Design of a 360° panoramic capture system based on a smart phone." *Chinese Optics Letter*, accepted, manuscript ID: COL-19-0016.

Zhenfeng Zhuang, Phil Surman, **Qijia Cheng**, Simon Thibault, Yuanjin Zheng, and Xiao Wei Sun. "Subpixel area-based evaluation for crosstalk suppression in quasi-three-dimensional displays." *Applied Optics* 56, no. 19 (2017): 5450-5457.

Zhenfeng Zhuang, **Qijia Cheng**, Phil Surman, Yuanjin Zheng, and Xiao Wei Sun. "A compact and lightweight off-axis lightguide prism in near to eye display." *Optics Communications* 393 (2017): 143-151.

Shizheng Wang, Xiangyu Zhang, **Qijia Cheng**, Kaviya Rajendran, Phil Surman, Junsong Yuan, and Xiao Wei Sun. "Glasses-free light field 3D display." In *Visual Communications and Image Processing (VCIP)*, 2015, pp. 1-1. IEEE, 2015.

Qijia Cheng, Zhuoteng Peng, Jie Lin, Shanshan Li, and Fei Wang. "Energy harvesting from human motion for wearable devices." In *Nano/Micro Engineered and Molecular Systems (NEMS)*, 2015 IEEE 10th International Conference on, pp. 409-412. IEEE, 2015.

Qijia Cheng, Minghao He, Ulf Ekenberg, Kezheng Li and Hongyu Yu.
"Ultra-thin CdTe Thin-film Solar Cell with Surface Texturing." China
Semiconductor International Conference 2015.

Bibliography

- [1]. Mazuryk, Tomasz, and Michael Gervautz. "Virtual reality-history, applications, technology and future." (1996).
- [2] Thomas Weynants, "From the Rarekiek, 17th. & 18th. Century optical entertainment, to the birth of Television", 2003.
- [3] Robinson, David (1996). From Peep Show to Palace: the Birth of American Film. New York: Columbia University Press. p. 56. ISBN 0-231-10338-7.
- [4] Dickson, William Kennedy-Laurie, and Antonia Dickson. History of the kinetograph, kinetoscope and kinetophonograph. The Museum of Modern Art, 2000.
- [5] Mutoscope, definition from Wikipedia, <https://en.wikipedia.org/wiki/Mutoscope> .
- [6] "Children Of All Nations, Their Homes, Their Schools And Their Playgrounds" by G. Clausen et al. Cassell & Company.
- [7] Edison, Thomas A. "Kinetoscope." U.S. Patent No. RE12,037. 30 Sep. 1902.
- [8] Burdea Grigore, C., and P. Coiffet. Virtual reality technology. London: Wiley-Interscience, 1994.
- [9] "Sensorama simulator." U.S. Patent 3,050,870, issued August 28, 1962.
- [10] Sutherland, Ivan E. "Sketchpad a man-machine graphical communication system." Transactions of the Society for Computer Simulation 2.5 (1964): R-3.
- [11] Tim Merel, "The reality of VR/AR growth", Tech Crunch, 11 Jan 2017.
- [12] The Venture Reality Fund <http://www.thevrfund.com/>
- [13] Costello, Patrick J. Health and safety issues associated with virtual reality: a review of current literature. Advisory Group on Computer Graphics, 1997.
- [14] Alexei Sourin, "Computer Graphics: From a Small Formula to Cyberworlds, 3rd edition" , Pearson Prentice Hall, Singapore (2012)

