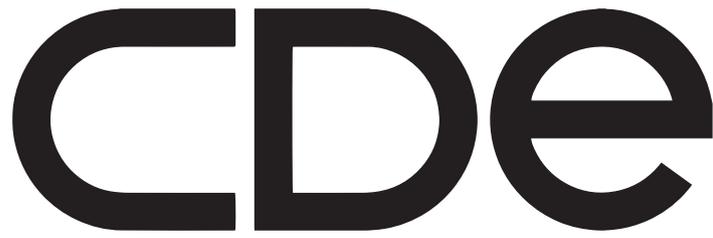


# **Media Production for Computer-Generated Holography**



**Centre for Digital  
Entertainment**

**Aaron Demolder**

**Academic Supervisors:** Dr. Hammadi Nait-Charif

Dr. Valery Adzhiev

Industry Supervisor: Dr. Andrzej Kaczorowski

Centre for Digital Entertainment

Bournemouth University

This dissertation is submitted for the degree of  
*Doctor of Engineering*

June 2024



## **Declaration**

I hereby declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Aaron Demolder

June 2024



## Acknowledgements

I would first like to thank my supervisors, Dr Hammadi Nait-Charif, Dr Valery Adzhiev and Dr Andrzej Kaczorowski, for putting up with me for several years. Thank you to all three of you for giving me feedback on papers I decided to write last minute, and for helping guide the chaos transferred from my mind into research which can be easily deciphered.

I'd like to thank everyone at VividQ for providing such a fun (and challenging) environment to work in. Thanks go to AK for continually encouraging me to prove him wrong, and especially to Alfie, Tom, Indi and Chris for enduring my questions. There are now too many people at VividQ to name individually, but special mentions also go to Ram, for suffering through the demo creation process with me, Alden, for being the most organised person in the universe, Camille, for being the IP expert behind submitting our patent, and to Rob, for always finding a way to get up to mischief.

I would like to thank the CDE for giving me the opportunity to complete a doctorate, if you had told me I would be doing a "doctorate of engineering" when I first went to university, I would have run far away. It's not something I had ever considered, and I still wonder how we got here - so thank you to Valery for running the research unit in the computer animation undergrad, and to Hammadi for supervising my project and encouraging me to apply to the EngD. I'd also like to thank Mike and Zoe, for making everything run smoothly and for making sure that I had the equipment I needed.

A big thank you of course to my family, friends and wife for listening to me talk mostly about holograms for several years. To my parents, Catrina and Damien, I'm extremely grateful for the constant support you've given me, for encouraging me to work hard but to also do things I was interested in doing - though I think we're all a bit surprised I ended up here by pursuing art! To my wife, Victoria, thank you for your daily encouragement, belief and passion, it really does mean the world.

A final thank you to those who not only read my paper submissions, but liked them enough to publish and award them in various ways. It has been really rewarding to have my efforts well received, and I'm especially thankful for the RTX6000 GPU awarded by NVIDIA at CVMP 2020, which has made my research, and consequently my life, a lot easier.



## Abstract

Computer Generated Holography (CGH) is an immersive next-generation display technology with significant potential. It is the only technique that allows the reconstruction of true three-dimensional imagery while not compromising display resolution - achieved through the diffraction and interference of light. When a hologram is illuminated with coherent light and the replay field is observed, all psychological and physiological depth cues of the image are preserved and the scene is interpreted as reality by the viewer.

CGH has recently received significant attention from researchers, with applications in Virtual/Augmented Reality, Head-Up Display, and larger panels. Providing natural depth perception (defocus, parallax, occlusion), high dynamic ranges and wide colour gamuts.

In CGH research there is a focus on algorithmic and hardware developments, but a failure to make use of advancements in image rendering made in Computer Graphics. Works largely implement their own basic renderers, or use only RGB + Depth information rendered from a single perspective. Neither of these can yet replicate the full effects of a realistic 3D scene with complex material properties, and do not match the feature set and quality of a 2D final-frame rendering and compositing process.

In collaboration with partner company VividQ, this work develops the relationship between renderer and display with a practical framework of methods that support industry practice. By utilising existing production renderers, extensive features and familiar working environments are enabled for both high quality 2D and 3D content delivery. This work introduces methods such as "holographic compositing", which composites imagery in frequency space to enable the viewing of multiple depth and view inputs with improved occlusions - enabling support for transparencies, motion-blur, refractions, reflections, volumetrics such as smoke, fire and fog, fine details such as fur and improved accuracy in edge details. A unique multi-view "discrete-hybrid" hologram generation algorithm is also introduced, with appropriate production tools, allowing an observer to experience view-dependent effects such as large parallax shifts and changes in perspective.

The developments presented in this work are primarily proved via computer simulations. Though, many of the methods described have already been implemented and proven in the partner company's software product.



# Table of contents

<b>List of figures</b>	<b>xiii</b>
<b>List of tables</b>	<b>xvii</b>
<b>Nomenclature</b>	<b>xix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Research Problems Overview . . . . .	1
1.2 Company Description . . . . .	3
1.3 Experience at VividQ . . . . .	3
1.4 Contributions . . . . .	4
1.5 Thesis Outline . . . . .	6
1.6 Publications . . . . .	6
<b>2 Background - An Introduction to Holography</b>	<b>9</b>
2.1 What is Holography? . . . . .	9
2.1.1 Fundamentals . . . . .	9
2.1.2 Holography with Photographic Film . . . . .	10
2.1.3 Computer-Generated Holography . . . . .	15
2.2 CGH in Practice . . . . .	18
2.2.1 Modern CGH methods - Compute Reduction . . . . .	18
2.2.2 Random Phase . . . . .	20
2.2.3 Display Hardware . . . . .	21
2.2.4 Example Display System . . . . .	23
<b>3 Production Workflow</b>	<b>25</b>
3.1 Introduction . . . . .	25
3.1.1 Technical Integration with Media Production . . . . .	25
3.1.2 Lack of Artistic Consideration in CGH . . . . .	28

