

DISS. ETH NO. 22745

# ADVANCED 3D DISPLAY TECHNOLOGIES

A thesis submitted to attain the degree of  
DOCTOR OF SCIENCES of ETH ZURICH  
(Dr. sc. ETH Zurich)

presented by

NICOLA RANIERI

Master of Science ETH in Computer Science, ETH Zurich

born on 07.08.1985

citizen of Bachenbülach (ZH), Switzerland

accepted on the recommendation of

Prof. Dr. Markus Gross, examiner  
Prof. Dr. Henry Fuchs, co-examiner  
Prof. Dr. Otmar Hilliges, co-examiner

2015



# Abstract

Displays capable of providing the illusion of depth have been emerged over the past decades and even found their way into home entertainment. These devices can be as simple as just separately exposing two different views for the left and the right eye but also can be as complex as reproducing a full 4D light field. While they succeed in creating the illusion of depth, they typically are lacking many visual cues which would be required to give an immersive and natural viewing experience.

In this thesis, we analyze different 3D display technologies. We assess their strengths and weaknesses and explore ways on how to improve them. With the gained knowledge, novel display designs are proposed, built and evaluated.

More specifically, we investigate three different types of displays. We first enhance stereoscopic displays by adding transparency as a feature. Also we propose a retroreflective system to serve independent content to multiple viewers on a large-scale display.

We then generalize the notation of multi-layered volumetric displays by creating a framework which describes the basic display primitives for a existing screens. Within this framework, we then provide the operations applied to light traversing each primitive and provide a decomposition algorithm for a given input light field.

Finally we also provide advanced versions of light field displays. We show how multiple layers can reduce the bandwidth required by a 4D light field display. Then we describe how a transparent light field display can be achieved. Further we describe a light field display with a significant higher bandwidth than state of the art displays, using the temporal domain.

Since many of the created prototypes use multiple layers, we also propose a method to calibrate multi-layered display architectures.

For each idea we have built novel hardware prototypes to proof soundness of our theories. Results are captured and critically discussed to verify the feasibility of the concepts. We show, that understanding both the underlying hardware as well as the theory is crucial and can lead to improved performance and realism of advanced 3D displays.



# Zusammenfassung

Bildschirme mit der Fähigkeit Tiefe darzustellen wurden in den letzten Jahren immer populärer und haben sogar ihren Weg in die Heimunterhaltung gefunden. Diese Geräte stellen zum Teil ganz einfach die unterschiedlichen Bilder für das linke und das rechte Auge separat dar, können aber auch ganz komplex sein und ein volles 4D Lichtfeld reproduzieren. Sie können zwar sehr gut Tiefe darstellen, jedoch fehlt es ihnen typischerweise an vielen anderen visuellen Signalen, welche für eine immersive und natürliche Tiefenwahrnehmung nötig wären.

In dieser Arbeit analysieren wir unterschiedliche 3D Display-Technologien. Wir beurteilen ihre Stärken und Schwächen und suchen Möglichkeiten sie zu verbessern. Mit dem gewonnenen Wissen entwerfen und bauen wir neue Displays und werten diese dann aus.

Inbesondere untersuchen wir drei verschiedene Arten von Displays. Zuerst verbessern wir stereoskopische Bildschirme indem wir sie transparent machen. Desweiteren stellen wir ein retroreflektives System vor, welches unterschiedlichen Inhalt für verschiedene Zuschauer darstellen kann.

Dann verallgemeinern wir die Notation für mehr-schichtige volumetrische Displays indem wir ein Modell entwerfen, welches die Grundbausteine existierender Bildschirme beschreibt. In diesem Modell beschreiben wir dann die Operationen, welche jede Schichtart auf Licht ausübt. Dazu beschreiben wir einen Algorithmus, welcher ein gegebenes Lichtfeld auf die Schichten verteilt.

Zuletzt stellen wir noch neuartige Lichtfeld-Bildschirme vor. Wir zeigen wie mehrere Schichten die benötigte Bandbreite für 4D Lichtfelder drastisch reduziert. Dann stellen wir ein transparentes Lichtfeld-Display vor. Desweiteren beschreiben wir ein Lichtfeld-Bildschirm mit einer signifikant höheren Bandbreite als herkömmliche Displays, welches die zeitliche Domäne nutzt.

Da viele unserer Prototypen aus mehreren Display-Schichten besteht, präsentieren wir einen Kalibrierungs-Algorithmus für mehr-schichtige Display-Architekturen.

Für jede Idee haben wir Prototypen gebaut, welche unsere Theorien bestätigen. Resultate wurden aufgenommen und kritisch hinterfragt um die Realisierbarkeit nachzuweisen. Wir zeigen dass das Verständnis über Hardware und Theorie von 3D Displays zu verbesserten Systemen mit erhöhtem Realismus führen kann.



# Acknowledgments

First I want to thank Prof. Dr. Markus Gross for creating this unique opportunity for me. His initial vision of a transparent 3D Display led to a series of interesting research topics. He gave me all the time, support and funding I required to investigate and work in this promising field of research. With help of his experience and expertise I was able to achieve the results presented in this work.

Then I would like to thank Prof. Dr. Henry Fuchs for the many good and inspirational chats about displays, projectors and virtual reality. His advice and input was and still is always highly appreciated and his ongoing interest in my work very welcome and flattering. Also I am very happy that he agreed to co-examine this thesis.

Also I would like to thank Prof. Dr. Otmar Hilliges for his time and efforts as second co-examiner of my work. His feedback improved this thesis and it was due to his help that my work got into the shape it is.

A very special thanks goes to Dr. Quinn Smithwick. He already helped me in many publications and acted as mentor and as supervisor in so many projects. I enjoyed a lot doing my two internships with him.

Another person with great influence on to my work is Dr. Simon Heinzle. He was my first mentor and taught me many lessons about research and the academic approach how to solve problems. His patience and support in many projects helped me a lot and I would like to thank him for all he has done for me.

Further I want to give special thanks to all my collaborators within the projects we did. I had the luck to have awesome team mates and enjoyed to talk and to discuss with them any matters. Each one shared his or her perspective with me, extending my knowledge and broadening my horizon for which I am deeply grateful. Especially I want to thank Hagen Seifert for his great efforts while working on my ideas during both his bachelor and master thesis.

I am deeply thankful to all my friends and family for their great support throughout my whole Ph.D. Especially during deadlines, I was happy to always be able to rely on them and get any help I needed. Without them, I would not be where I am now.



# Contents

<b>Introduction</b>	<b>1</b>
1.1 Overview . . . . .	3
1.1.1 Stereoscopic Displays . . . . .	3
1.1.2 Volumetric Displays . . . . .	4
1.1.3 Light Field Displays . . . . .	4
1.1.4 Calibration . . . . .	5
1.2 Principal Contributions . . . . .	5
1.3 Thesis Outline . . . . .	7
1.4 Publications . . . . .	7
<b>Related Work</b>	<b>11</b>
2.1 Stereoscopic Displays . . . . .	11
2.2 Volumetric Displays . . . . .	13
2.3 Lightfield Displays . . . . .	15
<b>Stereoscopic Displays</b>	<b>21</b>
3.1 Transparent Backprojection Foils . . . . .	22
3.1.1 Isotropic Back-Projection Screens . . . . .	24
3.1.2 Anisotropic Back-Projection Screens . . . . .	24
3.2 Motion Parallax . . . . .	26
3.2.1 Head Tracking . . . . .	26
3.2.2 Motion Prediction . . . . .	28
3.2.3 Prediction Quality . . . . .	29
3.3 Transparent Telecommunication System . . . . .	29
3.3.1 Display Prototypes . . . . .	29
3.3.2 Communication . . . . .	32
3.3.3 Discussion . . . . .	33
3.4 Large Autostereoscopic Retroreflection Screens . . . . .	33
3.4.1 Prototype . . . . .	36
3.4.2 Software . . . . .	40
3.4.3 Configurations . . . . .	42
3.4.4 Discussion . . . . .	48
3.4.5 Conclusion . . . . .	54

## Contents

<b>Volumetric Displays</b>	<b>55</b>
4.1 Mathematical Framework . . . . .	56
4.1.1 Emissive Layer . . . . .	57
4.1.2 Blocking Layer . . . . .	58
4.1.3 Automultiscopic Layer . . . . .	58
4.2 Light field Decomposition . . . . .	58
4.3 Decomposition Analysis . . . . .	62
4.3.1 Projective Error . . . . .	62
4.3.2 Light Field Sampling Error . . . . .	63
4.3.3 Quantitative Error . . . . .	64
4.4 Physical Prototype . . . . .	66
4.4.1 Temporally Multiplexed . . . . .	67
4.4.2 Spatially Multiplexed . . . . .	68
4.5 Discussion . . . . .	69
4.5.1 Results . . . . .	69
4.5.2 Component Limitations . . . . .	70
4.5.3 Algorithm Limitations . . . . .	70
<b>Lightfield Displays</b>	<b>73</b>
5.1 Multi-Layered Automultiscopic Displays . . . . .	74
5.1.1 Lightfield Distribution . . . . .	75
5.1.2 Algorithm . . . . .	76
5.1.3 Bandwidth Analysis . . . . .	76
5.1.4 Software Framework . . . . .	80
5.1.5 Hardware Setups . . . . .	81
5.1.6 Discussion . . . . .	84
5.2 Transparent Holographic Light Field Display . . . . .	87
5.2.1 Field-sequential Image Multiplexing . . . . .	89
5.2.2 Prototypes . . . . .	91
5.2.3 Discussion . . . . .	93
5.3 Beam Steered Light Field Display . . . . .	94
5.3.1 Gradient Index Model . . . . .	94
5.3.2 LC GRIN Cells . . . . .	96
5.3.3 System Design . . . . .	98
5.3.4 Bandwidth . . . . .	100
5.3.5 Discussion and Application . . . . .	102
<b>Vision Based Calibration</b>	<b>105</b>
6.1 Homography-based mapping . . . . .	108
6.2 Closed Form Solution . . . . .	109
6.3 Free View Transform . . . . .	110
6.4 Calibration . . . . .	112

## Contents

6.5	Implementation . . . . .	114
6.6	Physical Prototype . . . . .	115
6.7	Generalization and Discussion . . . . .	117
<b>Conclusion</b>		<b>119</b>
7.1	Review and Discussion of Principal Contributions . . . . .	120
7.2	Future Work . . . . .	122
<b>Bibliography</b>		<b>125</b>
<b>Curriculum Vitae</b>		<b>133</b>



---

# C H A P T E R

# 1

## Introduction

---



**Figure 1.1:** *Artistic rendering of a futuristic transparent 3D screen using back projection.*

---

## *Introduction*

The idea of a device capable of creating the illusion of depth has fascinated mankind for centuries. Since depth is perceived as disparity between features recognized in the left eye's view and the right eye's view, people realized early that it is sufficient to expose a pair of images taken from slightly different perspectives to the two eyes.

This effect called stereopsis or retinal disparity is a binocular cue which usually is sufficient to create the illusion of depth. However, as depth perception is based on many different binocular and also monocular cues, devices which are only capable of providing stereopsis can cause eye strain, fatigue and even nausea.

As consequence, researchers started investigating devices and methods on how to provide natural viewing experience. For stereoscopic displays, this has mainly be achieved by means of content preparation. User studies have revealed several rules on how content can be prepared to reduce eye strain. This ranges from applying defocus blur over proper clipping to depth compression. Exotic systems even include motion parallax or accommodation. However, many issues still remain in these systems, as they do not approximate reality very well.

To further increase realism, a new 3D display branch emerged, which tried to discretize a volume by covering it with emissive elements. Content which is restricted to this volume can be rasterized into a voxel grid and displayed. Since this approximates real objects very well, most of the perceptual cues for depth are fulfilled. However, drawback of most of these displays is that voxels emit light uniformly into all directions. Thus, together with the typically low resolution, displayed content is restricted to Lambertian surfaces and does not look very natural.

To provide all perceptual cues and display high quality content at the same time, the full light field of a scene has to be reconstructed. Any scene can be expressed as 5D function describing the colors of all possible rays defined by the 3D coordinates of their origin and the 2D coordinates of their direction (e.g. expressed in euclidean space and spherical coordinates). As we are interested only in those rays passing the display surface, the 5D light field can be reduced to four dimensions. Thus, ultimate goal is to create a device which can emit light of any color for any 2D position on the display surface and any 2D direction. Unfortunately, this third type of display suffers both from the discrete nature common to digital systems as well as the immense bandwidth required to process content data. Therefore, clever display systems and methods are required to make such screens feasible.

## 1.1 Overview

Even though light field displays are the most promising one, stereoscopic displays and volumetric displays have their justified area of application and thus too require development to evolve. In this thesis we investigate aspects of each of the three display types, describe and quantify their problems and propose solutions to address some of the issues.

### 1.1.1 Stereoscopic Displays

Stereoscopic displays have been analyzed thoroughly and thus, there are numerous different devices achieving this goal. Displays of this type share in common that their screen is capable of generating two different images for the left and right eye, where one image is only exposed to one eye and the other only to the other. This simple approach led to stereoscopic content being widely available, which supported the further development of such displays.

A broad area of application for such displays is augmented reality (AR) or virtual reality (VR). The real world is enriched with virtual information to assist, guide or simply entertain a viewer. However, most such applications use an opaque screen and overlay the virtual content with a captured image of the reality. Few attempts have been made so far to create a fully transparent stereoscopic display, which is first topic of this thesis.

While opaque stereoscopic displays have been improved to overcome many of the known issues causing eyestrain, transparent 3D displays impose a whole new range of problems to be solved. First of all, common display technologies like LCD or transparent OLED suffer from heavy light loss and thus can only be called semi-transparent and therefore do not provide deep immersion into the VR. Hence, a transparent display technology has to be found which is capable of multiplexing different images to different eyes. Once this has been achieved, special care has to be taken how virtual content is aligned with reality, as misalignments become easily visible due to the transparency.

In this thesis we present a solution based on a transparent backprojection foil as well as a method to render content on it, overcoming both mentioned issues. The system is not only transparent and 3D but also provides motion parallax for an enhanced viewing experience. We deploy our technology in a life-size telecommunication platform.

Besides our transparent system, we present a large autostereoscopic retroreflective display. The screen provides bright images and is capable of serving independent stereoscopic content to different users. In addition of giving each viewer the choice whether he want to see content in 2D or 3D, each viewer can watch his own movie, making the display suitable for many use cases.

### **1.1.2 Volumetric Displays**

There are many different approaches on how to create a volumetric display. In general, these kind of displays create a volume of emissive voxels. Hence they work independent from the viewer position and are capable of providing proper accommodation cues.

In most state of the art volumetric displays, the emissive voxels radiate a single color uniformly into all directions. Therefore, only Lambertian surfaces without view dependent effects can be displayed properly. Furthermore, voxels usually are not capable of occluding other voxels and thus displayed content has a ghostly look. Also, a common framework to describe and compare all these many existing volumetric displays is missing. This makes it especially hard to find a common algorithm on how to prepare content to be shown on these displays.

As part of this thesis we propose a framework describing layered volumetric displays in terms of basic display layer primitives. The framework allows to analyze and compare such displays and to provide a single rendering algorithm applicable to all such screens. We provide two different volumetric display prototypes to show how our theory is applied in practice. Our prototypes are capable of providing both view dependent effects as well as correct occlusions.

### **1.1.3 Light Field Displays**

To overcome many issues with perceptual cues and being able to show naturally looking virtual content, devices are required which are capable of approximating a 4D light field. So called light field displays try to reconstruct the complete light wavefront of a 3D scene, often discretized as ray geometry.

There is mainly one reason why light field displays are not widely available, although this principle is known since a long time. For data storage, transmission and displaying, the continuous 4D signal describing a light field has to be discretized both in the spatial and angular domain. The discrete

sampling of a continuous signal can lead to aliasing which is a very well known and described effect for 2D images. Related work [Chai *et al.*, 2000; Levin and Durand, 2010] has described this effect for 3D scenes and it was shown that the angular sampling of a light field directly influences the aliasing free displayable depth of a scene [Zwicker *et al.*, 2006]. Thus, enormous bandwidth would be required by any light field display to satisfy the Nyquist sampling theorem both in the spatial and the angular domain. If bandwidth is reduced, either the quality of the images is reduced drastically or the displayable depth, both making these systems very unpleasant for viewers to watch.

Hence, ways how to make better use or generally improve the display bandwidth of such systems have to be found. As part of this thesis, we present the math how multiple light field layers can drastically reduce the required bandwidth of the complete display system. We further present a way to create a transparent light field display combining the knowledge of different projects within this thesis. Last but not least we present a novel display technology which has a huge bandwidth compared to existing systems.

### 1.1.4 Calibration

Some of the proposed prototypes in this thesis use multiple layers and it is assumed that the display geometry, how the layers are positioned relative to each other, is known. However, only few methods for automatic calibration of multiple layers are available, most of them are very specific designed for a certain family or even kind of display. Thus this thesis also contains a chapter about calibration of multi-layer displays, which can be generalized to many different types of display systems.

## 1.2 Principal Contributions

The main contributions of this thesis are:

1. **A transparent stereoscopic system with full motion parallax.** We demonstrate a novel display type combining a holographic, anisotropic back-projection screen with a dual projector setup. The system is able to show stereoscopic 3D content while leaving unused parts of the display transparent. It supports viewer tracking and adapts the rendering perspective accordingly to provide full motion

parallax. The display can therefore be used to embed virtual content into the real world. The system is used in a 3D teleconference platform.

2. **A retroreflective large-scale autostereoscopic screen.** We present a novel type of display made of a 1D retroreflective material. Projectors placed over a viewer's head create images on the screen which are reflected back to the eyes located below the projectors. The system can show individual stereoscopic content for each viewer independently and provides good brightness. We evaluate the system for different viewer configurations.
3. **A mathematical framework describing a variety of multi-layered volumetric displays.** It can be used to compare many different volumetric display architectures. We provide a light field decomposition algorithm given a certain display geometry and an input light field. The method can be used to prepare content for any display which can be described by our framework. We show proof of concept on two novel volumetric display prototypes capable of view dependent effects and correct occlusions.
4. **The theory on bandwidth-reduction for multi-layered automulti-scopic displays.** We show how the required total bandwidth of a system can be reduced linearly with the number of layers deployed. The theory is motivated both in ray space as well as in the frequency domain. We analyze and discuss problems induced by the solution and show correctness of the approach on two multi-layered multi-view displays.
5. **A transparent holographic multi-view display.** We present the idea of how to create a display capable of showing glasses-free 3D content while still maintaining transparency. The system uses the full temporal bandwidth of digital micro-mirror devices and is the result of the findings in other contributions of this paper. A light field projector to be used in the final prototype has been built and validated.
6. **A novel architecture for a high bandwidth light field display.** As bandwidth is crucial for light field displays we evaluate a novel display architecture which would drastically enhance the capabilities of light field displays. It deploys a steerable GRIN cell to electronically change the angle of beams of light at real-time refresh rates. Together with a fast switching light source e.g. as a laser diode, an ultra high angular sampling can be achieved enabling 3D content at large depths.

7. **A calibration algorithm for multi-layered displays.** We present a method to compute the mapping of pixels between layers given a certain eye position. This vision based approach uses homographies and thus the pixel grid on a given layer can itself be a projection. The calibration can then be used to find all pixels intersected by a ray defined by the eye position and any pixel of any layer in the display setup. Based on these mappings, content can be prepared to accommodate for layer misalignment or manufacturing errors.

## 1.3 Thesis Outline

This thesis is structured in the following way: Chapter 2 lists related work in the field of stereoscopic, volumetric and light field displays. In Chapter 3 we present our transparent stereoscopic display with motion parallax and describe the teleconferencing system in which it is used. Also in this chapter we describe our large retroreflective display for independent stereoscopic content. A mathematical framework and corresponding light field decomposition algorithm for multi-layered volumetric displays is introduced in Chapter 4. Two prototypes demonstrate practicability of the method. We perform a bandwidth analysis of multi-layered light field displays in Chapter 5 and show how increasing the number of layers decreases the overall required bandwidth of such a system. We then present an idea for a transparent light field display as well as a novel architecture for a light field display with significantly increased bandwidth. In Chapter 6 we propose a method for vision based calibration of multi-layered displays. Chapter 7 concludes the thesis, discusses the contributions and achieved results as well as potential future work.

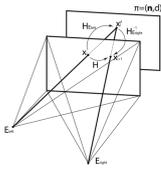
## 1.4 Publications

In the context of this thesis, four first author papers have been accepted.



**N. Ranieri**, H. Seifert, and M. Gross. Transparent Stereoscopic Display and Application, *Proceedings of SPIE 9011 (San Francisco, USA, 4-8 February, 2014)*, *Stereoscopic Displays and Applications XXV*, 90110P, 2014.

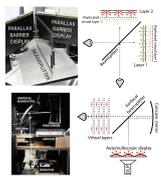
## Introduction



**N. Ranieri**, and M. Gross. Vision-based Calibration of Parallax Barrier Displays, *Proceedings of SPIE 9011 (San-Francisco, USA, 4-8 February, 2014), Stereoscopic Displays and Applications XXV*, 90111D, 2014.

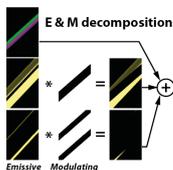


**N. Ranieri**, S. Heinzle, P. Barnum, W. Matusik, and M. Gross. Light-Field Approximation Using Basic Display Layer Primitives, *SID Symposium Digest of Technical Papers 2013 (Vancouver, Canada, 19-24 May, 2013), Display Week*, pp. 408-411, 2013.



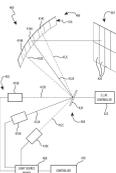
**N. Ranieri**, S. Heinzle, Q. Smithwick, D. Reetz, L. S. Smoot, W. Matusik, and M. Gross. Multi-Layered Automultiscopic Displays, *Proceedings of Pacific Graphics 2012 (Hong Kong, 12-14 September, 2012), Pacific Graphics*, pp. 2135-2143, 2012.

The paper *Light-Field Approximation Using Basic Display Layer Primitives* has been invited to be published as extended version in the Journal of the Society for Information Display.

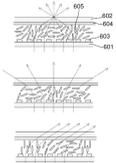


**N. Ranieri**, S. Heinzle, P. Barnum, W. Matusik, and M. Gross. Multi-planar plenoptic displays, *Journal of the Society for Information Display (Blackwell Publishing Ltd)*, **invited journal paper**, vol. 21, no. 10, pp. 451-459, 2013.

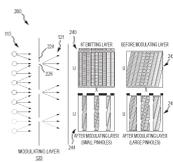
In addition, three inventions related to this thesis have been disclosed.



**N. Ranieri**, Q. Smithwick, H. Seifert, and M. Gross. SYSTEM AND METHOD FOR INCREASED SPATIAL RESOLUTION, *U.S. Patent US2015/250021 A1*, filed on 02.03.2015.

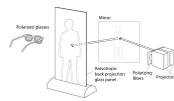


**N. Ranieri**, and M. Gross. DISPLAY DEVICE FOR TIME-SEQUENTIAL MULTI-VIEW CONTENT, *European Patent WO2014/180554 A1*, filed on 05.05.2014.



**N. Ranieri**, S. Heinzle, Q. Smithwick, D. Reetz, L. S. Smoot, W. Matusik, and M. Gross. MULTI-LAYER PLENOPTIC DISPLAYS THAT COMBINE MULTIPLE EMISSIVE AND LIGHT MODULATING PLANES, *U.S. Patent US2013/082905 A1*, filed on 04.04.2013.

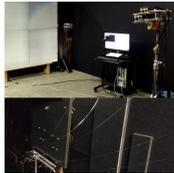
Also, parts of three second author publications have been used in this thesis.



C. Kuster, **N. Ranieri**, T. Martin, P.Y. Laffont, T. Popa, J.C. Bazin, and M. Gross. An Immersive Bidirectional System for Life-size 3D Communication, *submitted to 3DV*, 2015.



Q. Smithwick, and **N. Ranieri**. A Large 1-D Retroreflective Autostereoscopic Display, *Proceedings of SPIE 9391 (San Francisco, USA, 9-11 February, 2015), Stereoscopic Displays and Applications XXVI*, 9391-16, 2015.



C. Kuster, **N. Ranieri**, Agustina, H. Zimmer, J.C. Bazin, C. Sun, T. Popa, and M. Gross. Towards Next Generation 3D Teleconferencing Systems, *Proceedings of 3DTV-CON (Zurich, Switzerland, October 15-17, 2012), 3DTV-CON*, pp. 1-4, 2012.



## *Introduction*

---

# C H A P T E R

# 2

## Related Work

The vision of a device which can show 3D content has been driver of many different research projects and hence the amount of related work is tremendous. Also there is a huge diversity in ideas of how to achieve the goal and thus there are many ways on how to classify and collect them into sets of similar displays. In this thesis, we distinguish for simplicity between stereoscopic displays, volumetric displays and light field displays. However, boundaries are not strict. In fact many displays are a combination of these three classes drawing advantage of each.

Each of these classes has its own list of related publications and it would be beyond the scope of this thesis to list them all. In addition to the related work mentioned here, there are three recommendable surveys giving insight into recent developments. [Dodgson, 2005] describes the principle of autostereoscopic displays and lists some interesting approaches. A survey on volumetric displays is given by [Favalora, 2005]. Finally, [Kim *et al.*, 2010] gives a comprehensive overview over trends in lenticular based systems.

Displays especially important to this thesis are listed below, categorized into the three mentioned classes.

### 2.1 Stereoscopic Displays

Generally speaking, stereoscopic displays are capable of generating at least two distinct images and expose them to the corresponding left and right

## Related Work

eyes. The separation of the images with respect to left and right eye can be done in various ways. They can be temporally multiplexed, meaning the images are displayed time sequentially and blocked in some way for one eye or the other. A simple example are shutter glasses, blocking synchronized with the display the eye which must not see the current frame. Images can also be polarization-division multiplexed or bandwidth multiplexed, where both images are displayed at the same time but with different polarization or filtered colors. Glasses with corresponding filters are then used to select the right image for each eye. Another way of separating images is spatial multiplexing where the different images are exposed to different directions and locations where the right or left eyes are assumed. This approach can also be combined with eye tracking for dynamic update of displayed content. While the glasses based approaches are pretty simple and straight forward, there are many interesting approaches for spatial multiplexing.

[Ives, 1903] patented the idea of using a parallax barrier for spatial separation of left and right image. Small interleaved transparent and opaque stripes are put in front of an image with scrambled left and right views. While one eye can see pixels of one image through the transparent stripes, the pixels of the other image are blocked by the opaque stripes. While his invention described static barriers, the underlying principle was still used in commercial products a hundred years later.

[Isono *et al.*, 1993] extend the static idea of [Ives, 1903] by using a liquid crystal display (LCD) as barrier which dynamically changes the barrier pattern. This allows to adapt the display to a varying viewer position but also enables multi-viewer applications. Furthermore, as the barrier layer can be switched to be completely transparent, the display can change between 2D content at full resolution or 3D content at lower resolution, depending on the number of generated views.

Instead of using a regular pattern, Nashel and Fuchs propose to use a random hole pattern [Nashel and Fuchs, 2009]. Their design eliminates the repeating zones common to regular stereoscopic displays and enables multiple viewers to watch stereoscopic content simultaneously from arbitrary locations. This gives the system a clear advantage over others, as content can be adapted to more than just one tracked viewer. The idea has been taken further in [Ye *et al.*, 2010] to implement an autostereoscopic tabletop system.

Jacobs *et al.* from Sharp Laboratories also combine a LCD display and a retarder with stripe pattern as parallax barrier [Jacobs *et al.*, 2003]. The parallax barrier can be turned off so that the the display maintains its full spatial resolution in 2D mode. In contrast to [Isono *et al.*, 1993], Jacobs *et al.* only describe static applications to simplify manufacturing of those displays.

To maintain the spatial resolution of these displays in 3D mode, Kim et al. proposed to use the temporal domain [Kim *et al.*, 2007]. Instead of a parallax barrier in front of a display they use a pinhole array as back light of a regular LCD. In each frame they move the pinhole array and adapt the content to restore the spatial resolution at the cost of temporal resolution.

To make best use of the dynamic capabilities of LC parallax barriers, both Perlin et al. and Peterka et al. track a users eye position and adapt content as well as barrier accordingly. Perlin et al. propose an iterative method on how to compute the scrambling pattern of the image layer given the eye positions relative to the screen and under the assumption that the image and barrier layer are perfectly coplanar and with known distance in between [Perlin *et al.*, 2000]. [Peterka *et al.*, 2008] too use information of a viewers position to adapt the barrier. In addition they use slanted barrier patterns to equally distribute the loss of resolution to the horizontal and vertical dimension of the screen. Also they scramble multiple completely distinct images for each viewer.

As summarized by [Dodgson, 2005], lenses can be used instead of parallax barriers, which is also known as integral photography introduced by Lippmann in 1908 [Lippmann, 1908]. In contrast to parallax barrier displays, systems based on lenses do not block light and thus are brighter at the cost of reduced flexibility present in dynamic parallax barrier displays. This issue has been addressed by Kao et al. who propose to use liquid crystal structures to create tunable gradient index (GRIN) lenses [Kao *et al.*, 2009]. As liquid crystals can be used to create a phase delay it can be arranged such that a LC cell acts as a GRIN lens. This makes integral imaging devices as flexible as LCD based parallax barrier systems and the display can switch between full resolution 2D when the lens is turned off and low resolution 3D when the lens is turned on.

## 2.2 Volumetric Displays

As indicated by the name, volumetric displays work rather on a volume than images taken from certain perspectives of a scene. Instead of a 2D rasterisation like occurring in the image, content is rasterised in 3D space creating a voxel grid. The display then somehow emits proper colors in the corresponding voxel. This naturally approximates a given scene and thus, the viewing experience is immersive as many of the visual cues are fulfilled. However, most of these displays restrict their content to the working volume, are hard to scale and thus limited strictly. Furthermore, most of the displays deploy technologies which remain translucent also in voxels which are covered by

## Related Work

content, giving the displayed scene a ghostlike look. Finally, the emitted light usually has the same color and brightness emitted to all directions. Therefore, only Lambertian surfaces can be displayed properly and no view dependent effects are possible. Volumetric displays come in many exotic forms, some of them deploying only static or electrically switchable hardware and some including mechanically moving parts.

A completely static setup has been proposed by Tamura and Tanaka, who use beam splitter to overlay commercially available 2D displays [Tamura and Tanaka, 1982]. Through a slight offset of each display or parts of the same display to the half mirror, each image appears at a slightly different distance to the viewer. The spatial resolution of each layer corresponds to the resolution of the displays or used parts of the display. The resolution in depth is the number of displays or used parts of the display and thus the number of layers. Therefore, the depth resolution is very coarse.

In contrast to this multi-layer approach which displays each layer at the same time, [Buzak, 1985] developed a field-sequential display using a set of electrically switchable bi-state mirrors. The mirrors can either transmit or reflect light dynamically. Thus they can be used to change the apparent distance between viewer and display plane. If the switching is done fast enough, the individual layers will be perceived as volume and the system behaves the same like the approach presented in [Tamura and Tanaka, 1982].

Similar to that idea, Leung et al. use a stack of LCD where one layer shows the corresponding image while the others are set to a transmissive state [Leung *et al.*, 1998]. In each frame, the active layer is switched, building up the volume over time. Similar to [Buzak, 1985], the depth resolution is limited by the switching speed of each individual layer, as a certain frame rate needs to be maintained to perceive the display as one volume rather than consecutively shown layers. [Bell *et al.*, 2008] describe a similar system but in addition make use of the stacked LCD to increase the contrast and dynamic range of the displayed content.

Gold and Freeman patented another field-sequential approach but using bi-state shutter glasses instead of LCDs [Gold and Freeman, 1998]. This kind of special glass can electrically be switched between a state which is up to 80% transparent and a state where it is diffusing light. Similar to the multi-layered LCD approach, one of the glasses is turned opaque while the others are set to their transparent state. In combination with a high speed projector, a volume of emissive voxels can be created. Sullivan used the same principle to create a commercially available product [Sullivan, 2004] called DepthCube.

An advanced method how to span up a volume of emissive voxels is to

use a varifocal mirror. A varifocal mirror is a reflecting surface capable of being bent dynamically. Approximating a parabolic surface, changing the curvature of the mirror influences its focal length. Traub changes the curvature electronically to move a planar surface virtually forth and back [Traub, 1967]. Smithwick et al. took the idea even further using a varifocal beam splitter and an additional parabolic mirror to make the volume float in space in front of the viewer [Smithwick *et al.*, 2012]. Their approach would allow for viewer interaction, as users can touch the virtual content.

Suyama et al. present an approach that improves the problem of coarse depth resolution due to a relatively small number of layers [Suyama *et al.*, 2000; 2004]. They show that the pixels of different layers can be attenuated according to the depth of the element they represent. By doing so, the visual system of the human brain gets tricked and virtually the pixels are set in between the layers. This greatly enhances the perceived depth resolution.

Based on these insights, Holroyd et al. provide algorithms and methods to convert a 3D model into a stack of 2D layers [Holroyd *et al.*, 2011]. They show how to use these multiple layers to cheaply produce a volumetric display. Though the proposed product is static, their insights can be applied to other multi-layered volumetric displays.

A further approach to mention here has been presented by Akeley et al., who combine stereoscopic displays with the volumetric approach [Akeley *et al.*, 2004]. Each eye's view is led through a series of mirrors and beam splitters to a separate multi-layered volumetric display. This approach has all benefits of volumetric displays but also provides proper occlusions and view dependent effects. It is a good example for the fuzzy border between the three classes of display types presented here.

## 2.3 Lightfield Displays

Instead of rendering perspectives as done for stereoscopic displays or creating voxels for volumetric displays, light field displays directly approximate the plenoptic function. Any scene can be interpreted as 5D plenoptic function. The five dimensions can be interpreted as a 3D spatial coordinate and 2D spherical coordinate defining together the start point and direction of waves of light in the scene. In the discrete case, this corresponds to the coordinates of rays perpendicular to the wave front, and the function returns the color and intensity of the specific ray. For 3D displays, only those rays are of interest penetrating the display surface. This reduces the plenoptic function to four dimensions, often expressed as a pair of 2D coordinates describing

## Related Work

the intersection coordinates of the ray with two co-planar planes with unit distance [Zwicker *et al.*, 2006].

In a light field display, the continuous plenoptic function is usually sampled discretely. The sampling is done once in the spatial domain, corresponding to the sampling of any common 2D display, and once in the angular domain to make each spatial pixel emitting different colors to different directions. Chai *et al.* show, that there is a strong relation between the angular sampling frequency and the depth which can be displayed by the screen without aliasing [Chai *et al.*, 2000]. Thus, the higher the angular resolution of a light field display, the higher the depth which can be displayed. Content outside this depth of field must be filtered with proper methods proposed by [Zwicker *et al.*, 2006]. Similar to the blur in filtered 2D images, the displayed content is blurred according to the depth and is thus visually less appealing, but remains aliasing free.

Levin and Durand give further insights into the theory of light fields. In particular, they investigated the frequency spectrum of the 4D plenoptic function. They found out that nearly all the energy of a light field is covered by only a 3D volume out of the possible 4D frequency space [Levin and Durand, 2010]. This is especially interesting since a better use of this lost bandwidth could significantly increase the displayable depth of light field displays.

There are numerous displays capable of approximating the plenoptic function. Already in 1908, Lippmann presented a way how this can be done using a lens array [Lippmann, 1908]. Depending on the direction in which one looks at a collimator lens, a different point on the back of the lens is in focus. Thus, if an intensity function is generated behind each lens, the lens can be viewed as a pixel with view dependent intensity. If the intensity function behind the lens is discrete as e.g. in a pixel grid, this discretization corresponds to the angular sampling.

Thus light field displays typically trade in spatial resolution to create angular images. In the case of the approach proposed by [Lippmann, 1908], this linear decrease of spatial resolution with the number of angular views would affect horizontal resolution only. To make use and sacrifice horizontal and vertical resolution equally, Berkel suggests to slant lenses [Berkel, 1999]. He also shows how to prepare content and further make use of the individual RGB pixels.

Another way to increase bandwidth is to simply put more angular pixels behind each lens/spatial pixel. Takaki proposes to use an array of small LCD panels and optics to expose each of the images to a different direction [Takaki,

2006]. They have built a prototype which generates 64 angular views with a dense enough angular sampling to provide accommodation cues to the human eye.

Kao et al. presented a clever way of how to use liquid crystal cells as electronically switchable gradient index (GRIN) lens [Kao *et al.*, 2009]. The technology can be used as 3D display similar to how it was introduced by [Lippmann, 1908]. They provide methods for tracing rays through such a system and provide simulated results. Similarly, Oka et al. propose a GRIN based 3D system and provide crosstalk measurements [Oka *et al.*, 2013].

Itoh uses the very same principle of liquid crystal GRIN lenses in a commercial product [Itoh, 2012]. They further use the flexibility of these lenses to create a switchable 2D/3D display. In 3D mode, the display suffers from the common spatial resolution loss to create multiple views. However, when the lenses are switched off, the display maintains its full spatial resolution. In this way, no resolution is lost when content is viewed in 2D mode.

Another way to increase angular resolution is to use the temporal domain. If the display is capable of creating a higher frame rate than required by the human eye, the additional frame rate can be used to produce additional views. Jang and Javidi propose to do so by shifting a micro lens array [Jang and Javidi, 2002]. By shifting the lenses by only a small amount, the resulting views shift by half an angular sampling, doubling the angular resolution. In their work, they use a real scene with a strobing light to simulate a high speed display device. The same approach has been chosen by Bogaert et al. [Bogaert *et al.*, 2010]. They use a fast DLP projection device and a vibrating lens array to create a small offset in the directional views and thus create a digital version of [Jang and Javidi, 2002].

Another way to create additional views has been proposed by Fuchs et al. [Fuchs *et al.*, 2008]. They use directional passive back light which is modulated by an imaging device. Thus, they extend the 3D capability of an integral imaging device by the additional degrees of freedom of a directional back light.

Cossairt et al. expand the idea of volumetric display to create a light field display [Cossairt *et al.*, 2007]. A fast projection device is synchronized with a rotating diffuser, exposing the proper perspective corresponding to a certain diffuser orientation. Using a screen which does not diffuse light uniformly but direction dependent, they create a volumetric display capable of occlusions as well as view dependent effects. Thus their approach can be classified rather as a light field display than a volumetric display.

Jones et al. propose at the same time to use a mirror instead of a diffuser

## Related Work

which simply reflects the light instead of spreading it [Jones *et al.*, 2007]. Due to the sweep of the mirror, images can still be perceived by the human eye. Also, due to the discrete updates of the projection device, the angular sampling is still discrete, although the mirror is moving continuously. They present render methods and show the capabilities of their system in a live teleconferencing system.

This approach has been taken even further by Butler *et al.* [Butler *et al.*, 2011]. They use optical elements to make the 3D content visible outside the volume of the rotating mirror. Thus the system can become user interactive, since displayed content can be touched. A Kinect and other acquisition devices are used to capture viewer input.

Instead of the temporal domain, multiple image generating devices can be used instead. Nagano *et al.* use an array of pico-projectors, projecting images from different angles onto a retroreflective screen [Nagano *et al.*, 2013]. The reflected rays form a light field, since rays of different projectors hitting the same pixel on the screen are reflected into distinct directions. Thus the angular sampling corresponds to the number of used projectors divided by the field of view.

Based on a similar principle, Holografika introduces a scalable lightfield display system called HoloVizio [Balogh *et al.*, 2008]. They use an array of optical modules behind a holographic foil, assembling in each pixel the rays coming from the different modules. In their work they present both a small as well as a large prototype.