-
- [15] Moss, Jason D., and Eric R. Muth. "Characteristics of head-mounted displays and their effects on simulator sickness." *Human factors* 53.3 (2011): 308-319.
- [16] <https://www.x.company/glass/>
- [17] <https://www.oculus.com/>
- [18] Cakmakci, Ozan, and Jannick Rolland. "Head-worn displays: a review." *Journal of display technology* 2.3 (2006): 199-216.
- [19] Keller, Kurtis, and D'nardo Colucci. "Perception in HMDs: what is it in head-mounted displays (HMDs) that really make them all so terrible?." *Aerospace/Defense Sensing and Controls. International Society for Optics and Photonics*, 1998.
- [20] Melzer, James E., et al. "Overcoming the field of view: resolution invariant in head-mounted displays." *Proc. SPIE*. Vol. 3362. 1998.
- [21] Hoppe, Michael J., and James E. Melzer. "Optical tiling for wide-field-of-view head-mounted displays." *Proc. SPIE Conf. on Current Developments in Optical Design and Optical Engineering VIII*. 1999.
- [22] Yoshida, Akitoshi, Jannick P. Rolland, and John H. Reif. "Design and applications of a high-resolution insert head-mounted-display." *Virtual Reality Annual International Symposium, 1995. Proceedings.. IEEE*, 1995.
- [23] Melzer, James E. "Design evolution of a wide field of view head-mounted display for aircraft training and simulation." *Proc. SPIE*. Vol. 4361. 2001.
- [24] Rolland, Jannick P., Myron W. Krueger, and Alexei Goon. "Multifocal planes head-mounted displays." *Applied Optics* 39.19 (2000): 3209-3215.
- [25] K.E. Jachimowicz, K.E. "Projection Display Technologies," In M.A., Karim, ed., *Electra-Optical Displays*, Marcel Decker, New York, 1992
- [26] M. Thomas, H. Geltmacher, "Combat simulator display development," *Information Display*, 4&5,23-26, 1993
- [27] A. Fernie, "Helmet-mounted display with dual resolution," *Journal of the SID*, 3/4, 151-153, 1995

-
- [28] E.M. Howlett, "A wide field of view, high resolution compact virtual reality display," Proceedings of SID Symposium Digest, 1992
- [29] J.E. Melzer, K. Moffitt, "Partial binocular-overlap in helmet-mounted displays," Proc. SPZE 1117, 56-62, 1989
- [30] F.L. Kooi, "Binocular Configurations of a Night-Flight Head-Mounted Display," Displays, 14(1), 11-20, 1993
- [31] M.L. Thomas, G. Reining, G. Kelly, "The display for advanced research and training: an "inexpensive" answer tactical simulations," Proceedings of SPIE, Large-Screen Projection, Avionic and Helmet-Mounted Displays, 665-75, 1991
- [32] M.J. Tuck, "Display system having wide field of view," U.S. Patent 4,772,942, 1987
- [33] A.L. Berman, J.E. Melzer, "Optical Collimating Apparatus," U.S. Patent 4,859,031, 1989
- [34] J. Leigh, A. Johnson, L. Renambot, T. Peterka, B. Jeong, D. J. Sandin, J. Talandis, R. Jagodic, S. Nam, H. Hur, and Y. Sun, "Scalable resolution display walls," Proc. IEEE 101, 115–129 (2013).
- [35] B. Sajadi and A. Majumder, "Autocalibration of multi-projector CAVElike immersive environments," IEEE Trans. Vis. Comput. Graph. 18, 381–393 (2012).
- [36] R. Kooima, A. Prudhomme, J. Schulze, D. Sandin, and T. DeFanti, "A multi-viewer tiling autostereoscopic virtual reality display," in Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology, Hong Kong, 2010, pp. 172–174.
- [37] W. Li, H. Wang, M. Zhou, S. Wang, S. Jiao, X. Mei, and J. Kim, "Principal observation ray calibration for tiled-lens-array integral imaging display," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Portland, Oregon, 2013, pp. 1019–1026.
- [38] J. Hahn, H. Kim, Y. Lim, G. Park, and B. Lee, "Wide viewing angle dynamic holographic stereogram with a curved array of spatial light modulators," Opt. Express 16, 12372–12386 (2008).