3.2	Background . . . . .	28
3.2.1	Digital Representations of Light . . . . .	29
3.2.2	Digital Representations of Colour . . . . .	29
3.2.3	Delivery of Imagery . . . . .	30
3.3	Holographic Production Workflow . . . . .	31
3.3.1	Enabling High Quality Holographic Display . . . . .	31
3.3.2	3D Data Representation . . . . .	33
3.3.3	Defocus Control . . . . .	33
3.3.4	Multiple Depths . . . . .	34
3.3.5	Multiple Views . . . . .	34
3.3.6	Data Required . . . . .	35
3.4	High Dynamic Range Handling . . . . .	35
3.4.1	Existing HDR . . . . .	37
3.4.2	Holographic HDR . . . . .	37
3.4.3	Luminance Reserved for Highlights . . . . .	42
3.5	Flexible Wide Colour Handling . . . . .	44
3.5.1	Laser Primary Choices . . . . .	45
3.5.2	Academy Color Encoding System (ACES) . . . . .	47
3.5.3	Modern Color Matching Function . . . . .	48
3.5.4	Chroma Sub-Sampling Support . . . . .	48
3.5.5	Nearly Ultimate Gamut . . . . .	48
3.5.6	Proposed HQH Delivery Format . . . . .	51
3.6	Artistic Optical System Controls . . . . .	52
3.6.1	Virtual Aperture Control . . . . .	52
3.6.2	Lens Flare and Glare . . . . .	58
3.6.3	Field Curvature . . . . .	62
3.7	Implementation in Production Environment . . . . .	62
3.7.1	Tooling Configurations . . . . .	63
3.7.2	Compute Process . . . . .	66
3.8	Holographic Projector Calibration . . . . .	67
3.9	Production Workflow Results . . . . .	69
3.9.1	Improved Image Quality & Control . . . . .	70
3.9.2	Conclusion . . . . .	70
<b>4</b>	<b>Multi-Sampled Depth</b>	<b>75</b>
4.1	Limits of the Z Channel . . . . .	75
4.2	Background . . . . .	79

---

4.2.1	The Z-Buffer in Computer Graphics . . . . .	79
4.2.2	The Z-Buffer in Layer-Based CGH . . . . .	82
4.3	Holographic Compositing & AOVs . . . . .	86
4.3.1	Holographic Compositing Operations . . . . .	88
4.3.2	Ray Length Based ZDepth . . . . .	95
4.3.3	Limitations of AOVs . . . . .	101
4.4	Occlusion Synthesis . . . . .	102
4.4.1	Insufficient Scene Data . . . . .	102
4.4.2	Process Overview . . . . .	103
4.4.3	Worked Examples . . . . .	105
4.4.4	Limitations of Occlusion Synthesis . . . . .	109
4.5	Deep Holograms . . . . .	110
4.5.1	Layer based CGH from a Deep source . . . . .	113
4.5.2	Deep Occlusion . . . . .	116
4.6	Renderer Integration Streamlining . . . . .	121
4.6.1	Complete Scene Representation . . . . .	121
4.6.2	Context Based Layer Slicing . . . . .	122
4.6.3	Other Scene Data . . . . .	124
4.7	Multi-Sampled Depth Results . . . . .	127
<b>5</b>	<b>Multi-View Holographic Display</b>	<b>131</b>
5.1	Multi-Perspective CGH for Large Format Displays . . . . .	131
5.2	Related Work . . . . .	134
5.2.1	Lightfield Display . . . . .	134
5.2.2	The Holographic Stereogram - Comparison with Lightfield . . . . .	134
5.2.3	Holographic Field Sampling . . . . .	136
5.2.4	Holographic Field Sampling Reduction . . . . .	137
5.2.5	Multi-View via Eye Tracking . . . . .	138
5.2.6	Multi-plane with Multi-view . . . . .	140
5.3	Discrete Hybrid Holography . . . . .	141
5.3.1	Hogel Layout Design & Holographic Computation Algorithm . . . . .	143
5.3.2	View Handover & View Dependant Effects . . . . .	146
5.3.3	Compute Balancing . . . . .	151
5.3.4	View Rendering & Capture . . . . .	154
5.3.5	View Storage . . . . .	160
5.3.6	Considerations for Physical Display . . . . .	162
5.3.7	Limitations . . . . .	165

---

5.4	Multi-View Results . . . . .	167
5.4.1	Discussion & Results . . . . .	167
5.4.2	Conclusion . . . . .	177
<b>6</b>	<b>Discussion &amp; Further Work</b>	<b>179</b>
6.1	Discussion . . . . .	179
6.1.1	Production Workflow . . . . .	180
6.1.2	Multi-Sampled Depth . . . . .	182
6.1.3	Multi-View Rendering . . . . .	183
6.2	Further Work . . . . .	184
6.2.1	Layers and Other Scene Quantisation Schemes . . . . .	184
6.2.2	High Quality Display . . . . .	186
6.2.3	Occlusion . . . . .	187
6.2.4	Quality of Life . . . . .	188
<b>7</b>	<b>Conclusions &amp; Final Note</b>	<b>189</b>
7.1	Conclusion . . . . .	189
7.2	EngD Experience Final Note . . . . .	191
	<b>References</b>	<b>193</b>