A computational approach has been presented by Lanman *et al.* by stacking two LCDs [Lanman *et al.*, 2010]. Back light has to pass one pixel of each layer and gets modulated by both of them, forming a ray of a certain color. No two distinct rays share the same two pixels, but each pixel contributes to multiple rays. Thus an overdetermined linear system can be solved to allocate the different pixels in a way to approximate a certain light field.

Gotoda extended this idea by using multiple LCD layers [Gotoda, 2010]. Instead of working on the attenuation of a ray in each pixel, he describes the problem as a concatenation of rotations of the polarization direction of the ray, resulting in a certain intensity once the ray leaves the display system. Each additional layer adds degrees of freedom to the linear system to be solved and thus enhances the quality of the results. In his follow up work he presents methods to reduce the blur of the results even further [Gotoda, 2011]. A slightly different notation and solution to the same problem using the same approach has been presented one year later in [Lanman *et al.*, 2011].

The system implemented in [Lanman *et al.*, 2011] has further been improved

by Mamoine et al. who added eye tracking to the display [Maimone *et al.*, 2014]. Knowing a rough estimate of a viewers eye positions, only rays coming close to these eye positions are considered in the linear system. This significantly reduces the constraints and thus leads to better images and increased quality.

Another computational approach based on the tomographic principle has been introduced by Wetzstein et al. [Wetzstein *et al.*, 2011]. They formulate the light field problem as integral of a ray passing through an attenuation volume. They provide the method on how to compute the attenuation map to approximate a given input light field. Proof of concept is shown on stacked acrylic plates using colors as attenuation material.

They take their approach even further, combining it with the approach presented in [Lanman *et al.*, 2011] to create an advanced light field display system [Wetzstein *et al.*, 2012]. Directional back light as well as the light field factorization provide many degrees of freedom and thus the approach results in a display with high bandwidth. They give the method for rendering content as well as an analysis on bandwidth of these kind of displays.

The same concept has been extended by assuming known eye positions by Mamoine et al. [Maimone *et al.*, 2013]. They constrain the tensor factorization of the rays to only those rays being close to the eye position. Besides an improved display bandwidth through lesser constraints on the linear system they also increase the ray density around a viewers eyes. This provides potentially nearly correct accommodation cues besides the binocular and motion parallax cues.

## *Related Work*

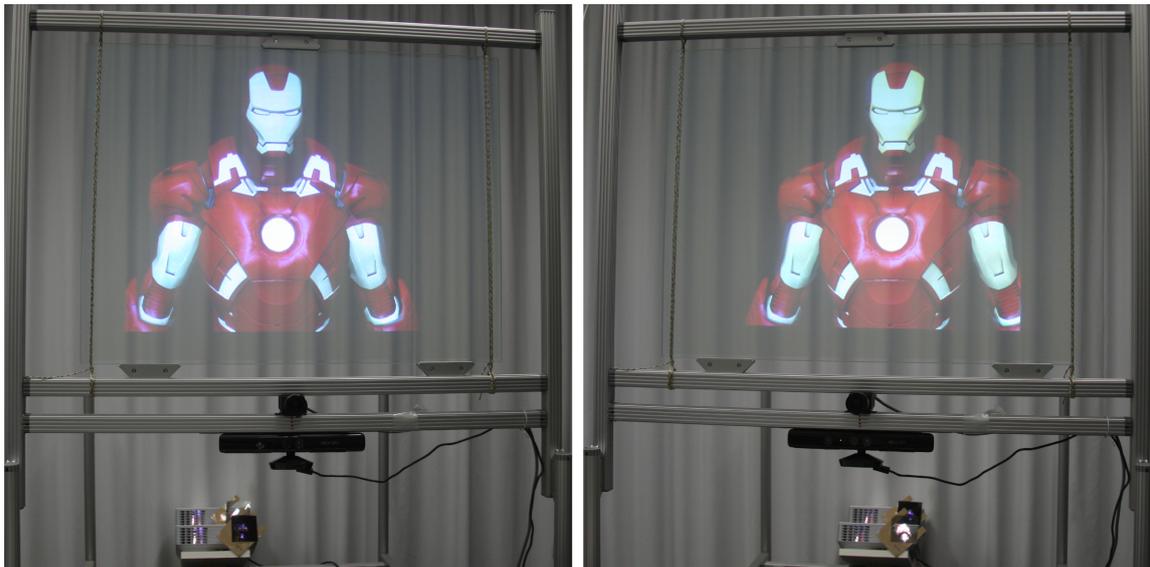
---

# C H A P T E R

# 3

## Stereoscopic Displays

---



**Figure 3.1:** *Results captured on our transparent stereoscopic screen. Left and right images are perspectives filtered with the corresponding linear polarizer, illustrating what a viewer would see when wearing matching polarized glasses.*

---

Stereoscopic displays have found wide application in home entertainment. Nearly any display sold nowadays is 3D enabled and can deliver 3D content when viewed with proper shutter or polarizing glasses. However, wearing glasses is not perceived well by customers and the trend is ceasing already. Furthermore, few attempts have been made to use transparent display technology for visualizing 3D content.

Glasses-free approaches have another problem: They typically have a certain viewing zone or a sweet spot where 3D will be perceived. If a viewer leaves this zone, he will not perceive content in proper 3D anymore. This problem can be solved by adding viewer tracking to the system.

Major problem of transparent 3D systems is how to align virtual content with the real world. This too can be solved by tracking a viewer's eye position relative to the screen and adapting the rendered content accordingly. Though, tracking devices usually suffer from low refresh rates and high latency. These issues can be overcome using a Kalman filter and prediction, which has been successfully done for human motion and opaque 3D screens in previous work [Azarbayejani *et al.*, 1993; Azuma, 1995; Jebara and Pentland, 1997; Machida *et al.*, 2012].

But transparent 3D displays demand a much higher accuracy and lower latency, as the viewer can see and compare real objects and virtual content at the same time with respect to their movement. Thus, these existing approaches have to be re-evaluated to see whether the quality still can achieve an immersive experience.

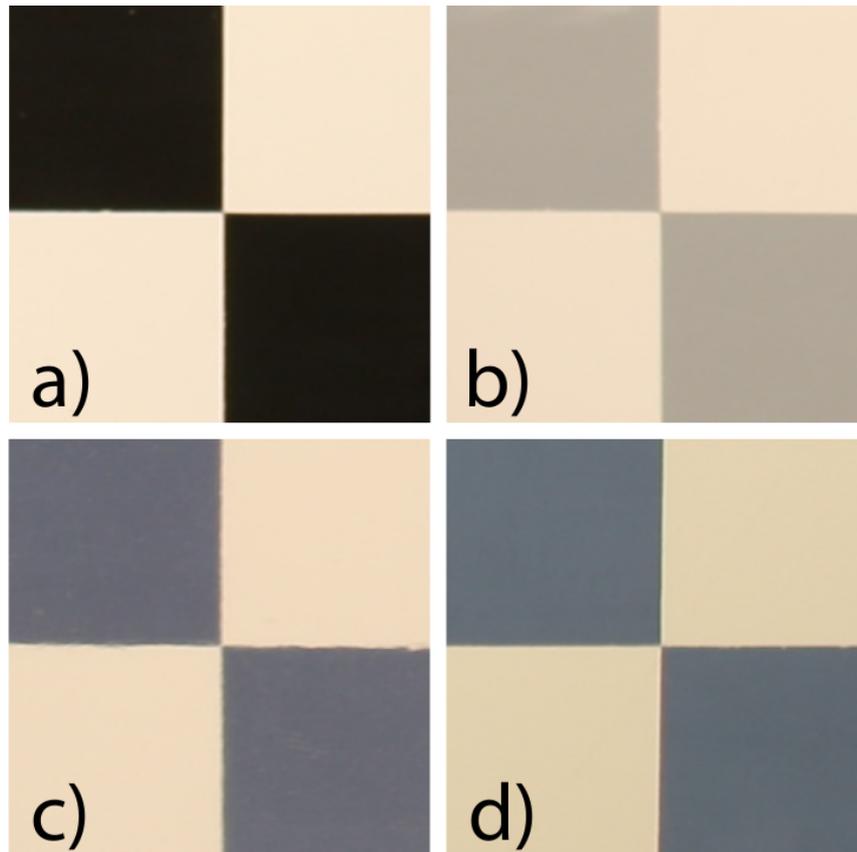
We therefore evaluate different transparent projection screens in combination with viewer adaptive 3D content to prove feasibility of transparent augmented reality displays [Ranieri *et al.*, 2014]. We then use our screen in an immersive teleconferencing system to prove validity of the concept [Kuster *et al.*, 2012; 2015]. Later in this chapter we then present a glasses-free approach with the further advantage that each viewer can watch his individual content [Smithwick and Ranieri, 2015].

### 3.1 Transparent Backprojection Foils

In applications with a static screen and sufficient space for a projector setup, transparent back-projection screens offer a simple and convenient way to overlay virtual 3D content with the real environment. Generally speaking, there are two kind of back-projection technologies: anisotropic and isotropic ones. Anisotropic back-projection screens redirect or diffuse light only if coming from a specific direction, where isotropic back-projection screens

### 3.1 Transparent Backprojection Foils

diffuse all incoming light from any direction. Both can be used either with active or passive 3D glasses. For active shutter glasses, consecutive frames show views for the left and right eyes, which are then blocked by the left or right shutter glass accordingly. For passive glasses, different polarizations are used for the left and right view, which are filtered in the corresponding polarizer of the glasses.



**Figure 3.2:** Comparison of different back-projection screens. *a) shows a test pattern with no screen in front used as ground truth. b) shows the test pattern with an isotropic back-projection screen in front. c) shows the effect of an anisotropic foil and d) the effect of an anisotropic glass. The low contrast common to isotropic screens is clearly visible when comparing a) and b). Anisotropic glass and foil perform equally well with the foil being a little brighter than the glass. All pictures are taken with the same camera, the same settings and under the same lighting conditions.*

### **3.1.1 Isotropic Back-Projection Screens**

Isotropic Back-Projection Screens diffuse incident light from all angles into all directions. Hence the projectors can be positioned anywhere towards the screen, giving the setup much flexibility. On the other hand it is weak against environment light as also light e.g. from a lamp will be diffused, lowering the quality of the displayed image.

Isotropic back-projection screens can be realized by particles or droplets embedded in a clear material. Light intersecting with these droplets are diffused into all directions. As light is diffused even within the screen and also is totally internally reflected, subsurface scattering occurs. This creates a blurred ghost around the displayed image. Also, the ratio between area covered by droplets and clear area defines the transparency as well as the brightness of displayed images. To be able to compete with the brightness of other technologies, this trade-off usually is chosen in a way which is resulting in poor contrast. A comparison between isotropic and anisotropic screens is illustrated in Figure 3.2.

Isotropic back-projection foils are sold as polarization preserving which is true for the directly diffused light. However, the light which is scattered screen internally loses its polarization and is hence perceived as crosstalk in a passive stereo approach, where it is perceived as low frequency/ blurred halo with active shutter glasses.

### **3.1.2 Anisotropic Back-Projection Screens**

Anisotropic Back-Projection Screens selectively redirect light coming from a specific direction. Usually they are optimized such that for each point on the screen only the light rays from a specific center of projection are first redirected to coincide with the screen surface normal and then spread to form a certain field of view. Hence, this approach is very robust against environment light, as incident light from other directions than the optimized center of projection can pass unhampered. As drawback, the setup is fix and must maintain distance as well as direction between screen and projector. Furthermore, much light passes these screens without being redirected or diffused and hence the approach is not very light efficient. Also, the light passing the screen creates a secondary image on the wall which is disturbing. It can be unacceptable in certain situations but can often be avoided by cleverly position the display setup.

Anisotropic back-projection screens can be realized with diffractive optical

### 3.1 Transparent Backprojection Foils

elements (DOE): A holographic grating is used to redirect and shape the incident beams. Therefore such screens also suffer from the rainbow effect, as different wavelengths are redirected differently.

The holographic film itself is usually polarization preserving and hence, both passive and active stereo can be used with such screens. However, the coating or the material in which the film is embedded might be depolarizing and thus this property has to be verified by the manufacturer. Due to the structure in the DOE, the polarizer for the left and right view do not only have to be aligned with the filters in the glasses but also towards the screen. In addition, crosstalk is higher at flat viewing angles, narrowing the field of view. The active approach mostly suffers from the known drawbacks as bulky glasses and required line of sight to the synchronization beacon.

Table 3.1 gives a summarized overview and lists the advantages and drawbacks of each approach in use with active or passive stereo.

	<b>anisotropic</b>	<b>isotropic</b>
<b>passive</b>	<ul style="list-style-type: none"> <li>+ cheap lightweight glasses</li> <li>– light inefficient</li> <li>+ robust against environment light</li> <li>– fix projector position</li> <li>– crosstalk at flat viewing angles</li> <li>– rainbow effect</li> </ul>	<ul style="list-style-type: none"> <li>+ cheap lightweight glasses</li> <li>– poor contrast</li> <li>– weak against environment light</li> <li>+ arbitrary projector position</li> <li>– crosstalk at flat viewing angles</li> <li>– subsurface scattering perceived as crosstalk</li> </ul>
<b>active</b>	<ul style="list-style-type: none"> <li>– heavy glasses and IR beacon</li> <li>– light inefficient</li> <li>+ robust against environment light</li> <li>– fix projector position</li> <li>– rainbow effect</li> </ul>	<ul style="list-style-type: none"> <li>– heavy glasses and IR beacon</li> <li>– poor contrast</li> <li>– weak against environment light</li> <li>+ arbitrary projector position</li> <li>– subsurface scattering perceived as halo</li> </ul>

**Table 3.1:** Comparison of anisotropic and isotropic back-projection screens in use with active or passive glasses.

## 3.2 Motion Parallax

An important monocular cue for 3D perception is motion parallax. For transparent 3D display systems, this perceptual cue becomes very important as the viewer can compare the motion of virtual objects with real objects behind the screen. Thus we track a viewer's head and eye position to render the virtual content from the proper perspective.

### 3.2.1 Head Tracking

To adapt the rendering perspective and to provide motion parallax to the user, head tracking is performed using the Microsoft Kinect. The Kinect was released by Microsoft as a motion sensing input device for the Xbox 360 in 2010. The subsequent release of Windows drivers and SDK, low cost and high availability make it interesting for research purposes as well. To overcome the Kinect's inherent latency and low frame rate, we apply motion prediction using a Kalman filter.

The sensor sends a color and a depth stream from its cameras to the host computer, which are used for tracking. The Kinect SDK and Developer Toolkit provide a set of useful libraries to process these data.

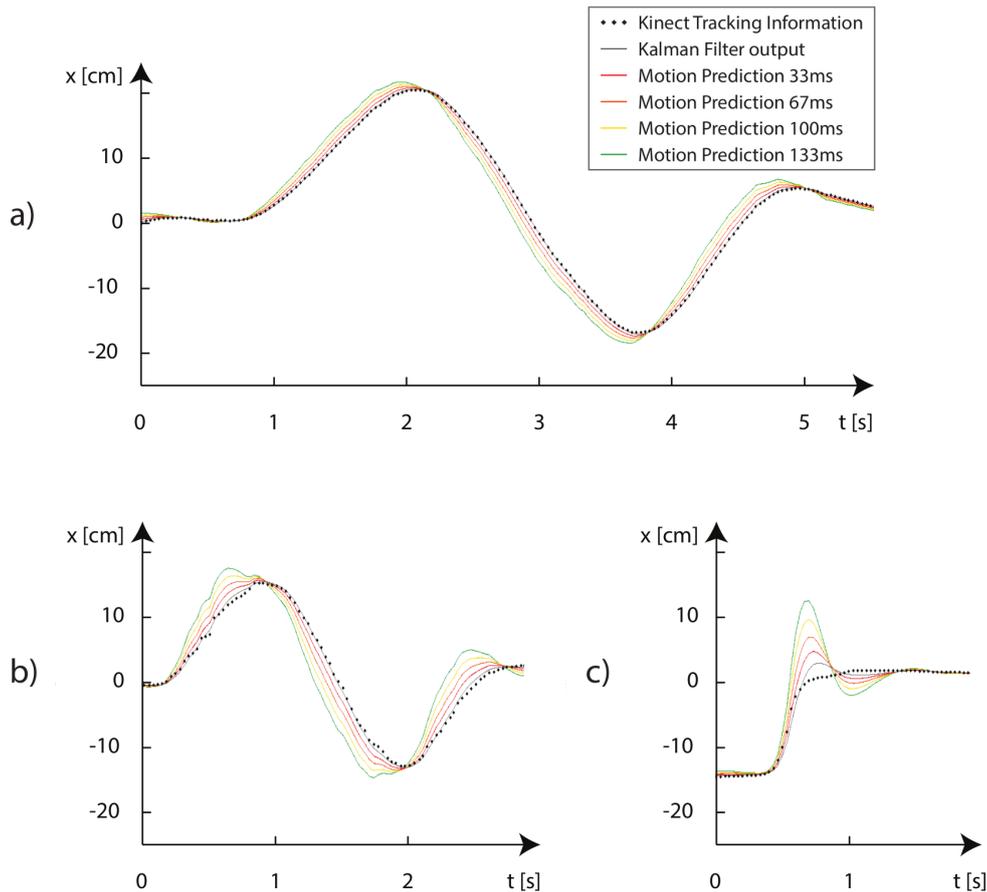
**Skeleton Tracking** provides the 3D positions of 21 joints of the tracked body. It is robust against almost any body orientation and even partial occlusion. However, the head is represented by only a single joint and its accuracy lies only in the centimeter range.

**Face Tracking** detects faces with high accuracy and returns 121 3D vertices for the face incorporating the user specific face shape and current mimics. Unfortunately the high accuracy comes at the cost of robustness. The face is often lost when it is tilted or rotated away from the Kinect too much, or if the distance between face and sensor exceeds  $1.5m$ .

Sudden losses of tracked faces greatly disrupt the immersion and thus, continuity is of high importance. Hence, *Skeleton Tracking* is the preferred tracking method, despite its shortcomings in accuracy. Alternatively, both approaches can be combined, such that face tracking is used when available and skeleton tracking used when the face cannot be detected.

A common problem in either tracking method is the delay imposed by the Kinect sensor which is reported to be about  $125ms$  [Livingston *et al.*, 2012]. If a user located  $1m$  in front of the screen moves sideways at  $1\frac{m}{s}$  and the virtual 3D content is  $1m$  behind the screen, this content appears to be offset by  $12.5cm$

in the opposite moving direction of the user. When the user stops his motion, content aligns in the following  $125ms$ , inhibiting the immersive experience. Furthermore, the sensor's frame rate is only 30 frames per second. This adds a varying delay between 0 and  $33ms$  which leads to a perceived jitter of the content of up to  $3cm$  using the same settings as above. Both the jitter and the delay can be counteracted by motion prediction using a Kalman filter.



**Figure 3.3:** *These plots compare predicted with measured eye positions of a viewer. The black dotted line represents the x-coordinates captured with full latency, where the dashed lines represent x-coordinates at different prediction time. Ideally, all colored curves are identical with an offset on the temporal axis  $t$ , corresponding to the time predicted into the future. a) First order motion prediction of face tracking leads to very smooth results. b) Second order motion prediction of skeleton tracking with noise decreases the quality of the predicted curves heavily. c) Illustration of overshoot and ringing artifacts caused by sudden motion changes and second order motion prediction.*

### 3.2.2 Motion Prediction

The Kalman filter is an algorithm used for stochastic state estimation from noisy sensor measurements and can also be used for motion prediction. It is used predominantly in military and civilian navigation systems such as GPS. More significantly for this work, it is also used extensively for tracking in interactive computer graphics [Welch and Bishop, 1995; Azuma, 1995].

The state of the tracked head is described by the following discrete time model:

$$x_k = Ax_{k-1} + w_{k-1} \quad (3.1)$$

where  $x$  represents the state vector,  $A$  the state transition matrix, and  $w$  the process noise which is assumed to be Gaussian white noise. The state vector contains the position, velocity and, optionally, the acceleration of the tracked feature. For position and velocity, the model can be found from Newton's laws of motion:

$$x = \begin{bmatrix} s \\ \dot{s} \end{bmatrix} \quad A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad (3.2)$$

Measured values  $z$  relate to the estimated state by the observation model  $H$  and the normally distributed measurement noise  $v$ :

$$z_k = Hx_k + v_k \quad H = [1 \ 0] \quad (3.3)$$

For the Kalman filter, the process and measurement noise covariances  $Q$  and  $R$  are required additionally, which are design parameters of the filter.

$$Q = E[w_k w_k^T] = \begin{bmatrix} \frac{1}{4}\Delta t^4 & \frac{1}{2}\Delta t^3 \\ \frac{1}{2}\Delta t^3 & \Delta t^2 \end{bmatrix} \sigma_a^2 \quad R = E[v_k v_k^T] = \sigma_m^2 \quad (3.4)$$

The execution can be divided into two steps. In the *prediction* step, the current state is estimated by propagating the previous one.

$$\hat{x}_k^- = A\hat{x}_{k-1} \quad (3.5)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (3.6)$$

In the *correction* step, the resulting *a priori* estimate  $\hat{x}_k^-$  is updated with the measurement results to receive the improved *a posteriori* estimate  $\hat{x}_k$ .

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (3.7)$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \quad (3.8)$$

$$P_k = (I - K_k H)P_k^- \quad (3.9)$$

### 3.3 Transparent Telecommunication System

A more detailed explanation of the Kalman algorithm and notation can be found in [Welch and Bishop, 1995].

Multiple prediction steps are performed for each received Kinect frame, increasing the refresh rate and thus lowering the jitter caused by the low Kinect frame rate. To improve the overall latency of  $125ms$ , a further prediction by a constant amount of time can be applied.

The prediction results from the Kalman filter largely depend on the choice of  $Q$  and  $R$ . Furthermore, larger prediction times lead to noisy results and overshoot. In our experiments, prediction times of  $67ms$  led to the best compromise in latency reduction and tracking quality, halving the Kinect's latency. The lowered latency and jiggling demonstrate the potential of motion prediction for the Kinect and for head tracking.

#### 3.2.3 Prediction Quality

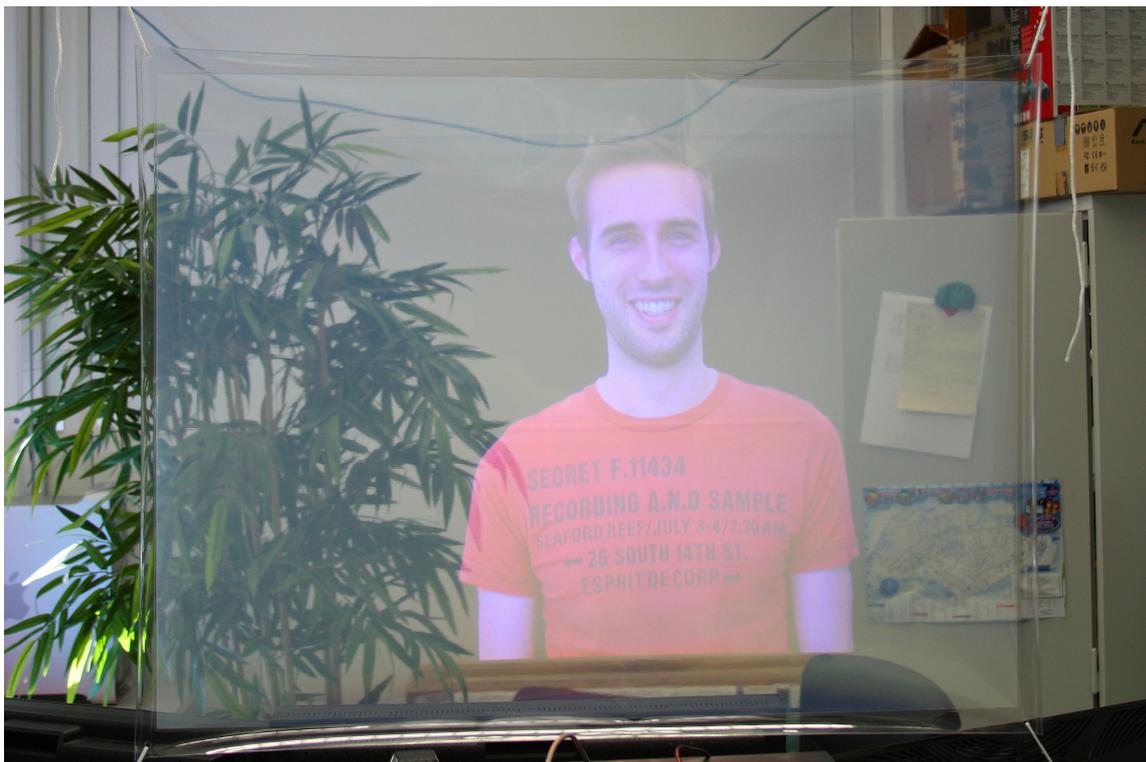
Quality of the improved tracking results are shown in Figure 3.3. The plots compare predicted, up-sampled and filtered signals with different prediction time (dashed lines) with the measured signal (dotted line). The predicted signal decreases in quality with increasing prediction time. Our system is able to provide motion parallax for slow viewer motion. If the viewer moves fast, further improvements or a better tracker would be required, as the prediction time of  $67ms$  becomes too large.

### 3.3 Transparent Telecommunication System

We use our transparent stereoscopic display technology in a telecommunication system for augmented telepresence [Kuster *et al.*, 2012; 2015]. For this purpose we have built different prototypes with different advantages and disadvantages presented in this section.

#### 3.3.1 Display Prototypes

We have built several prototypes for transparent stereoscopic 3D. For the screen we decided to use an anisotropic backprojection foil called HOPS.® from Visionoptics. The screen is robust against environment light and achieves high quality imaging, allowing immersion into the virtual reality. It comes both as light weight foil as the one used for the prototype shown



**Figure 3.4:** *A TV sized transparent screen for immersive teleconferencing. This prototype uses alternate frame sequencing and shutter glasses to provide a viewer with stereoscopic 3D. The transparent screen is made of the same material as our life-size version, but without the glass coating.*

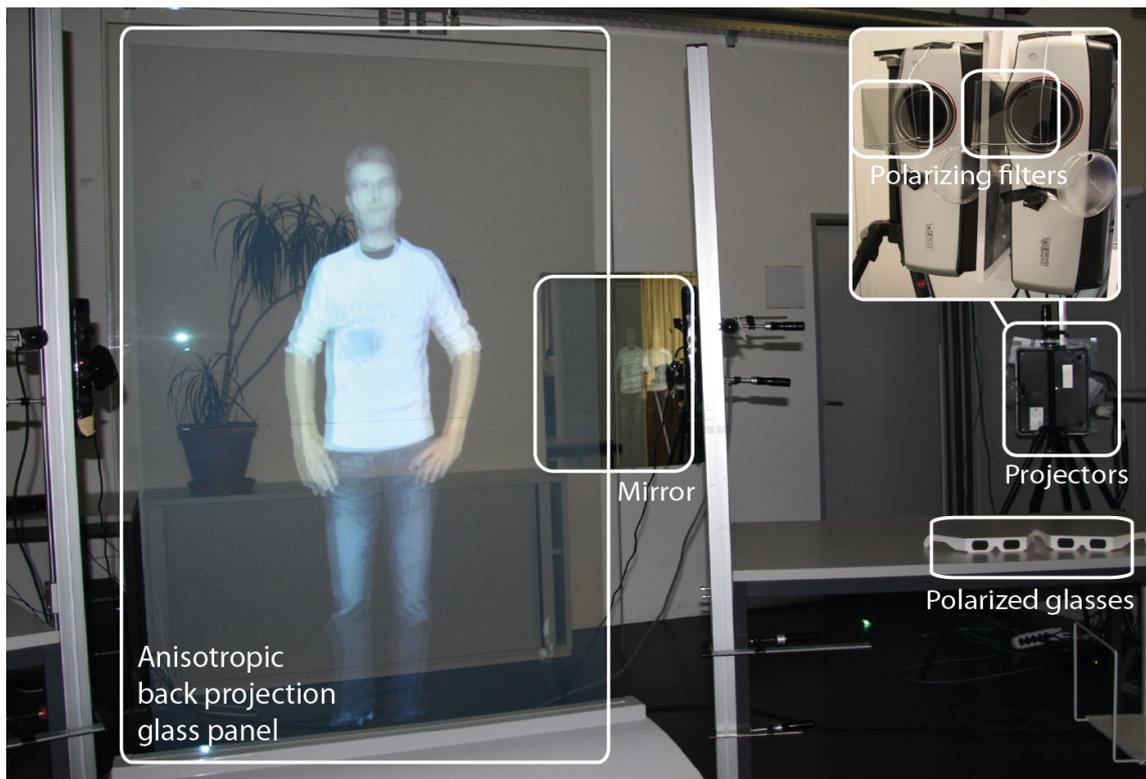
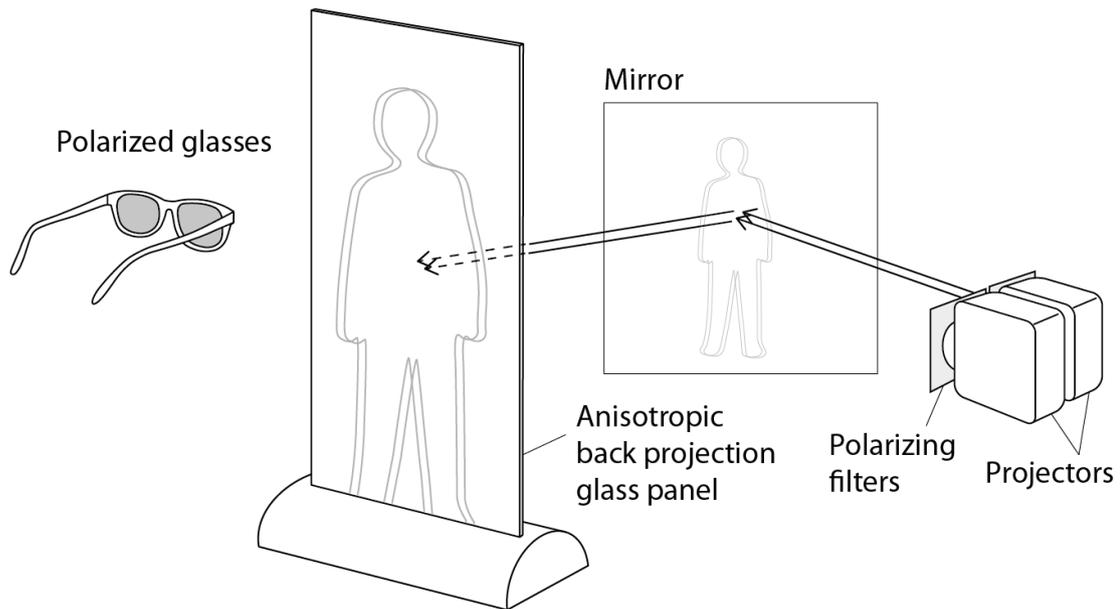
in Figure 3.4 or coated in glass as the one used for our life-size version shown in Figure 3.5.

In our first prototype visible in Figure 3.4 we use a HOPS.® foil. Since the foil is not polarization preserving we use the nVidia 3D Vision enabled projector model H5360 from acer with active shutter glasses, as described in [Kuster *et al.*, 2012]. Advantage is the low crosstalk, however, the glasses are bulky and the viewer requires clear sight to the synchronization beacon.

Thus, in our second prototype shown in Figure 3.1, we use the HOPS.® coated in glass, which preserves linear polarization. We deploy two Acer H5360 projectors with linear polarizers in front. Viewed through glasses with corresponding linear polarization, 3D can be seen. Crosstalk is acceptable and the light weight glasses provide better immersion into the augmented reality. The DualHead2Go multi-display adapter from Matrox is used to provide both projectors with perfectly synchronized images.

Finally we have built the same system in a life-size version depicted in

### 3.3 Transparent Telecommunication System



**Figure 3.5:** *Our life-size telecommunication system. Two projectors with crossed linear polarizers in front project a stereo image pair on to an anisotropic transparent backprojection screen. As the screen preserves polarization, stereoscopic 3D can be seen when viewing the display with proper polarized glasses. A mirror is used for optical path folding to reduce the projection distance and the overall size of the setup.*

Figure 3.5. The principle is the same as with the second prototype: We use two BenQ SH910 projectors with linear polarizers and matching passive glasses. However, due to its size, this prototype imposed some challenges. First of all the holographic foil cannot be manufactured in one piece at this size and hence there is a visible seam between the two foils embedded in glass. Furthermore, the redirection angle of holographic elements is restricted. But the angle between the eyes of a viewer standing close to the screen and the lower border of the display is very steep. Thus, a small fade out in the content can be observed in the lower part if the viewer comes too close to the screen. Finally, the light efficiency of the screen is not very good and since the surface is very large, the very bright projectors we deploy are required to create a proper image. However, in the presented configuration, a clear stereoscopic image with low crosstalk can be observed. The complete system has been used in [Kuster *et al.*, 2015].

Images of a stereoscopic display must perfectly coincide to avoid eye strain and fatigue of the viewer. Thus, latter two prototypes require calibration, as the two images from different projectors are not aligned. We correct the keystone and achieve rectification of the two images by calibrating the two projector coordinate spaces using homographies. A checkerboard pattern is displayed on each projector and captured from an external camera without changing camera position. Using four point correspondences between projector and camera coordinates, homographies can be computed which are then used to relate one projector image to the other.

### 3.3.2 Communication

Our display system accepts a stereo image pair to be displayed. They are rectified and aligned by our setup and thus the device generating the images must not take care of it.

To keep the system generic and maintain flexibility, we have created a communication framework using WebRTC, WebGL and HTML5. Both image source, e.g. a stereo image camera, and image sink, e.g. our display prototypes, access a certain web page. Image source can select whether a top-down or a side-by-side image stream is created while the image sink has to make the same choice matching its capability. While the image source can further select the camera to stream (which can be any camera accessible over a DirectShow driver, even a virtual one), the image sink starts displaying in fullscreen mode.

Both clients connect to a server for communication hand shake. After a source and a sink are connected, the stream sends data in peer to peer manner. Image

### 3.4 Large Autostereoscopic Retroreflection Screens

sources can stream data to multiple sinks while a sink can receive multiple streams to be able to switch between multiple senders very quickly. More details are provided in [Kuster *et al.*, 2015].

#### 3.3.3 Discussion

Combining virtual content with the real world is important to many upcoming applications. As presented in the previous chapters, it enables new ways for telecommunication, where a conversational partner is not seen through an opaque display acting as a window into his room, but where he is projected into our room as being present. This is not only of interest for private communication but for teaching, talking to remote experts or even in business meetings.

Also other virtual reality applications would benefit greatly from transparent display technologies. Wherever the real world needs to be augmented by information, transparent displays will find use. Furthermore, as they remain transparent in their off state, they are also more appealing for putting into a room as they do not absorb much light.

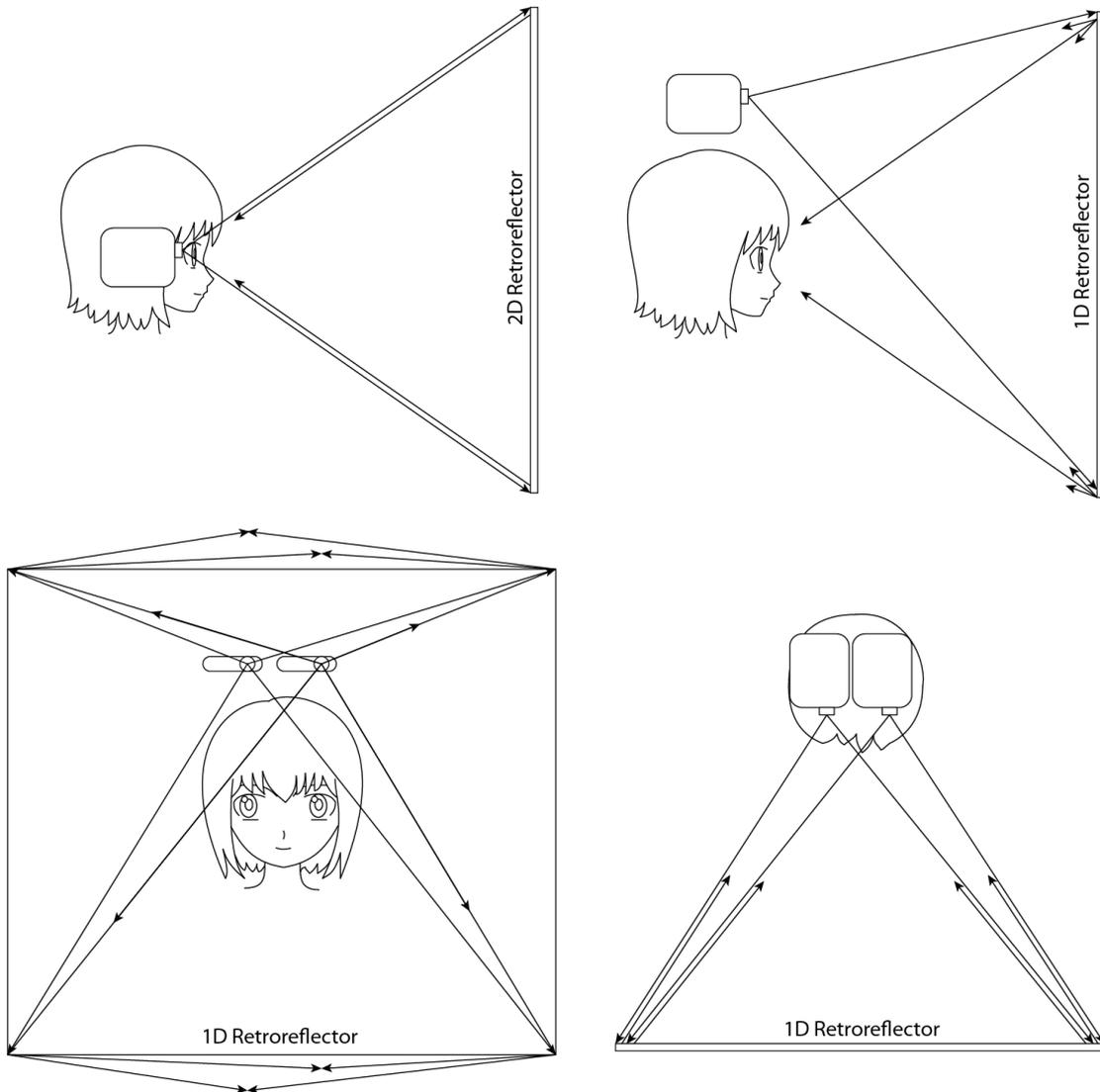
Another feature lacking in our system which would be important for natural viewing are background occlusions. All content presented will blend with the background and thus the virtual content looks ghost-like.

Introducing viewer tracking to transparent stereoscopic displays provides two important features. A viewer is not only fixed to a sweet spot where he is able to see 3D but also always receives the correct perspective which provides motion parallax cues.

Disadvantage is, that the system can only provide a single viewer with proper perspectives. There are systems which overcome this disadvantage as the one presented by [Nashel and Fuchs, 2009]. However, these systems have to be redesigned to also provide transparency, which is a difficult task due to the underlying display technique.

### 3.4 Large Autostereoscopic Retroreflection Screens

Besides transparency, a glasses-free viewing experience would be another desirable feature. Furthermore, the capability to show individual content to each viewer would add great value to a stereoscopic display. A screen where each viewer not only can decide whether the movie is displayed in 2D or 3D but even what content is shown independent of the others greatly enhances



**Figure 3.6:** Illustration of the working principle of a retroreflective screen. Top left shows a 2D retroreflector, which reflects light back into the same direction as the incident ray. Due to inaccuracies, light spreads out over a small angle and thus, images of a projector situated very close to a viewers eye can be seen in high brightness. The remaining three images show a 1D retroreflector from side (top right), front (bottom left) and top (bottom right) view. While light is reflected in the horizontal direction, it is diffused in the vertical direction. Thus, the projectors do not have to be close to the eyes but close to the vertical axis through the eyes and therefore can be located above a viewers head.