-
- [39] R. W. Massof, L. G. Brown, and M. D. Shapiro, "Head mounted display with full field of view and high resolution," U.S. patent 6,529,331 (4 March 2003).
- [40] M. Gutin and O. Gutin, "Automated design and fabrication of ocular optics," Proc. SPIE 7060, 70600H (2008).
- [41] L. G. Brown and Y. S. Boger, "Invited paper: applications of the Sensics panoramic HMD," SID Symp. Dig. Tech. Papers 39, 77–80 (2008).
- [42] Song, W., Cheng, D., Deng, Z., Liu, Y., & Wang, Y. (2015). Design and assessment of a wide FOV and high-resolution optical tiled head-mounted display. *Applied optics*, 54(28), E15-E22.
- [43] D. Cheng, Y. Wang, H. Hua, and J. Sasian, "Design of a wide-angle, lightweight head-mounted display using free-form optics tiling," *Opt. Lett.* 36, 2098–2100 (2011).
- [44] D. Cheng, Y. Wang, and H. Hua, "Large field-of-view and high resolution free-form head-mounted display," Proc. SPIE 7652, 76520D (2010).
- [45] D. Cheng, Y. Wang, H. Hua, and J. Sasian, "Free-form optics enable lightweight, high-performance head-mounted displays," *Laser Focus World* 48, 67–69 (2012).
- [46] AKENINE-MÖLLER, T. AND HAINES, E. 2002. *Real-Time Rendering*. A. K. Peters Ltd., Natick, MA.
- [47] Crow, Franklin C. "A comparison of antialiasing techniques." *IEEE Computer Graphics and Applications* 1.1 (1981): 40-48.
- [48] Cook, Robert L. "Stochastic sampling in computer graphics." *ACM Transactions on Graphics (TOG)* 5.1 (1986): 51-72.
- [49] Deering, Michael, et al. "The triangle processor and normal vector shader: a VLSI system for high performance graphics." *Acm siggraph computer graphics*. Vol. 22. No. 4. ACM, 1988.
- [50] FUCHS, H., GOLDFEATHER, J., HILLQUIST, J. P., SPACH, S., AUSTIN, J. D., BROOKS, F. P., EYLES, J. G., AND POULTON, J. 1985. Fast spheres, shadows, textures, transparencies, and image enhancements in pixelplanes. In *Proceedings of SIGGRAPH Computer Graphics Conference*, 111–120.

-
- [51] HAEBERLI, P. AND AKELEY, K. 1990. The accumulation buffer: Hardware support for high-quality rendering. In Proceedings of SIGGRAPH Computer Graphics Conference, 309–318.
- [52] Irani, Michal, and Shmuel Peleg. "Super resolution from image sequences." Pattern Recognition, 1990. Proceedings., 10th International Conference on. Vol. 2. IEEE, 1990.
- [53] Ben-Ezra, Moshe, Assaf Zomet, and Shree K. Nayar. "Jitter camera: High resolution video from a low-resolution detector." Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on. Vol. 2. IEEE, 2004.
- [54] Park, Sung Cheol, Min Kyu Park, and Moon Gi Kang. "Super-resolution image reconstruction: a technical overview." IEEE signal processing magazine 20.3 (2003): 21-36.
- [55] Damera-Venkata, Niranjan, and Nelson L. Chang. "Display supersampling." ACM Transactions on Graphics (TOG) 28.1 (2009): 9.
- [56] Jaynes, Christopher, and Divya Ramakrishnan. "Super-resolution composition in multi-projector displays." IEEE Int'l Workshop on Projector-Camera Systems. Vol. 8. 2003.
- [57] Allen, Will, and Robert Ulrich. "47.4: Invited paper: Wobulation: Doubling the addressed resolution of projection displays." SID Symposium Digest of Technical Papers. Vol. 36. No. 1. Blackwell Publishing Ltd, 2005.
- [58] R.M. Steele, S. Webb, and C. Jaynes, "Monitoring and Correction of Geometric Distortion in Projected Displays", International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Plzen, Czech Republic, 2002.
- [59] Zhang, S.,Huang P.,Novel method for structured light system calibration. Optical Engineering,2006,45(8):083601
- [60] Legarda-Saenz R., Bothe T.,Juptner W. Accurate Procedure for the calibration of a Structured light system. Optical Engineering,2004.43(2):464
- [61] Li, Z., Shi Y. Wang C. et al. Accurate calibration method for a structured light system. Optical Engineering, 2008.47(5):053604