# List of figures

1.1	Contributions Summary in Workflow . . . . .	6
2.1	Interference Pattern via Double Slit Experiment . . . . .	10
2.2	Diffraction Patterns and Replays . . . . .	11
2.3	The Holographic Recording Process . . . . .	11
2.4	The Holographic Replay Process . . . . .	12
2.5	Holographic Viewing Breakdown . . . . .	14
2.6	VISA Dove Security Hologram . . . . .	15
2.7	The Propagation of a Wavefront . . . . .	16
2.8	The Diffraction Field . . . . .	17
2.9	Points, Polygons and Planes Comparison . . . . .	19
2.10	Specular vs Diffuse Reflection Comparison . . . . .	20
2.11	LCD and SLM Cross-section Comparison . . . . .	22
2.12	Holographic Display Diagram, with Viewing Examples. . . . .	24
2.13	Hologram of a Dog on a simple projector . . . . .	24
3.1	An Overview of the Media Production Pipeline . . . . .	26
3.2	Image Quality to Display . . . . .	32
3.3	Data Values Representing Depth . . . . .	33
3.4	VFX Pipeline with Hologram Data Requirements Met . . . . .	36
3.5	The Barten Threshold . . . . .	38
3.6	Highlight Compression in Holographic HDR . . . . .	41
3.7	Holographic Display Coverage . . . . .	42
3.8	Coverage of Images . . . . .	44
3.9	BT.709 Gamut vs Display Gamut with Monochromatic Primaries . . . . .	46
3.10	Holographic Replays using Varying Colourspace Standards . . . . .	47
3.11	Variance in Colour Matching Functions . . . . .	49
3.12	Larger Display Gamuts with Monochromatic Primaries . . . . .	50

---

3.13	Artistic Bokeh in The Lego Batman Movie . . . . .	53
3.14	Bokeh Control with Eyebox and Iris Considerations . . . . .	54
3.15	Custom Defocus Applied to Holograms . . . . .	55
3.16	Controlling the Appearance of Pupil Dilation and Constriction in Defocus. . . . .	56
3.17	Bokeh shaped as a Butterfly and Animations . . . . .	57
3.18	Examples of Lens Flares and Other Lens Based Effects in Existing Media. . . . .	58
3.19	Flare Defocus Control . . . . .	59
3.20	Holographic Lens Glares. . . . .	61
3.21	Field Curvature Control in Holographic Imagery. . . . .	62
3.22	Generation Pipeline Configuration Layout . . . . .	64
3.23	Hologram Enabled Workflow On-Premises. . . . .	66
3.24	Hologram Enabled Workflow in a Cloud Environment. . . . .	67
3.25	Holographic Display Colour Calibration. . . . .	68
3.26	Comparison of HDR vs SDR Holographic Replays. . . . .	70
3.27	Fast Holographic Media Production. . . . .	71
3.28	Holographic Replays from Existing Productions. . . . .	72
4.1	Complex Scene Example Image. . . . .	76
4.2	Single vs Multi-Depth Representation . . . . .	78
4.3	Pinhole vs Non-Zero Aperture . . . . .	81
4.4	Occlusion Issues In CGH - Viewing Position . . . . .	83
4.5	Occlusion Issues In CGH - Defocus . . . . .	84
4.6	Point Based Occlusion Method from Shi et al. 2021 . . . . .	85
4.7	Occlusion Artefacts in Current State-of-the-Art . . . . .	86
4.8	Z Depth Issues . . . . .	87
4.9	AOV Compositing Process . . . . .	89
4.10	Holographic Plus Compared to Holographic Over . . . . .	92
4.11	A Layer with a Partial Occlusion . . . . .	93
4.12	Holographic Image Multiplication . . . . .	94
4.13	Depth Passes Encode Only First-hit Geometry . . . . .	96
4.14	Optical Depth Ray Diagram . . . . .	97
4.15	Reflection and Refraction Holographic Compositing . . . . .	98
4.16	Compositing of Reflection and Refraction Results . . . . .	99
4.17	Multiple Focal Points Through a Window Pane . . . . .	100
4.18	Occlusion Synthesis Requirement Regions . . . . .	103
4.19	Occlusion Synthesis Overview . . . . .	104
4.20	Inputs to Occlusion Synthesis . . . . .	106

---

4.21	No Occlusion Synthesis vs. Synthesised Layer Info . . . . .	107
4.22	Holographic Replay with Occlusion Synthesis . . . . .	107
4.23	Occlusion Alone vs Occlusion Synthesis . . . . .	108
4.24	Before and After Occlusion Synthesis on a Projector . . . . .	109
4.25	Z Depth vs Deep Rendering for Depth Passes . . . . .	111
4.26	Comparison of ZDepth and Deep . . . . .	112
4.27	Deep Compositing Example . . . . .	112
4.28	Overview of Deep Sample Layer Allocation . . . . .	114
4.29	Layer Slices from Deep EXR . . . . .	117
4.30	Volumetric Scene Render Input and Simulated Holographic Replays . . . . .	118
4.31	Deep Occlusion Synthesis Requirement Cases . . . . .	119
4.32	Field of View Defined Occluded Regions . . . . .	120
4.33	The Disney Multiplane Camera . . . . .	123
4.34	Motion Blur via Convolved FFT . . . . .	126
4.35	Holographic Motion Blur Post-Process . . . . .	126
4.36	Reflection and Refraction Holographic Replay Simulated . . . . .	128
4.37	Reflection and Refraction Holographic Replay on Physical Display . . . . .	129
5.1	Eyebox Size Comparisons . . . . .	132
5.2	Hologram Hogel Basics . . . . .	133
5.3	Lightfield Capture & Display . . . . .	135
5.4	Multi-View in a Multi-SLM Panel via Eye Tracking . . . . .	139
5.5	Comparison of Conventional, Holographic & Discrete Hybrid Rendering . . . . .	142
5.6	Discrete Hybrid Holography Processing Overview . . . . .	144
5.7	Zern Shifting and Tiling System . . . . .	145
5.8	Off Axis Perception of Tiles . . . . .	146
5.9	Continuous Surface Perception Variations . . . . .	148
5.10	Reflection Complexity Comparison . . . . .	149
5.11	Occlusion of Reflections . . . . .	150
5.12	Diffuse, Reflective and Refractive View Sampling Variance . . . . .	151
5.13	View Layer Allocation . . . . .	152
5.14	Pixel vs Speckle Resolution Comparison . . . . .	153
5.15	Defining Near Clipping Safety for Render . . . . .	155
5.16	Focal Length, Angle of View and Camera Back Size Relationship . . . . .	156
5.17	Virtual Camera Rig . . . . .	157
5.18	Temporal vs Spatial View Traversal . . . . .	157
5.19	Sparse Input View Test Image . . . . .	158