### 3.4 Large Autostereoscopic Retroreflection Screens

attractiveness of such systems. We therefore present a system which is able to present individual images for each eye of each viewer without requiring them to wear any head mounted gear [Smithwick and Ranieri, 2015].

An effective way to produce glasses-free stereoscopic imagery for each viewer is the use of two projectors in close proximity to the viewer's eyes and a collaborating retroreflective screen [Fergason, 1997; Krum *et al.*, 2012] as shown in Figure 3.6 in the top left image. Left/right eye images projected towards a retroreflective screen are sent back towards the projector (rather than scattering the light like a diffuser or reflecting the light like a mirror). Slight inaccuracies in the retroreflected direction allows each eye to view the bright retroreflected image of its corresponding projector. The retroreflected direction is often accurate enough that the brightness drops significantly when the eye is not nearly collocated with the projector. Also, in contrast to back-projected approaches as the one presented in [Balogh *et al.*, 2008], our retroreflective version is space saving, since the room is used for unfolding the projected images rather than some space behind the screen.

The placement of projectors next to the viewer's head is cumbersome and would need adjustment for each viewer depending on height. Fortunately, only horizontal differences in left/right eye images contribute to the perception stereoscopic 3D vision. Only horizontal retroreflection is needed to redirect the images back to the eye. The screen can be vertically scattering, so 3D images are viewable below the projectors as depicted in Figure 3.6. This allows the viewers to perceive 3D imagery as long as they are in the same vertical plane as the projectors.

Nguyen *et al.* created a retroreflecting teleconferencing system in which multiple users would have their own correct perspective view of the teleconferenced participants [Nguyen and Canny, 2005]. Each participant had their own projector connected to a paired remote camera. Each projector was placed below the viewer on the conference table. For each teleconference site, each remote view was projected onto a common 1D retroscreen consisting of a retroreflective screen with a lenticular vertical diffuser. Each viewed image was a different perspective but not a stereo image.

Gao and Xiao patented Retro-Reflective Light Diffusing Display Systems [Gao and Xiao, 2010] in which they combine the retroreflected head-mounted stereo display (as in [Fergason, 1997]) with a 1D retroreflective screen and vertically offset projectors (as in [Nguyen and Canny, 2005]). The stereo images are retroreflected horizontally to the viewer, but diffused vertically allowing stereoscopic viewing with projectors placed below the viewer.

Surman, *et al.* recently implemented such as system with motorized pro-

jectors that move with the tracked viewer's head [Surman *et al.*, 2013]. In their implementation, the projectors' optical axes are always aligned with the screen normal, so homography calculations were not used and the images were translated in software to stay aligned with the screen at the expense of field of view and resolution.

In the same time frame as Gao and Xiao, we independently created and investigated automultiscopic projection using a 1D retroreflective screen and a horizontal array of five vertically offset projectors. We investigated various compositions and configurations of the 1D retroreflective screen, including the combination of 1D horizontal only retroreflective material and 1D vertically diffusing material [Smithwick *et al.*, 2011]. From our preliminary research with automultiscopic versions of the system, we found the field of view of personal projectors and hence the field of view of the retroreflected image too small ( $15^\circ$  full angle) to provide an immersive experience. Furthermore, the projectors need to be calibrated, and a manual corner-push pin homography is too slow and imprecise to be done for multiple projectors.

For our preliminary and current research, personal microprojectors were chosen because they can be easily mounted overhead without interfering with each other, can be readily modified to provide wide-angle projection, and can provide bright images with retroreflection. Optoma desktop projectors (e.g. Optoma GT750E), staggered vertically, were considered, but the retroreflected imagery still did not have the desired field of view for an immersive experience ( $> 36^\circ$ ), and their large complicated optics made modifying these projectors for wide field of view projection a daunting task. Extremely short throw projectors (e.g. Optoma W307USTi) were also considered, however, their large bounce mirrors interfered with neighboring projection (even when mounted staggered), and their small throw ratios led to extreme keystoneing which wasted a large number of pixels off-screen. Thus, microprojectors in combination with wide-angle lenses seemed to be the best choice to achieve our goal.

### 3.4.1 Prototype

We aim to produce a multi-viewer wide field-of-view large stereoscopic display based upon overhead auto-calibrated projectors and a collaborative 1D retroreflective screen with an embedded fiber optic array.

To achieve a wide field of view, we modify sets of personal microprojectors with additional wide angle converter lenses. A set of projectors are mounted over each viewer's seating location. The use of wide angle converter lenses

### 3.4 Large Autostereoscopic Retroreflection Screens



**Figure 3.7:** *Illustration of the modified Optoma microprojectors. Two concatenated wide angle lenses greatly enhance the field of view. The top image shows a single projector while the bottom image shows a stereo projector configuration.*

causes pincushion distortion in the images that must be corrected. The projected images also need to be rectified and aligned to each other to allow proper and comfortable stereo fusion. Once undistorted, rectified, and aligned, crosstalk between viewed images can then be compensated for in software, if needed.

The 1D retroreflective projection screen is created using micro-cornercube retroreflective sheets layered with an anisotropic diffuser. The projection screen also has an embedded fiber-optic array feeding optical sensors. Projected structured lighting from each projector is detected by the fiber optic and sensor array to facilitate projector calibration and image rectification and alignment. Each projector projects gray code structured light detected by the screen's photodetector array. From this, each projector's distortion coefficients and homography matrix elements are estimated. Crosstalk cancellation is also implemented by subtracting weighted neighboring views [Doutre and Nasiopoulos, 2011].

We investigate various arrangements of multiple projectors for single person automultiscopic projection and projector sets for multi-person stereoscopic projection of common and independent stereoscopic content.

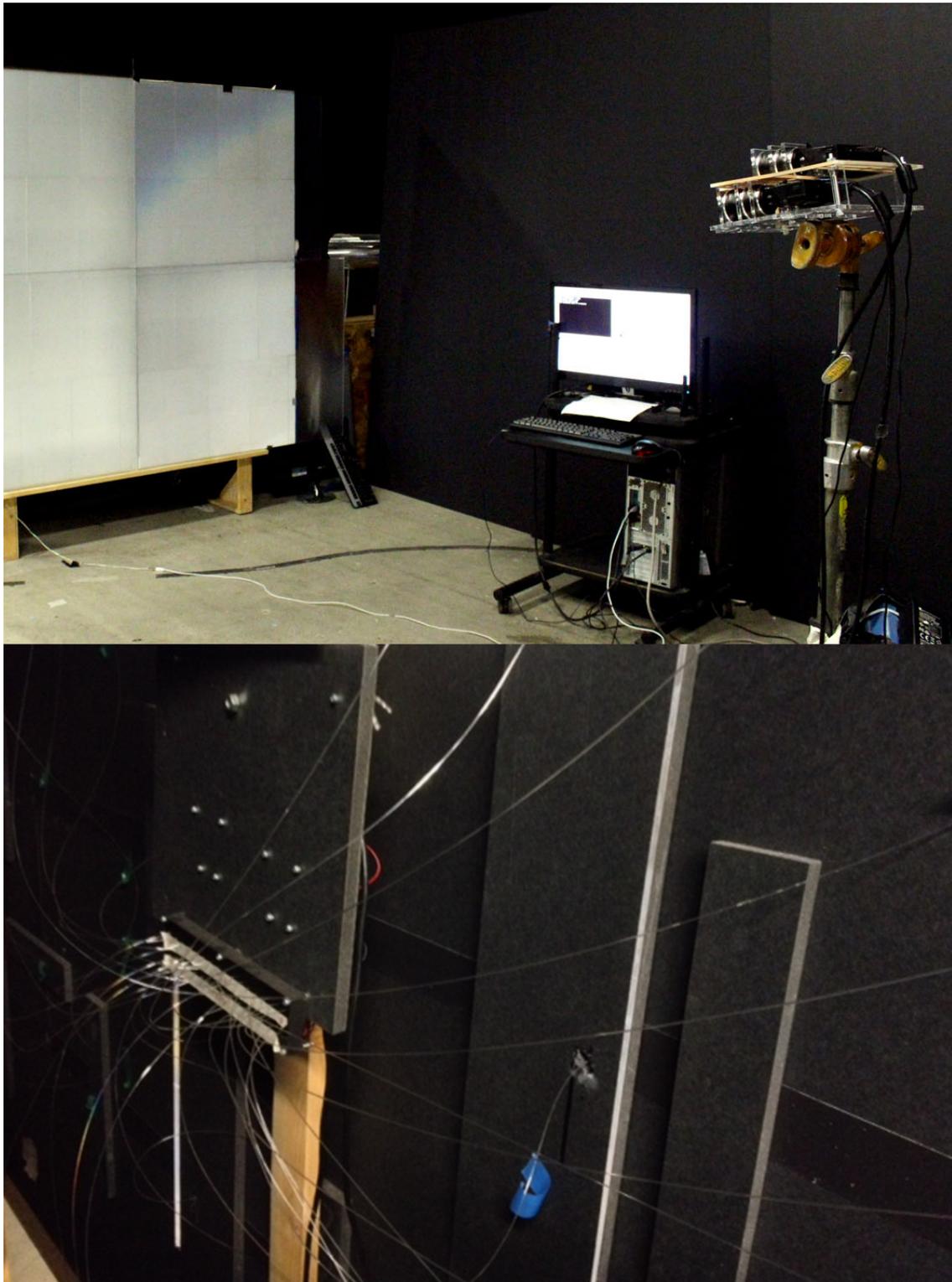
### **Modified Microprojectors**

Microprojectors ( Optoma PK320 ) are modified by re-centering the projection lenses and adding two 0.5x Super wide angle lenses (Crystal Optics), as shown in Figure 3.7. The projectors must be modified because the optical axes of the unmodified microprojectors' lenses are offset from the spatial light modulator's (SLM) normal so the projector may be laid on a table and project upwards onto a screen. Using an unmodified projector with the two wide angle converter lenses causes vignetting and additional aberrations in the projection. The microprojector's case and supporting structure must be cut and removed to make room for the recentered projection lens pack. Once modified, the focus is adjusted by controlling the spacing between the modified projector and the two 0.5x converter lenses. A simple structure made from laser-cut acrylic supports the modified projector and the two (abutting) wide angle converter lenses. Two such structures with modified projectors and converter lenses are mounted over-under and offset horizontally slightly more than an eye-width apart. A nominal viewing distance of 15 feet from the 1D retroreflective was chosen to provide more than a 36° horizontal field of view as recommended by THX specifications for an immersive experience.

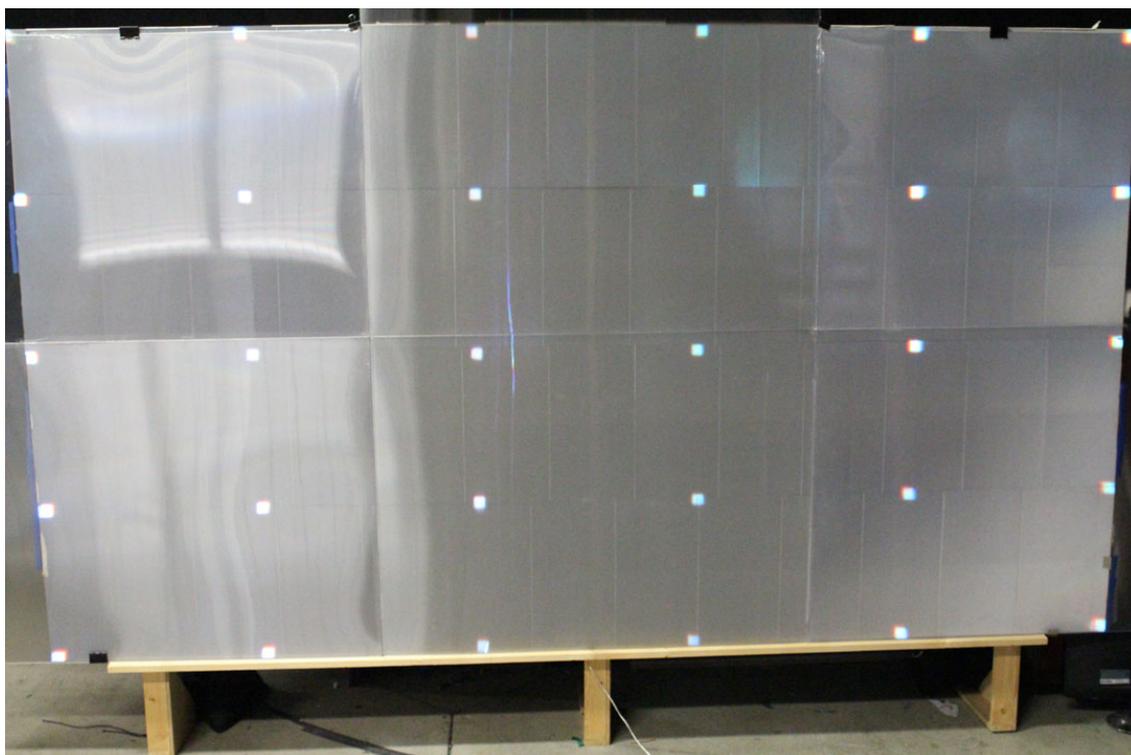
### **1D Retroreflective Screen and Fiber-Optic Array**

A 10' tall by 6' wide projection screen is constructed by tiling strips of cornercube retroreflective sheets (Reflexite) with an overlay of six 4' x 6' lenticular sheets (Microlens Technology) acting as an anisotropic vertical diffuser. The

### 3.4 Large Autostereoscopic Retroreflection Screens



**Figure 3.8:** *The retroreflective screen used in our prototype system. The top image shows the screen from the front with a single stereo projector. The bottom image shows the back of the screen with embedded optical fibers and sensors to read out intensity values.*



**Figure 3.9:** *Reprojected/Illuminated fiber positions as found by our calibration method.*

screen is shown in Figure 3.8 in the top image. An array of thirty optical fibers are arranged in an evenly spaced rectangular grid piercing the backside of the projection surface emerging behind the vertical diffuser as shown in Figure 3.8 in the bottom image. Fibers are coupled to photodetectors (TSL13S, Taos) whose voltage outputs are sampled by a pair of microcontrollers (Arduino) communicating to a computer (Dell) via USB serial communication. As part of the calibration procedure, each projector projects gray code structured light patterns detected by the screen's photodetector array. From the collection of detected codes at each fiber's screen location, a Matlab (Mathworks) program estimates each projector's homography matrix elements via SVD least squares, and determines the nonlinear distortion coefficients using a gradient descent method.

### 3.4.2 Software

#### Light Sensor Server

The optical fibers embedded in the screen are connected to light sensors, which are read out by an Arduino MEGA 2560. The Arduino is connected to a computer by USB which reads out the sensor data. To keep our implemen-

### 3.4 Large Autostereoscopic Retroreflection Screens

tation flexible and modular, we implemented a server application to which an arbitrary number of clients can connect. Each registered client will receive UDP packages containing the light sensor data.

#### **Light Sensor Visualizer**

For debug and maintenance purposes, we implemented a client application that registers with the Light Sensor Server, reads out sensor data on the fly and displays the intensity of each sensor. This allows a convenient way to check if each sensor is functional and also to validate the mapping from screen coordinates to the corresponding sensor by covering one sensor after the other and observe whose intensity drops.

#### **Light Sensor Mapping**

As we use modified projectors with wide angle converter lenses, the projected images are heavily distorted as visible in Figure 3.10 on the top. Also, images of different projectors are neither rectified nor aligned. We implemented another client application which performs a sophisticated calibration without requiring any manual interaction. Binary stripe patterns are projected with increasing frequency. While the first pattern covers half of the picture with white and the other half with black, the second already shows interleaved two black and two white stripes, while the third pattern consists of eight black and eight white stripes. Thus  $\log_2(\text{resolution}X) + \log_2(\text{resolution}Y)$  patterns are required to determine the projector pixel covering a specific sensor. The entire process works as follows: For each pattern the brightness of each sensor is read out to determine whether the sensor is currently lit or not. Thus, with the first pattern it is determined whether a sensor is in the left or right half of the projected image. The next pattern is then used to determine if the sensor is in the right or left half of the previously determined half. Or in other words, each pattern can be used to determine one bit of the image coordinate of the sensor's location. The procedure is executed once for the horizontal coordinate using vertical patterns and once for the vertical coordinate using horizontal patterns. After Light Sensor Mapping, for each sensor, the coordinates of the pixel covering that sensor are known as visible in Figure 3.9.

#### **Projector Calibration**

The previously discovered pixels' coordinates and the absolute positions of the sensors on the screen can be used to both compute homographies to rectify and overlay the images [Lee *et al.*, 2004], as well as to correct for distortions. To do so, we implement a Matlab program performing the calibration. To correct for the different kinds of distortions, we use the following Brown-Conrady model

$$x_d = x_u(1 + K_1r^2 + K_2r^4) + (P_2(r^2 + 2x_u^2) + 2P_1x_uy_u)(1 + P_3r^2) \quad (3.10)$$

$$y_d = y_u(1 + K_1r^2 + K_2r^4) + (P_1(r^2 + 2y_u^2) + 2P_2x_uy_u)(1 + P_3r^2) \quad (3.11)$$

with  $(x_u, y_u)$  as the undistorted screen coordinates,  $(x_d, y_d)$  as the distorted image coordinates and

$$r = \text{sqrt}((x_u - x_c)^2 + (y_u - y_c)^2) \quad (3.12)$$

as the distortion radius with distortion center  $(x_c, y_c)$ . Regular homographies are used to account for rectification and alignment of the images. An initial estimation of the homography mapping pixel coordinates to corresponding sensor coordinates on the screen is made in a least squares manner using SVD. A gradient descent method is then used to refine the homography and to find the best distortion parameters  $K_i$  and  $P_i$  as well as the distortion center  $(x_c, y_c)$ . The results of our calibration procedure are shown in Figure 3.10.

### Video Player

OpenGL/GLSL vertex and fragment shaders implemented on the computer graphics cards' GPU (NVIDIA) pre-distort and correct frames decoded from standard side-by-side SBS stereo movies (C++, VLC). The fragment shader also implements a simple manually adjustable crosstalk cancellation algorithm by subtracting weighted neighboring views according to [Doutre and Nasiopoulos, 2011] using the following formulas:

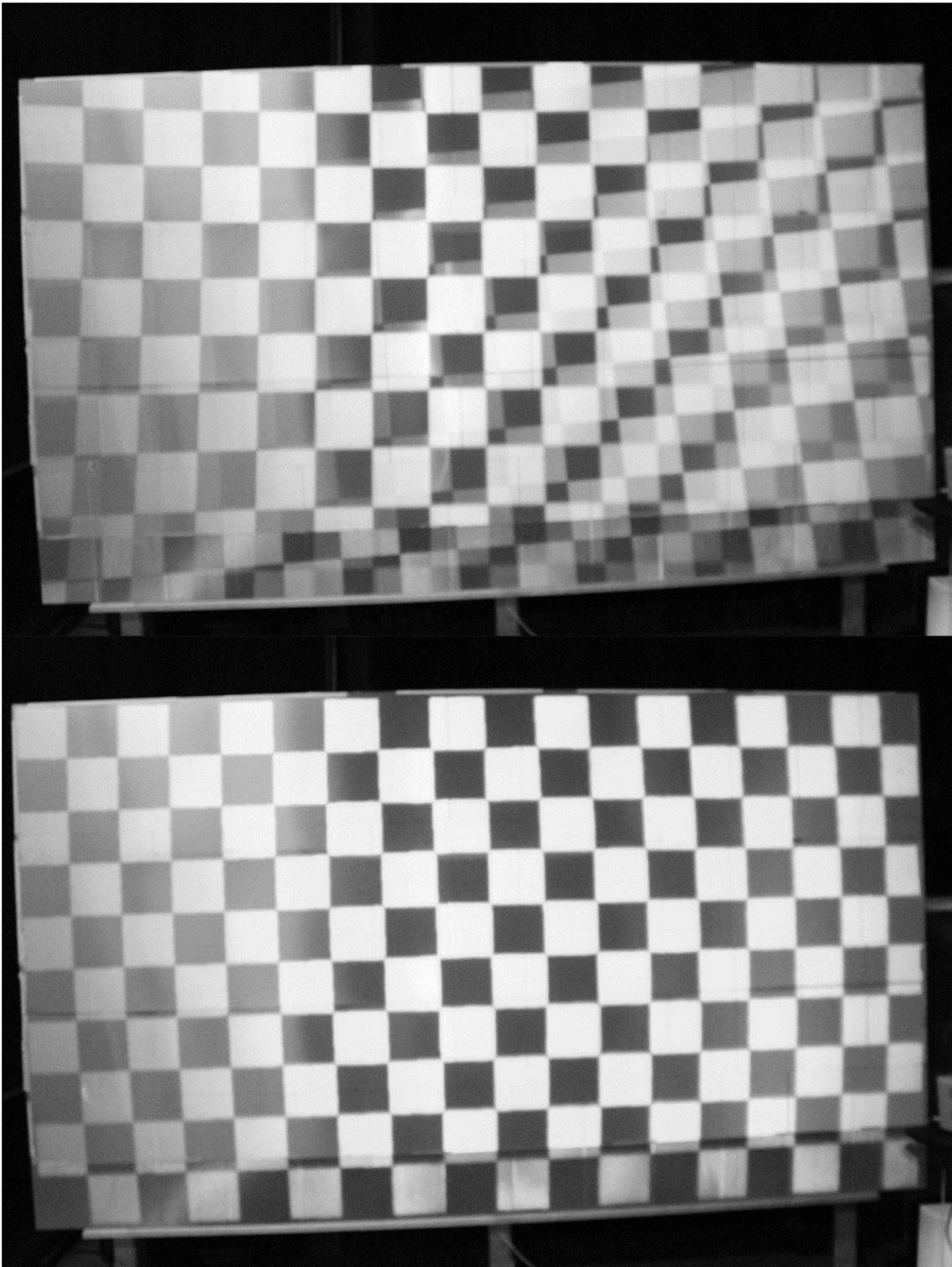
$$i_{L,disp}(x,y) = \frac{i_L(x,y) - c \cdot i_R(x,y)}{1 - c^2} \quad (3.13)$$

$$i_{R,disp}(x,y) = \frac{i_R(x,y) - c \cdot i_L(x,y)}{1 - c^2} \quad (3.14)$$

where  $i_L$  and  $i_R$  are the original left and right image brightnesses,  $c$  is the crosstalk fraction, and  $i_{L,disp}$  and  $i_{R,disp}$  are the displayed left and right image brightnesses. The two projectors in a stereo projector set are controlled from a single video output via a Triple-Head-2-Go desktop extender (Matrox).

### 3.4.3 Configurations

Several projector configurations are tested representing typical arrangements that may be expected in practice.



**Figure 3.10:** *Illustration of the impact of our calibration. Top image shows projected checkerboard patterns of two projectors. Due to the wide angle lens, the images are heavily distorted and do not align at all. Bottom image shows the two pattern after our calibration and with applied undistortion.*



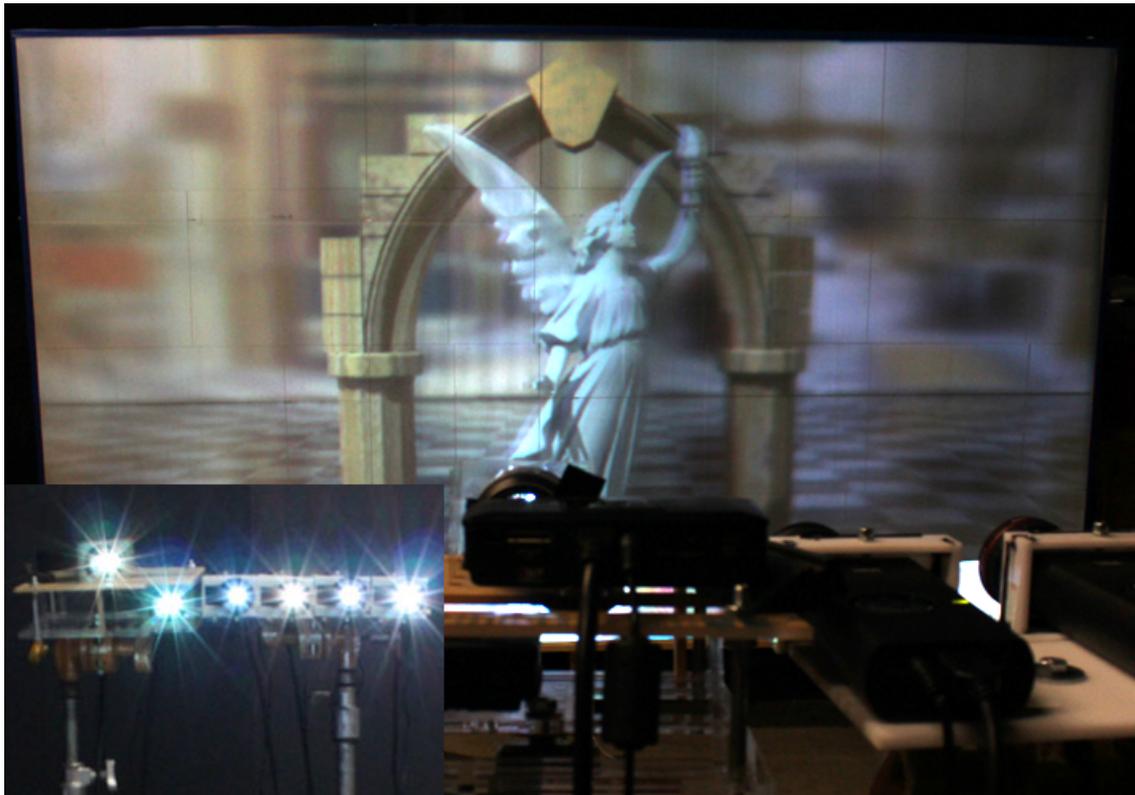
**Figure 3.11:** *The left eye image of a single stereo projector pair. Two projectors are arranged side by side with a spacing matching the average human eye distance. The image is bright due to the retroreflection and shows only very low crosstalk from the right eye image.*

### **Single viewer stereo images**

This configuration represents a single person viewing an stereoscopic movie. A projector set consisting of two Optoma projectors and matching lens holders are placed on a stand 54" tall and centrally placed 15' from the 1D retro-reflective screen. The projectors are placed 2.5" apart, equal to the average eye spacing, visible in Figure 3.11 in the inset. The projectors are automatically calibrated using the aforementioned projected structured lighting calibration routine. To measure crosstalk, a test pattern consisting of black and white squares against a black background are projected from the right projector, and the inverse pattern with a black background projected from the left projector. This enables measurement of direct white, cross-over white, and ambient light levels from measurements taken with a luminance meter (Konica Minolta, LS-110). For subjective judgment, the calibrated projectors display left/right views from a standard stereoscopic movie called Big Buck Bunny.

### **Single viewer multi-view images**

This configuration represents a single person viewing an automultiscopic movie. Three projector sets consisting of a total of six projectors and matching lens holders are placed on stands as illustrated in Figure 3.12 in the inset. The rightmost projector is centrally placed 12' away from the 1D retro-



**Figure 3.12:** *One image of a multi-view setup. Six projectors are arranged side by side to cover a viewers head and some space to the left and right for head motion.*

flective screen and uses the modified Optoma microprojectors (as described in Section 3.4.1). The other two projector sets use 3M microprojectors each with an additional single 0.5x converter lens (with similar performance to the modified Optoma projectors). The projectors are placed 2.5" apart from each other, equal to the average adult's eye spacing. They are individually calibrated, with the Optoma projector set using automated projected light calibration, and the 3M projector set calibrated manually using a standard homography for keystone correction (i.e. four corner push-pin correction) and Brown-Conrady model for radial distortion. The calibrated projectors each display one of six different corresponding neighboring views of a rendered 3D scene [Weissman and Woods, 2011]. Resulting images captured from positions just below each projector are shown in Figure 3.13.

#### **Multiple viewer side-by-side stereo pair imagery**

This configuration represents multiple people in the same row viewing the same common or different independent stereoscopic movies. Two projector sets each consisting of two projectors and matching lens holders are placed on a stand 54" tall. One projector set (using modified Optoma microprojectors)

*Stereoscopic Displays*



**Figure 3.13:** *The six images from our multi-view setup as seen just below each projector.*

### 3.4 Large Autostereoscopic Retroreflection Screens



**Figure 3.14:** A stereo setup for two viewers. Two projector pairs are placed side by side. Two different stereo image pairs are exposed to each viewer. The left images of both viewers are shown in the top and in the bottom image.

is centrally placed 12' away from the 1D retroreflective screen. The other projector set (3M microprojectors each with an additional single 0.5x converter lens ) is placed 22" apart on the side of the first centrally placed set, equal to a shoulder's width apart, representing two viewers sitting side-by-side. The setup is shown in the inset of Figure 3.14. The calibrated projectors display the same common left/right stereo frame from a standard stereoscopic movie, *Big Buck Bunny*, or two different independent stereo views with one projector set projecting *Big Buck Bunny* and the other a rendered 3D scene [Weissman and Woods, 2011]. The left images of both viewers when watching independent content are shown in Figure 3.14.

### **Multiple viewer forward-behind stereo pairs**

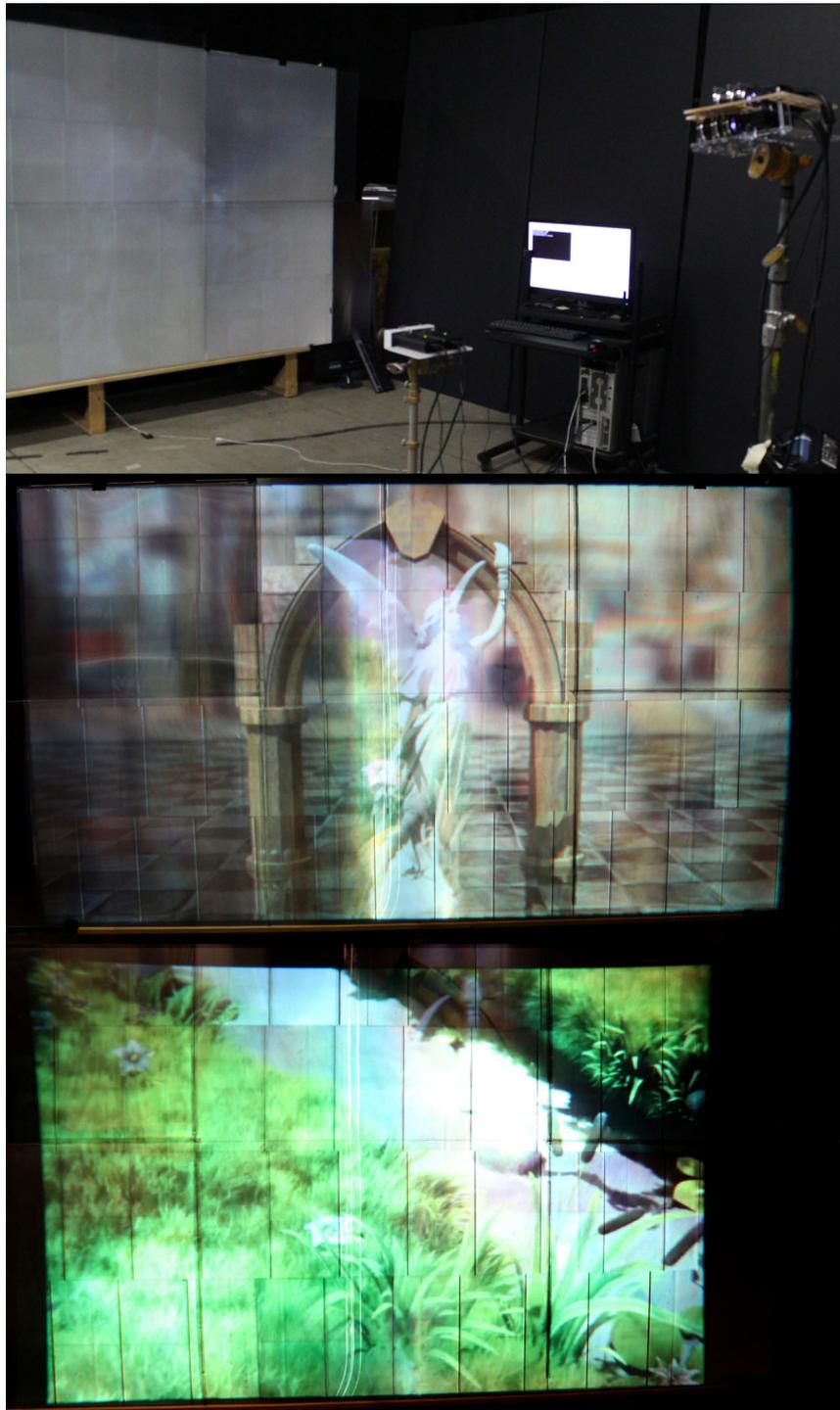
This configuration represents multiple people in different rows viewing the same common or different independent stereoscopic movies. Two projector sets each consisting of two projectors and matching lens holders are positioned one in front of the other. One projector set (using modified Optoma microprojectors) on a stand 54" tall is centrally placed 12' away from the 1D retroreflective screen. The other projector set (3M microprojectors each with an additional single 0.5x converter lens ) is placed on a 3.5" tall stand, 36" directly in front of the first centrally placed set, representing two viewers sitting in the center seats of two different rows. The setup is shown in the inset of Figure 3.15. The calibrated projectors display the same common left/right stereo frame from a standard stereoscopic movie, *Big Buck Bunny*, or two different independent stereo views with one projector set projecting *Big Buck Bunny* and the other a rendered 3D scene [Weissman and Woods, 2011]. The left images of both viewers when watching independent content are shown in Figure 3.15.

## **3.4.4 Discussion**

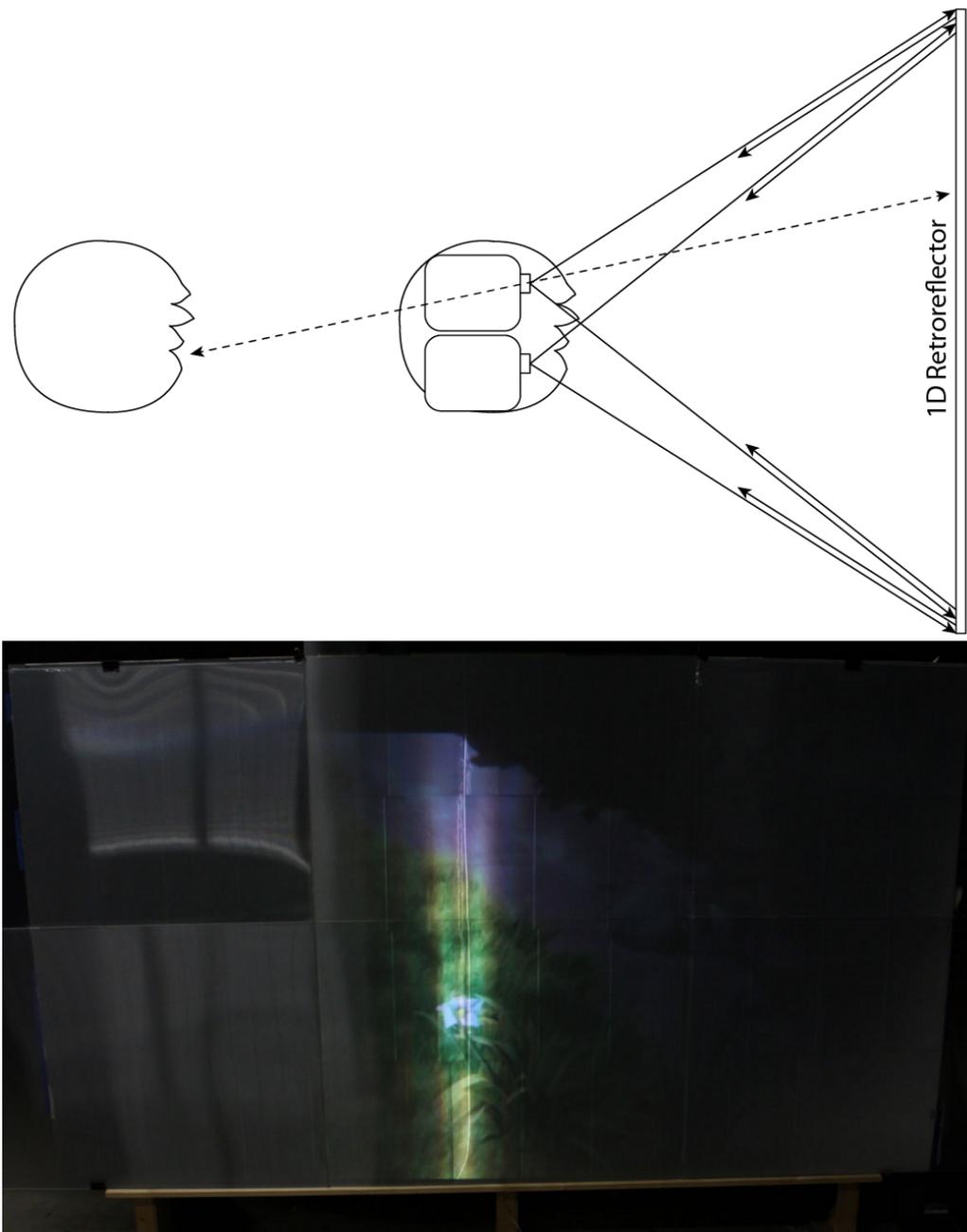
### **Single viewer stereo images**

Modifying the Optoma microprojectors increased the projectors' horizontal fields-of-view (fov) from  $26^\circ$  to  $46^\circ$ , providing full screen coverage even with keystoneing. With the 10' wide screen 12' away, the viewer has a  $45^\circ$  field of view of the stereo image. Projector pair calibration and image rectification/alignment was successful. Each projector reprojects light to identified fiber positions (Figure 3.9) and produces undistorted image pairs (Figure 3.10). Using the test pattern, the crosstalk between paired projectors was measured to be 3% – 5%. Crosstalk reduction was not necessary. Wide fov immersive movies were viewable with easily fuseable bright stereo imagery. Screen seams and parquet lines in the retroreflective and anisotropic diffuser ma-

### 3.4 Large Autostereoscopic Retroreflection Screens



**Figure 3.15:** *A stereo setup for two viewers. Two projector pairs are placed one behind each other to simulate the rows of e.g. a cinema. The configuration is shown in the top image. Two different stereo image pairs are exposed to each viewer. Left image of both viewers are shown in the middle and in the bottom image. Stripes of strong crosstalk are visible in both images, an effect discussed in Section 3.4.4*



**Figure 3.16:** Results showing crosstalk visible in the forward-behind stereo configuration. Problematic are those regions where the left eye of one viewer is behind the right eye of the other viewer or vice versa as depicted in the illustration at the top. For better visibility, content for one viewer corresponding to the camera position is turned off in the bottom image, revealing crosstalk caused by content of the other viewer. Reason is the vertical diffusion of the 1D retroreflector, making both images visible in one vertical plane of view.

### 3.4 Large Autostereoscopic Retroreflection Screens

material are noticeable but do not affect stereo fusion. The viewing conditions are sensitive to head position both side-to-side and front-back from directly under the projectors. Away from the nominal viewing location, there is noticeable picket fencing (moving dark stripes) caused by luminance fall-off between projection frustums and rainbow artifacts from the retroreflector material.