-
- [62] Hartley, R. Estimation of relative camera positions for uncalibrated cameras. NLCS,1992.
- [63] Hartley, R. Self-calibration of stationary cameras. International Journal of Computer Vision,1997.22(1): 5-23.
- [64] Faugeras, O., Luong Q., Maybank S. Camera self-calibration: Theory and Experiments. Springer,1992
- [65] Maybank S., Faugeras O. A theory of self-calibration of a moving camera. International Journal of Computer Vision,1992.8(2):123-151.
- [66] B. Achtner, "HMD device with imaging optics comprising an aspheric surface," U. S. Patent 6903875.
- [67] M. Kaschke, "Technology introduction," <http://www.kaschke-medtec.de/vorlWS09.html>.
- [68] H. Mukawa, K. Akutsu, I. Matsumura, S. Nakano, T. Yoshida, M. Kuwahara, and K. Aiki, "A full color eyewear display using planar waveguides with reflection volume holograms," J. Soc. Inf. Disp. 17(3), 185–193 (2009).
- [69] <http://www.sony.co.uk/hub/personal-3d-viewer>.
- [70] K. Aiki and S. Nakano, "Illumination optical device and virtual image display," U. S. Patent 2010/0027289.
- [71] H. Mukawa, "Head-mounted display," U. S. Patent 2010/0046070
- [72] M. J. Heinrich and M. I. Olsson, "Wearable display device," U. S. Patent D659741.
- [73] B. Kress and T. Starner, "A review of head-mounted displays (HMD) technologies and applications for consumer electronics," Proc. SPIE 8720, 87200A (2013).
- [74] N. Owano, "Epson's 3-d glasses simulate 80-inch screen," <http://phys.org/news/2012-04-epson-d-glassessimulate-inch.html>
- [75] M. Takagi, T. Miyao, T. Totani, A. Komatsu, and T. Takeda, "Light guide plate and virtual image display apparatus having the same," U. S. Patent 2012/0057253

-
- [76] P. Äyräs, P. Saarikko, and T. Levola, "Exit pupil expander with a large field of view based on diffractive optics," *J. Soc. Inf. Disp.* 17(8), 659–664 (2009).
- [77] D. Cheng, Y. Wang, H. Hua, and M. M. Talha, "Design of an optical see-through head-mounted display with a low f-number and large field of view using a freeform prism," *Appl. Opt.* 48(14), 2655–2668 (2009).
- [78] S. C. McQuaide, E. J. Seibel, J. P. Kelly, B. T. Schowengerdt, and T. A. Furness III, "A retinal scanning display system that produces multiple focal planes with a deformable membrane mirror," *Displays* 24(2), 65–72 (2003).
- [79] H. Urey, "Diffractive exit-pupil expander for display applications," *Appl. Opt.* 40(32), 5840–5851 (2001).
- [80] T. Levola, "Novel diffractive optical components for near to eye displays," in *SID International Symposium Digest of Technical Papers (Society for Information Display, 2006)*, Vol. 37, No. 1, pp. 64–67.
- [81] R. Shi, J. Liu, J. Xu, D. Liu, Y. Pan, J. Xie, and Y. Wang, "Designing and fabricating diffractive optical elements with a complex profile by interference," *Opt. Lett.* 36(20), 4053–4055 (2011).
- [82] Y. Amitai, A. A. Friesem, and V. Weiss, "Holographic elements with high efficiency and low aberrations for helmet displays," *Appl. Opt.* 28(16), 3405–3416 (1989).
- [83] Z. Yan, W. Li, Y. Zhou, M. Kang, and Z. Zheng, "Virtual display design using waveguide hologram in conical mounting configuration," *Opt. Eng.* 50(9), 094001 (2011).
- [84] Y. Amitai, "Extremely compact high-performance HMDs based on substrate-guided optical element," in *Society for Information Display (SID) International Symposium Digest (Society for Information Display, 2004)*, Vol. 35, No. 1, pp. 310–313.
- [85] Z. R. Zheng, X. Liu, H. F. Li, and L. Xu, "Design and fabrication of an off-axis see-through head-mounted display with an x-y polynomial surface," *Appl. Opt.* 49(19), 3661–3668 (2010).
- [86] Y. Amitai, M. Hadad, and N. Friedmann, "Collimating optical device and system," U.S. 0126051 A1 (8 May 2014).