---

5.20	New View Synthesis Overview . . . . .	159
5.21	Multi-View Compositing in Nuke . . . . .	161
5.22	Tile Multiplexing and Beaming Display . . . . .	165
5.23	Planar vs Curved Holographic Display . . . . .	166
5.24	Example Input Image Channels . . . . .	168
5.25	Input Images vs Output Results . . . . .	169
5.26	Example Display System Design . . . . .	170
5.27	Source Render vs At-Eye Holographic Defocus . . . . .	170
5.28	Full View and Focus Range . . . . .	171
5.29	Stereograms of Varying Focus . . . . .	172
5.30	Iris Size Comparisons . . . . .	173
5.31	Ball Lens Focal Length . . . . .	174
5.32	Layers Exploded View . . . . .	175
5.33	Image Doubling & Occlusion Issue . . . . .	176

# List of tables

3.1	An outline of Holographic HDR benefits and considerations. . . . .	39
3.2	Production Environment Processes . . . . .	63
4.1	Holographic Compositing Operations Enabled via Occlusion. . . . .	92
4.2	Render Passes for Composite . . . . .	101
4.3	In-painting Methods and Quality of Output . . . . .	102
4.4	Summary of Scene Source Considerations for Layer Slicing . . . . .	125
5.1	Comparison of Display Configurations . . . . .	164



# Nomenclature

## Acronyms / Abbreviations

2.5D Two and a Half Dimensional

2D Two Dimensional

3D Three Dimensional

ACES Academy Color Encoding System

AOV Arbitrary Output Variable

AoV Angle of View

APAS Accurate Phase Added Stereograms

AR Augmented Reality

ASIC Application Specific Integrated Circuit

BFL Back Focal Length

CCT Correlated Colour Temperature

CG Computer Graphics

CGH Computer-Generated Holography

CIE International Commission on Illumination

CIHD Coarse Integral Holographic Display

CLI Command-Line Interface

CMF Colour Matching Function

DCC	Digital Content Creation
DCP	Digital Cinema Package
EDR	Entertainment Dynamic Range
EFL	Effective Focal Length
EOTF	Electro-Optical Transfer Function
EXR	Extended Range (Image Format)
FFT	Fast Fourier Transform
FoV	Field of View
FPAS	Fast Phase Added Stereograms
FPGA	Field Programmable Gate Array
GPU	Graphics Processing Unit
HDR	High Dynamic Range
HFH	Hogel Free Holography
HLG	Hybrid-Log Gamma
HMD	Head Mounted Display
HQH	High-Quality Hologram
HS	Holographic Stereograms
HUD	Heads Up Display
IDT	Input Display Transform
IMF	Interoperable Master Format
IO	Input/Output
ITU	International Telecommunication Union
JPEG	Joint Photographic Experts Group
LCD	Liquid-Crystal Display

---

LC	Liquid Crystal
LCoS	Liquid Crystal on Silicon
LED	Light Emitting Diode
MCGD	Measured Colour Gamut Data
MLA	Micro-Lens Array
SXR	Multi-view EXR
NeRF	Neural Radiance Fields
OCIO	OpenColorIO
ODT	Output Display Transform
OETF	Opto-Electrical Transfer Function
OIO	OpenImageIO
OTT	Over-the-Top
PAS	Phase Added Stereograms
PPD	Pixels Per Degree
PQ	Perceptual Quantizer
RCA	Radio Corporation of America
RGB	Red, Green and Blue
RGBZ	Red, Green, Blue and Z [axis] (depth)
RRT	Reference Rendering Transform
SDR	Standard Dynamic Range
SLM	Spatial Light Modulator
SMTPE	Society of Motion Picture and Television Engineers
SXR	Stereo EXR
TCGD	Target Colour Gamut Data

TV Television

UI User Interface

VFX Visual Effects

VHS Video Home System

VR Virtual Reality

WCG Wide Colour Gamut

# Chapter 1

## Introduction

Computer-Generated Holography (CGH) is a technique for the reconstruction and display of three-dimensional imagery through the diffraction and interference of light. A hologram can be thought of as a generalised diffraction grating - which when illuminated by coherent light of specific wavelengths, will “replay” a full 3D scene. CGH enables a hologram to be generated digitally by calculating light propagation, which can then be displayed on a device called a Spatial Light Modulator (SLM). When this display system is viewed, a full natural 3D scene with visual cues such as defocus and parallax [32] can be perceived - this has been described as “the most complete and visually satisfying 2D record of a 3D scene we know how to make” [13].

Whilst holography has received a significant amount of recent attention from both academic and industrial bodies [165, 183, 177], there is little in the way of work done to practically integrate into existing media production and computer graphics techniques, and even less work done taking into account artistic considerations. This thesis bridges the gap between production practice and holographic generation techniques. Ultimately, holographic display, as any other display, is a window into a specific artistic intent - and this work focuses on introducing advanced techniques to improve and develop the viewing experience whilst ensuring control over the result remains with the artist. Though holography is not new by any means, first developed by Dennis Gabor [68] in the late 1940s, it is still yet to mature - and this work presents a route into what digital holography could be as a widely adopted display medium in the future.

### 1.1 Research Problems Overview

Advancements in computer graphics (CG) have built up over many years to provide widely used tools and techniques that realistically simulate how light illuminates 3D scenes. Whether

this is via ray-tracing, rasterisation or otherwise - images which appear to be indistinguishable from real scenes can be created, with increasingly reduced render times [20], and at increasing output resolutions. The data presented by this rendering process may exist in some kind of 3D representation until it reaches the final stages of the imaging pipeline, where it is flattened and output to a 2D display, either directly or via a form of delivery network.