#### **Single viewer multi-view images**

With multiple (calibrated) projector sets projecting different views of the same 3D scene (Figure 3.12), the viewer is able to move his/her head from side to side and perceive 3D images with motion parallax (Figure 3.13). In this configuration, head-tracking is not needed to see motion parallax. The retroreflection has a narrow Gaussian lobe that falls off quickly, so crosstalk between non-neighboring projectors is as small as  $< 2\%$ . For a coherent 3D scene, the neighboring views are similar, so crosstalk mainly limits the depth of the 3D scene. However, motion parallax occurs in discrete steps, there is noticeable picket fencing (moving dark stripes) caused by luminance fall-off between projector frustums and rainbow artifacts from the retroreflector material. Color calibration and luminance balance among the projector sets is necessary but was not implemented in this work. By using multiple projectors overhead, a retroreflective light field display is created. The viewing location is not restricted to directly below the projectors, but 3D views may be seen in a limited region in front of and behind the projector plane.

#### **Multiple viewer side-by-side stereo pair imagery**

At a shoulder's width away (7") from the neighboring projector pair, the crosstalk between neighboring projector pairs has dropped to less than 1%. Two separate independent movies could be played on side-by-side separated projectors with little interference (Figure 3.14), except in dark scenes or scene transitions where the screen appears black. However, in those dark areas even the dim crosstalk between projector sets can become apparent and distracting, especially with dynamic scenes.

With multiple (calibrated) projectors sets projecting the same stereo pairs, the corresponding right and left eye images are aligned and overlap. There is not an increase in crosstalk from additional projector sets. Images from one projector set that are visible from another projector set's viewpoint are mainly the result of low-level diffuse scattering from the screen. There is a very narrow specular reflection that may be visible, but this is ignored for now, as they can be avoided by tilting the screen downwards [Gao and Xiao, 2010; Surman *et al.*, 2013]. These diffuse images equally contribute to the corresponding view and the crosstalk view (both signals are raised by approximately the same amount).

Following an analysis similar to Weissman and Woods [Weissman and Woods, 2011] for crosstalk measurement using observed white, cross-over, and ambient levels, we can examine the effect of scattered aligned stereo views. We can define observed white  $O_{WL}$ , ghost  $O_{GL}$  and black  $O_{BL}$  levels as follows:

$$O_{WL} = T_L(W) + X_{RL}(B) + A + c \cdot T_L(W) + c \cdot X_{RL}(B) \quad (3.15)$$

$$O_{GL} = T_L(B) + X_{RL}(W) + A + c \cdot T_L(B) + c \cdot X_{RL}(W) \quad (3.16)$$

$$O_{BL} = T_L(B) + X_{RL}(B) + A + c \cdot T_L(B) + c \cdot X_{RL}(B) \quad (3.17)$$

where  $W$  is the light intensity from the white source image,  $B$  is the intensity of the black source image,  $c$  is the diffuse scattering fraction contributions from aligned images,  $A$  is the ambient light,  $X_{RL}$  is the cross-over light, and  $T_L$  is the direct retroreflected light.

Subtracting Equation 3.17 from Equation 3.15 and Equation 3.17 from Equation 3.16 gives

$$O_{WL} - O_{BL} = T_L(W - B) \cdot (1 + c) \quad (3.18)$$

$$O_{GL} - O_{BL} = X_{RL}(W - B) \cdot (1 + c) \quad (3.19)$$

taking the ratio of Equation 3.18 and Equation 3.19 and solving for the crosstalk  $X_{RL}$  reveals

$$X_{RL}(W - B) = T_L(W - B) \cdot \frac{O_{GL} - O_{BL}}{O_{WL} - O_{BL}} \quad (3.20)$$

Dividing by the retroreflected fraction  $T_L$  gives the observed crosstalk from the right to the left image

$$OCT_{RL} = \frac{X_{RL}(W - B)}{T_L(W - B)} = \frac{O_{GL} - O_{BL}}{O_{WL} - O_{BL}} \quad (3.21)$$

A similar derivation can be made for the observed crosstalk from the left to the right image. Notice the diffuse components do not change the ratio. With the addition of aligned stereo pairs from multiple projector sets, the observed crosstalk does not change. The black levels do increase slightly, but so do the white luminances. As an additional experiment, the homography and radial distortion correction was also performed successfully (both automated and manually) with a projector 5" off-center representing a viewer on an aisle seat in-line with the edge of the screen. The 3M projectors' 4:3 format

### 3.4 Large Autostereoscopic Retroreflection Screens

SLMs are useful as the extra vertical pixels can be used by the homography to compensate for the keystone of the 16:9 image caused by horizontal off-axis projection.

#### **Multiple viewer forward-behind stereo pairs**

For two projector sets positioned at different depths from the screen, one in front of the other (Figure 3.15), there was crosstalk between images for independent projected images, but also when projecting the same stereo pair. The front projectors' stereo pair images are retroreflected back to their projectors, but then continue to project behind the projectors' locations. Even with the front viewer's head blocking some of the retroreflected light at the projector location, the vertical diffusion is large enough that the rear viewer behind (and slightly above) the front viewer will receive some of the front viewer's retroreflected light. A schematic illustration of this effect is given in Figure 3.16 in the top image. The second viewer will see his/her direct retroreflected image, but also portions of both of the front viewer's stereo pair in each eye as visible in Figure 3.16 in the bottom image.

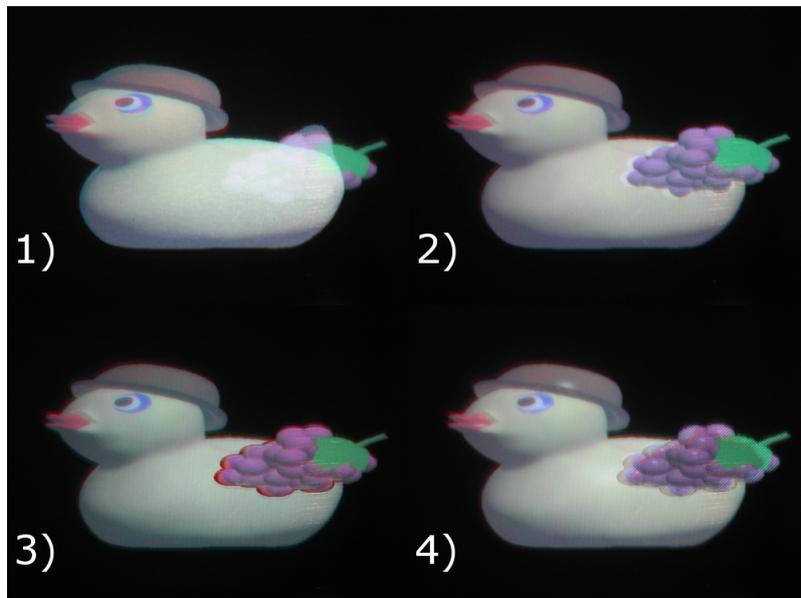
For two rows of viewers, even staggering the front and back row half-a-shoulder's width apart would not reduce this crosstalk because the front viewer's left/right eyes will be in either of the rear viewer's view frustums. Hence the front viewer will still see a portion of the rear viewer's. For a screen which uses a 2D retroreflector (e.g. beaded or cornercubes) and 1D diffuser (holographic or lenticular), the diffusion is around both row's projectors above the viewers' heads and thus a large vertical diffusion angle is required. Any vertical diffusion that is sufficient to allow the lower front row viewer to see will also allow the higher back row viewer to see as well. Although the crosstalk due to vertical diffusion and projection spill cannot be avoided in this configuration, the projected image artifacts (e.g. visible seams, brightness bands and rainbow artifacts, sensitivity to head position) can be reduced using the micro-corner-cube retroreflector with a holographic vertical diffuser (Luminit  $40^\circ \times 0.2^\circ$ ) and a slight horizontal diffusion instead of the lenticular vertical diffuser. However, even with a slight horizontal diffusion ( $0.2^\circ$ ), crosstalk increased to 12% – 15%. Separating the projectors by slightly more than an eye width (3" - 3.5") still produced bright 3D images but with reduced crosstalk, due to the neighboring projector being further down the retroreflected lobe's Gaussian falloff. Computational crosstalk cancellation was effective with slight contrast reduction, but the sense of stereoscopic 3D was not as strong as when using lenticular sheets for vertical diffusion (and with almost no horizontal diffusion).

### **3.4.5 Conclusion**

We produced wide field of view stereoscopic imagery on a large 1D retroreflective screen by modifying overhead microprojectors with wide angle converter lenses. The screen's embedded sensor array allowed automated system calibration needed for multiple projectors. Projected structured light calibration for homography estimation was extended to computed distortion correction introduced by the wide angle converter lenses. The ability of the system to provide each viewer with independent content and the choice whether to watch content in 2D or 3D makes this configuration interesting for many use cases. The use of multiple overhead projectors creates a retroreflective light field display, allowing viewers to experience motion parallax without the need for head-tracking or motorized projectors. In the current configuration, the vertical diffusion likely prevents the use of 1-D retroreflective screen and mounted microprojectors for multiple rows (front-behind) of viewers. While this represents the cases for a small multi-row theater, there are cases where a single row of side-by-side viewers is acceptable.

## Volumetric Displays

---



**Figure 4.1:** *Illustration of the steps of our light field decomposition presented in this chapter. We perform volumetric rendering (1), view-independent (2) and view-dependent (3) occlusion culling and finally view-dependent rendering on a parallax type layer (4). Photographs are taken from our multi-layered display prototype.*

---

Multi-planar volumetric displays have been suggested to provide natural ways to show 3D scenes at nearly correct accommodation cues with increased display bandwidth and hence higher angular and spatial resolution. There are many different designs as listed in Section 2, however a common framework to describe these systems is lacking. Thus, in this work we generalize these concepts to multi-planar plenoptic displays by defining basic display layer primitives, consisting of emissive and modulating layers. We provide a mathematical framework to describe light transport through any combination of such layers. Based on this framework, we then provide a method to distribute an input light field according to a given display configuration. Furthermore, a quantitative error analysis for different layer configurations is provided. Finally, we present two physical prototypes, capable of rendering volumetric content with view dependent effects, proper occlusions and better accommodation cues and we show results for both of them [Ranieri *et al.*, 2013a; 2013b]. The idea has been protected in [Ranieri *et al.*, 2013c].

## 4.1 Mathematical Framework

Our model assumes co-planar display layers which are aligned with the  $xy$ -plane. Each of the layers can either be emitting or modulating, performing a certain operation on the overall light transport which will be described in the following. Our model is using a similar notation and concepts as presented in [Durand *et al.*, 2005]. In general, the light field  $\ell$  describes the radiance of light rays passing through points  $(x, y)$  and  $(u, v)$  at distance  $z$  from the  $xy$  plane, and is denoted as  $\ell(x, y, z, u, v)$ . For simplicity, we will only consider light rays traveling along the positive  $z$  direction, as our displays will only be viewed from the front.

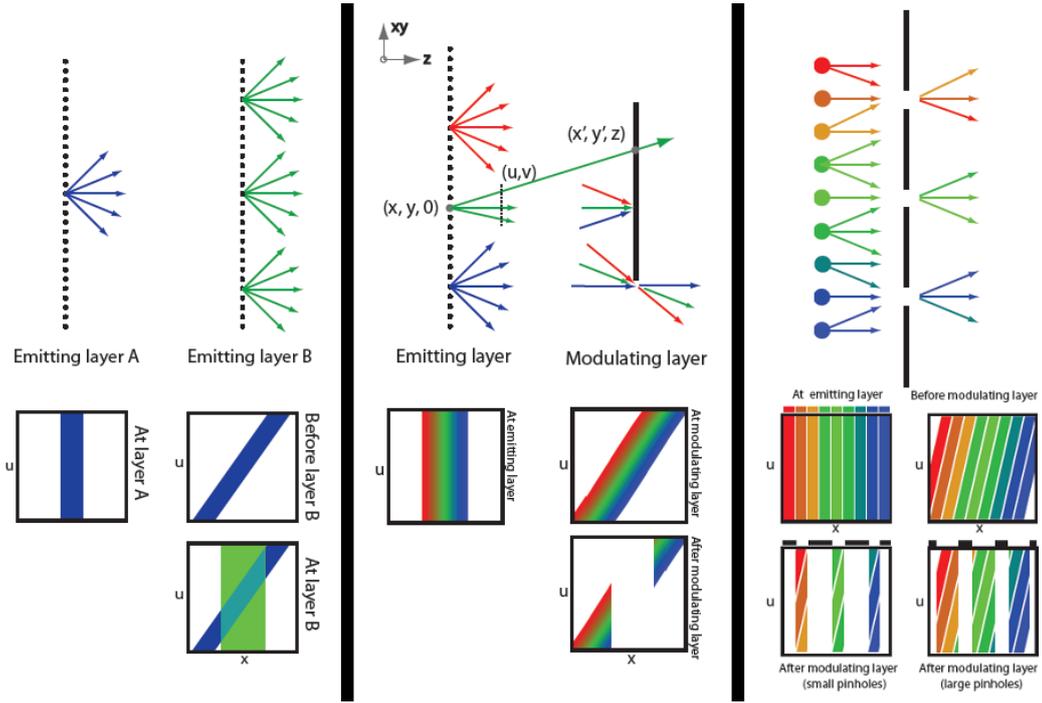
The basic light transport is illustrated in Figure 4.2 (middle). A ray starting at position  $(x, y, z)$  passing through  $(u, v)$  traverses in free space to

$$x' = x + \Delta z \cdot u \quad (4.1)$$

$$y' = y + \Delta z \cdot v \quad (4.2)$$

As the ray moves in depth, its position will change to  $(x', y', z + \Delta z)$  while keeping its original traveling direction  $(u, v)$ .

Many of the display layer types deployed in modern systems can be generalized to two categories. We describe both categories as basic display primitive and provide a mathematical notation for the light transport operator. Together



**Figure 4.2:** *The three basic display layer primitives. Overlay of two emissive layers is shown on the left, impact of a modulating layer in the middle and the principle of a parallax barrier layer on the right.*

with the notation of light transport in free space, more complex systems made of any combination of such display primitive layers can be described.

### 4.1.1 Emissive Layer

An emissive layer  $E$  acts like an array of point light sources, emitting constant spherical light. We use the notation  $E_z(x, y)$  as the light portion at  $x$  and  $y$  on the plane at depth  $z$ , radiating into all directions  $(u, v)$ . The emissive layer adds light to an input light field  $\ell_{IN}$ , and yields the output light field  $\ell_{OUT}$ , as illustrated in Figure 4.2 on the left:

$$\ell_{OUT}(x, y, z, u, v) = \ell_{IN}(x, y, z, u, v) + E_z(x, y) \quad (4.3)$$

Opaque emissive layers can be found in any 2D display consisting of e.g. a back light with a modulating color LCD. However, to optically overlay them with other display layer primitives, transparent emitters are preferred. This can be implemented using the upcoming transparent OLED technology, by transparent back-projection foils or, as in our prototype, polymer dispersed liquid crystal (PDLC) layers in combination with a projector.

### 4.1.2 Blocking Layer

A spatial modulating plane  $M$  will gradually attenuate all rays  $(u, v)$  passing through a certain pixel  $(x, y)$ . The modulating layer  $M_z(x, y)$  is therefore represented as scalar between zero and one, and the output light field can be described as:

$$\ell_{OUT}(x, y, z, u, v) = \ell_{IN}(x, y, z, u, v) \cdot M_z(x, y) \quad (4.4)$$

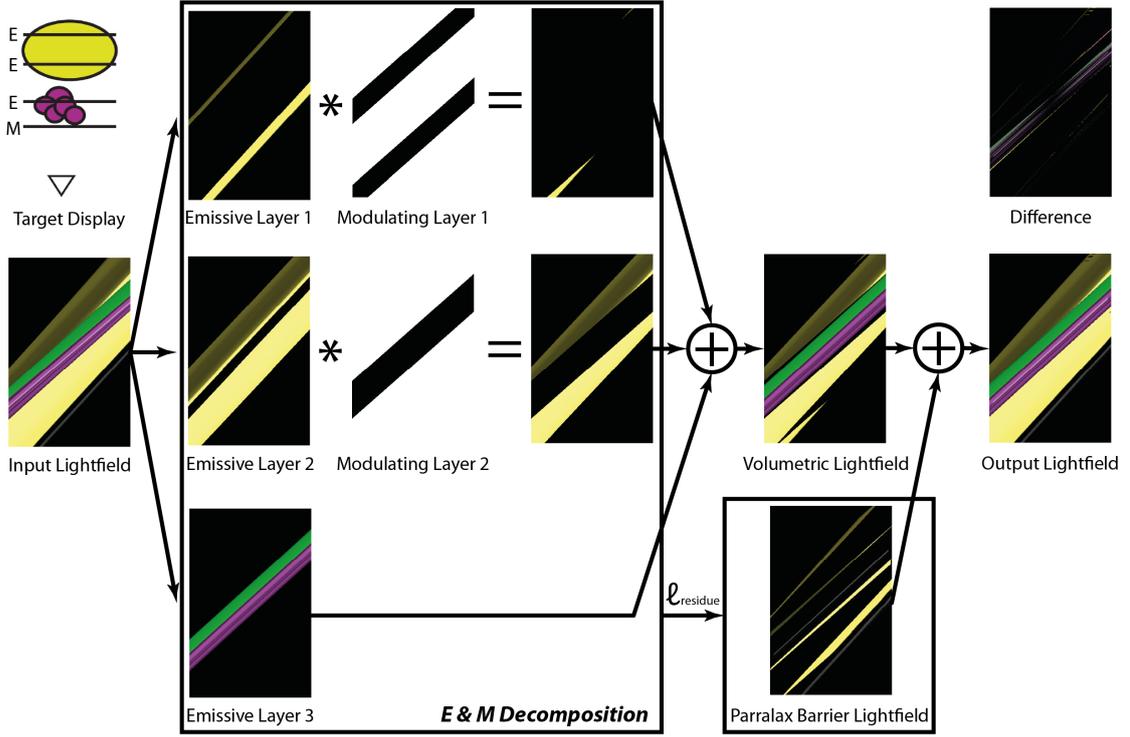
The operation is shown in Figure 4.2 in the middle. Modulating layers can be implemented by (grayscale) liquid crystal displays, as the polarization rotation capability of twisted nematic liquid crystals can be used to block light when combined with two polarizers.

### 4.1.3 Automultiscopic Layer

Automultiscopic layers are capable of emitting light depending on the angle. In practice, this can be achieved e.g. using parallax barriers. Parallax barrier displays are basically a combination of an emissive layer and a modulating layer with a special modulating pattern and small spacing  $\Delta z$  between the two layers. Therefore, the same light transport operators as for above layers are used. Since this pairing is fundamental, we define it as third basic primitive shown in Figure 4.2 on the right. The modulating layer is used to achieve ray separation by displaying a vertical slit, diagonal slit or pinhole pattern, while the emissive layer displays the different rays that pass through the slits/pinholes. As a consequence, an observer will see different rays from different directions.  $N$  pixels on the emitter plane can be partitioned into any number of spatial and angular samples  $(x, y, u, v)$ , such that  $N > x \cdot y \cdot u \cdot v$ . In practice, such displays trade a substantial reduction of spatial resolution for a relatively small amount of rays and apparent depth. The work of Levin et al. provides a good background on this subject [Levin and Durand, 2010].

## 4.2 Light field Decomposition

Based on the analysis in the previous section, we will describe an algorithm that approximates an input light field  $\ell_{IN}(x, y, z, u, v)$  as an output light field  $\ell_{OUT}(x, y, z, u, v)$  targeted for a given multi-layer plenoptic display  $D_{IN}$ . Our algorithm decomposes the input light field into a number of components. Each component is then displayed on one or multiple display



**Figure 4.3:** Illustration of the decomposition algorithm for a display layer configuration of one front-most modulating layer followed by three emissive layers. View independent emissive elements are assigned, occlusions computed and the residue added using parallax barrier rendering.

primitives. In order to aid the decomposition process, we assume that for all rays  $R(x, y, z, u, v)$  of the input light field we know the depth  $z$  of the closest object, the diffuse component  $R_{\text{diffuse}}$ , and the specular/glossy component  $R_{\text{specular}}$ .

- 1  $\ell_{VIV} \leftarrow \text{assignViewIndependentVolumetric}(\ell_{IN}, D_{IN})$
- 2  $\ell_{\text{residue}} \leftarrow \ell_{IN} - \ell_{VIV}$
- 3  $\ell_{VDV} \leftarrow \text{assignViewDependentVolumetric}(\ell_{IN}, \ell_{\text{residue}}, D_{IN})$
- 4  $\ell_{\text{residue}} \leftarrow \ell_{\text{residue}} - \ell_{VDV}$
- 5  $\ell_{VDL} \leftarrow \text{assignViewDependentLightField}(\ell_{\text{residue}}, D_{IN})$
- 6  $\ell_{\text{residue}} \leftarrow \ell_{\text{residue}} - \ell_{VDL}$

**Algorithm 4.1:** High level overview of the rendering algorithm

More specifically, the algorithm separates the light field data into planar components. First, the view-independent volumetric components  $\ell_{VIV}$  are extracted from the light field, i.e.  $\ell_{VIV}$  will contain the rays that are not occluded at any angle. Next, a view-dependent partially occluded volumetric part

## Volumetric Displays

$\ell_{VDV}$  is extracted. Finally, the remaining light field  $\ell_{VDL}$  is extracted for rendering with parallax barrier layers. Algorithm 4.1 gives a high level overview of how to generate  $\ell_{OUT}(x, y, z, u, v)$  from an input light field  $\ell_{IN}(x, y, z, u, v)$  for a given display setup  $D_{IN}$ . Example decomposition is given in Figure 4.3.

```

1 for emissiveLayer  $\in D_{IN}$  do
2   for  $x, y, z, R_{diffuse} \in \ell_{IN}$  do
3      $dz \leftarrow distance([x, y, z], emissiveLayer)$ 
4     if  $R_{diffuse} \neq 0$  and  $dz < z_{thresh}$  and  $notOccluded(x, y, z)$  then
5        $emissiveLayer[x, y] \leftarrow R_{diffuse}$ 
6        $\ell_{VIV}.add([x, y, z, R_{diffuse}])$ 
7     end
8   end
9 end

```

**Algorithm 4.2:**  $\ell_{VIV} \leftarrow assignViewIndependentVolumetric(\ell_{IN}, D_{IN})$

In a first step all diffuse components not occluded from any viewing angle are extracted from the light field. The extracted components are then distributed onto the available emitting layers. Only components that are spatially close enough to the emitting layers are considered for display, to minimize the re-projection error. More specifically, each part of  $R_{diffuse}$  that is within a distance  $z_{thresh}$  from any layer is assigned to the nearest layer in the display setup  $D_{IN}$ . Assignment is performed by parallel projection. The parts of  $R_{diffuse}$  that are further than  $z_{thresh}$  from all layers are not processed and left as residue for the automultiscopic display layers. Algorithm 4.2 summarizes this procedure.

```

1 for emissiveLayer  $\in D_{IN}$  do
2   for  $x, y, z, R_{diffuse} \in \ell_{residue}$  do
3      $dz \leftarrow distance([x, y, z], emissiveLayer)$ 
4     if  $R_{diffuse} \neq 0$  and  $dz < z_{thresh}$  then
5        $emissiveLayer[x, y] \leftarrow R_{diffuse}$ 
6        $\ell_{VDV}.add([x, y, z, R_{diffuse}])$ 
7       for  $x', y', z' \in \ell_{IN}$  do
8         if  $occludes([x, y, z], [x', y', z'])$  then
9            $modulator \leftarrow getClosestModulator([x', y', z'])$ 
10           $modulator.occlude([x, y, z], [x', y', z'])$ 
11        end
12      end
13    end
14  end
15 end
16 for modulatingLayer  $\in D_{IN}$  do
17   for  $u, v \in modulatingLayer.OccludingPixels()$  do
18      $\ell_{VDV}.removePartsOccludedBy(modulator[u, v], \ell_{IN})$ 
19   end
20 end

```

**Algorithm 4.3:**  $\ell_{VDV} \leftarrow assignViewDependentVolumetric(\ell_{IN}, \ell_{residue}, D_{IN})$

In a second step shown in Algorithm 4.3, diffuse components which are partially occluded are assigned to emissive layers and properly occluded by a modulating layer. Optimally, each emissive layer is preceded by a modulating layer to provide correct occlusions. However, as in practice modulating layers often absorb much light also in their transparent state, a fewer number of modulating layers is desired. For each emissive pixel, occlusions for all emissive layers in front are detected, and the modulating layer closest to but behind the occluding layer is used for masking. The occlusion mask is retrieved by intersecting the ray from the emissive pixel to its occluding pixel with the chosen modulating layer. This step creates black borders (illustrated in the third image of Figure 4.1 or in the bottom image of Figure 4.5) since occlusions are detected conservatively over the whole viewing angle. Therefore, such over-occlusions have to be added back to the residue, as they have to be rendered as view dependent light field.

```

1 for  $x, y, z, R_{diffuse}, R_{specular} \in \ell_{residue}$  do
2   | if  $R_{diffuse}(x, y, z)$  on emissiveLayer  $\in D_{IN}$  then
3   |   | project  $R_{specular}(x, y, z, u, v)$  on  $R_{specular}(x, y, emissiveLayer.z, u, v)$ 
4   |   end
5   end
6 Perform hole filling for continuous surfaces:  $WarpResidue(\ell_{residue})$ 
7 for  $x, y, z, R_{diffuse}, R_{specular} \in \ell_{residue}$  do
8   |  $lightfieldLayer \leftarrow getClosestLightFieldLayer(x, y, z, D_{IN})$ 
9   |  $lightfieldLayer[x, y, u, v] \leftarrow R_{diffuse} + R_{specular}$ 
10  |  $\ell_{VDL}.add([x, y, z, u, v, R_{diffuse}, R_{specular}])$ 
11 end

```

**Algorithm 4.4:**  $\ell_{VDL} \leftarrow assignViewDependentLightField(\ell_{residue}, D_{IN})$

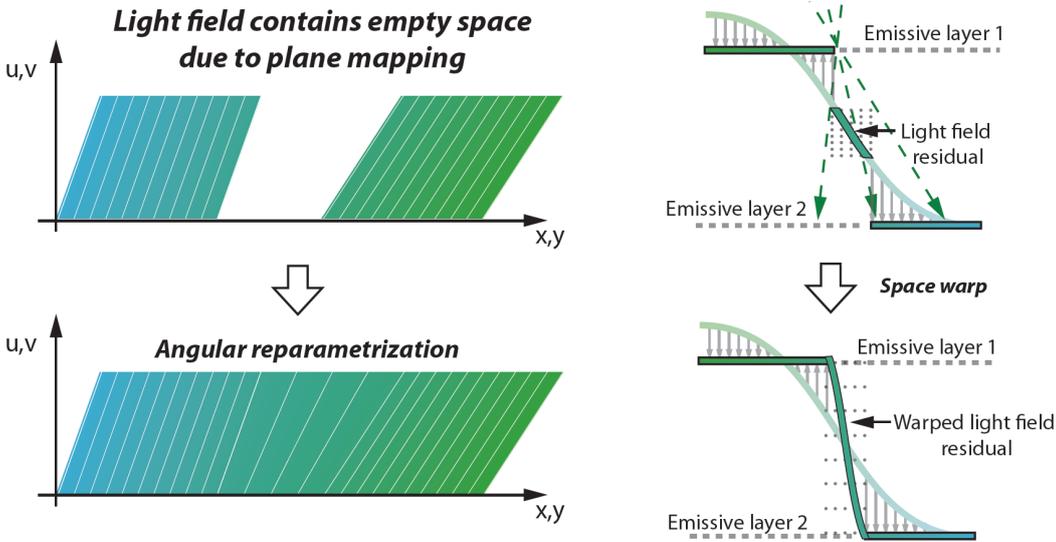
In a last step, the residue has to be rendered, filling the black borders and adding other view-dependent light field portions. Due to the planar mapping to the emissive layers, the holes cannot be filled naively. Our solution stretches the occluded residual to match the gap borders as illustrated on the left hand side of Figure 4.4, which corresponds to a scale in depth as shown on the right hand side. The same has to be considered for the glossy parts: They are mapped to the same plane as the view independent/diffuse part before being rendered. All components of this third step shown in Algorithm 4.4 are rendered using the closest parallax barrier layer and have to be filtered accordingly using existing approaches [Zwicker *et al.*, 2006]. A result of the complete algorithm is shown on the right of Figure 4.1.

### 4.3 Decomposition Analysis

We analyze our system with respect to the projective error, the error introduced by the sampling in the light field layer and a quantitative error measurement.

#### 4.3.1 Projective Error

The projection onto the planar emitters inherently produces an approximation of the motion parallax. The motion parallax  $d$  produced by an object at distance  $z$  to a viewer with focal length  $f$  that moves along a baseline at a distance  $b$  can be expressed as



**Figure 4.4:** Adaptive hole filling between emissive layers illustrated in ray space on the left and in geometry space on the right.

$$d = -\frac{f \cdot b}{z} \tag{4.5}$$

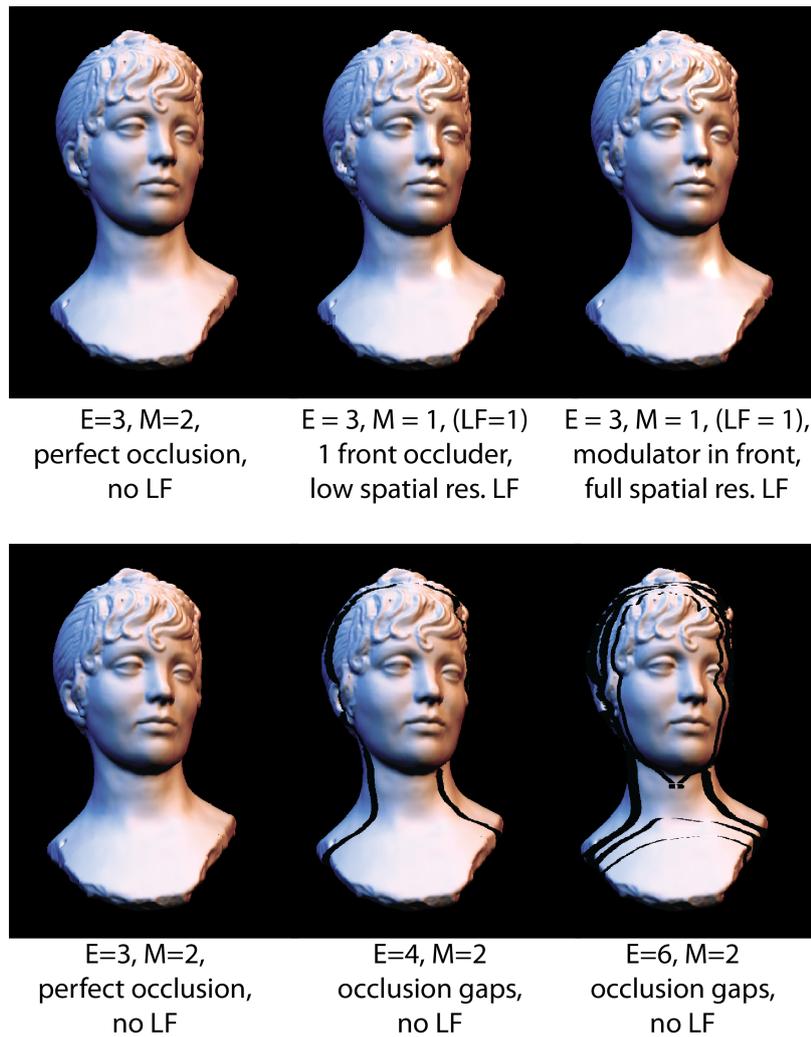
We express the relative projective approximation error of an object at distance \$z\$ projected on a plane at distance \$z\_0\$ by

$$e(z, z_0) = \left| \frac{1}{z_0} - \frac{1}{z} \right| \tag{4.6}$$

The emissive layers should therefore ideally be placed near dense occurrences of objects in depth, and the occluding layers should be placed as near as possible to the respective layers that need occlusion. Furthermore, fewer display elements are needed the farther the scene is with respect to the viewer’s position. This error can be used to determine the optimal display configuration using a suited optimization method, in cases where a display configuration is optimized for a given scene.

**4.3.2 Light Field Sampling Error**

Light field layers usually trade off spatial against angular resolution, and the approximation error is directly proportional to the loss in spatial resolution. However, additional errors are introduced if aliasing occurs when the angular frequencies are too high. These problems can be overcome by either using

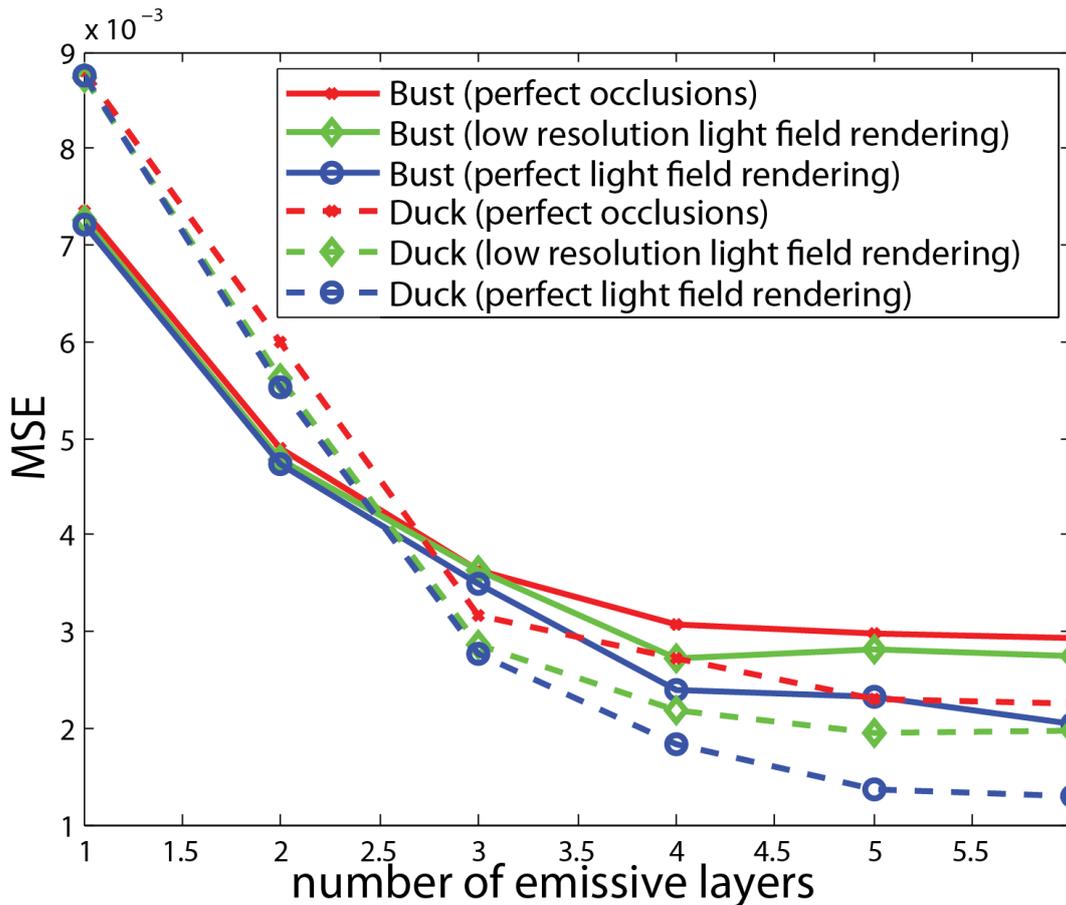


**Figure 4.5:** *Simulated results used for the quantitative error analysis.*

time multiplexing for the parallax barrier display, or by combining multiple parallax barrier displays superimposed onto the same optical path. Latter approach has further the advantage, that the required bandwidth for each light field layer is decreased by  $\frac{1}{n^2}$ , with  $n$  the number of layers, resulting in a total required bandwidth of  $\frac{1}{n}$ , as described in Section 5.1.

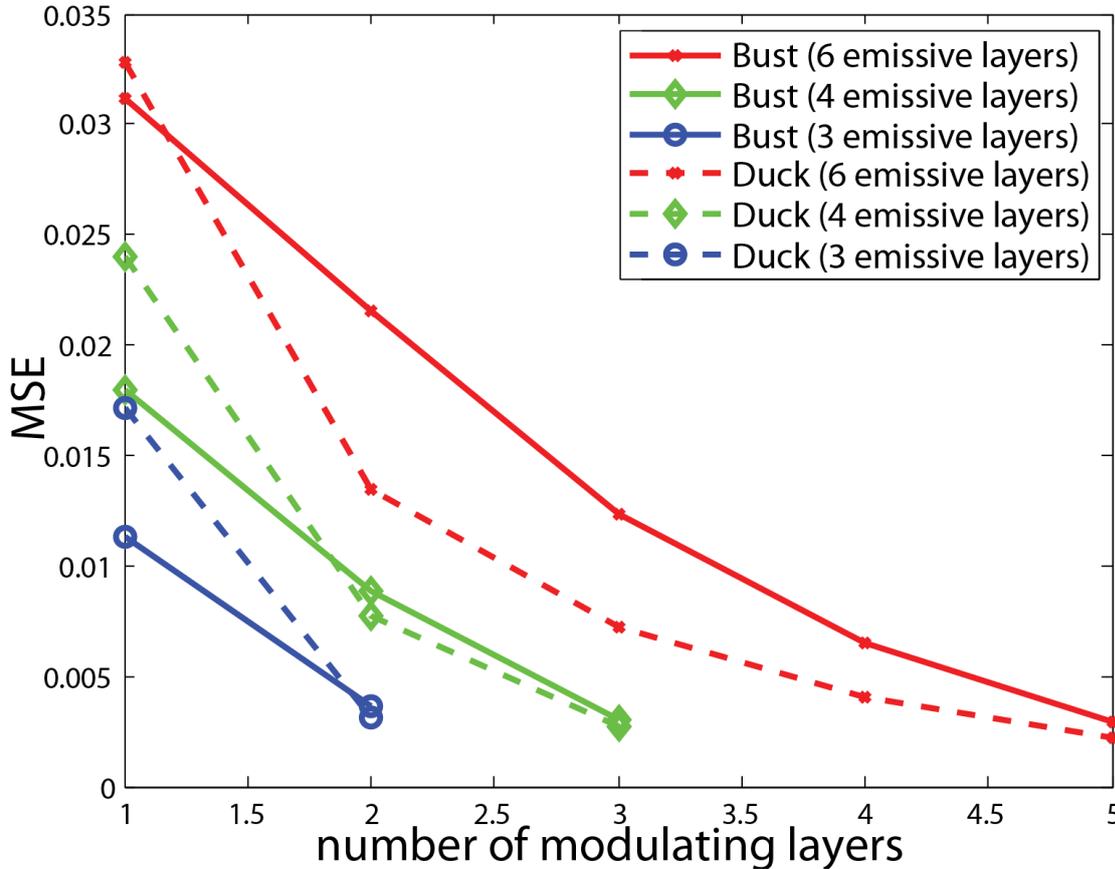
### 4.3.3 Quantitative Error

To analyze the impact of number of layer primitives, we compare the resulting reprojection errors using our software simulation. We simulated two different scenes: a duck scene containing two objects at different depths with occlusion, and a bust scene depicting a continuous surface. Both scenes contain a small



**Figure 4.6:** *Quantitative error analysis using simulated results. Shown is the impact of an increasing number of emissive layers. The mean squared error decreases fast for the first few layers but is not greatly affected by a high number of layers.*

amount of specular highlights. The simulated results are compared to a perfect rendering, and the MSE between the simulated and perfect images are computed for a number of views in a field of view of  $15^\circ$ . The resulting error plots are shown in Figure 4.6 and Figure 4.7. In Figure 4.6, the impact of an increasing number of emissive layers is depicted for three cases: In a first case, each emissive layer is preceded by a modulating layer, providing perfect occlusions for scene content in the back (red plots). In a second and third case, only one modulating layer is deployed front most and view dependent content is added by low resolution (green plots) or full resolution (blue plots) light field rendering. The error decreases fast when adding the first few layers, however, adding more than 4 layers does not reduce the error as drastically anymore. In Figure 4.7, the impact of an increasing number of modulating layers is assessed for a fix number of six (red plots), four

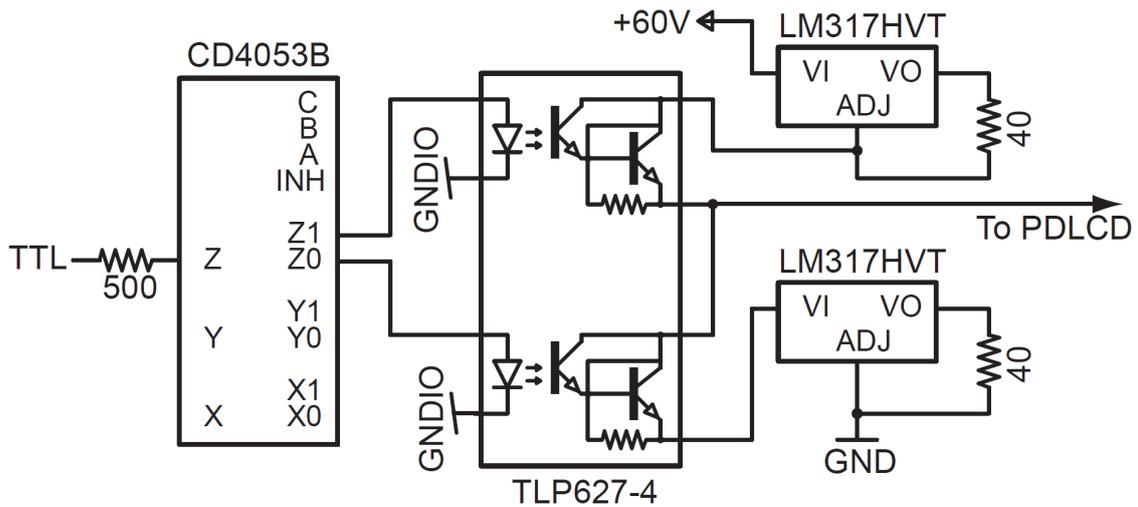


**Figure 4.7:** Illustration of the impact of an increasing number of modulating layers. No more modulating layers are required than the number of emissive layers minus one.