-
- [87] D. W. Cheng, Y. T. Wang, C. Xu, W. T. Song, and G. F. Jin, "Design of an ultra-thin near-eye display with geometrical waveguide and freeform optics," *Opt. Express*, 22(17), 20705-20719 (2014).
- [88] H. Mukawa, K. Akutsu, I. Matsumura, S. Nakano, T. Yoshida, M. Kuwahara, K. Aiki, and M. Ogawa, "A full color eyewear display using holographic planar waveguides," *SID*, 39(1), 89-92 (2008).
- [89] O. Cakmakci, and J. Rolland, "Design and fabrication of a dual-element off-axis near-eye optical magnifier," *Opt. Lett.*, 32(11), 1363-1365 (2007).
- [90] O. Cakmakci, G. E. Fasshauer, H. Foroosh, K. P. Thompson, and J. P. Rolland, "Meshfree approximation methods for freeform surface representation in optical design with applications to head-worn displays," *Proc. SPIE*, 7061, 70610D (2008).
- [91] O. Cakmakci, K. Thompson, P. Vallee, J. Cote, and J. P. Rolland, "Design of a freeform single-element head-worn display," *Proc. SPIE*, 7618, 761803 (2010).
- [92] A. Bauer, and J. P. Rolland, "Visual space assessment of two all-reflective, freeform, optical see-through head-worn displays," *Opt. Express*, 22(11), 13155-13163 (2014).
- [93] H. Hoshi, N. Taniguchi, H. Morishima, and T. Akiyama, "Off-axial HMD optical system consisting of aspherical surfaces without rotational symmetry," *Proc. SPIE*, 2653, 234-242 (1996).
- [94] T. Togino, "Prism optical system," U.S. 5991103 (23 Nov 1999).
- [95] H. Mukawa, and Y. Sugano, "Image display apparatus," U.S. 0064310 A1 (22 Mar 2007).
- [96] T. Togino, T. Takeyama, and A. Sakurai, "Image-forming optical system and viewing optical system," U.S. 6201646 B1 (13 Mar 2001).
- [97] Q. F. Wang, D. W. Cheng, Y. T. Wang, H. Hua, and G. F. Jin, "Design, tolerance, and fabrication of an optical see-through head-mounted display with freeform surface elements," *Appl. Opt.*, 52(7), C88-C98 (2010).

-
- [98] D. W. Cheng, Y. T. Wang, H. Hua, and J. Sasian, "Design of a wide-angle, lightweight head-mounted display using freeform optics tiling," *Opt. Lett.*, 36(11), 2098–2100 (2011).
- [99] O. Cakmakci, A. Gupta, and O. A. Martinez, "See-through eyepiece for head wearable display," U.S. 0177519 A1 (25 Jun 2015).
- [100] H. Hua, "Enabling focus cues in head-mounted displays," *Proc. IEEE* 105(5), 805–824 (2017).
- [101] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *J. Vis.* 8(3) (2008).
- [102] S. Liu, Y. Li, P. Zhou, Q. Chen, and Y. Su, "Reverse-mode PSLC multi-plane optical see-through display for AR applications," *Opt. Express* 26(3), 3394–3403 (2018).
- [103] S. Liu, H. Hua, and D. Cheng, "A novel prototype for an optical see-through head-mounted display with addressable focus cues," *IEEE Trans. Vis. Comput. Graph.* 16(3), 381–393 (2010).
- [104] J. P. Rolland, M. W. Krueger, and A. Goon, "Multifocal planes head-mounted displays," *Appl. Opt.* 39(19), 3209–3215 (2000).
- [105] D. Cheng, Q. Wang, Y. Wang, and G. Jin, "Lightweight spatial-multiplexed dual focal-plane head-mounted display using two freeform prisms," *Chin. Opt. Lett.* 11, 031201 (2013).
- [106] Y. Takaki and N. Fujimoto, "Flexible retinal image formation by holographic Maxwellian-view display," *Opt. Express* 26(18), 22985–22999 (2018).
- [107] J. S. Lee, Y. K. Kim, and Y. H. Won, "Time multiplexing technique of holographic view and Maxwellian view using a liquid lens in the optical see-through head mounted display," *Opt. Express* 26(2), 2149–2159 (2018).
- [108] C. Jang, K. Bang, S. Moon, J. Kim, S. Lee, and B. Lee, "Retinal 3D: augmented reality near-eye display via pupil-tracked light field projection on retina," *ACM Trans. Graph.* 36(6), 190 (2017).