Advancements in hologram rendering have similarly tried to replicate the success of physically correct light simulation in CG, and attempt to turn 3D input scenes into realistic imagery - but this time preserving the additional dimension of depth for display. Research has so far focused on creating light simulations via wave-based [229] and ray-based [91, 32, 29] models. When the results of the images produced via these methods are compared to the state-of-the-art in computer graphics, these attempts clearly do not appear to meet the quality or feature-set expectations of a modern production renderer [178], with limited shaders and geometry type support, increased render times and therefore more powerful hardware requirements [162, 186].

This is where “layer-based” or “multi-plane” holograms have succeeded in the research - rather than simulate directly from a 3D scene, layer-based holograms are composed of inputs from the traditional renderer, be this a 2.5D RGBZ representation, or a more discrete selection of image depth planes. This produces imagery which, given a hologram generation algorithm and display hardware of sufficient quality, can now reproduce as much realism as the production renderer can provide at interactive or real-time frame rates. This is, of course, not without a catch.

Layer-based holograms, using a single 2.5D image, rendered from a single perspective per frame as their inputs, can provide a reasonable approximation of the input scene, but it remains exactly that - an approximation. Problems with holograms generated with this technique, even when not limiting the number of layers for compute/speed purposes, include:

1. Visible gaps between layers, visible in both parallax and defocus - a failure to correctly handle occlusions.
2. No support for semi-transparencies - which includes semi-transparent objects, solid object edges with anti-aliasing (also inclusive of fur/hair/grass), motion blur, volumes such as smoke/fog.
3. No support for optical effects such as depths via reflective or refractive surfaces.
4. No support for view-dependent effects or multiple viewing perspectives.

This work responds to these drawbacks, but in order to do so, must first present a complete workflow method for media production for CGH. Features which are considered basic in

media production and Visual Effects (VFX) production, and will provide the foundation to which to build any developments, are missing from the current state of the art in CGH. At the outline of these research problems, CGH research has seemingly never utilised standard production practices, only recently considered how High Dynamic Range (HDR) or Wide Colour Gamut (WCG) may be implemented, and not considered the possibility of artistic expression provided by the additional technical controls. This work proposes solutions for the workflow requirements of a successful production for holographic display, and this is then used to develop advancements in layer-based hologram generation.

This thesis, by improving the viewing experience and creation experience of holographic display, and providing improved and increased amounts of data to the hologram generation algorithm, with additional tools for artistic control - allows the most advanced form of 3D display known to look better than ever.

## 1.2 Company Description

VividQ is a software startup in Cambridge, UK, with expertise in computer generated holography. It offers a complete software platform for enabling holographic display on consumer devices that have only standard computing power, which for the first time brings the possibility of widespread CGH of any kind. Using a network of partnerships and collaborators including manufacturers and academic institutions alike, VividQ has been developing both state-of-the-art algorithms and hardware to bring realistic 3D display to market, be it in Augmented or Virtual Reality (AR/VR) Head Mounted Display (HMD), Heads-Up Display (HUD), or larger format displays. VividQ's software, which provides Graphic Processing Unit (GPU) accelerated image/layer based holography, can deliver real-time frame rates to the viewer, and has demonstrated such via various successful prototype projectors and headsets.

## 1.3 Experience at VividQ

Working at VividQ has been certainly both exciting and a steep learning curve. I had little in the way of knowledge of real diffractive holography when I first began the process of joining, and holograms, as I write this at the end of the 3 years of placement, still seem like magic.

Coming from an undergraduate in Computer Animation & Visualisation at the NCCA, it was clear from day one that despite the infinite wisdom and knowledge of the talented team at VividQ, which I am extremely thankful for - no one really had any idea how animation, film or games actually get made, and this is something that is also reflected in the research.

So as much as I hate to admit that my supervisor at VividQ, Andrzej, was right, holography could learn a lot from computer graphics and media production. I was thrown in the deep end during my first few months, tasked with building a selection of interactive experiences in the game engine Unity to display on a new prototype AR headset, “headsetv2”. This gave great insight into VividQ’s tools, practices, assumptions, and what worked well about holography on real hardware - but mostly what didn’t work well. This experience greatly proved out a lot of the issues described in the initial research proposal and helped to visualise the problems and prompt me to come up with solutions.

From there on, having a good understanding of how the issues behaved in practice, and having learnt the internal tools, I started by making file format conversion tools to take multi-channel renders and turn them into VividQ’s 2.5D format “CDM” (colour depth map). During this process, and seeing what happened if I changed the RGB channels for the reflection pass from the renderer; the lightbulb moment happened: “Oh, layer-based holographic display is a compositing problem!?”. If your favourite compositing software of choice operated in Fourier space - you’ve basically got a highly configurable hologram generation algorithm. After outlining the end-to-end process, comparing what happened in VividQ’s 8bit Standard Dynamic Range (SDR) pipeline against how a full scene is composited for high-end delivery, from plate to display, I could then decide what to do with my selection of VFX-inspired tools to make the display problems outlined go away.

Ultimately this work has led to a number of papers, a patent, and some techniques I’m quite pleased with - which have been implemented in various forms into VividQ’s software offering.

## 1.4 Contributions

This work solves a number of the issues described in section 1.1. The contributions offered by this thesis consist of:

1. The proposal of a complete theoretical framework: involving the creation of a production pipeline workflow and the associated appropriate tools, that allow ingestion of digital media from various sources whilst automating the preparation of image data for hologram generation which conforms with current media standards and practice. This includes:
  - (a) HDR handling recommendations, high luminance support
  - (b) Colour management recommendations, wide gamut support
  - (c) “High-Quality Hologram” (HQH) working and delivery format specification

- (d) Artistic optical controls (aperture shape, lens speed, aberrations)
- (e) Renderfarm integration
- (f) Holographic display calibration

All of which allow the further sections to be developed, and collectively assist in improving the creation and viewing experience. Previously, there had been no practical considerations for production implementation of CGH, and this is the first work to consider colour and HDR management for CGH, as well as the first to consider the control of qualities of the display for artistic purpose.