(green plots) and three (blue plots) emissive layers. For this analysis, the modulating layers are placed after each emissive layer, starting from the front-most layer. An increasing number of modulating layers helps to reduce the error significantly. The plots furthermore show that increasing the number of emissive layers without increasing the number of modulating layers leads to significant high errors due to incorrect occlusions, perceived as black gaps as shown in Figure 4.5 on the bottom.

#### 4.4 Physical Prototype

Based on the analysis presented in the previous section, we implemented two types of multi-layer plenoptic displays. Our two setups are shown in Figure 4.9 and Figure 4.10. The first prototype uses temporal multiplexing to



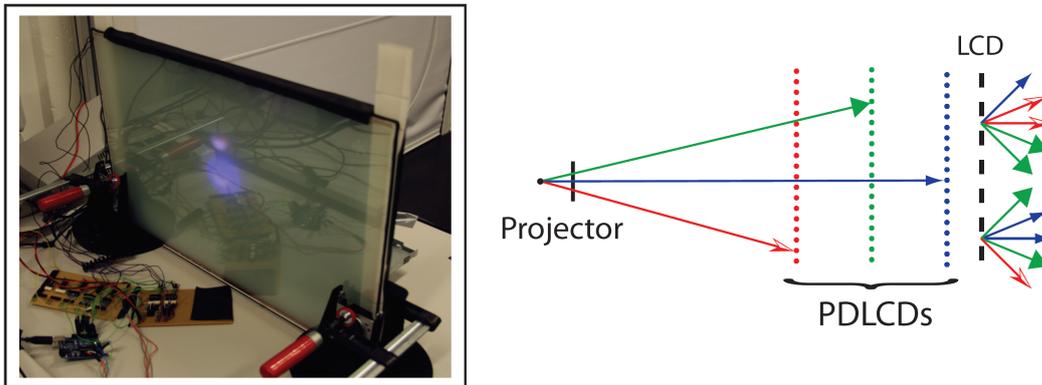
**Figure 4.8:** Circuitry used to create an alternating square wave function to drive the PDLC layers.

superimpose the different layers, while the second uses spatial multiplexing. Temporal multiplexing is performed by combining a projector with multiple bi-state scattering planes. Spatial multiplexing is performed by combining two automultiscopic displays using a beam splitter.

#### 4.4.1 Temporally Multiplexed

We provide an example configuration in a first, temporally multiplexed prototype. It deploys three PDLC layers in combination with a projector to approximate the emissive layer primitive, and a LCD in the front as modulating layer. In each frame, one of the PDLC layers is opaque and diffuses incoming light while the others are transparent. This allows showing different images on different layers in time multiplexed manner, similar to [Sullivan, 2004]. The PDLC layers are driven by the circuitry shown in Figure 4.8 which creates an alternating square wave function, preserving damages to the liquid crystal structures. The circuitry is synchronized with the v-sync signal of the projector such that the opaque layer is switched with each newly projected frame. As modulating layer, we deploy an LCD with non-diffusing polarizing films. This layer is front most and used both to provide approximated occlusion as well as to render the light field portion of our decomposition algorithm. The PDLC layers are spaced at 4mm, 10mm and 16mm from the front LCD. The projector has XGA resolution and a refresh rate of 60Hz, matching the refresh rate of the used PDLC layers. The front LCD renders 12 views in a  $10^\circ$  field of view when used as parallax barrier layer. The effective refresh rate of the display is 15Hz, since the front most emissive layer is used

both for volumetric rendering as well as for parallax barrier rendering. The complete setup is shown in Figure 4.9. Multi-planar plenoptic displays with homogeneous and well-aligned optical elements do not require calibration, since pixels are stacked directly behind each other in a one-to-one correspondence. However, in our setup we combine heterogeneous elements such as the projector and the LCD, which makes software calibration necessary. We propose a variant of the calibration scheme proposed in [Annen *et al.*, 2006] and perform homography estimation based on photographs of projected checkerboard patterns.

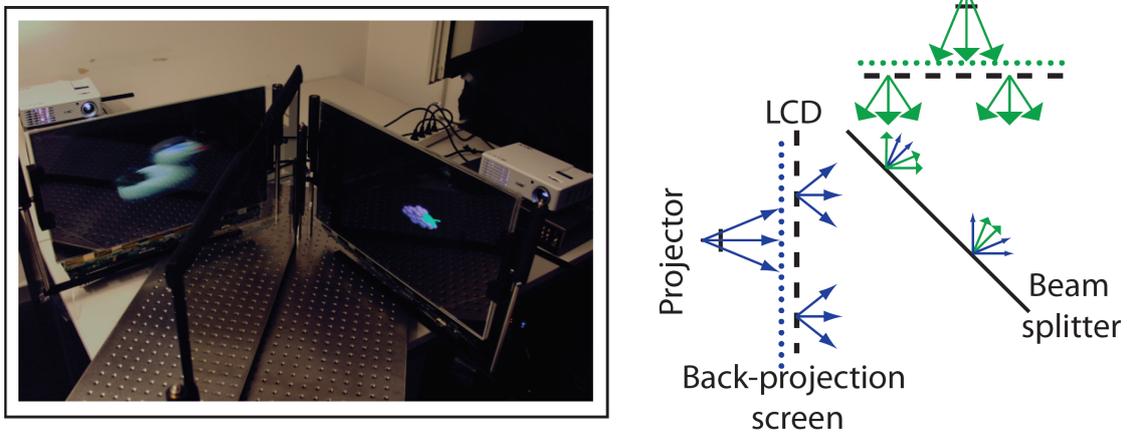


**Figure 4.9:** *Our temporally multiplexed prototype consists of three emissive layers and a front most modulating layer.*

#### 4.4.2 Spatially Multiplexed

Our second prototype shown in Figure 4.10 combines two automultiscopic parallax barrier displays using spatial multiplexing. The parallax barriers are only used for 3D light-fields, i.e. the barriers only provide distinct rays aligned with the horizontal plane. Both displays are combined onto the same optical path using a beam-splitter mirror. Each of the displays is placed at a different distance from the beam splitter and is then used to display different parts of the light field to achieve increased depth range. Each parallax barrier display is composed of a 120 Hz projector paired with a diffuse back-projection layer for the emissive primitive and a TN-LCD displaying a parallax barrier pattern as modulating primitive. Parallax barrier displays usually require big spacing between the barrier stripes to achieve an acceptable angular resolution. They therefore result in spatially under-sampled images that additionally lack a considerable amount of brightness due to the pin-hole nature of the barrier. We therefore employ temporal multiplexing for each barrier display: multiple spatially offset barrier patterns are projected

in short sequence, with the respective light field content on the emissive primitive to achieve higher perceived spatial resolution. Both projectors and displays have a native resolution of  $1920 \times 1080$ , the pixel spacing is approximately  $0.27\text{mm}$ . The parallax barrier and the emissive plane are at a distance of  $10\text{mm}$ . The spacing between the parallax barrier slits is between 9 and 12 pixels for one frame, and is adjusted according to the displayed scene. Three consecutive frames are used for temporal multiplexing, resulting in perceived parallax barrier spacing of 3 to 4 pixels. The corresponding 9 to 12 views of each parallax barrier consequently can be used to achieve a field of view of  $7^\circ$  to  $9.5^\circ$ . The virtual distances between the two parallax barrier displays are  $100\text{mm}$  and  $200\text{mm}$ .



**Figure 4.10:** *Our spatially multiplexed prototype uses a beamsplitter to overlay two parallax barrier displays.*

## 4.5 Discussion

We critically assess our achieved results and show limitations induced by both the hardware and the algorithm.

### 4.5.1 Results

We demonstrate several examples for both the spatially and temporally multiplexed setups. Figure 4.1 shows all steps of our algorithm captured on the temporally multiplexed prototype. Figure 4.5 shows some of the simulation results used for our quantitative analysis. Figure 4.11 shows real results displayed on the spatially multiplexed prototype with visible parallax between the two views. Our displays are quite unique since an observer is

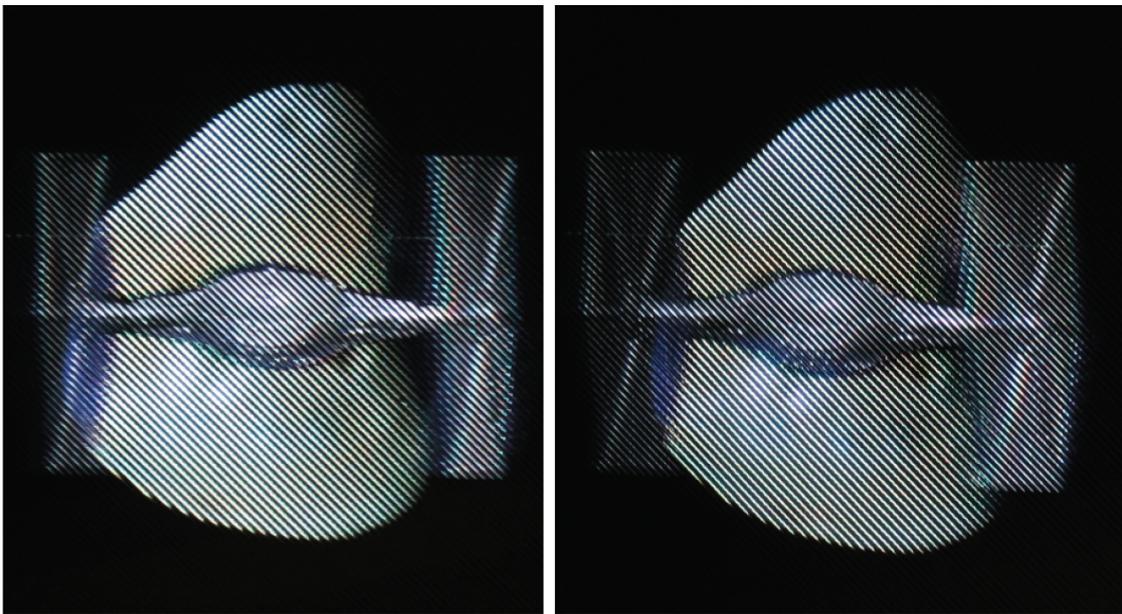
provided with accommodation cues, visible in Figure 4.12, as well as binocular cues and motion parallax. In addition to this, view-dependent effects can be observed clearly and they add to the depth perception. We believe that the superimposed spatial modulator does not significantly influence the accommodation cues.

### 4.5.2 Component Limitations

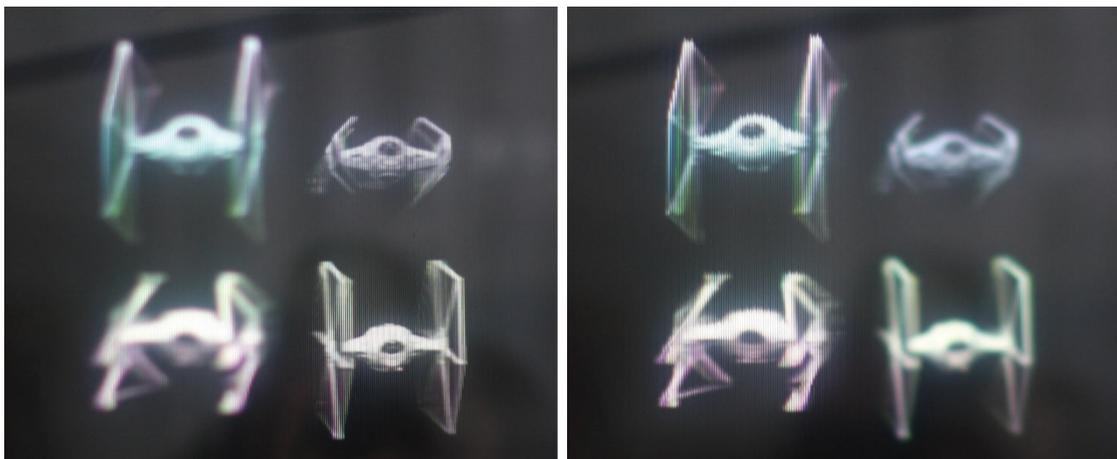
Our beam-splitter setup as well as the time multiplexed setup is limited in terms of possible size and scalability. As shown by [Barnum *et al.*, 2010], the maximum possible number of layers is limited by the finite switching speed of the layers, and the projector refresh rate. In addition, the dynamic range of the display is limited. In order to achieve constant brightness across the displayed light field, the lines in parallax patterns must be brighter than the parts rendered using the emissive layers only. Furthermore, the PDLCD panels are maximally 80% transparent and can only switch at 60Hz. These could be replaced by more-transparent and faster switching panels, as shown by [Sullivan, 2004]. In addition, common LCD panels take several milliseconds to switch from white to black, causing shadowing. Finally, transmission of LCD layers is typically less than 10%, which makes the stacking of many LCD layers impractical. Much faster switching panels exist, such as the -cell, and there is research on large-size, sub-millisecond switching modulators.

### 4.5.3 Algorithm Limitations

Our light field decomposition assumes knowledge of the scene depth, and is so far restricted to synthetic scenes. Our algorithm could work well with pre-recorded light fields, as long as a sufficiently accurate and dense depth map can be computed. Furthermore, we currently cannot handle transparent or semi-transparent scene elements. Finally, the spatially-varying modulators assign fully blocked or transparent states only. Taking advantage of the intermediate states could, in principle, improve the overall bandwidth.



**Figure 4.11:** *Two views captured on our spatially multiplexed prototype with clearly visible parallax and proper occlusions. The spaceship is rendered on the front layer where the moon is displayed by the layer in the back.*



**Figure 4.12:** *Multi-planar plenoptic displays provide good accommodation cues as visible in these images, captured on our spatially multiplexed prototype with different camera focus.*

## *Volumetric Displays*

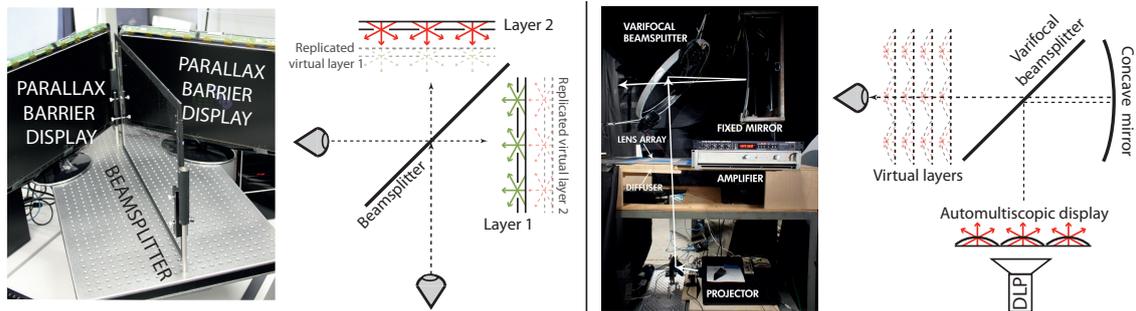
---

# CHAPTER

# 5

## Lightfield Displays

---



**Figure 5.1:** *The two different hardware prototypes used to proof our bandwidth theory. The first hardware implementation uses a beam splitter to superimpose two custom-built parallax barrier color displays. The second hardware implementation uses a varifocal beam splitter setup: the central beam splitter is vibrated to virtually position a lenticular-based automultiscopic monochrome display onto many depth planes. Both displays support horizontal and vertical parallax, wide viewing angles and better accommodation cues as single layer displays..*

---

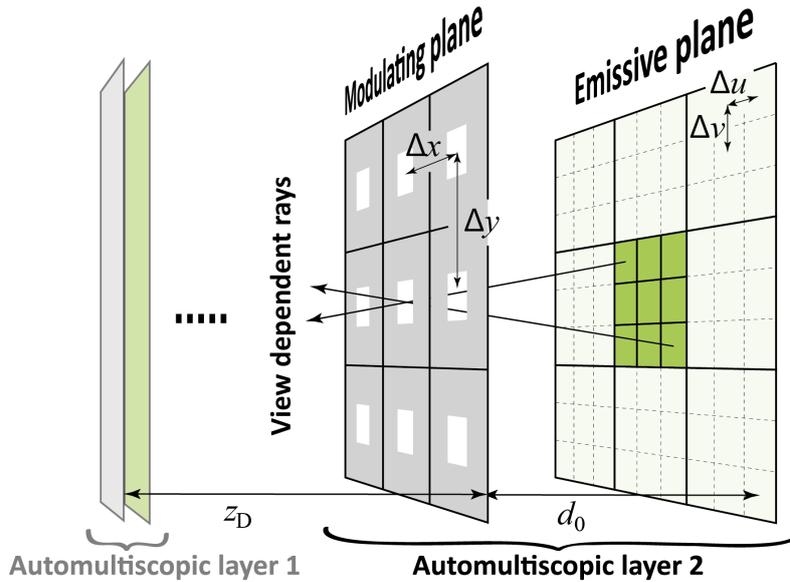
The light field display class is the most promising one, yielding potential to revolutionize 3D display technologies. Such systems do not require a viewer to wear glasses, are not restricted to their display volume and theoretically are capable to provide all visual cues for a comfortable viewing experience. However, to do so they require an enormous bandwidth to be capable to sample the plenoptic function dense enough. If the sampling is too low, the accommodation cue is lost, if it is decreased even further aliasing occurs which restricts the displayable depth.

We therefore present a bandwidth theory to show that multiple light field layers can drastically increase the bandwidth. Also, we present a novel idea how light field displays can be realized as transparent versions allowing to overlay virtual content over a real scene. Finally we present a new technology which has a significant higher bandwidth than state of the art light field display technologies.

### 5.1 Multi-Layered Automultiscopic Displays

To increase the bandwidth of light field displays, a novel display architecture has to be designed. Thus we introduce multi-layered automultiscopic displays, a hybrid display model that combines the benefits of volumetric and parallax-based displays [Ranieri *et al.*, 2012]. The idea is very simple: multiple translucent display layers at different depths are combined onto the same optical path. In contrast to previous light field display designs, each of the layers is comprised of an automultiscopic layer to emit true view dependent rays. Our display model is therefore very similar to volumetric displays but capable of view-dependent occlusion. In addition to that, we can show that this design has a significant higher bandwidth than just the sum of the bandwidth of each individual layer, which is mathematically derived in Section 5.1.3.

Figure 5.2 illustrates this concept for a dual-layer configuration. Each automultiscopic layer consists of two planes. Rays are generated on the *emissive plane* in the back with angular sampling  $(\Delta u, \Delta v)$ . The emissive pixels are spatially separated into view-dependent rays, using pinholes on the *modulating plane* in the front, with spatial sampling  $(\Delta x, \Delta y)$ . The emissive and modulating plane of each layer are separated by a distance  $d_0$ , the layers themselves are positioned at a distance  $z_D$  apart. The individual automultiscopic layers are then superimposed onto the same optical path, at different depths. Note that we assume that our layers cannot support true occlusion, i.e., a display layer cannot block light from any back layer.



**Figure 5.2:** Multi-layered automultiscopic display for 4D light fields. Each layer consists of an automultiscopic display, e.g. using parallax barriers (shown here) or lenslet arrays. The layers are multiplexed on the same optical path. Rays are generated on an emissive plane, and spatially separated into view-dependent rays on the modulating plane.

### 5.1.1 Lightfield Distribution

For a given display configuration, the input light field  $\ell_{IN}(x, y, u, v)$  needs to be distributed to the individual display layers  $Layer_i$ . In principle, our distribution algorithm is very simple. For each output light field  $\ell_{OUT_i}$ , assign each ray from  $\ell_{IN}$  to the display layer  $Layer_i$  closest to the ray origin.

Due to the nature of our display, front layers cannot block incoming light from the back layers. Such occlusions can implicitly be handled by the light field distribution: occlusions are represented as rays with zero luminance in the respective output light fields  $\ell_{OUT_i}$ . Intuitively, these occlusions correspond to a shadow that is cast by an occluder on all the following display planes. Note, that rays generated outside the depth range of the layers (including object shadows) can lead to aliasing. In order to avoid such aliasing, the output light fields need to be filtered accordingly, e.g. using a method similar to [Zwicker *et al.*, 2006]. Algorithm 5.1 summarizes the light field distribution.

```

1 for  $\forall ray \in \ell_{IN}$  do
2   |  $layer_i \leftarrow closestLayer(ray.origin, Layers)$ 
3   |  $assignRay(layer_i, ray)$ 
4 end
5 for  $\forall layer_i \in Layers$  do
6   |  $\ell_{OUT_i} \leftarrow prefilter(layer_i, \ell)$ 
7 end

```

**Algorithm 5.1:** Light field distribution overview.

### 5.1.2 Algorithm

Algorithm 5.1 assumes known depth information for the input light field. For synthetic scenes, this information is already given, and the algorithm can be reformulated as simple modification to any given ray tracing framework. For each layer, the respective depth range is determined. Then, ray tracing is performed within this depth range only. To achieve anti-aliasing comparable to [Zwicker *et al.*, 2006], we employ spatio-angular multi-sampling: instead of casting one ray only, we stochastically sample the original  $(x, y, u, v)$  sampling locations. Algorithm 5.2 summarizes our extension. Note that the depth range of the outermost layers is extended to  $\pm\infty$ , in order to capture the whole scene.

```

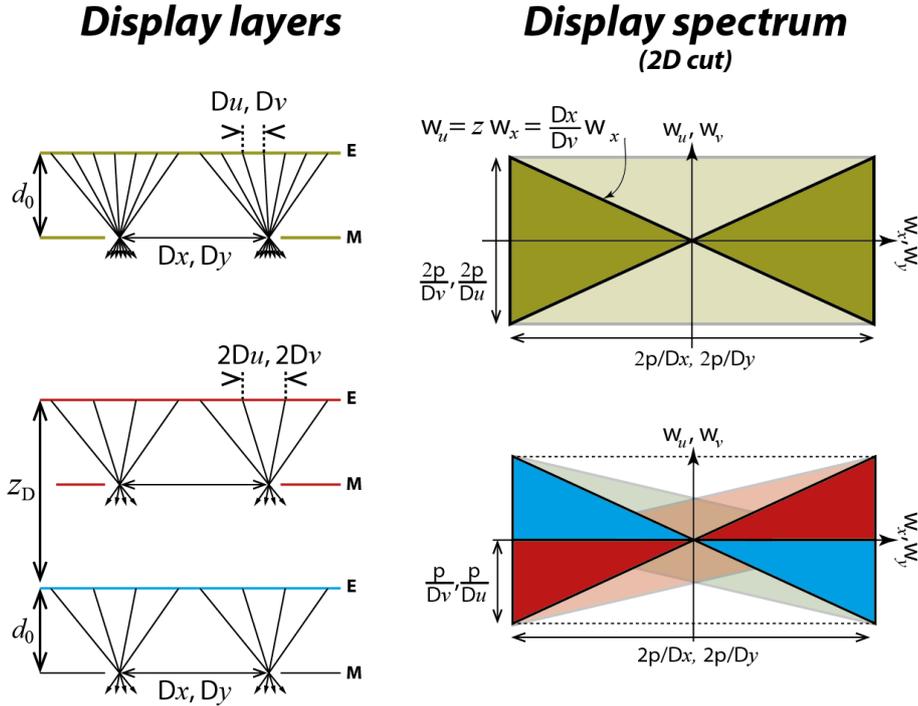
1 for  $\forall layer_i \in Layers$  do
2   | for  $\forall ray \in layer_i$  do
3     |  $generat\ n\ rays\ around\ (ray.x, ray.y, ray.u, ray.v)$ 
4     |  $intersect\ rays\ with\ scene$ 
5     |  $assign\ zero\ luminance\ if\ intersection\ is\ out\ of\ range\ of\ layer_i$ 
6     |  $filter\ rays$ 
7   | end
8 end

```

**Algorithm 5.2:** Light field ray tracing overview.

### 5.1.3 Bandwidth Analysis

In this section we will evaluate the effective bandwidth usage of multi-layered automultiscopic displays. We will start by introducing the bandwidth of a single layer, and will show that much of the available display bandwidth for 4D light fields is unused. We will then show that multiple layers can represent the same frequency content using less overall bandwidth.

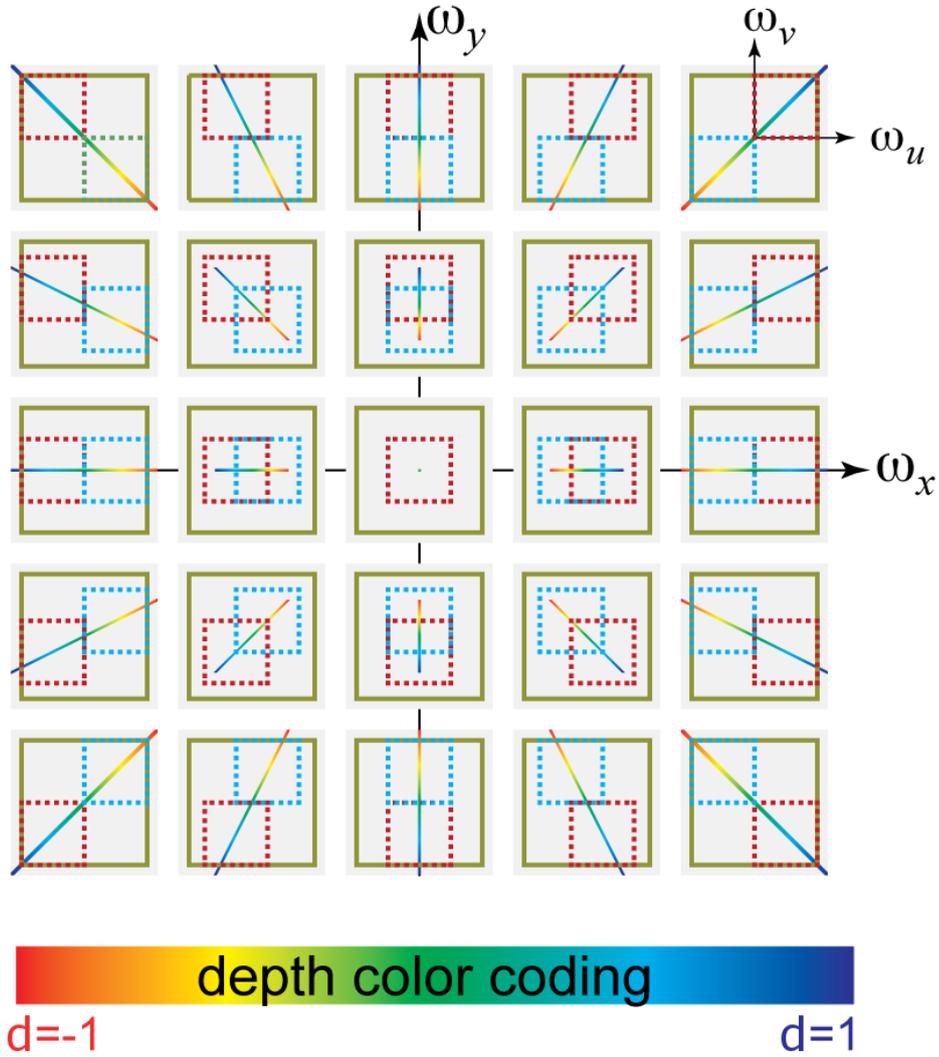


**Figure 5.3:** Bandwidth analysis for multi-layered automultiscopic displays, illustrated as a 2D cut through the 4D light field. A single layer display is illustrated in the top row, a dual layer configuration is illustrated in the bottom row. While both configurations share the same spatial resolution, each layer in the dual configuration only requires  $\frac{1}{4}$  angular resolution for horizontal and vertical parallax. The solid wedges delimit the frequencies that can be displayed without aliasing at full spatial resolution. The (translucent) rectangular boxes show the overall displayable spatial and angular frequencies for each layer. Both displays exhibit the same depth of field for frequencies at full spatial resolution.

Without loss of generality, we will assume a display with uniform spatial sampling  $\Delta x = \Delta y$  and uniform angular sampling  $\Delta u = \Delta v$  as well as unit spacing  $d_0 = 1$  between the  $(x, y)$  and  $(u, v)$  planes for this derivation.

### Single Layer

The bandwidth of an automultiscopic display is defined as the range of all possible frequencies that can be represented by the display. As noted by Zwicker et al. [Zwicker *et al.*, 2006], the maximum spatial and angular frequencies are delimited by  $\frac{\pi}{\Delta x}$  and  $\frac{\pi}{\Delta u}$ . The bandwidth of a display is then



**Figure 5.4:** The frequency support of a 4D light field corresponding to the 2D cut shown in 5.3. The outer axis represents the spatial frequencies while the axis of the inner sub plots stand for the angular frequencies. Covered frequencies are shown as rainbow colored lines where the color encodes the corresponding depth. The green boxes illustrated the bandwidth of a single layer while the red and blue dotted boxes represent the bandwidth of two layers in a dual layer configuration. The gain of bandwidth is easily visible as the smaller boxes cover less of the spectra but still support all frequencies covered by the light field.

## 5.1 Multi-Layered Automultiscopic Displays

defined as

$$H(\omega_x, \omega_y, \omega_u, \omega_v) = \begin{cases} 1, & \text{for } |\omega_{x,y}| \leq \frac{\pi}{\Delta x}, |\omega_{u,v}| \leq \frac{\pi}{\Delta u}, \\ 0, & \text{otherwise.} \end{cases}$$

The bandwidth corresponds to a 4D hyperbox with volume  $(2\pi/\Delta x)^2 \cdot (2\pi/\Delta u)^2$ . See Figure 5.3 for an illustration using a 2D cut. As also noted in [Zwicker *et al.*, 2006], much of the available bandwidth will be unused.

According to [Chai *et al.*, 2000; Levin and Durand, 2010], the light field  $\ell$  of a Lambertian plane parallel to the display plane will only exhibit frequency entries on a 2D plane. Therefore, the light field spectrum  $\hat{\ell}$  will contain non-zero entries only on  $\hat{\ell}(\omega_x, \omega_y, z\omega_x, z\omega_y)$ , where  $z$  is the distance of the plane to display. As a consequence, only frequency content lying within a 3D wedge in 4D space can be displayed without aliasing at full spatial resolution, illustrated as rainbow colored lines in the 4D plot shown in Figure 5.4. This implies that only a very small subset of the total bandwidth can effectively be used. Furthermore, the bandwidth box delimits the maximum depths that can be displayed without aliasing to  $z_{\max} = \pm(\Delta x/\Delta u)$ . More specifically, the maximum displayable spatial frequency  $\tilde{\omega}_x$  for a plane at distance  $z$  can be described as

$$\tilde{\omega}_x(z) = \pm \min\left(\frac{\pi/\Delta u}{z}, \frac{\pi}{\Delta x}\right). \quad (5.1)$$

The maximum frequency  $\tilde{\omega}_y$  follows directly. See Figure 5.3 for an illustration.

### Multiple Layers

We will first show the bandwidth usage of two automultiscopic layers. Assume that each layer has a quarter of the overall angular sampling  $\Delta u' = 2\Delta u$  and  $\Delta v' = 2\Delta v$  of a single layer display, and the spatial sampling remains constant. Therefore, the maximum depth that can be displayed without aliasing on a single layer reduces to  $z'_{\max} = \pm\Delta x/(2\Delta u) = z_{\max}/2$ . In order to cover the same depth range compared to the single layer display, the two layers need to be positioned at  $\pm z'_{\max}$  respectively.

The bandwidth of these two layers is smaller than the bandwidth of the single layer display, although it is able to represent the same content within  $z_{\max}$ . More specifically, the angular sampling is reduced by  $\frac{1}{2}$  in both  $u$  and  $v$  direction for both display layers. Therefore, the bandwidth of the system reduces to  $2 \cdot (2\pi/\Delta x)^2 \cdot (\pi/\Delta u)^2 = \frac{1}{2} \cdot (2\pi/\Delta x)^2 \cdot (2\pi/\Delta u)^2$ . As a consequence, two layers can display the same content compared to a single layer, but using only half the number of rays.

By the same geometric construction, the overall bandwidth and ray count for  $n$  display layers reduces to  $\frac{1}{n}$  compared to a single layer configuration. The maximum depth  $z_{\max}$  that can be displayed at full spatial resolution remains identical to the single layer configuration. The maximum displayable spatial frequency at a distance  $z$  from one individual layer  $\tilde{\omega}_x^{(l)}$  (positioned at  $z_l$ ) reduces to  $\tilde{\omega}_x^{(l)}(z) = \pm \min(\frac{\pi/(n\Delta u)}{z-z_l}, \frac{\pi}{\Delta x})$ . Each layer therefore has its own range of displayable frequencies. All layers are additively superimposed, and therefore the individual layers do not influence each other. Due to the overlap of the bandwidth boxes, however, some of frequencies can be displayed on both layers and therefore do not add additional information. The overlay of the bandwidth boxes are illustrated both in Figure 5.3 and Figure 5.4. The maximum displayable frequencies of the combined depth of fields can then be described by using the outer layer:

$$\tilde{\omega}_x^{(n)}(z) = \pm \min(\frac{\pi/(n\Delta u)}{z - \frac{n-1}{n}z_{\max}}, \frac{\pi}{\Delta x}),$$

As long as the light field content is within  $\pm z_{\max}$ , our multiple layer approach can display the same light field content. Outside this depth of field, the possible frequencies are reduced compared to the single layer configuration. Note, that in the limit ( $n \rightarrow \infty$ ), each layer would only emit one angular sample, and the multi-layered automultiscopic display would resemble to a volumetric display without view dependency. In order to implicitly support occlusions for additive layers, view dependent pixels need to be supported, and currently occlusions can only be approximated by filtered object shadows.

If the frequency spectrum is much sparser, the overall bandwidth could be optimized even further for a given scene. For example, a scene could be comprised of a foreground object and a background object separated by a large distance. Then, the display layers could be separated by the same distance in order to optimally use the respective bandwidths. If the scene or the depth separation changes, however, the displays would have to be reconfigured.

#### 5.1.4 Software Framework

We implemented our light field distribution algorithm for ray tracing within Optix [Parker *et al.*, 2010], a framework for general purpose ray tracing on the GPU. Our implementation performs ray tracing for all layers in parallel, and all rays are generated according to the ray sampling of the layers. Anti-aliasing is achieved by stochastic super-sampling around the original sampling locations, and all samples are interpolated using a box filter. Our

## 5.1 Multi-Layered Automultiscopic Displays

ray tracer supports only basic ray-casting only but it could easily be extended for more realistic image generation.

**Layer borders.** The layer distribution can introduce high frequencies, especially when continuous surfaces are separated by two layers. In order to avoid noisy artifacts, a high number of multi-samples would be necessary. We mitigate these issue using a different strategy. Instead of using binary cuts, our implementation employs 'fuzzy' layer borders: we slightly overlap neighboring depth of fields, and we linearly weight the corresponding samples that fall within the overlapping region, with respect to the actual layer borders. This strategy greatly reduces these artifacts without the need for very high sampling densities.

**Optimized field of view.** Using regular sampling, many rays will fall outside the field of view for close viewing positions. In order to increase the effective ray utilization for a given viewing position, the emitted view rays can be sheared along the angular direction  $u' = u + sx$  and  $v' = v + sy'$ , where  $s$  is dependent on the display parameters and viewing distance. In our implementation, we round the sheared rays to the nearest sampling location determined by the pixel grid.

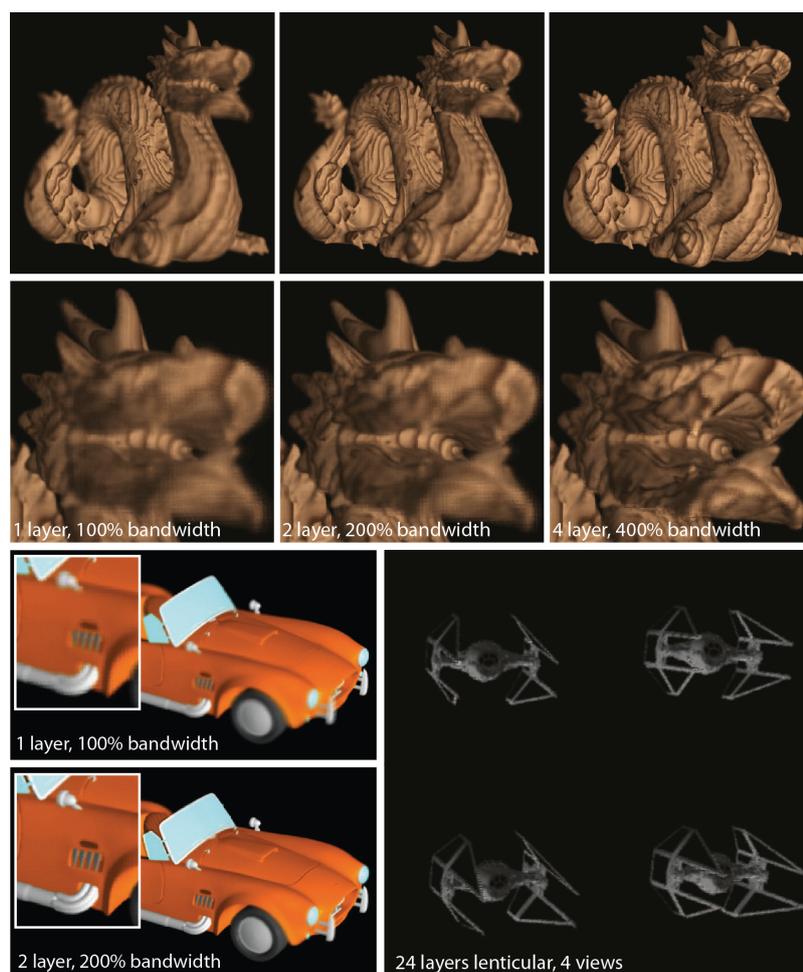
**Display simulation.** The simulated results have been generated using a custom ray tracer implemented within the Optix framework. Each display layer is represented by two planes. The emitting plane is assigned a luminance texture corresponding to the generated pattern from the light field distribution. The modulating plane is assigned a transmission texture corresponding to the spatial sampling. For the simulation, we apply super-sampling of the viewing rays to approximate cross-talk.