-
- [109] 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 Trans. Graph.* 34(4), 60 (2010).
- [110] 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 Trans. Graph.* 29(6), 1–10 (2010).
- [111] D. Chen, X. Sang, X. Yu, X. Zeng, S. Xie, and N. Guo, "Performance improvement of compressive light field display with the viewing-position-dependent weight distribution," *Opt. Express* 24(26), 29781–29793 (2016).
- [112] D. Lanman and D. Luebke, "Near-eye light field displays," *ACM Trans. Graph.* 32(6), 1–10 (2013).
- [113] Y. Takaki and Y. Yamaguchi, "Flat-panel see-through three-dimensional display based on integral imaging," *Opt. Lett.* 40(8), 1873–1876 (2015).
- [114] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs, "Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources," *ACM Trans. Graph.* 33(4), 89 (2014).
- [115] W. Song, Y. Wang, D. Cheng, and Y. Liu, "Light field head-mounted display with correct focus cue using micro structure array," *Chin. Opt. Lett.* 12(6) (2014).
- [116] H. Hua and B. Javidi, "A 3D integral imaging optical see-through head-mounted display," *Opt. Express* 22(11), 13484–13491 (2014).
- [117] H. Huang and H. Hua, "Systematic characterization and optimization of 3D light field displays," *Opt. Express* 25(16), 18508–18525 (2017).
- [118] Takaki, Yasuhiro, and Nichiyo Nago. "Multi-projection of lenticular displays to construct a 256-view super multi-view display." *Optics express* 18, no. 9 (2010): 8824-8835.
- [119] Takaki, Yasuhiro. "Super multi-view display with 128 viewpoints and viewpoint formation." In *Stereoscopic Displays and Applications XX*, vol. 7237, p. 72371T. International Society for Optics and Photonics, 2009.

-
- [120] Takaki, Yasuhiro. "Development of super multi-view displays." *ITE Transactions on Media Technology and Applications* 2.1 (2014): 8-14.
- [121] C. Yao, D. Cheng, T. Yang, and Y. Wang, "Design of an optical see-through light-field near-eye display using a discrete lenslet array," *Opt. Express* 26(14), 18292–18301 (2018).
- [122] K. Akşit, J. Kautz, and D. Luebke, "Slim near-eye display using pinhole aperture arrays," *Appl. Opt.* 54(11), 3422–3427 (2015).
- [123] H. Urey, K. V. Chellappan, E. Erden, and P. Surman, "State of the art in stereoscopic and autostereoscopic displays," *Proc. IEEE* 99(4), 540–555 (2011).
- [124] J. Geng, "Three-dimensional display technologies," *Adv. Opt. Photonics* 5(4), 456–535 (2013).
- [125] W. Song, D. Cheng, Z. Deng, Y. Liu, and Y. Wang, "Design and assessment of a wide FOV and high-resolution optical tiled head-mounted display," *Appl. Opt.* 54(28), E15–E22 (2015).
- [126] L. Wei, Y. Li, J. Jing, L. Feng, and J. Zhou, "Design and fabrication of a compact off-axis see-through head mounted display using a freeform surface," *Opt. Express* 26(7), 8550–8565 (2018).
- [127] D. Cheng, Y. Wang, H. Hua, and M. M. Talha, "Design of an optical see-through head-mounted display with a low f-number and large field of view using a freeform prism," *Appl. Opt.* 48(14), 2655–2668 (2009).
- [128] J.-H. Park, K. Hong, and B. Lee, "Recent progress in three-dimensional information processing based on integral imaging," *Appl. Opt.* 48(34), H77–H94 (2009).
- [129] Á. Tolosa, R. Martinez-Cuenca, H. Navarro, G. Saavedra, M. Martínez-Corral, B. Javidi, and A. Pons, "Enhanced field-of-view integral imaging display using multi-Köhler illumination," *Opt. Express* 22(26), 31853–31863 (2014).
- [130] A. Nashel and H. Fuchs, "Random hole display: A non-uniform barrier autostereoscopic display", 2009 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video. IEEE, 2009: 1–4.