2. The creation of novel methods for supporting multi-sampled depths in holographic display from production renderers and compositing packages. Developing use of explicit layer rendering, layer synthesis and Fourier-space "Holographic Compositing" to allow the physically correct display of complex elements such as: semi-transparencies, materials with optical properties and of volumetrics via Deep formats - all with holographic occlusions. Artefacts exhibited by layer-based holography and issues shown in the literature are discussed and consequently resolved by unique methods proposed here.
3. The demonstration of a novel display algorithm for the input of multiple render perspectives for large-format hologram generation, which allows omni-directional viewing with reduced computational requirements and natural view interpolation. The development of a holographic multi-camera rendering rig in Digital Content Creation (DCC) packages to enable this method is also detailed, as well as view image storage techniques following the proposed HQH format. This standard format support ensures all of the other developments in this thesis can be applied to future large-format holographic displays.

The sum of the contributions presented here bring the renderer and viewer closer together than ever by integrating computer graphics techniques more deeply with computer-generated holography methods. An overview of the novel contributions of this thesis are outlined in figure 1.1, along with their associated chapter.

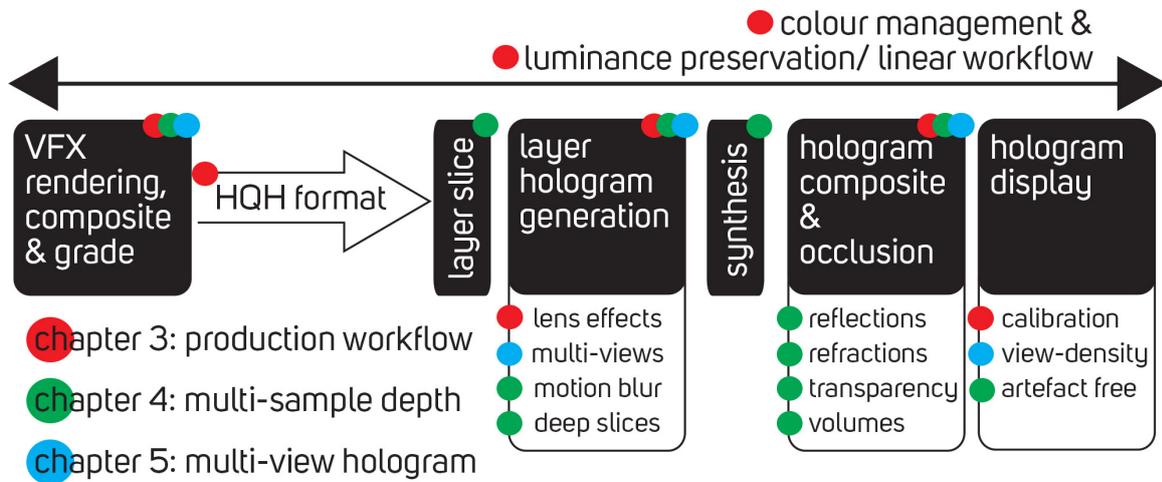


Fig. 1.1 An overview of the contributions of this work and where they fit in the holographic media production process.

## 1.5 Thesis Outline

This thesis is structured as follows. Chapter 1 provides an introduction to the problem area and the projects completed. Chapter 2 presents the background required to understand fundamentally what holography is, how it looks, and what it can do for digital media. Following this are three chapters, 3, 4 and 5, each analysing a specific problem area, which present the project work and results - these encompass projects covering production workflow and getting holography up to date with 2D practice, multi-depth support for layer based CGH, and multi-view holography respectively. Chapter 6 discusses the results overall along with potential future works. The final chapter, Chapter 7, provides a brief conclusion.

## 1.6 Publications

Publications which resulted from this work, listed in order of thesis appearance:

1. Demolder A., "Toward the Standardization of High-Quality Computer-Generated Holography Media Production Workflow", in SMPTE Motion Imaging Journal, vol. 131, no. 1, pp. 48-58, Jan.-Feb. 2022, doi: <https://doi.org/10.5594/JMI.2021.3130941> [47]

Society of Motion Picture and Television Engineers (SMPTE) Motion Imaging Journal (IEEE) Technical Paper. Awarded Best Student Paper 2021. This paper is part of the topic covered in section 3.2.3, which presents techniques related to the end-to-end

production workflow for CGH. A portion of this work was also invited to be presented to the Moving Picture Experts Group (MPEG) Ad-hoc Group "Render-based 3D Technologies" to help define immersive display requirements.

2. Demolder A., Kaczorowski A., Newman A., Nait-Charif H., "Artistic Control of Defocus in Computer-Generated Holography", presented at ACM CVMP '20: European Conference on Visual Media Production, 2020. <https://www.cvmc-conference.org/files/2020/short/11.pdf> [49]

CVMP 2020 Short Paper and Presentation. Awarded Best Short Paper. The topic of artistic optical control, which this paper is a part of, is presented in section 3.6. Paper presentation available: <https://www.youtube.com/watch?v=xk0gjvL9IAY&t=408s>

3. Demolder A., Newman A., Durrant T., Nait-Charif H., Adzhiev V., Kaczorowski A., "Enabling Reflective & Refractive Depth Representation in Computer-Generated Holography" Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH '21 Posters), August 09–13, 2021, ACM: New York, NY, USA, 2 pp. doi: <https://doi.org/10.1145/3450618.3469177> [50]

ACM SIGGRAPH 2021 Poster & Presentation. Invited panel speaker to the "Perception & Displays" Technical Papers session. Awarded as winner of the ACM SIGGRAPH Graduate Student Research Competition 2021. This paper is part of the topic covered in section 4.3.2, which presents methods of enabling holograms generated with multi-sampled depths. Presentation and supplementary material available via doi link above. This work was also invited to be presented in an SPIE (International Society for Optics and Photonics) AR/VR/MR Webinar [63].