### 5.1.5 Hardware Setups

We evaluate the effectiveness of our algorithm using two hardware prototypes. The first prototype uses two automultiscopic layers superimposed onto the same optical path using a beam-splitter with high resolution, while the second prototype allows us to demonstrate our technique with up to 24 depth planes, but at lower resolution.

#### Beam-Splitter Prototype

The setup of the first prototype consists of two automultiscopic layers which are combined onto the same optical path using a beam-splitter, as illustrated in Figure 5.1 on the left. Our custom automultiscopic layers are constructed



**Figure 5.5:** *Simulated display results. The overall depth of field increases greatly when using multiple layers, supporting large parallax movement.*

using two LCD layers stacked on top of each other. The back emissive layer is comprised of a regular LCD display with backlight. The front modulating layer consists of a disassembled and modified LCD panel from a regular LCD display. The diffusing front polarizer and the back polarizer have been removed and replaced with non-diffusing and matching polarizers, rotated by 90 degrees. In order to reduce Moire patterns, an additional diffuser with a small point spread function of approximately one pixel is placed in front of the back LCD. Both LCDs are then stacked on top of each other and physically separated using a layer of acrylic glass. The assembly of one individual layer is very similar to [Lanman *et al.*, 2010], with the following differences. In their setup, the diffuser exhibits a much larger point spread function, effectively reducing their available display resolution. Furthermore, their modulating LCD features only a front polarizer but no back polarizer. Unfortunately, the diffuser to reduce the Moire pattern destroys much of the polarization

## 5.1 Multi-Layered Automultiscopic Displays

from the emissive layer, and using a front polarizer only would not result in sufficient contrast. The LCD panels have been taken from a Acer HN274H display (27", 1920x1080, 120Hz). The panels are driven from a dual-head NVIDIA GTX 580 graphics card.

In order to increase the perceived spatial resolution, we employ time multiplexing of the parallax barriers similar to Kim et al. [Kim *et al.*, 2007]. Synchronization for one automultiscopic layer is performed implicitly by the graphics board. Synchronization between the two automultiscopic layers would require higher end graphics boards, and thus displaying dynamic scenes is not possible with the current prototype. The alignment of the stacked LCDs as well as the automultiscopic layers was performed using careful manual adjustment, and could be improved using more advanced assembly setups and by using automated calibration techniques as presented in Chapter 6.

### Varifocal Multi-Plane Prototype

For the second verification of our algorithm we use a setup similar to a recently proposed volumetric display [Smithwick *et al.*, 2012]. In its original form, a 60Hz display is used in conjunction with a high-speed DLP projector used as back-light. The depth extrusion is achieved using a large vibrating beam-splitter, which is comprised of a metalized Mylar polyester film membrane stretched over a circular hoop. Three equidistant transducers are mounted to the edge of the hoop in order to vibrate the beam-splitter axially. The beam-splitter surface tension is tuned to vibrate at an eigen frequency of 30Hz with a high Q-factor, and as such, its surface becomes alternatively convex and concave. The display surface is relayed by the vibrating beam-splitter towards a fixed concave mirror. The returning light passes through the beamsplitter and forms a real 2.5D stack of 2D images in front of the apparatus. In its current setup, the system is able to create a layered volume of up to  $16.7 \times 12.5 \times 18.8$  cm, without support for occlusions.

In our setup shown in Figure 5.1 on the right, we exchanged the 2D display with an integral imaging-display by substituting the LCD with a rear-projected emissive screen in conjunction with a micro-lens array placed on top. Our display is therefore able to support volumetric layers with occlusion effects, in contrast to [Smithwick *et al.*, 2012]. A high-speed DLP projector (Light Commander from Logic PD) provides monochrome images that are synchronized to the vibration of the mirror. Using the projector's monochrome mode, we are able to achieve up to 24 images per stroke of the mirror at an aggregate frame rate of 1440 frames per second, forming up to 24 image planes at 60 Hz. The microlens array is comprised of staggered 2D

fly's eye lenslets in a close-packed hexagonal format, in order to support both horizontal and vertical ray separation of the underlying image. The Light Commander's image size (at a resolution of 1024x768 pixels) is chosen to provide 13 pixels horizontally and 11.3 pixels vertically under each lenslet. The field of view of the lenslets ( $41^\circ$ ) exceeds the field of view of the volumetric display ( $19^\circ$ ). We therefore pad the outer viewing rays to reduce light transmission through the seams between the individual lenses, which leads to effects similar to cross-talk. A total number of 6x6 rays is used per lenslet.

### 5.1.6 Discussion

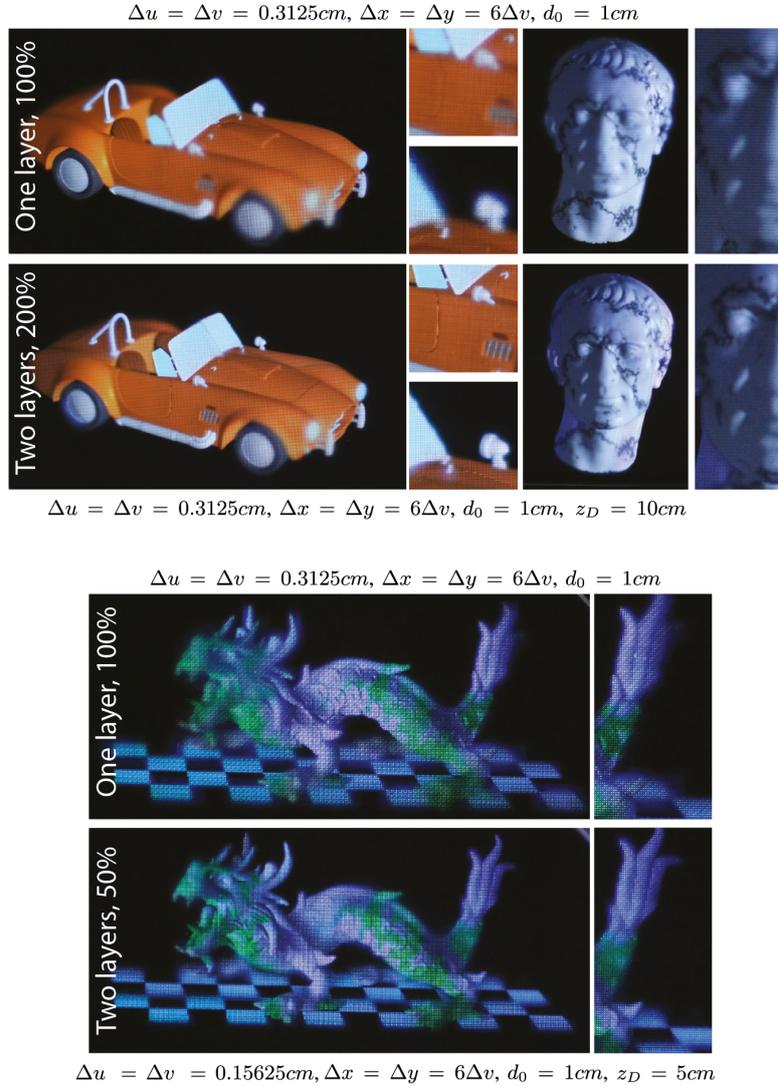
We evaluate the effectiveness of our approach using simulated results and results from both prototypes. The simulation results shown in Figure 5.5 illustrate the effect of adding additional layers. With every additional layer, the depth of field increases, leading to much sharper images for the outer depth ranges. In our simulation, we employed 36 time multiplexing steps to show the results in full spatial resolution. Note, although a high number of multiplexing steps is extremely difficult to achieve with current display technology, we intend to demonstrate the effect of the increased depth of field without distracting resolution artifacts.

Figure 5.6 shows results captured with our beam-splitter prototype. Although only using two layers, the depth of field already enhances noticeably when using more bandwidth than a single layer display. The results also show that the same depth of field of a single layer display can be reproduced by two displays using only half of the overall ray count.

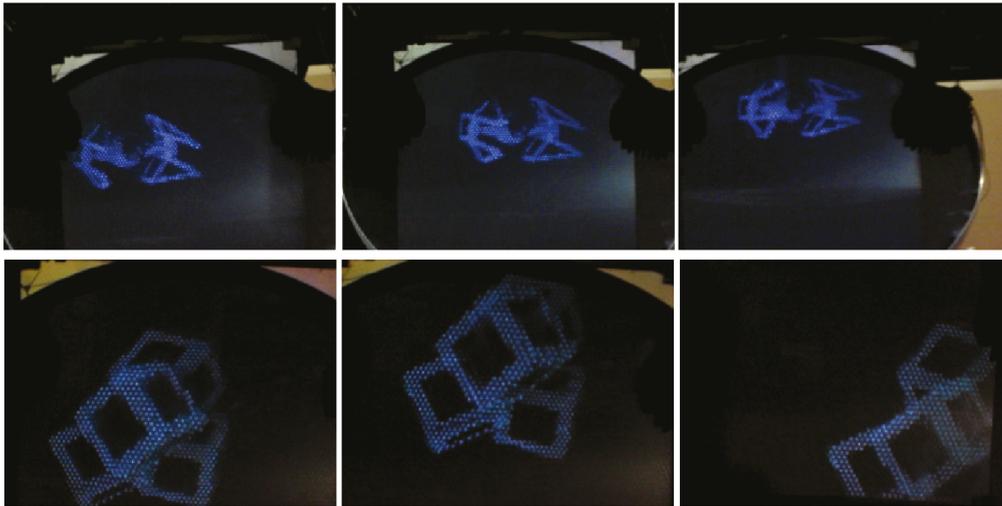
The results of the varifocal display prototype are shown in Figure 5.7. The possible depth of field (18.8cm) at wide viewing angles of  $18^\circ$  is huge compared to existing displays. Furthermore, by using up to 24 layers, the display provides nearly correct accommodation cues. Although the prototype is able to support monochrome images only, the resulting parallax movement can be perceived quite well in the accompanying video. Note, that the varifocal prototype therefore exceeds the limits of human depth resolution, see [Akeley *et al.*, 2004] for more details.

Our display model shares the same trade-off between spatial and angular resolution as all automultiscopic displays. For a next prototype version, we plan to use smaller lenses with higher image resolutions. While using a parallax barrier approach reduces the brightness considerably, using lenticular arrays avoids this brightness loss at the cost of increased crosstalk. In addition, both prototypes employ optical stacking of multiple primitives which leads to a

## 5.1 Multi-Layered Automultiscopic Displays



**Figure 5.6:** Results from our dual layer beamsplitter prototypes. The car and the bust model are shown on single layer configuration with 100% bandwidth (top row), and on a dual-layer configuration with 200% bandwidth (bottom row). The dragon model is shown on a single layer configuration with 100% bandwidth (top row), and a dual layer configuration with 50% bandwidth (bottom row). Note the increased sharpness around edges and high-frequency textures. We employ 9 time multiplexing steps to increase the spatial resolution. Photos of the dual-layer prototype are taken with a Canon Eos 1D Mark III, ISO 3200 and  $\frac{1}{4}$ s exposure time.



$$\Delta u = \Delta v = 0.18mm, \Delta x = \Delta y = 2.33mm, \text{ and } d_0 = 3mm$$

**Figure 5.7:** Results from our varifocal display prototype. Using 24 layers, we can achieve a very wide depth of field of approximately 19 cm at wide viewing angles. Note that by using many layers, our system is able to provide nearly correct accommodation cues. Photos of the varifocal results are taken with a small hand-held PowerShot camera.

brightness reduction of  $\frac{1}{n}$  for each individual layer, compared to using one layer only. Due to the stacking of multiple layers, our displays inherently do not support view repetitions that are commonly found in parallax barrier or integral imaging displays.

The varifocal mirror prototype only supports monochrome images, and we hope to extend the system to a gray-scale version using e.g. temporal super-sampling. Combining multiple pixels into one luminance value could be another option, by using higher resolution projectors in conjunction with diffusers that support respective blur kernels. High speed color images might become possible in the future, for example by combining multiple DLP chips into one optical system.

In our prototypes, the individual layers cannot block incoming light from other layers and therefore occlusions cannot be handled correctly by the hardware. Using the implicit occlusion handling of our light field distribution algorithm, the resulting filtered object shadows will be blurred out. This can lead to slightly noticeable black halos around the occluder, which becomes more pronounced when increasing the distance between the layers. However, the shadow is barely noticeable in our experiments in cases where anti-alias filters can be applied. The varifocal mirror system only supports monochrome

images, and unfortunately no such filtering strategies can be applied unless multiple luminance values would be supported.

Both display prototypes exhibit a substantial amount of cross-talk, which results in additional blur. The customized LCD layers furthermore exhibit a slight color difference, and low contrast, which is most likely due to imprecise alignment of the emissive and modulating planes, as well as some slight shifts between the respective polarizers. Furthermore, alignment of the individual layers is performed manually at the moment, and some object seams are visible due to the imperfect alignment. Using the calibration technique presented in Chapter 6, this issue could be mitigated.

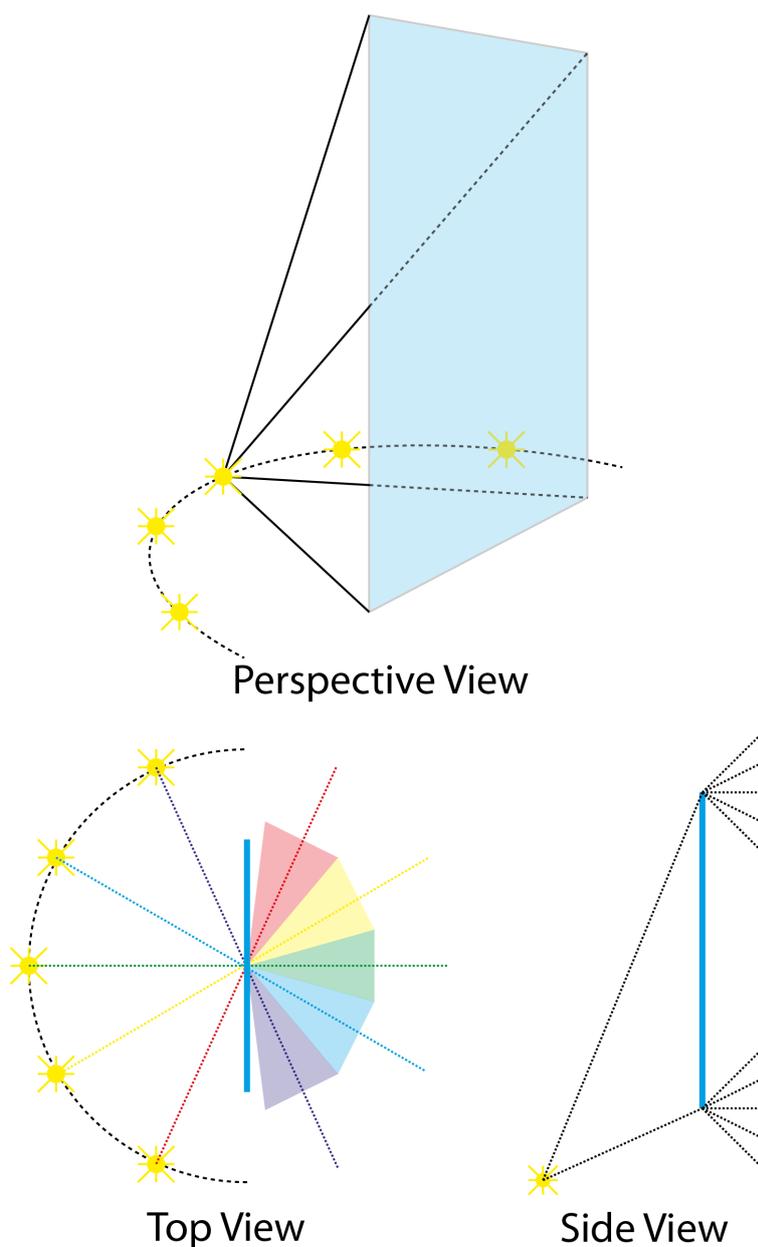
## 5.2 Transparent Holographic Light Field Display

Several approaches for light field displays have been proposed over the past decades. A few of them are semi-translucent as a result of the chosen technology, rather than a desired effect aimed for. To create a transparent light field display, the technology used in Section 3.3.1 for the transparent stereoscopic system can be paired with the approach presented in [Nagano *et al.*, 2013].

The screen used for our life-size telecommunication system is transparent but capable of diffusing rays of light from certain directions over a specified horizontal and vertical angle. This optical effect is generated using holographic elements diffracting light in the desired way. Thus, the angle of diffusion is a design parameter which can be chosen by a customer. If the angle of horizontal diffusion is chosen relatively narrow, the screen becomes a transparent 1D diffuser.

A transparent 1D diffuser can be used to replace the 1D retroreflector used in [Nagano *et al.*, 2013]. Instead of rays being reflected on the horizontal axis while being diffused over the vertical axis, rays simply pass the screen and only get diffused in the vertical plane. The principle is shown in Figure 5.8. Depending on how dense the projectors are packed, the vertical diffusion angle of the transparent holographic screen have to be chosen wide enough such that views from neighboring projectors are blended seamlessly.

Also, as so many projectors as used in the system of [Nagano *et al.*, 2013] are difficult to maintain and handle, we suggest to use a single high speed projector and field sequential multiplexing of images towards the screen which is described in the next section. The idea of using the temporal domain to build up a light field has been protected by the patent [Ranieri and Gross, 2014a].



**Figure 5.8:** Schematic illustration of an anisotropic transparent 1D diffuser. Similar to the screen used in [Kuster et al., 2012; 2015; Ranieri et al., 2014], only incident light from specific directions is diffused. Top image shows an arrangement where light coming from a semi arc is accepted by the screen while other light rays pass unhampered. In contrast to existing anisotropic foils, the horizontal angle of diffusion is kept small, as visible in the bottom left image. Projectors are placed next to each other with an angular spacing matching the horizontal diffusion angle. Thus, the image of different projectors is visible depending on the viewing angle, enabling automultiscopic 3D. Vertical diffusion behaves the same as in existing products, as shown in the bottom right image.

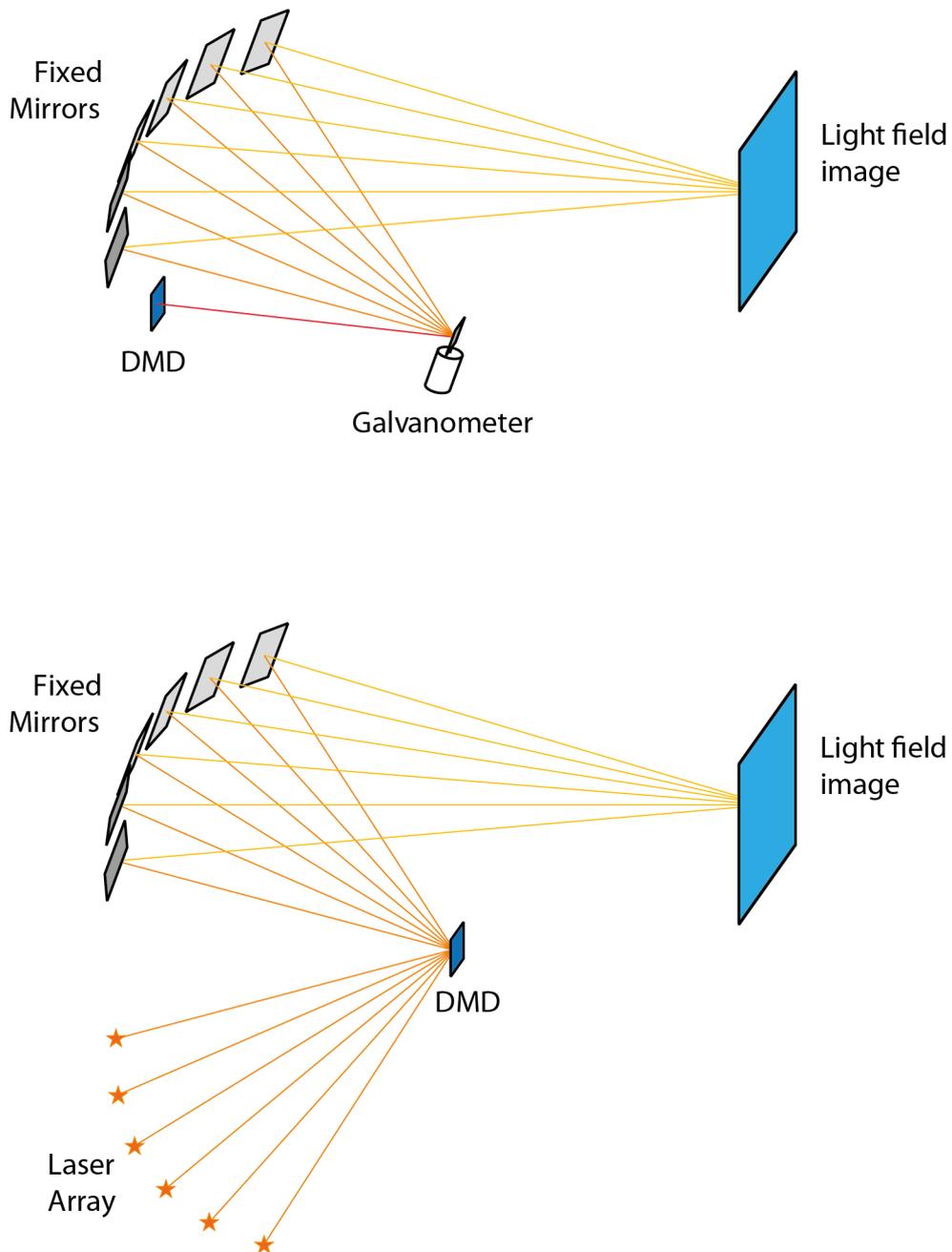
### 5.2.1 Field-sequential Image Multiplexing

Digital Micromirror Devices (DMD) are fast imaging devices capable of creating binary images at several kilo hertz. Consecutive binary frames can be turned into colored images using at least three different color primaries with tunable intensities, synchronized with the image flips of the binary frames. Example given, eight consecutive frames using a red light source with always half the intensity of the previous frame results in 256 different shades of red. Doing this for green and blue too, only 24 frames are required to create an image with over 16 million colors.

Repeating this procedure at 60Hz to enable motion pictures, the bandwidth of a DMD is still not completely used. In other words, images can be generated at e.g. 600Hz which can be interpreted as 10 different images each refreshed at 60Hz. If these images can be redirected to optical distinct paths, they can further be redirected to hit the holographic transparent 1D diffuser under different incident angles, achieving the same result as the multi-pico-projector setup suggested by [Nagano *et al.*, 2013]. We propose two different methods to do so, both illustrated in Figure 5.9.

**Galvonomic mirror:** A galvonomic mirror scanner is a small device, capable of changing the orientation of a small mirror very fast. The optics of a projector can be designed such that the projected image first converges to a small area before it expands again. If the mirror of the galvoscaner is placed at this point where the image is narrow, it can redirect images to different angles over time. The basic principle is shown in Figure 5.9 at the top. The redirected images can be captured by fixed mirrors to redirect them a second time back to the light field screen. As the galvo scanner is faster when the mirror surface is smaller, the design of the optics for image convergence and focusing is crucial. Finally, the scanner needs to be able to reposition fast enough to redirect all images at a decent frame rate.

**Multiple light sources:** The galvonomic mirror is a mechanically moving component, fragile and also not very reliable when operated at such high frequencies required in this system. Thus we propose in a second embodiment to use multiple light sources instead. As the DMD acts like a simple mirror, light sources at different locations create different projection angles as illustrated in Figure 5.9 in the bottom image. Each projection direction would need its own focusing optics if not point light sources are used. Therefore we suggest to use laser light sources in this version of the light field projector, as the so produced images always stay in focus. As the individual light sources can be turned on and off very quickly, the multiplexing of the images is not the bottleneck of the system as it is the case with the galvonomic mirror.



**Figure 5.9:** Illustration of two different embodiments on how to create a light field image using a single projector. Idea is to use a fast imaging device as a DMD and multiplex the individual images to different directions, where they are redirected by fixed mirrors to hit the 1D diffuser from different angles. In the top image, a galvanomic mirror is used to field-sequentially multiplex the images into different directions. In the bottom image, multiple light sources replace the mechanically moving mirror scanner achieving the same result.

### 5.2.2 Prototypes

For both approaches we provide prototypes for proof of concept. As the holographic 1D diffuser could not have been built within the time frame of this thesis, the system was used to increase spatial resolution of a 2D projection screen as described in [Ranieri *et al.*, 2015] rather than turning the additional images into angular resolution. The principles of the two projection systems stay the same. But instead of redirecting the different images to the same spot on a transparent 1D diffuse, images are projected adjacent to each other onto a projection screen.

**Galvonomic mirror:** We have built a prototype consisting of a LED light engine, a DMD to form an image, optics to focus the image as well as a galvonomic mirror to redirect the images to different directions. The light engines consist of a red, green and a blue high power LED from Luminus. They are overlaid onto the same optical path using an x-cube which is an optical element consisting of three dichroic mirrors used as beam combiner for different wavelengths.

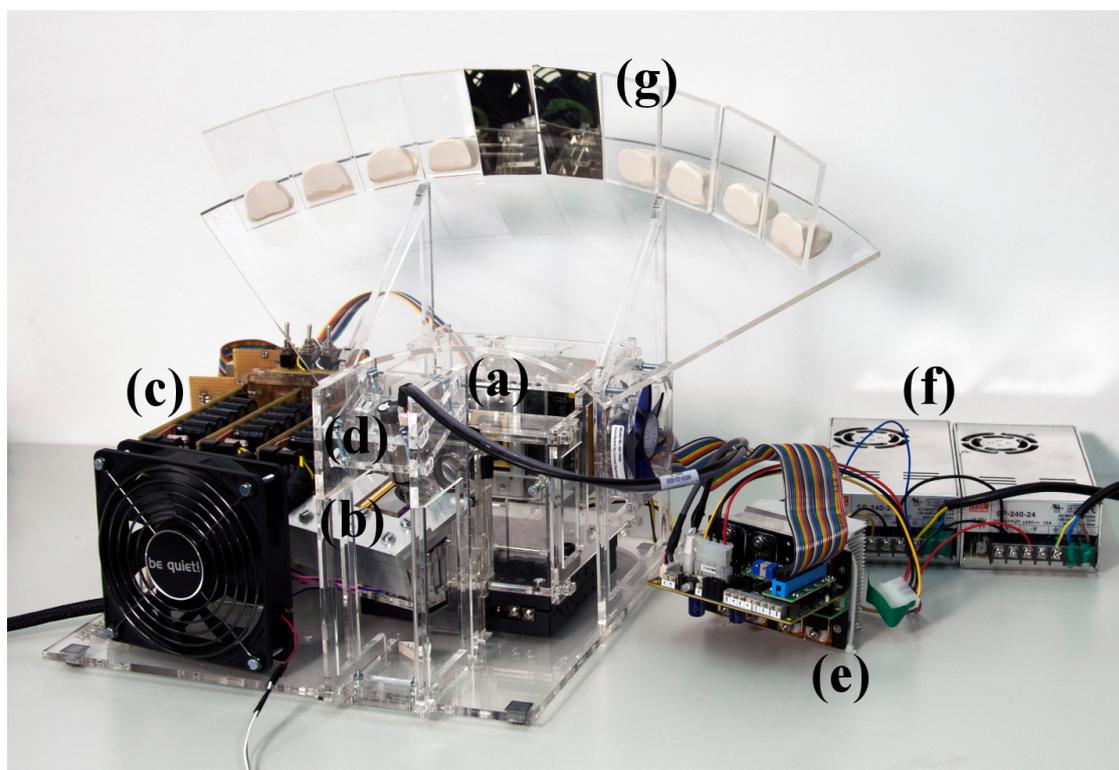
We use the DLP V-7000 platform from Vialux which has a DMD with a resolution of 1024 x 768 and refresh rate of 22kHz for binary patterns. Due to a missing link for real time image transmission to the platform, we download static images over USB to make use of the full bandwidth of the DMD. Ten distinct 24bit images are created and displayed at 60Hz each. Each individual image only requires a little more than 1ms to build up. Thus, either more images or higher frame rates would theoretically be possible, however the approach is limited by the speed of the galvonomic mirror.

The DMD board provides general purpose output pins which can be configured to trigger a signal whenever a binary image flip occurs. This trigger signal is be used to synchronize both LED and the galvonomic mirror position.

As galvo scanner we use the model 6220H from Cambridge Technology in combination with a beryllium mirror with a clear aperture of 10mm. It can be driven fast enough to re-orient to 10 different orientations at 60Hz which enough to multiplex the images generated by the DMD as described above.

The individual images are redirected a second time by ten fixed mirrors. Like this, only the fixed mirrors have to be moved to align the images on the projection screen. The complete setup for this light field projector is shown in Figure 5.10.

**Multiple light sources:** The galvonomic mirror is the slowest component in the system presented above. The DMD itself would be capable of providing

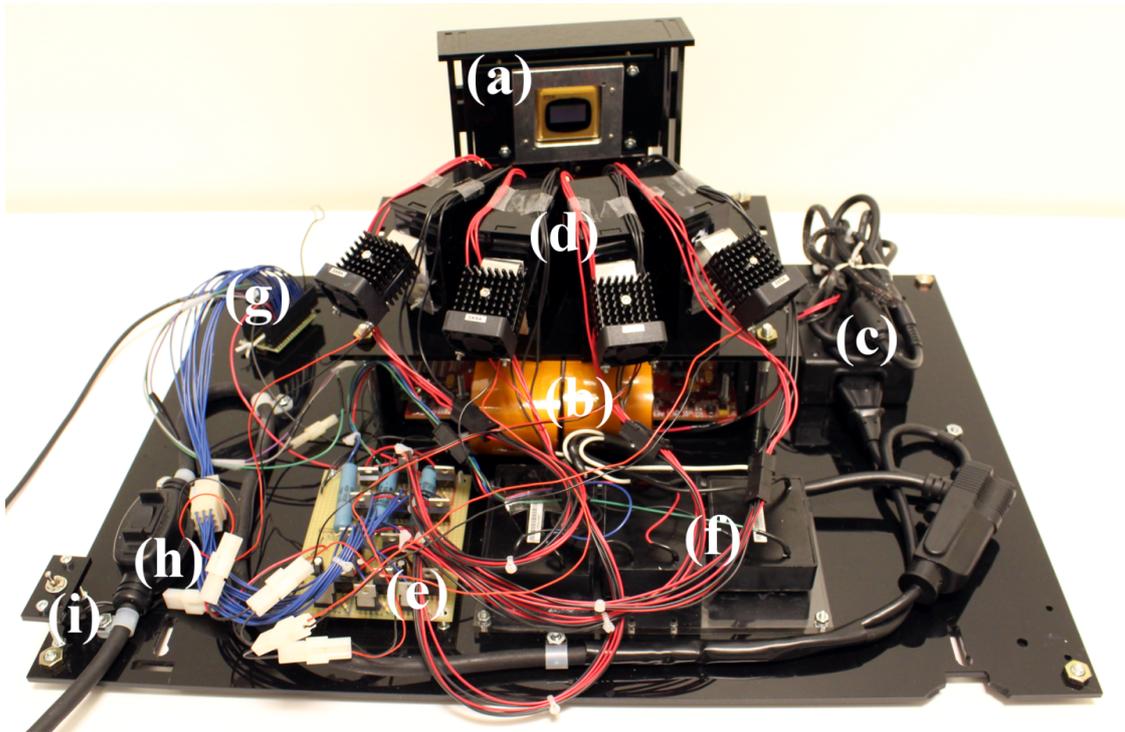


**Figure 5.10:** *Our setup for a light-field projector using a galvonic mirror to multiplex the individual images. The system consists of (a) the DMD, (b) the light source combining a red, a green and a blue LED using dichroic beam splitters, (c) the driving circuit for the LED, (d) the galvonic mirror, (e) the galvonic mirror driver, (f) the power supply for the setup and (g) the fixed mirrors to redirect the individual images. Only two out of the ten mirrors are coated in this image.*

twice as much bandwidth. However, the mirror cannot be repositioned fast enough to make use of it. Thus we have built a second prototype using multiple light sources instead of a mechanically moving part.

As light source we use red, green and blue laser diodes. Laser diodes become more readily available, also with increased optical power, and thus are suitable for projectors requiring high brightness. Furthermore, due to their small emissive area, they stay in focus even over large projection distances. This makes additional focusing optics obsolete which is another big advantage of this approach.

In this prototype we use the DLI4120 developer's kit using the 1080p DMD chip. The higher resolution comes at cost of a lower frame rate. Thus, we only create four distinct images. However, this still matches the total bandwidth



**Figure 5.11:** *Our setup for a light-field projector using multiple light sources consisting of (a) the DMD, (b) the DMD FPGA with (c) power supply, (d) four towers containing the lasers and optics, (e) the laser driver circuit and (f) laser power supply, (g) an FPGA for laser signal generation and power switches for the complete setup (h) as well as the lasers (i). The tower with fixed mirrors to redirect individual images to proper locations on the screen is not shown here.*

of the above described prototype. Synchronization is again done using the general purpose I/O pins of the DMD board.

To avoid dichroic mirrors and therefore increase optical throughput, each color channel of each image has its own optical path. Hence, a fixed mirror array of four times three is required to multiplex the three color channels of the four images generated by the DMD. The complete setup is shown in Figure 5.11. The mirror array is missing in this illustration.

### 5.2.3 Discussion

We have implemented two different designs for light field projectors using a single DMD chip. Advantage of our approach over a multi projector configuration is the better usage of the bandwidth of DMD, naturally synchronized

images as well as reduced system complexity. Both designs have advantages and disadvantages but work with the principle presented in [Nagano *et al.*, 2013]. Though the transparent holographic 1D diffuser could not have been built within the time of this thesis, we strongly believe that this approach can lead to a transparent automultiscopic 3D experience.

### 5.3 Beam Steered Light Field Display

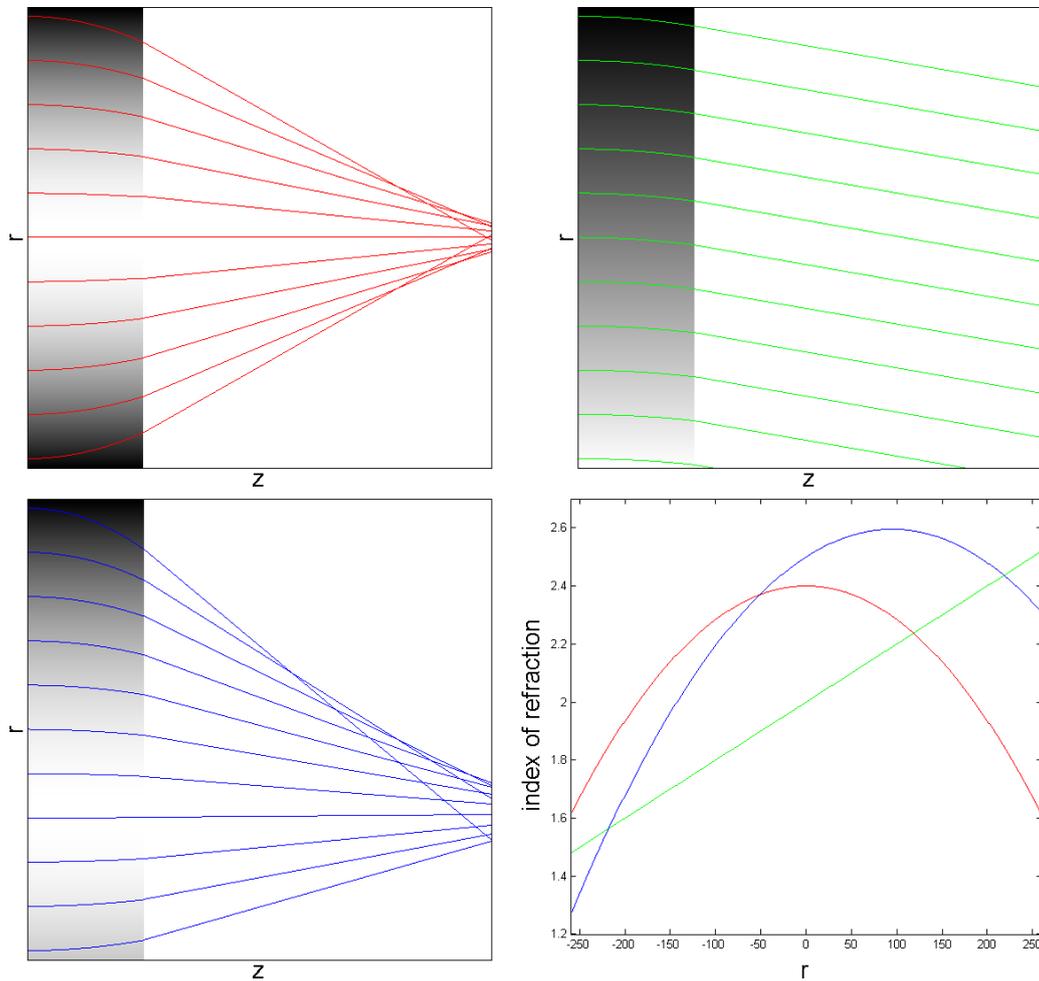
Major drawback of light field displays is the enormous bandwidth required to achieve an acceptable angular resolution. Usually, a significant amount of spatial resolution is traded in for some angular resolution [Dodgson, 2005]. Angular resolution is directly related to the depth which can be displayed aliasing-free [Chai *et al.*, 2000] and thus these displays can show content with only shallow depth [Zwicker *et al.*, 2006]. Therefore, current light field and multi-view displays are always unsatisfactory regarding either display resolution or perceived depth. Furthermore, if the bandwidth and thus the angular sampling is high enough, light field displays could even provide correct accommodation cues, as multiple rays generated by the display hit an eye's pupil and thus make it possible to focus on a given scene element. These are important reasons to find ways to significantly improve the bandwidth of such displays and not just in small steps as done in industries.

Fast switching image generators like DMDs, laser diode arrays or LED-matrix displays offer a different way to improve the bandwidth and thus the displayable depth of plenoptic displays. Instead of sacrificing spatial resolution, a certain number of consecutive frames are multiplexed to different directions over time. As such devices naturally work at high frame rates, bandwidth is extended without further cutbacks, an idea protected by [Ranieri and Gross, 2014a].

Therefore, we propose a system based on liquid crystal (LC) gradient index (GRIN) cells, which is able to temporally multiplex different images to different viewing angles. The basic idea is to focus and steer collimated beams of light over time to different points on a focal plane of a lenticular lens display for angle magnification, as outlined in Figure 5.12.