-
- [131] O. Cakmakci, and J. Rolland, "Head-Worn Displays: A Review." *Journal of Display Technology* 2(3),199-216 (2006).
- [132] H. Hua, L. D. Brown, and C. Gao, "Scape: supporting stereoscopic collaboration in augmented and projective environments," *IEEE Comp. Graph. Appl.* 24, 66–75 (2004).
- [133] J. P. Rolland and H. Fuchs, "Optical versus video see-through head mounted displays in medical visualization," *Presence* 9, 287–309 (2000).
- [134] J. P. Rolland and K. Thompson, "See-through head worn displays for mobile augmented reality," *Commun. China Comp. Fed.* 7, 28–37 (2011).
- [135] D. Cheng, Y. Wang, H. Hua, and M. M. Talha, "Design of an optical see-through head-mounted display with a low f-number and large field of view using a free form prism," *Appl. Opt.* 48, 2655–2668 (2009).
- [136] X. Hu and H. Hua, "High-resolution optical see-through multi-focal plane head-mounted display using freeform optics," *Opt. Express* 22, 13896–13903 (2014).
- [137] 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 Trans. Graph.* 34(4), 60 (2010).
- [138] 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 Trans. Graph.* 29(6), 1–10 (2010).
- [139] M. Liu, C. Lu, H. Li, and X. Liu, "Near eye light field display based on human visual features," *Optics Express*, 25(9), 9886 (2017).
- [140] Miaomiao Xu and Hong Hua, "High dynamic range head mounted display based on dual-layer spatial modulation," *Opt. Express* 25, 23320-23333 (2017)
- [141] Chen, Ai Hong, Daniel J. O'Leary, and Edwin R. Howell. "Near visual function in young children. Part I: Near point of convergence. Part II: Amplitude of accommodation. Part III: Near heterophoria." *Ophthalmic and Physiological Optics* 20, no. 3 (2000): 185-198.

-
- [142] Heide, Felix, Douglas Lanman, Dikpal Reddy, Jan Kautz, Kari Pulli, and David Luebke. "Cascaded displays: spatiotemporal superresolution using offset pixel layers." *ACM Transactions on Graphics (TOG)* 33, no. 4 (2014): 1-11.
- [143] Damera-Venkata, Niranjan, and Nelson L. Chang. "On the resolution limits of superimposed projection." In *2007 IEEE International Conference on Image Processing*, vol. 5, pp. V-373. IEEE, 2007.
- [144] Chao Zuo, Lei Huang, Minliang Zhang, Qian Chen, and Anand Asundi. "Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review." *Optics and Lasers in Engineering* 85 (2016): 84-103.
- [145] Li, Zhongwei, Yusheng Shi, Congjun Wang, and Yuanyuan Wang. "Accurate calibration method for a structured light system." *Optical Engineering* 47, no. 5 (2008): 053604.
- [146] Weng, Jing-Feng, and Yu-Lung Lo. "Novel rotation algorithm for phase unwrapping applications." *Optics Express* 20, no. 15 (2012): 16838-16860.
- [147] H. Hua, X. D. Hu, and C. Y. Gao, "A high-resolution optical see-through head-mounted display with eyetracking capability," *Opt. Express*, 21(25), 30993-30998 (2013).
- [148] Haiwei Chen, Guanjun Tan, and Shin-Tson Wu. "Ambient contrast ratio of LCDs and OLED displays." *Optics Express* 25, no. 26 (2017): 33643-33656.
- [149] S. C. Hsu, "The Kell Factor: Past and Present," in *SMPTE Journal*, vol. 95, no. 2, pp. 206-214, Feb. 1986, doi: 10.5594/J07847.
- [150] J. W. Pan, S. H. Tu, C. M. Wang, and J. Y. Chang, "High efficiency pocket-size projector with a compact projection lens and a light emitting diode-based light source system," *Appl. Opt.* 47, 3406–3414 (2008).
- [151] Atchison, David, and George Smith. *Optics of the human eye*. Butterworth Heinemann, 2000.
- [152] Watson, Andrew B., and John I. Yellott. "A unified formula for light-adapted pupil size." *Journal of vision* 12, no. 10 (2012): 12-12.

-
- [153] Nashel, Andrew, and Henry Fuchs. "Random hole display: A non-uniform barrier autostereoscopic display." In 2009 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video, pp. 1-4. IEEE, 2009.
- [154] Kim, Sung-Kyu, Ki-Hyuk Yoon, Seon Kyu Yoon, and Heongkyu Ju. "Parallax barrier engineering for image quality improvement in an autostereoscopic 3D display." *Optics express* 23, no. 10 (2015): 13230-13244.
- [155] Xing, Shujun, Shuo Liu, Xinzhu Sang, Huilong Cui, Xunbo Yu, and Duo Chen. "Ground autostereoscopic display with new type non-uniform barrier." *Optik* 127, no. 13 (2016): 5515-5519.