4. Demolder A., "Occlusion Synthesis", Method, apparatus, computer program and computer-readable medium for determining a CGH. Patent Pending: GB2314566.7. [48]

A patent, filed by VividQ Ltd. detailing the method and apparatus by which holograms can be generated from sparse 2.5D inputs whilst preventing visual artefacts in both single-perspective and multi-perspective display systems. This patent combines techniques detailed in section 4.4.2, where input imagery has additional scene information synthesised to fill gaps left by occlusions, and Chapter 5, where a method of inputting multiple views of a scene allows a large viewing window. Combined, this allows artefact-free viewing whilst providing sensible controls over the quality of the viewing experience versus time required for compute.



# Chapter 2

## Background - An Introduction to Holography

### 2.1 What is Holography?

Holography is a technique which enables a wavefront to be recorded and later reconstructed. It presents a method of encoding 3D images [81], that when illuminated by coherent or semi-coherent light of specific wavelengths, will diffract the input light in such a way to reconstruct an image in the resultant field. Imagery forms when the diffraction and propagation of this light creates a wavefront that constructively and destructively interferes with itself. The resulting wavefront of light is natural in that the expected visual cues such as accommodation, parallax and occlusion are preserved [32] - resulting in a scene projection that the viewer can focus on and move within freely.

When viewing the perfect holographic display, if designed with a sufficiently wide-spectrum illumination technique, high resolution and high refresh rate, the viewing experience would be no different to observing the real-world. Objects viewed via the hologram will appear at a physical location in space, defocus via the same mechanism as the real-world defocuses, and occlude other holographic objects correctly.

#### 2.1.1 Fundamentals

Holography has existed since Dennis Gabor proposed it in the late 1940s [68]. From first using film and a two-beam recording method, the concept has been slowly developing since - but the fundamentals remain the same by making use of the wave-like properties of light, which makes use of the effects demonstrated by Young's double slit experiment [226].

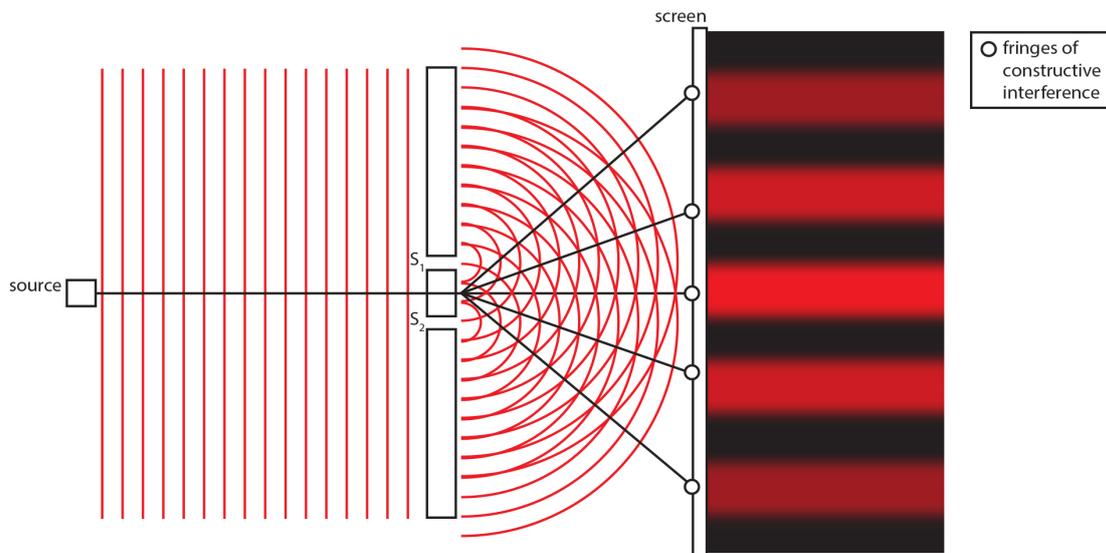


Fig. 2.1 The Double Slit Experiment. An illustration of the diffraction and interference of light produced by two narrow slits  $S_1$  and  $S_2$ . Bright and dark interference fringes result on the screen.

Figure 2.1 demonstrates this interaction between two wavefronts, diffracted from a coherent source, where interference between the waves causes “fringes” of light to appear as the light is diffracted by the two apertures. Diffraction is the process by which a wave spreads out after passing through a narrow aperture, provided the wavelength of light was small enough to do so. Bright spots appear where constructive interference occurs, summing their amplitudes, and dark spots where it is destructive, cancelling out the amplitudes - resulting in no light.

The image of the fringes has been encoded by the two slits, and reconstructed when illuminated with coherent light. Holography works under this same principle - a number of these slits/apertures together are referred to as a diffraction grating, a diffraction pattern, or simply a hologram. By adding more apertures, more complex images can be formed, shown in figure 2.2.

### 2.1.2 Holography with Photographic Film

Earlier examples of holography use the fundamental principles described in the previous section to capture the interference patterns of physical scenes onto a film plate, so that they may be replayed by analogue means.

Figure 2.3 shows a typical recording process setup. Coherent light illuminates a 3D scene via the illumination wave, which when reflected from the scene forms the “object wave”, this