#### 5.3.1 Gradient Index Model

Fermat's principle can be used to describe the optical path of light through a medium with varying index of refraction. In the context of calculus of varia-



**Figure 5.12:** Simulations according to Equation 5.4 of light through GRIN cells with three different gradient index functions described by Equation 5.3. The top left plot corresponds to the traditional GRIN lens with linear term  $b$  equals zero. The top right plot shows how light can be redirected using the linear term  $b$  but with quadratic term  $a$  equals zero. A combination of both terms can describe a cell able to focus and steer a focal point as visible in the bottom left plot. The bottom right plot shows the used GRIN functions in the corresponding colors, which are also illustrated by the gray-scale gradient in each cell. Note, while the spread of light in the focal point degenerates with increasing steering angle it is reduced with increasing focal length. Thus, a large focal length is desired and an additional lenslet array in front is used for angle magnification (not shown here).

tions, Fermat's principle for the optical path length of such a ray traveling from  $A$  to  $B$  within such a GRIN cell can be written as

$$\delta S = \delta \int_A^B n(z,r) ds = 0 \quad (5.2)$$

$n(z,r)$  describes the function defining the index of refraction in a point  $(z,r)$  within the cell. [Jones *et al.*, 1988] use for  $n(z,r)$  a function with quadratic and constant term, which allows to describe a GRIN lens capable of focusing light on a point on the optical axis  $z$ . We extend this gradient index function by a linear term to

$$n(r) = a \cdot r^2 + b \cdot r + c \quad (5.3)$$

The calculus of variations for Equation 5.2 using Equation 5.3 leads to

$$\frac{2 \cdot a \cdot r + b}{a \cdot r^2 + b \cdot r + c} = \frac{r''}{r'^2 + 1} \quad (5.4)$$

This second order differential equation can be used to simulate the path of light traveling along the positive  $z$  axis through a GRIN cell with index of refraction defined by Equation 5.3. The required initial condition  $r_0$  and  $r'_0$  are the entry position of the ray into the cell and the incident angle respectively.

Based on Equation 5.4, Figure 5.12 shows simulations of light through three GRIN cells with different gradient index function. The corresponding GRIN functions are drawn in the rightmost plot. While the parameter  $a$  in Equation 5.3 influences the curvature in the GRIN function and thus the focusing behavior of the cell (red lines), parameter  $b$  can be used to redirect light (green lines). Combining both parameters in the same gradient index function allows to design a cell capable of focusing light also off axis in nearly any point in front of the cell (blue lines).

### 5.3.2 LC GRIN Cells

Liquid crystals can be designed and produced to have rod-like molecule structures with dipoles. Within an electric field, the molecules align their main axis with the field lines. In addition to that, liquid crystal materials are birefringent. Depending on the orientation of the molecules, linearly

polarized light passing through them will experience different refractive indexes.

Together, these two features can be used to realize a dynamically changeable gradient in index of refraction. The principle has been widely used in beam steering devices based on linear gradients [Xu *et al.*, 2009]. Recently, it has also been discovered by the display community to create switchable lenses based on quadratic gradients [Oka *et al.*, 2013; Takagi *et al.*, 2013]. Expectedly it can also be used to create combined linear and quadratic gradients as described in Section 5.3.1. Therefore, such a LC GRIN cell is capable to dynamically change and move a focal point in space.

A specific liquid crystal material can be characterized by its birefringence which is the ordinary and the extraordinary index of refraction  $n_o$  and  $n_e$ . Depending on the strength of the electric field, smooth refractive indexes between  $n_o$  and  $n_e$  can be achieved. However, there are some constraints induced by these material properties.

Values of the gradient function described by Equation 5.3 cannot exceed  $n_e$  and also not fall below  $n_o$ . Therefore, there is a lower limit on the achievable focal length, as not arbitrarily curved gradient functions can be realized. The shortest focal length without steering the beam is defined by the GRIN function with parameters

$$a = \frac{4 \cdot (n_o - n_e)}{\text{cellWidth}^2}, \quad b = 0 \quad \text{and} \quad c = n_e \quad (5.5)$$

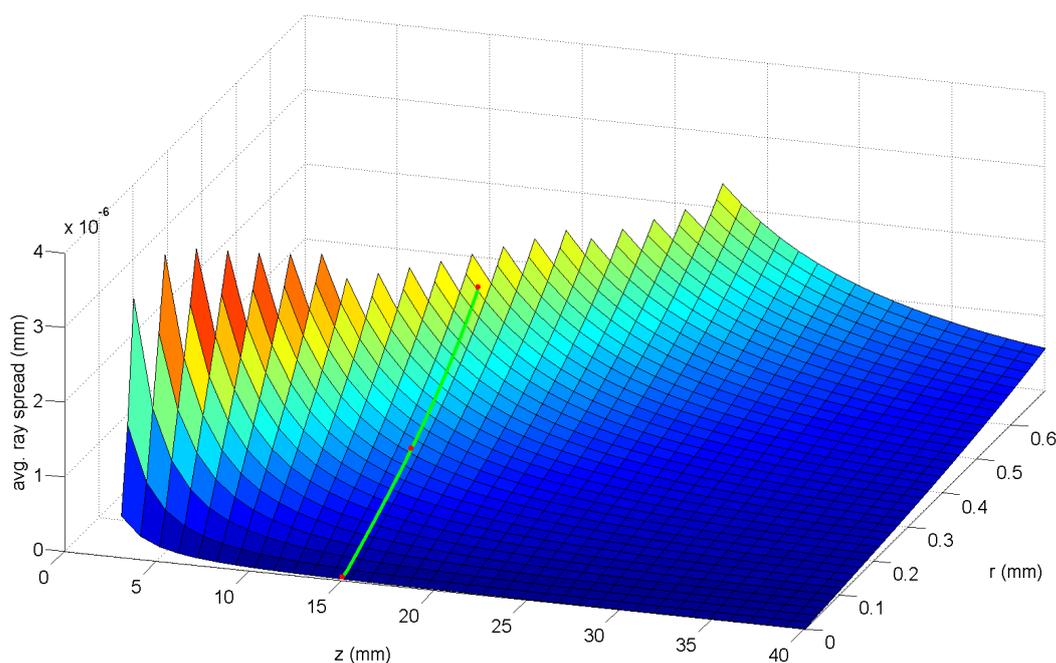
Also, as light is longer exposed to the gradient in thicker cells, shorter focal lengths can be achieved with cells of higher thickness. Similarly, the maximum steering angle without focusing the beam can be achieved with the parameters

$$a = 0, \quad b = \frac{n_e - n_o}{\text{cellWidth}} \quad \text{and} \quad c = \frac{n_e + n_o}{2}. \quad (5.6)$$

leading to larger steering angles achieved with smaller cell width.

Derived from these relations, a high cell strength and a small cell width would be desired to give maximum steering and focusing capability. However, discontinuous changes in the gradient function cannot be achieved by liquid crystal cells, as described by [Xu *et al.*, 2009]. The so called flyback region imposes a hard constraint: As it has been reported to be roughly as wide as the cell thickness, the cell must be wider than twice its thickness.

In addition, the cell thickness influences the strength of the electric field and thus cannot be chosen arbitrarily thick. Therefore, design parameters of such a LC GRIN cell have to be chosen carefully.



**Figure 5.13:** *The average spread of rays focused on a point  $(z,r)$  by a LC GRIN cell located at  $(0,0)$ . The parameters  $a$ ,  $b$  and  $c$  of the GRIN function described by Equation 5.3 have been optimized with a gradient descent method to minimize the plotted spread. LC GRIN cells cannot focus in all points due to the limitation imposed by the ordinary and extraordinary index of refraction of the liquid crystal material. There is a minimal focal length which is decreasing with increasing steering angle and thus, values in the upper left of the plot are missing. The spread of the rays increases with higher steering angle and decreases with increasing focal length. The green line indicates the focal plane chosen in the system design presented in Section 5.3.3. The ray intersections with the focal plane for the three points marked in red are shown in Figure 5.14.*

### 5.3.3 System Design

A single pixel of our proposed display system compromises three components. The first component is a light generator to create collimated beams of light. As images will be temporally multiplexed to different directions, the on/off time of the generator directly determines the angular frequency of the display and thus has to be high.

DMD offer high frame rates for binary patterns, which can be multiplexed to form gray-scale and color images. Even after this bandwidth reduction, they leave sufficient temporal resolution, which can be converted to enough angular resolution to compete with state of the art multi-view displays [Bogaert *et*

*al.*, 2010]. Alternatively, a color LED array can be used as image generator as it also offers good refresh rates. However, both need additional optical elements to collimate the emitted light, which would make manufacturing more difficult and expensive.

Laser diodes on the other hand have an extremely high switching speed and can emit collimated light. Furthermore they have a narrow color bandwidth, offer therefore a huge display color gamut and thus are the preferred embodiment of a light generator.

The second component of a pixel in the proposed display design is a LC GRIN cell as described in Section 5.3.2. For our simulations we have chosen a cell strength of 0.015mm and a cell width of 0.1mm which are common values in commercially available liquid crystal displays. Taking into account the flyback region of the liquid crystal material, the clear aperture of the cell and thus the required diameter of the laser diode is 0.07mm. The birefringence parameters of the liquid crystal material E7 are used for  $n_o$  and  $n_e$  in our simulations.

We evaluate the performance of these choices by means of simulations: Given a point  $(z, r)$  in front of the LC GRIN cell, we use a gradient descent method to optimize the parameters  $a$ ,  $b$  and  $c$  of Equation 5.3, such that when simulating a number of rays by Equation 5.4, the average spread of the focused rays around  $(z, r)$  is minimal.

Figure 5.13 shows the results of our simulations with above mentioned parameters for a number of points  $(z, r)$  and a LC GRIN cell located at  $(0, 0)$ . As can be seen, the spread of the rays and thus the focusing quality of the cell improves with increased focal length  $z$  but decreases with higher steering angle. Also visible in the plot is that the LC GRIN cell cannot focus on all points: For large steering angles and short focal lengths, the required gradient function would exceed the boundaries  $n_o$  and  $n_e$  imposed by the liquid crystal material. Note that the values are symmetric for positive and negative  $r$  and thus values for negative  $r$  are omitted.

Based on this evaluation, a proper focal length can be chosen. The longer it is, the smaller the overall ray spread and the larger is the steering range on  $r$ , but also the thicker the display will be. The ray spread determines how many different angular pixels can be generated per spatial pixel and thus determines the displayable depth. The steering range must be at least as large as the cell width and determines the spatial resolution of the display. In our design we have chosen a focal length of 15mm, marked by a green line in Figure 5.13, which is sufficiently long to allow for a steering range of  $\pm 0.4mm$ , leading to a spatial resolution of 0.8mm per pixel.

As the steering angle achieved by LC GRIN cells typically is small, an optical element for angle magnification is used to cover the whole field of view of the display. For this purpose, a diffuser is placed at the chosen focal length and a lens placed in front of it as third component of a display pixel. The focal length of the lens can be chosen similar to the design choices of integral imaging displays. Longer focal length lead to smaller field of views, more view repetitions but also a denser angular sampling compared to shorter focal lengths. Hence this parameter has to be chosen depending on the application of the display and is not discussed further here. Note, as the focal point created by the LC GRIN cell can be moved freely through space, the magnification lens does not need to have a flat focal plane. Therefore, special lenslet with better optical quality can be designed and used for angle magnification.

### 5.3.4 Bandwidth

[Chai *et al.*, 2000] have shown a direct relation between angular sampling and depth of a light field and [Zwicker *et al.*, 2006] have applied this theory to displays. They show that content displayed at full spatial resolution suffers from aliasing when exceeding a certain depth and thus has to be filtered and blurred. The aliasing-free depth is thus a direct quality measurement of light field displays and given by

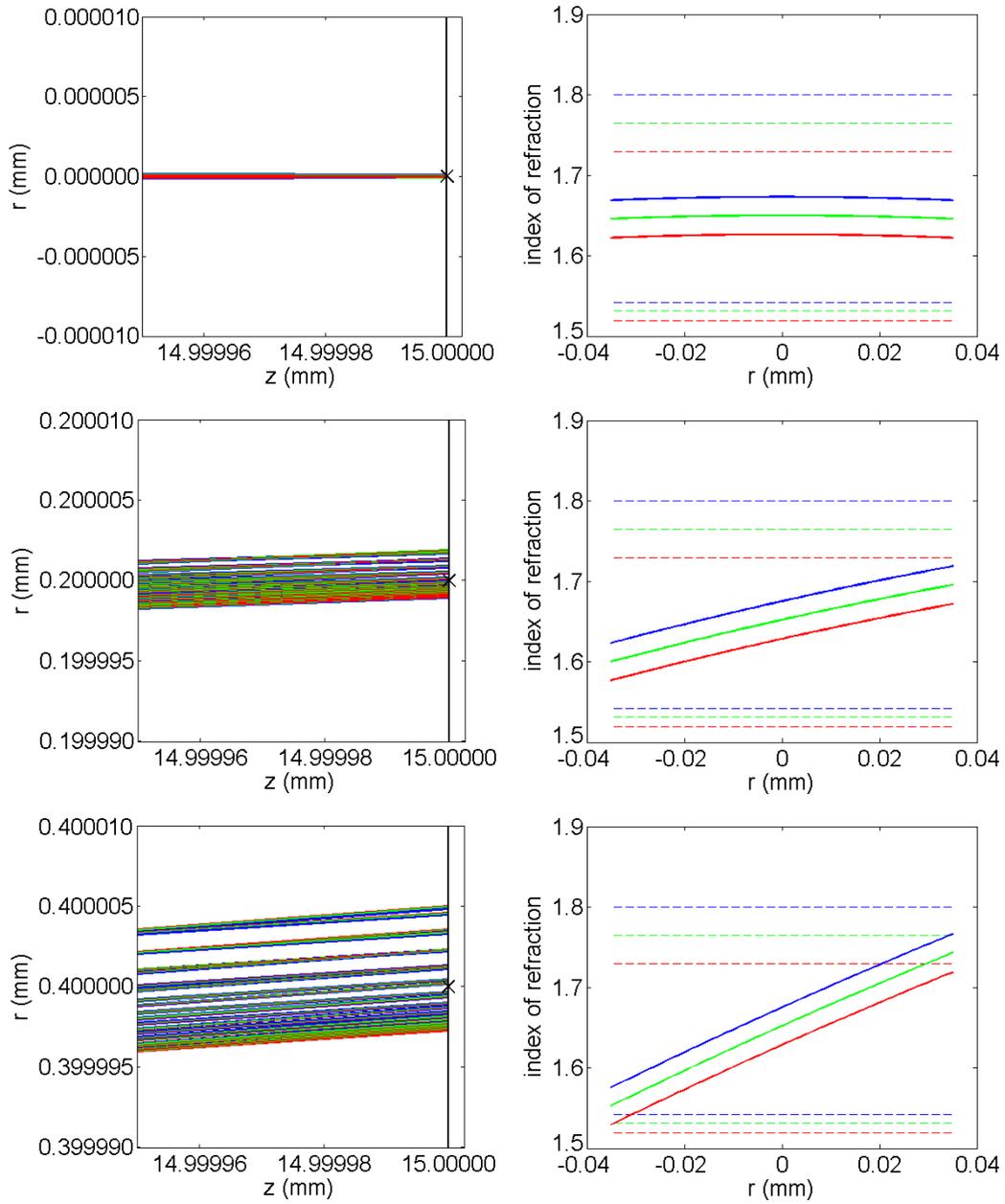
$$|z| < \frac{f \cdot \Delta t}{\Delta v} \quad (5.7)$$

$\Delta t$  is determined by the spatial resolution which is in our system the lenticular lens pitch and chosen to be 0.8mm.  $\Delta v$  is the angular pixel pitch and  $f$  the focal length of the lenslet. To determine  $\Delta v$ , we evaluate the theoretically achievable angular pixel pitch using our proposed system parameters and divide by the focal length of the lenticular lens.

The angular pixel pitch is directly determined by the ray spread of the focused light on the focal plane. To avoid crosstalk, the distance between two focused beams must be larger than the ray spread. Thus, the maximum ray spread is an upper bound on the angular pixel pitch for a zero crosstalk configuration.

Note, as the sweep of the focused light over the focal plane is continuous, the bandwidth analysis can be performed using a convolution kernel rather than the discrete approach proposed here. However, for simplicity and as it is sufficient for a lower bound of the display performance we just take the maximum ray spread as angular pixel pitch.

### 5.3 Beam Steered Light Field Display



**Figure 5.14:** Plots of the intersections of rays simulated through a LC GRIN cell with the focal plane of a lenticular lens are shown in the left column. The parameters  $a$ ,  $b$  and  $c$  of Equation 5.3 have been optimized for three different wavelengths for red (635nm), green (510nm) and blue (450nm) lasers and the corresponding gradient functions are shown in the right column. The dashed lines indicate the ordinary and extraordinary index of refraction  $n_o$  and  $n_e$  of the LC GRIN cell for each color. While different wavelengths experience different refractive indexes, optimizing for each one individually results in minimal color aberrations on the focal plane. Maximum ray spread occurs at the largest steering angle in the right plot and amounts less than 10nm for the configuration presented in Section 5.3.3.

To get an estimate of the maximum ray spread, the GRIN function parameters for several points on the chosen focal plane are optimized. Three of these points are marked by red dots in Figure 5.13. Their optical path close to the focal plane are plotted in Figure 5.14, together with the corresponding optimized GRIN function for three colors with wavelength 450nm, 510nm and 635nm, corresponding to common laser colors.

The ray spread is maximal at the largest steering angle and smaller than 10nm. To cover a field of view of  $\pm 45^\circ$ , a lenticular lens with focal length equals 0.4mm is required. Hence, when substituting into Equation 5.7, our exemplary display design can achieve an aliasing-free displayable depth of  $\pm 32m$ .

Limited by the ray spread of 10nm, the proposed design is theoretically capable to put 80000 different views below one lens in contrast to the 28 views found in current multi-view displays. Through the linear relation between number of angular views and displayable depth, the proposed design can show significantly more depth than other spatially multiplexed multi-view displays. To generate these views, the light generator must be able to switch at 4.8 megahertz which is possible with laser diodes. The LC GRIN cell on the other hand has to be driven at a refresh rate corresponding to the frame rate of the display, which is usually only 60Hz or 120Hz for desktop screens and achievable with common liquid crystal materials.

Thus, the bandwidth of the display is not limited any more by the display technology itself but by other aspects as e.g. the transmission bandwidth or availability of content. This gives the proposed design a clear advantage over other light field and multi-view approaches.

### 5.3.5 Discussion and Application

The simulation of a sample configuration presented in Section 5.3.3 yields promising results. Beams of light can be focused on an area less than 10nm and thus 80000 different views can be generated below each lens. For a field of view of  $\pm 45^\circ$ , content at up to  $\pm 32m$  depth can be shown aliasing free. Furthermore, the display will provide correct accommodation cues, as the angular sampling is dense enough [Takaki, 2006]. Therefore, the display can compete with the comfort zone of stereoscopic displays as well as with displayable depth of existing multi-view and light field displays.

To make the display attractive for the consumer market, it should be able to display already available stereoscopic content. For this purpose, an eye tracker can be attached to the display in a first application. While the LC

### 5.3 Beam Steered Light Field Display

GRIN cell is sweeping the focal point on the focal plane back and forth, colors from the left image or right image are displayed depending on the eye located in the direction corresponding to the current focal point. Due to the dense angular sampling, conflicts between different viewers are very unlikely. Thus, this display version can replace all glasses based stereoscopic home entertainment systems without any drawbacks.

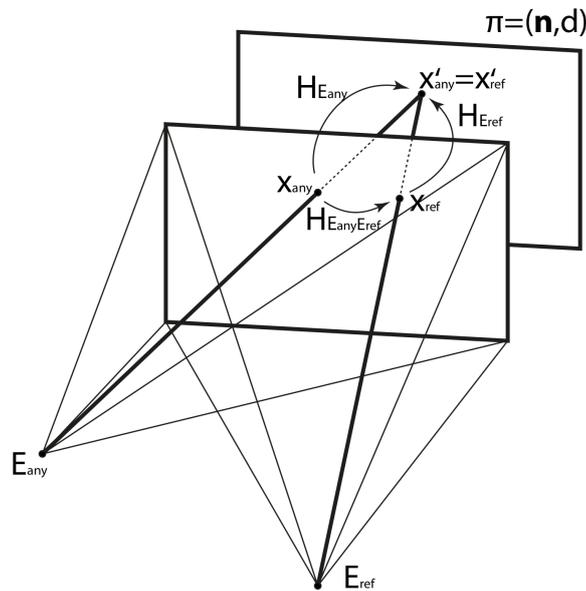
In a second version, view interpolation can be done between left and right images of stereoscopic content [Schaffner *et al.*, 2013]. In this way, viewers experience the illusion to be able to look around displayed objects, while still only stereoscopic content is required.

Instead of left and right image, more views can be used as input and view interpolation can be applied to approximate the light field. With every improvement on transmission bandwidth, the display can receive more accurate approximations of light field data. In a last and most advanced version, the full power of the display is used by rendering all the views or the continuous plenoptic function at full bandwidth of the display.

## *Lightfield Displays*

**Vision Based Calibration**

---



**Figure 6.1:** *The mapping  $H_{E_{any}}$  from barrier layer pixel to image layer pixel (or vice versa) can be computed for any eye position  $E_{any}$  using the mapping  $H_{E_{any}E_{ref}}$  defined in Equation 6.10, given a known reference eye position  $E_{ref}$ , known plane parameters of the image layer  $\pi = (\mathbf{n}^T, d)$  and known mapping for the reference eye position  $H_{E_{ref}}$ .*

---

In the previous chapters we have presented different advanced 3D display systems. While many stereoscopic displays consist of a single layer, simple homographies can be used to rectify and align the displayed content. The other systems consist of multiple layers, each one planar but with unknown spacing and orientation to each other. Therefore a method is required, to map pixels of one layer to corresponding pixels of other layers forming together a line which intersects a certain eye position. In this section we propose a vision based calibration method achieving exactly this. We demonstrate the idea on a parallax barrier display consisting of two layers but it easily can be generalized to any number and type of layers.

Stereoscopic parallax barrier displays deploy an image layer with two scrambled views and a barrier layer to multiplex the two views to distinct eye positions. A special translucent/opaque pattern is shown on the barrier layer, which exposes pixels of one view to one eye but blocks their sight to the other eye and vice versa. Careful alignment of barrier pattern and scrambled image pattern is required and crucial to avoid crosstalk and aliasing.

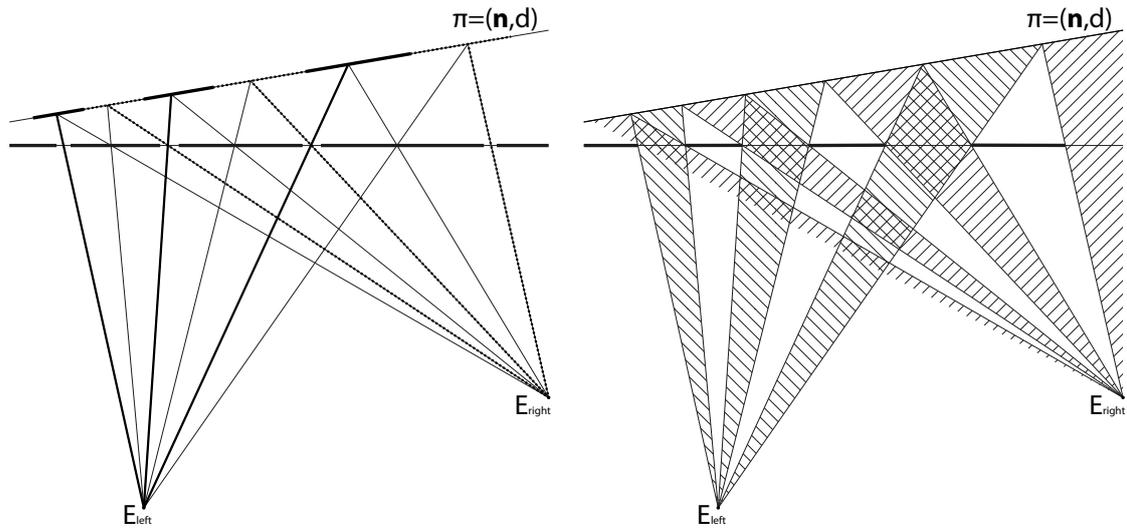
Furthermore, computation of the barrier pattern in existing methods often relies on known display geometry. Imperfections and misalignment caused by the manufacturing process can thus lead to Moiré patterns, crosstalk and other artifacts. Also, the iterative nature common to these methods induces a growing latency with increasing display resolution.

Perlin et al. [Perlin *et al.*, 2000] use known display geometry to compute the positions of the barrier slits and the positions of pixels for the left and right view, given the location of an eye pair. They are tracing rays from one eye to barrier pixels to get corresponding pixels on the image layer and vice versa with the other eye. This well known approach giving an optimal barrier pattern is iterative: The preceding barrier position has to be known to compute the following one.

Sakamoto et al. [Sakamoto *et al.*, 2005] use the very same procedure to determine corresponding positions on the barrier layer and the image layer. But instead of using a barrier slit pattern acting as pinhole, they use the whole area between two barrier pattern positions as clear or blocking patch.

For both approaches, refresh rate drops and latency increases with higher display resolution due to the iterative nature of the algorithm. In addition, inaccurate knowledge about the display geometry can impose challenges: Tiny aberrations caused by the manufacturing process lower the quality of the displayed stereo content which becomes even more severe for high resolution displays with small pixel size.

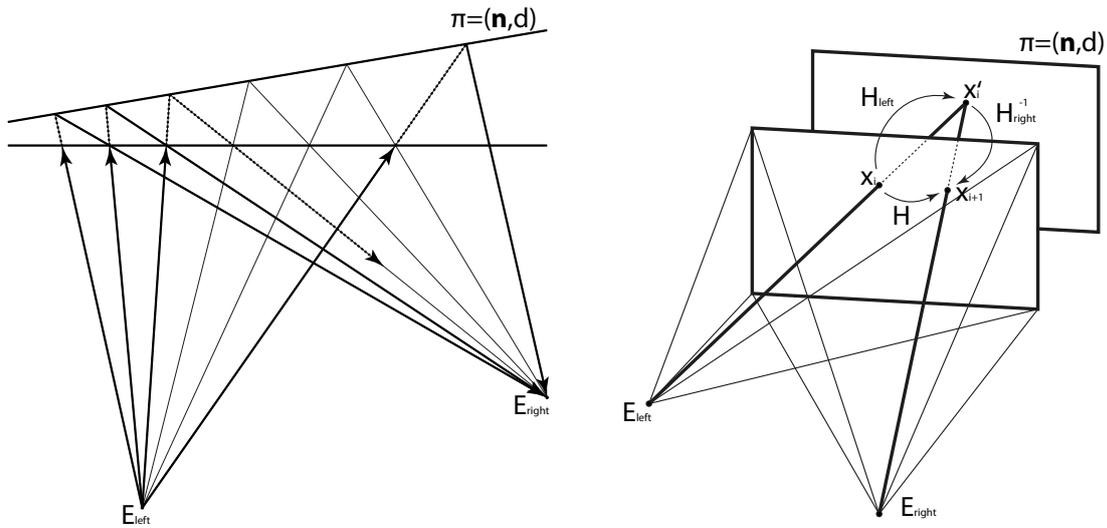
A computer vision based approach was used in the work of Annen et al.



**Figure 6.2:** *Illustration of two optimal barrier patterns. On the left side, barrier positions are used to define slits acting as pinhole, exposing left view's pixels (dashed lines) to the left eye  $E_{left}$  and the right view's pixels (solid lines) to the right eye  $E_{right}$  only. On the right side, the barrier positions define the border of alternating blocking and translucent barrier patches, providing much more brightness. Latter approach can further be improved by swapping the translucent and opaque patches each alternating frame, providing full spatial resolution.*

[Annen *et al.*, 2006] to calibrate a multi-view parallax barrier setup: They place a camera in front of the setup and capture images to compute for each barrier slit the corresponding pixels for one view point. Interpolation is then used to compute the corresponding pixels for novel views. The approach is very suitable to compute the scrambling pattern of a multi-view display but would require further adaptations for a free viewpoint stereoscopic display.

The advantage of such vision based methods is that they do not require knowledge about the display geometry. They do scale well with the resolution of the camera sensor, which is usually a multiple higher than the resolution of a display. Furthermore, the use of projective transformations allows the image layer to be a projected image with keystone instead of a perfectly rectangular screen. Thus, we present a computer vision based approach to calibrate any multi-layered display setup [Ranieri and Gross, 2014b].



**Figure 6.3:** Two computation methods for an optimal barrier pattern. On the left, rays are traced from one eye through a barrier layer position onto the image layer and back to the other eye, resulting in the next barrier position. The same can be achieved using one homography for each direction as illustrated on the right.  $H_{left}$  maps any barrier pixel to its corresponding image pixel as seen by the eye  $E_{left}$ . Similarly,  $H_{right}^{-1}$  maps image pixels to barrier pixels as seen by the eye  $E_{right}$ . Concatenating both homographies results in a homography  $H$ , which can be used to determine all barrier positions both at the top and the bottom border of the screen, using Equation 6.3.

## 6.1 Homography-based mapping

To compute an optimal barrier pattern for a given viewer, a relation between image layer and barrier layer has to be computed for each eye position. The relation must map a barrier pixel to the image pixel which is seen by the eye when looking through the barrier pixel and vice versa.

An optimal barrier slit pattern as computed by Perlin et al. [Perlin *et al.*, 2000] is shown in Figure 6.2 on the left. The barrier slits are as close together as possible without exposing left eye pixels to the right eye or the other way around. Exactly the same barrier positions can be used to create an optimal stripe pattern as used by Sakamoto et al. [Sakamoto *et al.*, 2005], which is illustrated in Figure 6.2 on the right.

Figure 6.3 left hand side shows a geometrical interpretation of the approach proposed by Perlin et al. [Perlin *et al.*, 2000] on how to compute the required positions using ray tracing. To compute a new barrier position, a ray from left eye to the preceding barrier position is intersected with the image layer.

Another ray from the intersection to the right eye intersects the barrier layer in the new barrier position. The same relation can also be expressed by homographies, a 3-by-3 matrix defining a mapping between two planes with respect to a center of projection.

For known homography  $H_{E_{left}}$  with center of projection in the left eye  $E_{left}$  and any pixel position  $x_i$  on the barrier layer, the corresponding pixel  $x'_i$  on the image layer can be computed by

$$\mathbf{x}'_i = H_{E_{left}} \cdot \mathbf{x}_i \quad (6.1)$$

as illustrated in Figure 6.3 on the right side. The iteration presented by Perlin et al. [Perlin *et al.*, 2000] can thus be expressed as

$$\mathbf{x}_{i+1} = H_{E_{right}}^{-1} \cdot H_{E_{left}} \cdot \mathbf{x}_i \quad (6.2)$$

using the inverse mapping  $H_{E_{right}}^{-1}$  for the right eye. The complete iteration for the  $i$ -th barrier position is given by

$$\mathbf{x}_i = H^i \cdot \mathbf{x}_0 \quad (6.3)$$

with

$$H = H_{E_{right}}^{-1} \cdot H_{E_{left}} \quad (6.4)$$

and  $\mathbf{x}_0$  some initial barrier point. Similarly, the image positions can be computed by

$$\mathbf{x}'_i = H_{E_{left}} \cdot H^i \cdot \mathbf{x}_0 \quad (6.5)$$

Thus, for each viewer position it is sufficient to know  $H_{E_{left}}$  and  $H_{E_{right}}^{-1}$  to compute all required barrier and image positions.

## 6.2 Closed Form Solution

Computing  $H^i$  in Equation 6.3 and Equation 6.5 would involve an iteration which is especially unsuited for parallel architectures such as GPUs and

would decrease refresh rate and increase latency for parallax barrier updates. Thus we propose to decompose  $H$  into

$$H = V \cdot D \cdot V^{-1} \quad (6.6)$$

using the eigendecomposition.  $V$  is a 3-by-3 matrix consisting of the eigenvectors of  $H$  and  $D$  is a diagonal matrix containing the corresponding three eigenvalues  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$ . The  $i$ -th power of  $H$  can now be rewritten as

$$H^i = \underbrace{\underbrace{V \cdot D \cdot V^{-1}}_H \cdot \underbrace{V \cdot D \cdot V^{-1}}_H \cdot \dots \cdot \underbrace{V \cdot D \cdot V^{-1}}_H}_{i \times} \quad (6.7)$$

As  $V^{-1} \cdot V$  cancels out, this leads to the more compact form

$$H^i = V \cdot D^i \cdot V^{-1} = V \cdot \begin{bmatrix} \lambda_1^i & 0 & 0 \\ 0 & \lambda_2^i & 0 \\ 0 & 0 & \lambda_3^i \end{bmatrix} \cdot V^{-1} \quad (6.8)$$

Equation (6.8) can be computed without the need of costly iterations. In combination with Equation (6.3) and Equation (6.5), any barrier or image layer position can be computed independently of other positions and in constant time, both important characteristics for a fast GPU accelerated implementation.

### 6.3 Free View Transform

A specific eye position  $E_{ref}$  together with the barrier layer can be seen as a virtual pinhole camera, where the eye is the center of projection and the barrier layer the virtual image plane. Together with any second eye positions  $E_{any}$ , a virtual stereo camera pair with shared camera image plane is defined, illustrated in Figure 6.1. As described in [Hartley and Zisserman, 2000, p. 325], the display image plane induces a homography  $H_{E_{any}E_{ref}}$  between these two virtual cameras. This means that the homography  $H_{E_{any}E_{ref}}$  relates two points, one corresponding to each camera, whose rays would share the intersection point on the display image plane, also shown in Figure 6.1.

Assuming that the homography  $H_{E_{ref}}$  relating barrier pixels to display image pixels for some reference eye position is known, the homography  $H_{E_{any}}$  for any other eye position can be computed by

$$H_{E_{any}} = H_{E_{ref}} \cdot H_{E_{any}E_{ref}} \quad (6.9)$$

According to [Hartley and Zisserman, 2000], the homography  $H_{E_{any}E_{ref}}$  can be computed by

$$H_{E_{any}E_{ref}} = K_{E_{ref}}(R - \mathbf{t}\mathbf{n}^T/d)K_{E_{any}}^{-1} \quad (6.10)$$

As the virtual cameras share the same image plane, namely the parallax barrier plane, the relative camera rotation  $R$  is the identity matrix. The relative translation can be computed by

$$\mathbf{t} = E_{ref} - E_{any} \quad (6.11)$$

where  $E_{any}$  is the given input position of the eye and  $E_{ref}$  the position of the reference camera computed during calibration.  $K_{E_{ref}}$  and  $K_{E_{any}}^{-1}$  denote the intrinsic parameters of the virtual cameras and can be computed by simply putting the eye center into the projection matrix form

$$K_{E_{ref}} = \begin{bmatrix} -E_{ref}.z & 0 & E_{ref}.x \\ 0 & -E_{ref}.z & E_{ref}.y \\ 0 & 0 & 1 \end{bmatrix} \quad (6.12)$$

Note, the negative sign comes from the positive z-axis being the viewing direction in the used coordinate system. Similarly, the inverse projection matrix can be derived by

$$K_{E_{any}}^{-1} = \begin{bmatrix} -\frac{1}{E_{any}.z} & 0 & \frac{E_{any}.x}{E_{any}.z} \\ 0 & -\frac{1}{E_{any}.z} & \frac{E_{any}.y}{E_{any}.z} \\ 0 & 0 & 1 \end{bmatrix} \quad (6.13)$$

The remaining unknowns  $\mathbf{n}$  and  $d$  in Equation (6.10) are the plane parameters of the display image layer

$$\pi = (\mathbf{n}^T, d) \quad (6.14)$$

which will be computed during calibration. As for Equation (6.10) the eye position  $E_{any}$  needs to be at the origin, the plane parameters might need to be shifted accordingly. With this information, all the required relations can be computed by replacing  $E_{any}$  by  $E_{left}$  and  $E_{right}$  for any viewer position.

Noticeable is, that the homography  $H_{E_{ref}}$  can consist of any combination of other homographies, including e.g. the projection and keystone of a projected image. This is useful for projector-based parallax barrier displays as the one given by Sakamoto et al. [Sakamoto *et al.*, 2005].

Hence, if the plane parameters  $\pi$  and the homography  $H_{E_{ref}}$  can be calibrated for one reference position  $E_{ref}$ , then the homographies for any new eye position can be easily computed using Equation (6.9) and Equation (6.10). These derived homographies for left and right eye can then be used in Equation (6.4) to compose the required homography for the barrier computation.

### 6.4 Calibration

In standard camera calibration procedures, images of a planar calibration pattern with known pattern coordinates are taken from different camera positions [Zhang, 1999]. For each image, a homography is computed, relating the coordinates in the calibration pattern to image coordinates of the camera sensor. Based on this, intrinsic and extrinsic camera parameters can be computed. Intrinsic parameters contain camera characteristics as focal length and principal point and extrinsic parameters describe the position and orientation of the camera for each image.

To calibrate a parallax barrier display, the same algorithms can be used. One calibration pattern is shown on the barrier layer and one on the image layer, both being visible from different viewing positions at the same time. Example of such a calibration pattern as used in this work is shown in Figure 6.4 in the bottom image. Similar to camera calibration, pictures from different view positions are taken and the pattern on the barrier layer is used to compute the extrinsics of the camera. This gives for each image the corresponding center of projection, expressed in pixel coordinates relative to the pixel positions of the barrier layer. These positions can be interpreted as candidates for the reference eye position  $E_{ref}$  used in Section 6.3.

Also, using the known calibration pattern corner coordinates and their corresponding coordinates detected in the camera image, one homography that maps barrier layer pixels to camera sensor pixels and one homography that maps camera sensor pixels to display image layer pixels can be computed, using the standard direct linear transform method [Hartley and Zisserman, 2000, p. 90]. Multiplying both matrices results in the homography that maps directly barrier pixels to display image pixels for each camera position. Hence, each such combined homography is candidate for  $H_{E_{ref}}$  required in Section 6.3.



**Figure 6.4:** *Our transparent stereoscopic display prototype once without displaying anything to illustrate transparency (top) and once showing the calibration pattern that was used in our calibration algorithms (bottom). A RGB liquid crystal display is used as barrier layer losing much transparency in the embedded color filters. Using a gray-scale liquid crystal display as deployed in e.g. medical screens would improve transparency significantly.*

The same homographies are also used to compute the plane parameters  $(\mathbf{n}, d)$  of the display image layer. For each camera center the corresponding mapping from image layer pixels to barrier layer pixels is applied to a single image layer pixel. This gives a number of barrier layer pixels which define together with their corresponding camera center a set of rays. These rays share a common intersection point on the physical image layer which can be found using known least squares methods. For three or more of such points, the plane parameters  $\pi = (\mathbf{n}^T, d)$  expressed in pixel coordinates relative to the barrier layer can be fitted.

As the homographies are only used to find corresponding pixels on the barrier layer, the image of the display image layer does not have to be rectangular and can e.g. be a projection with keystone or warped by any other homography. Also, the intrinsics derived during camera calibration are not used at all. Only the radial distortion parameters are used to undistort the images before any step of our algorithms.

The camera center closest to the preferred viewing position can be used as reference eye position.  $E_{ref}$ ,  $H_{E_{ref}}$  and  $\pi = (\mathbf{n}^T, d)$  can be further improved with a non-linear optimization, minimizing the re-projection error of Equation 6.10 when applied to all the other cameras and detected calibration pattern corners.

## 6.5 Implementation

For the implementation of our calibration algorithm we use the libraries ARToolKitPlus [Wagner and Schmalstieg, 2007] and OpenCV. ARToolKitPlus provides a convenient way to automatically detect virtual markers and extract their corners in our captured images. These initial guesses are then refined using sub-pixel accurate corner detection algorithms provided by OpenCV. Also, the inbuilt camera calibration and homography computation functions of OpenCV were used to get the camera centers and corresponding pixel mapping homographies.

We further use Matlab to fit the plane parameters and for a non-linear optimization of the reference camera parameters. As reference eye position we choose a virtual one positioned at the centroid of the viewing volume. We first compute it's homography based on the proposed method using the closest camera position. We then use a non-linear, gradient descent method to further optimize this initial guess of  $H_{E_{ref}}$ . For this purpose, we compute for each captured camera image the mapping homography based on our reference camera using Equation 6.10. Then, some barrier pixels are mapped

to image pixels, once using this homography and once using the homography derived during calibration. The summed distance between corresponding image pixels is used as error metric in the gradient descent method, as it optimally should be zero.

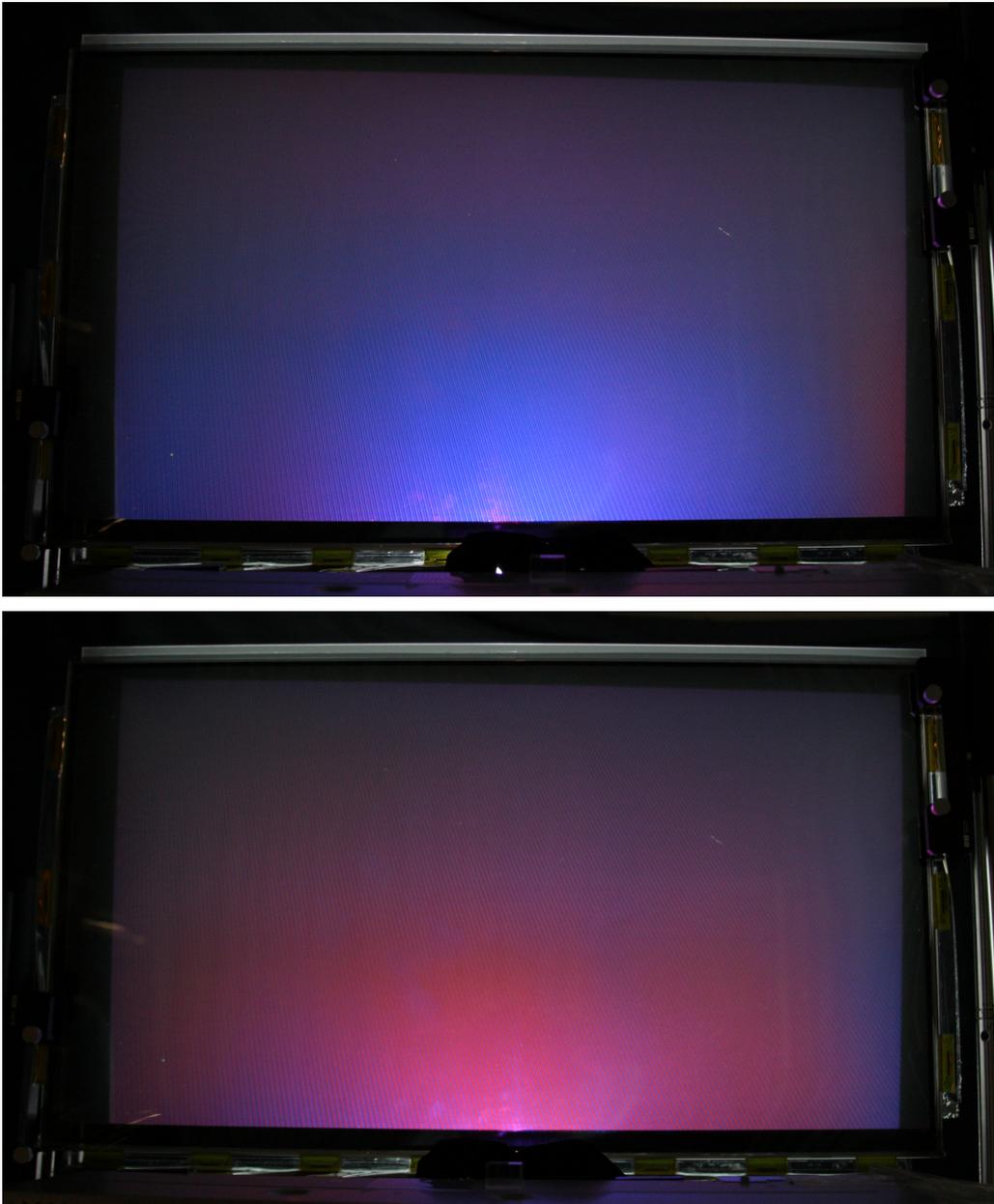
Our real-time barrier renderer is based on DirectX and also uses OpenCV. Each barrier stripe is represented as quad, with the  $y$ -coordinate as screen coordinate and  $x$  as barrier index. The vertex buffer for the barrier has to be uploaded to GPU only once as it remains constant. In each frame, the input eye coordinates are used to compute the homography  $H$  as described in Section 6.1 and Section 6.3. OpenCV is then used for the Eigendecomposition of  $H$ . We use the pseudo-eigendecomposition to avoid complex eigenvalues which will save GPU operations. In each frame,  $H_{E_{right}}^{-1}$ ,  $V$ ,  $V^{-1}$  and  $D$  are uploaded to the GPU shader. In the shader, these informations and the index of each vertex found in the  $x$  coordinate are used with Equation 6.8 to compute  $H^i$  and thus the new barrier position or, together with  $H_{E_{right}}^{-1}$ , the image position.

## 6.6 Physical Prototype

To demonstrate the potential of our algorithms we implemented a simplified version of the back-projected transparent stereoscopic screen suggested by Sakamoto et al. [Sakamoto *et al.*, 2005]. The isotropic back-projection screen ST-Professional-Trans from Screen-Tech® was used in combination with a BenQ SH910 Projector to create a transparent image layer which preserves polarization. As barrier layer we use an Acer HN274H with removed diffusing polarizer. The barrier layer has a theoretical transparency of only 16% as only a third of the light passes the color filters and further half the light is lost in the polarizer. As the barrier layer does not require colors, the system's transparency could further be improved by using a gray-scale screen as available e.g. in medical high-contrast displays. Instead of using two projectors with two polarizations as proposed in [Sakamoto *et al.*, 2005] we only use one projector and time multiplexing to swap the barrier pattern in each alternating frame.

A polarizer in front of the projector is used to create two polarized scrambled views on the image layer. The light then passes the twisted nematic liquid crystal barrier layer, and is, depending on the rotating state of the liquid crystal, either blocked by a polarizer in front of the setup or transmitted to the viewer.

Both image and barrier layer provide FullHD resolution and are separated by



**Figure 6.5:** Results captured on our prototype using the barrier pattern illustrated in Figure 6.2 on the right and calibrated with the proposed method. Left eye's view (top) and right eye's view (bottom) show the image separation. Time multiplexing for alternating transparent/opaque patches was applied to regain full spatial resolution. Crosstalk at the right border of the display is due to the slightly non-planar projection surface. The strong radial fall-off comes from the projector and is not subject of our work.

an average spacing of 2.5cm. The average barrier width at a viewing distance of one meter is 5 pixels. The back-projection screen is not co-planar to the barrier layer, the projected image is slightly rotated and suffers from keystone distortion. All these geometric unknowns are addressed by our algorithms. The complete setup is shown in Figure 6.4.

### 6.7 Generalization and Discussion

To assess quality of our calibration we use the pixel distance between calibrated homography and homography computed by our algorithms as described in Section 6.5. Sub-pixel accuracy in display pixel space has been achieved for completely planar barrier and image layer. We have observed an increasing error for slightly bent or non-planar layers as they do not fit the assumption of planar surfaces in our algorithm.

Our implementation is able to run at 120Hz for FullHD on a state of the art desktop graphics card, tested with an opaque parallax barrier display. The 60Hz refresh rate (limited by the projector) in our presented prototype could be achieved without any problems. The crucial Eigendecomposition can be computed in less than a millisecond due to the small and fixed size of the homography. Barrier computation is as fast as rendering at most 1920 quads on GPU for full HD.

Results captured on our prototype are shown in Figure 6.5. The two images show left and right eye's view with clear color separation. Slight crosstalk is visible at the right border of the screen, coming from the slightly non-planar back-projection screen.

The whole algorithm has been derived with respect to parallax barrier displays. However, it can easily be extended to any other multi-layered display system. As plane parameters and homographies for all layers and any possible eye positions are derived during calibration, the method can be used in algorithms other than parallax barrier computation. This makes the proposed approach applicable to many different advanced 3D display designs.



---

# C H A P T E R

# 7

## Conclusion

In this work we tried to push the boundaries of 3D display technologies. We first grouped displays into three different classes. For each class we then decided how this kind of display can be improved or what features can be added to make the viewing experience more immersive.

We improved stereoscopic displays by adding transparency as a feature. Further we analyzed the usability of state of the art eye tracking to include motion parallax in our transparent display systems. The whole system is used in a telecommunication platform to prove validity of the approach. Also we proposed a large-scale system based on a retroreflective material capable of creating autostereoscopic 3D for multiple viewers.

For volumetric displays we created a generic framework which can be used to describe a whole variety of multi-layered volumetric displays. The framework uses display primitives as light emitters and light attenuators which make up most of the current state of the art displays. We then presented a light field decomposition algorithm dividing different contributions of the plenoptic function of a scene to the different layers. Our proof of concept prototypes are capable to show view dependent effects and handle occlusions.

Also, we performed a bandwidth analysis on multi-layered light field displays. Using knowledge of the dimensionality gap in light fields we proposed a system to linearly reduce the required bandwidth with the number of layers. Then we extended the idea of our transparent stereoscopic display to be able to deliver automultiscopic content, maintaining its transparent nature. We

## Conclusion

then introduced a novel approach to approximate a plenoptic function using beam steering and performed several simulations to evaluate the efficiency.

Finally, we presented a method to calibrate multi-layered display configurations using computer vision. A simple camera was used to compute relations between different layers of a display setup by means of homographies. For any given eye position and any chosen pixel, these homographies were used to determine all pixels lying on the line between the eye and the input pixel. This can be used to calibrate many different multi-layered display prototypes.

In the next section we review the principal contributions of this work and critically discuss them.

### 7.1 Review and Discussion of Principal Contributions

**A transparent stereoscopic system with full motion parallax:** To our knowledge we were the first to use anisotropic backprojection foils to create stereoscopic 3D. These systems are able to provide the illusion of depth for content showed in certain parts of the display while others remain transparent. They also contain low crosstalk. However objects appear ghostly due to the additive blending common to this technology. Motion parallax was added using viewer tracking to additionally improve the viewing experience. We found out that current eye tracker are good enough to create the desired effect when being improved with motion prediction and when the viewer is moving slow. For faster viewer motion, faster eye tracker would be required.

**A retroreflective large-scale autostereoscopic screen:** We presented a novel projection screen to address the needs to serve multiple viewers with independent stereoscopic content. The system was suitable for viewers with fixed seating, preferably side by side of each other. We configured it for stereoscopic content as well as multi-view content. While the crosstalk is acceptable for viewers situated next to each other, the approach has difficulties when viewers sit behind each other. However, due to the high brightness and the fact that different viewers can watch different movies at the same time makes the system suitable for many different applications.

**A mathematical framework describing a variety of multi-layered volumetric displays:** We introduced a mathematical model which can be used to describe the basic display primitives of multi-layered volumetric displays. Along with the framework we also present a decomposition algorithm for light field data which can be used to prepare any content for any display configuration. The feasibility of the approach has been demonstrated on two physical prototypes. Also we performed a quantitative analysis on simulated

## 7.1 Review and Discussion of Principal Contributions

results. As even light field displays can be described by basic display primitives, the framework possibly can be extended to more exotic display designs. As common to most volumetric displays, the prototypes we have built are restricted to content fitting mostly within the display volume with just few depth outside, being realized by a lightfield display primitive.

**The theory on bandwidth reduction for multi-layered automultiscopic displays:** We applied the knowledge about the dimensionality gap in the 4D plenoptic function of real scenes to light field displays. We could show that the required bandwidth can be reduced linearly with the number of deployed light field layers. Disadvantage is that each additional layer reduces the depth which can be shown outside the display. However, the immensely increased bandwidth certainly justifies the approach for a range of use case scenarios. Furthermore, the insights in our theory motivate further research in this topic to gain a better understanding of light field approximation using display technologies.

**A transparent holographic multi-view display:** Based on our previously invented transparent stereoscopic display, we proposed a system which gets rid of the 3D glasses as well as taking the display from stereoscopic to a light field display. The system uses the same technology as in existing backprojection foils but in a different configuration. In combination with a special projection system which makes better use of DMD bandwidth to create additional views, the system would be capable of creating an immersive viewing experience. However, at the time of this thesis, a physical sample of the required backprojection foil was not available. The capabilities of our special projector was shown by creating additional spatial resolution instead of multiple views.

**A novel architecture for a high bandwidth light field display:** As the depth which can be shown on state of the art light field display is enormously shallow, huge improvements need to be made to make those systems more appealing. We motivated the use of the temporal domain to significantly improve the bandwidth of light field displays. Instead of using the relatively slow switching liquid crystal technology for color attenuation, we propose to use it for steering beams of light from left to right. In combination with a light emitter as e.g. a laser diode which can be attenuated much faster, the angular sampling and thus the bandwidth is increased drastically. Due to the required but unavailable infrastructure to produce such a prototype, we show feasibility of the approach in simulations only. The presented configuration would achieve a aliasing free depth of field of over  $64m$ .

**A calibration algorithm for multi-layered displays:** Finally we presented a method to calibrate multiple layers against each other. Only assumption

## *Conclusion*

to the system is that each layer is planar with a rectangular or projected pixel grid. During the calibration using a camera and principles known from computer vision, the geometry of each layer is computed. Furthermore, a homography for any given eye position can be computed, relating a pixel on one layer to the pixel behind or in front of it on another layer. This can be used to dynamically compute the barrier pattern of a parallax barrier system, but also to assist decomposition methods for multi-layered displays. The method achieves subpixel accuracy when the assumption of planar layers is met.

## **7.2 Future Work**

In this work, we clearly extended the boundaries of different kinds of 3D displays. However, with the fast development in existing technologies as e.g. improvements in laser diodes as well as the invention of novel optical elements as e.g. LC GRIN lenses many new possibilities for even better displays capable of creating the illusion of depth are opening up. Thus, in all three classes, the limits are not reached yet.

We showed that transparent stereoscopic 3D can easily be achieved. Also we present ways to get rid of the cumbersome glasses to enhance the viewing experience. An open issue however is still how to occlude background. As the proposed screen is self-emissive, its light contribution adds up with the background which lets the displayed object appear ghostly. For deep immersion and increased realism, a display needs to be invented which can block background selectively in certain regions of the screen, optimally even based on the viewing direction.

Volumetric displays mainly suffer from the fact that they typically cannot display content outside the working volume. Combining them with light field displays, they can overcome this disadvantage. Thus, future work in this field of research may include hybrid display types e.g. to reduce the required bandwidth of light field displays in regions within the display volume.

Most promising display type for inducing a new 3D trend similar to the one when red/cyan glasses came up as well as the one evoked by active and passive 3D glasses are light field displays. Thus it can be expected that researchers will focus mainly on these types of displays. A trend which can be seen in current research is that efforts are being made to bridge the gap between ray based light field displays and wavefront based displays as holograms.

In any case, bandwidth will remain the critical key performance indicator of light field displays. Thus, deeper knowledge about bandwidth usage of light field data is required. This will help to develop 3D displays systems capable of producing more depth at lower data rates.

At the same time, the bandwidth capacity of light field displays should be improved. This includes the development of new techniques as few have been presented in this work but also involves advances on the engineering side. Example given, the full bandwidth of DMD chips is still not used due to lacking data links between GPU and projection device. Thus it could make sense to put efforts in building a development platform where the data link is not the bottleneck of the bandwidth, which could act as enabler for future research.

All in all the development of advanced 3D displays has still not come to an end and many issues and questions are left open. Any advances made will have impact both on the market as well as our society and thus the topic remains an interesting field of research.

## *Conclusion*

## Bibliography

- [Akeley *et al.*, 2004] Kurt Akeley, Simon J. Watt, Ahna Reza Girshick, and Martin S. Banks. A stereo display prototype with multiple focal distances. *ACM Transactions on Graphics*, 23(3):804–813, 2004.
- [Annen *et al.*, 2006] Thomas Annen, Wojciech Matusik, Hanspeter Pfister, H-P Seidel, and Matthias Zwicker. Distributed rendering for multiview parallax displays. In *Electronic Imaging 2006*, pages 60550P–60550P. International Society for Optics and Photonics, 2006.
- [Azarbayejani *et al.*, 1993] A. Azarbayejani, T. Starner, B. Horowitz, and A. Pentland. Visually controlled graphics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(6):602–605, Jun 1993.
- [Azuma, 1995] Ronald Tadao Azuma. *Predictive tracking for augmented reality*. PhD thesis, Citeseer, 1995.
- [Balogh *et al.*, 2008] Tibor Balogh, Péter Tamás Kovács, Zsuzsa Dobrányi, Attila Barsi, Zoltán Megyesi, Zoltán Gaál, and Gáspár Balogh. The holovizio system—new opportunity offered by 3d displays. In *Proceedings of the TMCE*, pages 79–89, 2008.
- [Barnum *et al.*, 2010] Peter C Barnum, Srinivasa G Narasimhan, and Takeo Kanade. A multi-layered display with water drops. *ACM Transactions on Graphics (TOG)*, 29(4):76, 2010.
- [Bell *et al.*, 2008] Gareth P. Bell, Robert Craig, Robert Paxton, Gordon Wong, and

## Bibliography

- Duncan Galbraith. Beyond flat panels: Multi-layered displays with real depth. In *SID Symposium Digest of Technical Papers*, 39, 1, pages 352–355, 2008.
- [Berkel, 1999] Cees Van Berkel. Image preparation for 3D-LCD. *SPIE Stereoscopic Displays and Virtual Reality Systems*, 3639:84–91, 1999.
- [Bogaert *et al.*, 2010] Lawrence Bogaert, Youri Meuret, Stijn Roelandt, Aykut Avci, Herbert De Smet, and Hugo Thienpont. Demonstration of a multiview projection display using decentered microlens arrays. *Opt. Express*, 18(25):26092–26106, Dec 2010.
- [Butler *et al.*, 2011] Alex Butler, Otmar Hilliges, Shahram Izadi, Steve Hodges, David Molyneaux, David Kim, and Danny Kong. Vermeer: direct interaction with a 360° viewable 3d display. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST '11, pages 569–576, New York, NY, USA, 2011. ACM.
- [Buzak, 1985] T. S. Buzak. A field-sequential discrete-depth-plane three-dimensional display. *Society of Information Display Digest*, pages 345–347, 1985.
- [Chai *et al.*, 2000] Jin-Xiang Chai, Xin Tong, Shing-Chow Chan, and Heung-Yeung Shum. Plenoptic sampling. In *Proc. SIGGRAPH*, pages 307–318, 2000.
- [Cossairt *et al.*, 2007] Oliver S. Cossairt, Joshua Napoli, Samuel L. Hill, Rick K. Dorval, and Gregg E. Favalora. Occlusion-capable multiview volumetric three-dimensional display. *Applied Optics*, 46:1244–1250, 2007.
- [Dodgson, 2005] N.A. Dodgson. Autostereoscopic 3d displays. *Computer*, 38(8):31–36, 2005.
- [Doutre and Nasiopoulos, 2011] Colin Doutre and Panos Nasiopoulos. Crosstalk cancellation in 3d video with local contrast reduction. In *Proceedings of European Signal Processing Conference*, 2011.
- [Durand *et al.*, 2005] Frédo Durand, Nicolas Holzschuch, Cyril Soler, Eric Chan, and François X. Sillion. A frequency analysis of light transport. In *ACM SIGGRAPH 2005 Papers*, SIGGRAPH '05, pages 1115–1126, New York, NY, USA, 2005. ACM.
- [Favalora, 2005] G. Favalora. Volumetric 3d displays and application infrastructure. *IEEE Computer*, 38(8):37–44, 2005.
- [Ferguson, 1997] James L Ferguson. Patent: Optical system for a head mounted display using a retro-reflector and method of displaying an image, 1997.
- [Fuchs *et al.*, 2008] Martin Fuchs, Ramesh Raskar, Hans-Peter Seidel, and Hendrik P. A. Lensch. Towards passive 6D reflectance field displays. *ACM Transactions on Graphics*, 27(3), 2008.

- [Gao and Xiao, 2010] Chunyu Gao and Jing Xiao. Patent: Retro-reflective light diffusing display systems, 2010.
- [Gold and Freeman, 1998] R.S. Gold and J.E. Freeman. Layered display system and method for volumetric presentation. U.S. patent 5,813,742, 1998.
- [Gotoda, 2010] Hironobu Gotoda. A multilayer liquid crystal display for autostereoscopic 3D viewing. In *Stereoscopic Displays and Virtual Reality Systems XI*, volume 7524, 2010.
- [Gotoda, 2011] Hironobu Gotoda. Reduction of image blurring in an autostereoscopic multilayer liquid crystal display, 2011.
- [Hartley and Zisserman, 2000] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2000.
- [Holroyd *et al.*, 2011] Michael Holroyd, Ilya Baran, Jason Lawrence, and Wojciech Matusik. Computing and fabricating multilayer models. *ACM Transactions on Graphics*, 30(6):187:1–187:8, 2011.
- [Isono *et al.*, 1993] Haruo Isono, Minoru Yasuda, and Hideaki Sasazawa. Autostereoscopic 3-d display using lcd-generated parallax barrier. *Electronics and Communications in Japan (Part II: Electronics)*, 76(7):77–84, 1993.
- [Itoh, 2012] Goh Itoh. 17.1: Hardware and software technologies for glasses-free 3d tvs and pcs. *SID Symposium Digest of Technical Papers*, 43(1):210–213, 2012.
- [Ives, 1903] Frederic Ives. Parallax stereogram and process for making same. U.S. Patent No. 725,567, 1903.
- [Jacobs *et al.*, 2003] Adrian Jacobs, Jonathan Mather, Robert Winlow, David Montgomery, Graham Jones, Morgan Willis, Martin Tillin, Lyndon Hill, Marina Khazova, Heather Stevenson, and Grant Bourhill. 2d/3d switchable displays. *Sharp Technical Journal*, (4):15–18, 2003.
- [Jang and Javidi, 2002] Ju-Seog Jang and Bahram Javidi. Improved viewing resolution of three-dimensional integral imaging by use of nonstationary micro-optics. *Optics Letters*, 27(5):324–326, 2002.
- [Jebara and Pentland, 1997] T.S. Jebara and A. Pentland. Parametrized structure from motion for 3d adaptive feedback tracking of faces. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 144–150, Jun 1997.
- [Jones *et al.*, 1988] Kevin M. Jones, Scott Lundgren, and Alak Chakravorty. A calculus of variations demonstration: The gradient index lens. *American Journal of Physics*, 56(12):1099–1099, 1988.

## Bibliography

- [Jones *et al.*, 2007] Andrew Jones, Ian McDowall, Hideshi Yamada, Mark Bolas, and Paul Debevec. Rendering for an interactive 360 light field display. *ACM Transactions on Graphics*, 26(3):40:1–40:10, July 2007.
- [Kao *et al.*, 2009] Yung-Yuan Kao, Yan-Pean Huang, Kai-Xian Yang, Paul C.-P. Chao, Chi-Chung Tsai, and Chi-Neng Mo. An auto-stereoscopic 3D display using tunable liquid crystal lens array that mimics effects of grin lenticular lens array. *SID Symposium Digest of Technical Papers*, 40:111–114, 2009.
- [Kim *et al.*, 2007] Yunhee Kim, Joohwan Kim, Jin-Mo Kang, Jae-Hyun Jung, Heejin Choi, , and Byoungcho Lee. Point light source integral imaging with improved resolution and viewing angle by the use of electrically movable pinhole array. *Optics Express*, 15(26):18253–18267, 2007.
- [Kim *et al.*, 2010] Youngmin Kim, Keehoon Hong, and Byoungcho Lee. Recent researches based on integral imaging display method. *3D Research*, 1:17–27, 2010.
- [Krum *et al.*, 2012] David M. Krum, Evan A. Suma, and Mark Bolas. Augmented reality using personal projection and retroreflection. *Personal and Ubiquitous Computing*, 16(1):17–26, 2012.
- [Kuster *et al.*, 2012] Claudia Kuster, Nicola Ranieri, Agustina, Henning Zimmer, Jean-Charles Bazin, Chengzheng Sun, Tiberiu Popa, and Markus Gross. Towards next generation 3d teleconferencing systems. In *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, 2012.
- [Kuster *et al.*, 2015] Claudia Kuster, Nicola Ranieri, Tobias Martin, Pierre-Yves Laffont, Tiberiu Popa, Jean-Charles Bazin, and Markus Gross. An immersive bidirectional system for life-size 3d communication. In *submitted to 3DV*, 2015.
- [Lanman *et al.*, 2010] Douglas Lanman, Matthew Hirsch, Yunhee Kim, and Ramesh Raskar. Content-adaptive parallax barriers: optimizing dual-layer 3d displays using low-rank light field factorization. *ACM Trans. Graph.*, 29(6):163:1–163:10, 2010.
- [Lanman *et al.*, 2011] Douglas Lanman, Gordon Wetzstein, Matthew Hirsch, Wolfgang Heidrich, and Ramesh Raskar. Polarization fields: dynamic light field display using multi-layer LCDs. *ACM Transactions on Graphics*, 30(6), 2011.
- [Lee *et al.*, 2004] Johnny C Lee, Paul H Dietz, Dan Maynes-Aminzade, Ramesh Raskar, and Scott E Hudson. Automatic projector calibration with embedded light sensors. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 123–126. ACM, 2004.
- [Leung *et al.*, 1998] M. S. Leung, N. A. Ives, and G. Eng. Three-dimensional real-image volumetric display system and method. U.S. patent 5,745,197, 1998.

- [Levin and Durand, 2010] A. Levin and F. Durand. Linear view synthesis using a dimensionality gap light field prior. In *CVPR*, pages 1831–1838, 2010.
- [Lippmann, 1908] Gabriel M. Lippmann. La photographie integrale. *Comptes-Rendus*, 146:446–451, 1908.
- [Livingston *et al.*, 2012] M.A. Livingston, J. Sebastian, Zhuming Ai, and J.W. Decker. Performance measurements for the microsoft kinect skeleton. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, pages 119–120, March 2012.
- [Machida *et al.*, 2012] E. Machida, Meifen Cao, T. Murao, and H. Hashimoto. Human motion tracking of mobile robot with kinect 3d sensor. In *SICE Annual Conference (SICE), 2012 Proceedings of*, pages 2207–2211, Aug 2012.
- [Maimone *et al.*, 2013] Andrew Maimone, Gordon Wetzstein, Matthew Hirsch, Douglas Lanman, Ramesh Raskar, and Henry Fuchs. Focus 3d: Compressive accommodation display. *ACM Trans. Graph.*, 32(5):153:1–153:13, October 2013.
- [Maimone *et al.*, 2014] Andrew Maimone, Renjie Chen, Henry Fuchs, Ramesh Raskar, and Gordon Wetzstein. 36.1: Wide field of view compressive light field display using a multilayer architecture and tracked viewers. *SID Symposium Digest of Technical Papers*, 45(1):509–512, 2014.
- [Nagano *et al.*, 2013] Koki Nagano, Andrew Jones, Jing Liu, Jay Busch, Xueming Yu, Mark Bolas, and Paul Debevec. An autostereoscopic projector array optimized for 3d facial display. In *ACM SIGGRAPH 2013 Emerging Technologies*, page 3. ACM, 2013.
- [Nashel and Fuchs, 2009] Andrew Nashel and Henry Fuchs. Random hole display: A non-uniform barrier autostereoscopic display. In *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video, 2009*, pages 1–4. IEEE, 2009.
- [Nguyen and Canny, 2005] David Nguyen and John Canny. Multiview: Spatially faithful group video conferencing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05*, pages 799–808, New York, NY, USA, 2005. ACM.
- [Oka *et al.*, 2013] Shinichiro Oka, Tomohiko Naganuma, Takeo Koito, Yingbao Yang, and Shinichi Komura. 15.5: Invited paper: High performance autostereoscopic 2d/3d switchable display using liquid crystal lens. *SID Symposium Digest of Technical Papers*, 44(1):150–153, 2013.
- [Parker *et al.*, 2010] Steven G Parker, James Bigler, Andreas Dietrich, Heiko Friedrich, Jared Hoberock, David Luebke, David McAllister, Morgan McGuire, Keith Morley, Austin Robison, et al. Optix: a general purpose ray tracing engine. *ACM Transactions on Graphics (TOG)*, 29(4):66, 2010.

## Bibliography

- [Perlin *et al.*, 2000] Ken Perlin, Salvatore Paxia, and Joel S. Kollin. An autostereoscopic display. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, pages 319–326, July 2000.
- [Peterka *et al.*, 2008] Tom Peterka, Robert L. Kooima, Daniel J. Sandin, Andrew Johnson, Jason Leigh, and Thomas A. DeFanti. Advances in the dynallax solid-state dynamic parallax barrier autostereoscopic visualization display system. *IEEE Transactions on Visualization and Computer Graphics*, 14(3):487–499, May/June 2008.
- [Ranieri and Gross, 2014a] Nicola Ranieri and Markus Gross. Patent: Display device for time-sequential multi-view content, 2014.
- [Ranieri and Gross, 2014b] Nicola Ranieri and Markus Gross. Vision-based calibration of parallax barrier displays. In *Proceedings of SPIE 9011*, pages 9011–1D. International Society for Optics and Photonics, 2014.
- [Ranieri *et al.*, 2012] Nicola Ranieri, Simon Heinzle, Quinn Smithwick, Daniel Reetz, Lanny S. Smoot, Wojciech Matusik, and Markus Gross. Multi-layered automultiscopic displays. *Computer Graphics Forum*, 31(7pt2):2135–2143, 2012.
- [Ranieri *et al.*, 2013a] Nicola Ranieri, Simon Heinzle, Peter Barnum, Wojciech Matusik, and Markus Gross. Light-field approximation using basic display layer primitives. *SID Symposium Digest of Technical Papers*, 44(1):408–411, 2013.
- [Ranieri *et al.*, 2013b] Nicola Ranieri, Simon Heinzle, Peter Barnum, Wojciech Matusik, and Markus Gross. Multi-planar plenoptic displays. *Journal of the Society for Information Display*, 21(10):451–459, 2013.
- [Ranieri *et al.*, 2013c] Nicola Ranieri, Simon Heinzle, Quinn Smithwick, Daniel Reetz, Lanny S. Smoot, Wojciech Matusik, and Markus Gross. Patent: Multi-layer plenoptic displays that combine multiple emissive and light modulating planes, 2013.
- [Ranieri *et al.*, 2014] Nicola Ranieri, Hagen Seifert, and Markus Gross. Transparent stereoscopic display and application. In *Proceedings of SPIE 9011*, pages 9011–0P. International Society for Optics and Photonics, 2014.
- [Ranieri *et al.*, 2015] Nicola Ranieri, Hagen Seifert, Quinn Smithwick, and Markus Gross. Patent: System and method for increased spatial resolution, 2015.
- [Sakamoto *et al.*, 2005] K. Sakamoto, R. Kimura, and M. Takaki. Parallax polarizer barrier stereoscopic 3d display systems. In *Active Media Technology, 2005. (AMT 2005). Proceedings of the 2005 International Conference on*, pages 469–474, May 2005.

- [Schaffner *et al.*, 2013] M. Schaffner, P. Greisen, S. Heinzle, F.K. Gurkaynak, H. Kaeslin, and A. Smolic. Madmax: A 1080p stereo-to-multiview rendering ASIC in 65 nm CMOS based on image domain warping. In *ESSCIRC (ESSCIRC), 2013 Proceedings of the*, pages 61–64, 2013.
- [Smithwick and Ranieri, 2015] Quinn Smithwick and Nicola Ranieri. A large 1-d retroreflective autostereoscopic display. In *Proceedings of SPIE 9391*, pages 9391–16. International Society for Optics and Photonics, 2015.
- [Smithwick *et al.*, 2011] Quinn Smithwick, Lanny S. Smoot, and Daniel Reetz. Autostereoscopic display system with one dimensional (1d) retroreflective screen, 2011.
- [Smithwick *et al.*, 2012] Quinn Smithwick, Lanny S. Smoot, and Daniel Reetz. 59.1: A volumetric display using a rim-driven varifocal beamsplitter and high-speed DLP backlit LCD. *SID Symposium Digest of Technical Papers*, 43(1):792–795, 2012.
- [Sullivan, 2004] Alan Sullivan. DepthCube solid-state 3D volumetric display. In *Stereoscopic Displays and Virtual Reality Systems XI*, volume 5291, pages 279–284, 2004.
- [Surman *et al.*, 2013] Phil Surman, Sally Day, Bonny Bobby, Hao Chen, Hakan Urey, and Kaan Aksit. Head-tracked retroreflecting 3d display. *SID Symposium Digest of Technical Papers*, 44:247–250, 2013.
- [Suyama *et al.*, 2000] S. Suyama, H. Takada, K. Uehira, and S. Sakai. A novel direct-vision 3-d display using luminance-modulated two 2-d images displayed at different depths. *SID Digest*, pages 1208–1211, 2000.
- [Suyama *et al.*, 2004] Shiro Suyama, Sakuichi, Ohtsuka Hideaki, Takada Kazutake, and Uehira Sigenobu Sakai. Apparent 3-d image perceived from luminance-modulated two 2-d images displayed at different depths. *Vision Research*, (44):785–793, 2004.
- [Takagi *et al.*, 2013] Ayako Takagi, Shinichi Uehara, Masako Kashiwagi, and Masahiro Baba. 15.4: Function integrated LC GRIN lens for partially switchable 2d/3d display. *SID Symposium Digest of Technical Papers*, 44(1):162–165, 2013.
- [Takaki, 2006] Y. Takaki. High-density directional display for generating natural three-dimensional images. *Proceedings of the IEEE*, 94(3):654–663, 2006.
- [Tamura and Tanaka, 1982] S. Tamura and K. Tanaka. Multilayer 3-d display by multidirectional beam splitter. *Applied Optics*, 21:3659–3663, 1982.
- [Traub, 1967] A. Traub. Stereoscopic display using rapid varifocal mirror oscillations. *Applied Optics*, 6(6):1505–1511, 1967.

## Bibliography

- [Wagner and Schmalstieg, 2007] Daniel Wagner and Dieter Schmalstieg. Artoolkitplus for pose tracking on mobile devices. In *Proceedings of 12th Computer Vision Winter Workshop, 2007*.
- [Weissman and Woods, 2011] Michael A Weissman and Andrew J Woods. A simple method for measuring crosstalk in stereoscopic displays. In *IS&T/SPIE Electronic Imaging*, pages 786310–786310. International Society for Optics and Photonics, 2011.
- [Welch and Bishop, 1995] Greg Welch and Gary Bishop. An introduction to the kalman filter. 1995.
- [Wetzstein *et al.*, 2011] G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar. Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.*, 30(4), 2011.
- [Wetzstein *et al.*, 2012] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting. *ACM Transactions on Graphics*, 31(4):1–11, 2012.
- [Xu *et al.*, 2009] Lin Xu, Jian Zhang, and Li Ying Wu. Influence of phase delay profile on diffraction efficiency of liquid crystal optical phased array. *Optics and Laser Technology*, 41(4):509 – 516, 2009.
- [Ye *et al.*, 2010] Gu Ye, A State, and Henry Fuchs. A practical multi-viewer tabletop autostereoscopic display. In *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*, pages 147–156. IEEE, 2010.
- [Zhang, 1999] Zhengyou Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 666–673. IEEE, 1999.
- [Zwicker *et al.*, 2006] M. Zwicker, W. Matusik, F. Durand, and H. Pfister. Antialiasing for automultiscopic 3D displays. In *Eurographics Symposium on Rendering*, 2006.

# Curriculum Vitae

**Nicola Ranieri**

## Personal Data

Aug. 7, 1985                      Born in Glarus, Switzerland  
Nationality                      Swiss

## Education

May 26, 2015                      **Ph.D. Defense**

Feb. 2010 – May 2015           **Fast-Track Doctoral Studies**  
*Computer Graphics*  
ETH Zürich, Switzerland

Feb. 2009 – Feb. 2011           **MSc in Computer Science**  
*Visual Computing*  
ETH Zürich, Switzerland

Oct. 2005 – Oct. 2009           **BSc in Computer Science**  
ETH Zurich, Switzerland

Aug. 1998 – July 2004           **Swiss Matura**  
*Applied Mathematics and Physics*  
Kantonsschule Glarus, Switzerland

## Scientific Publications

C. KUSTER, N. RANIERI, T. MARTIN, P.Y. LAFFONT, T. POPA, J.C. BAZIN, and M. GROSS. An Immersive Bidirectional System for Life-size 3D Communication, *submitted to 3DV*, 2015.

Q. SMITHWICK, and N. RANIERI. A Large 1-D Retroreflective Autostereoscopic Display, *Proceedings of SPIE 9391 (SanFrancisco, USA, 9-11 February, 2015), Stereoscopic Displays and Applications XXVI*, 9391-16, 2015.

N. RANIERI, and M. GROSS. Vision-based Calibration of Parallax Barrier Displays, *Proceedings of SPIE 9011 (SanFrancisco, USA, 4-8 February, 2014), Stereoscopic Displays and Applications XXV*, 90111D, 2014.

N. RANIERI, H. SEIFERT, and M. GROSS. Transparent Stereoscopic Display and Application, *Proceedings of SPIE 9011 (SanFrancisco, USA, 4-8 February, 2014), Stereoscopic Displays and Applications XXV*, 90110P, 2014.

N. RANIERI, S. HEINZLE, P. BARNUM, W. MATUSIK, and M. GROSS. Multi-planar plenoptic displays, *Journal of the Society for Information Display (Blackwell Publishing Ltd)*, **invited journal paper**, vol. 21, no. 10, pp. 451-459, 2013.

N. RANIERI, S. HEINZLE, P. BARNUM, W. MATUSIK, and M. GROSS. Light-Field Approximation Using Basic Display Layer Primitives, *SID Symposium Digest of Technical Papers 2013 (Vancouver, Canada, 19-24 May, 2013), Display Week*, pp. 408-411, 2013.

C. KUSTER, N. RANIERI, AGUSTINA, H. ZIMMER, J.C. BAZIN, C. SUN, T. POPA, and M. GROSS. Towards Next Generation 3D Teleconferencing Systems, *Proceedings of 3DTV-CON 2012 (Zurich, Switzerland, 15-17 October, 2012), 3DTV-CON*, pp. 1-4, 2012.

N. RANIERI, S. HEINZLE, Q. SMITHWICK, D. REETZ, L. S. SMOOT, W. MATUSIK, and M. GROSS. Multi-Layered Automultiscopic Displays, *Proceedings of Pacific Graphics 2012 (Hong Kong, 12-14 September, 2012), Pacific Graphics*, pp. 2135-2143, 2012.

## Patents

N. RANIERI, Q. SMITHWICK, H. SEIFERT, and M. GROSS. SYSTEM AND METHOD FOR INCREASED SPATIAL RESOLUTION, *U.S. Patent US2015/250021 A1*, filed on 02.03.2015

N. RANIERI, and M. GROSS. DISPLAY DEVICE FOR TIME-SEQUENTIAL MULTI-VIEW CONTENT, *European Patent WO2014/180554 A1*, filed on 05.05.2014

N. RANIERI, S. HEINZLE, Q. SMITHWICK, D. REETZ, L. S. SMOOT, W. MATUSIK, and M. GROSS. MULTI-LAYER PLENOPTIC DISPLAYS THAT COMBINE MULTIPLE EMISSIVE AND LIGHT MODULATING PLANES, *U.S. Patent US2013/082905 A1*, filed on 04.04.2013

## **Employment**

- July 2014 – Oct. 2014 Internship at Walt Disney Imagineering Research & Dev., Disney Research, Los Angeles CA, United States of America
- July 2012 – Aug. 2012 Internship at Walt Disney Imagineering Research & Dev., Disney Research, Los Angeles CA, United States of America
- Oct. 2009 – Jan. 2009 Research assistant at ETH Zurich, Zurich, Switzerland.
- Jun. 2007 – Aug. 2007 Software Developer at MA Systems and Control Limited, United Kingdom.

*Curriculum Vitae*