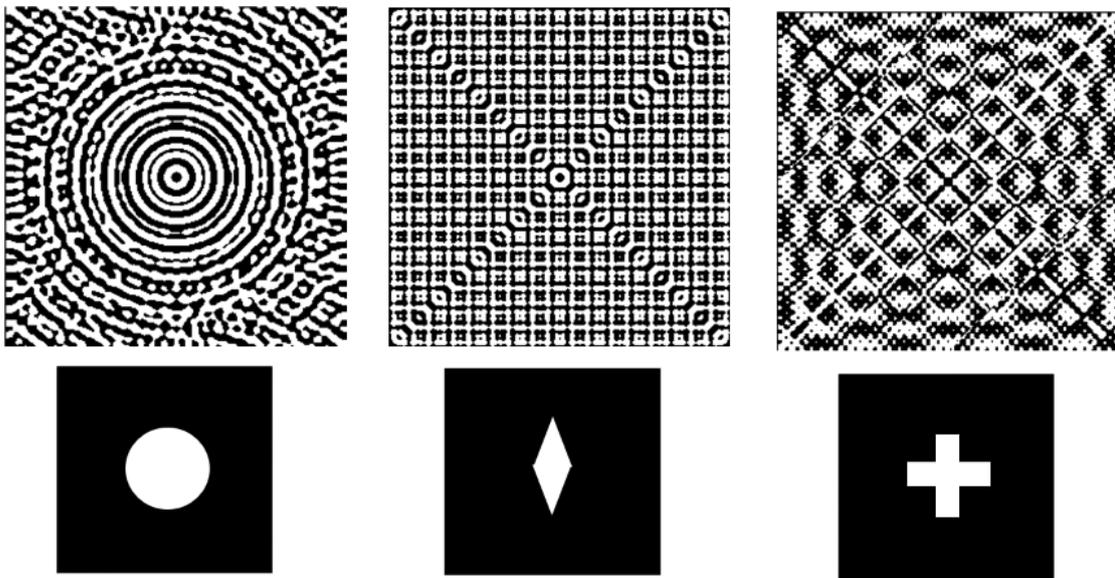


Fig. 2.2 Simple diffraction patterns (holograms) and their resulting replay images. The pattern must be illuminated by a coherent source to form the replay field - here, images of shapes. Taken from Prof. Tim Wilkinson's presentation "Holographic Projection Displays: Beyond Star Wars" [218].

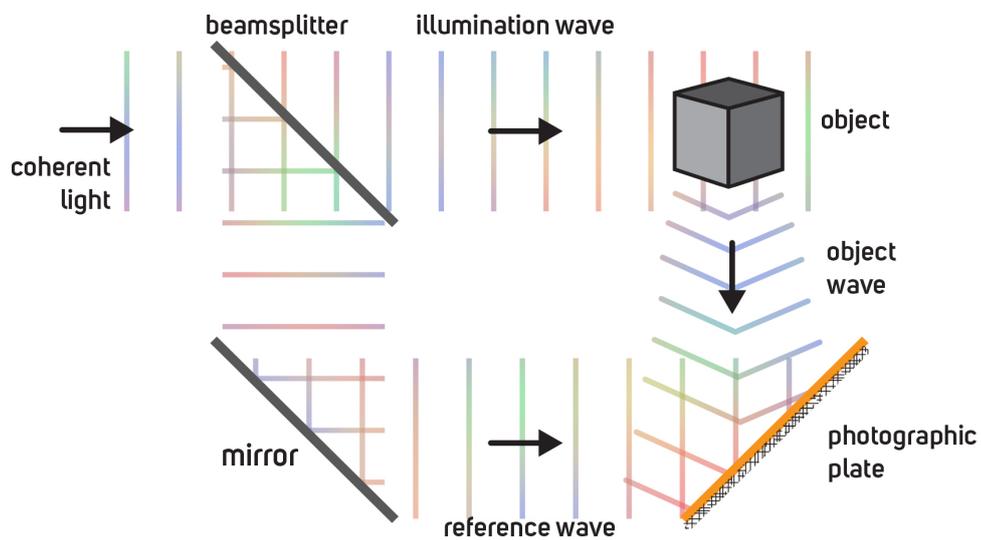


Fig. 2.3 An overview of the holographic recording process, showing how the reference wave interferes with the object wave and is recorded on the photographic plate. Diagram adapted from [163, 145].

wave contains the amplitude and phase of the scene. The object wavefront then interferes with the wavefront reflected from the source via the reference beam. These two light sources form the name of the method “two-beam recording” - the interference between the two waves, or beams, is then recorded by the photographic film plate. Phase variations of the wavefront are translated into amplitude variations [106]. The film, after being fixed and developed, allows later illumination via a copy of the reference beam - this beam diffracts through the physical interference pattern formed by the film grains attached to the plate, shown in figure 2.4. In practice, several copies/components of the intended image are formed (known as diffraction orders), but are unwanted and discarded/blocked by the optical design [73].

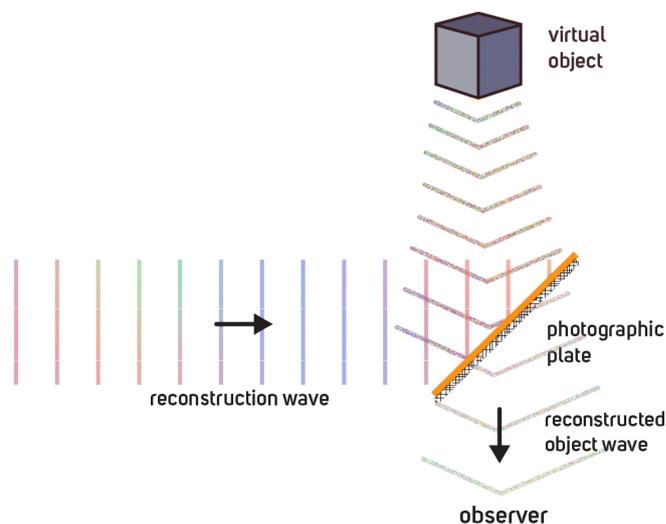


Fig. 2.4 An overview of the holography replay process, showing that when the photographic plate is be illuminated by a copy of the reference beam, as shown in figure 2.3, now named the reconstruction wave - a reconstructed object wave can be observed. Diagram adapted from [163, 145].

When viewed, this film provides a full three-dimensional reconstruction of the original scene, re-creating the original object wavefront. This may be done with a single wavelength, or multiple recording steps may occur for multiple wavelengths in order to make RGB metamer [66] holograms, or occur for wider spectrums/white light. An overview of the recording and reconstruction process is provided in figure 2.5, showing a scene (the cat) and a recording plane. The recording plane may be treated as a window to the scene, and captures the scene information. Given sufficient illumination of the recording medium, the original scene which has been captured will be recreated by the interference recorded in the plate, and is observable by the viewer. Given that the wavefront provided by the plate is identical to

the original, the viewer can focus freely on various objects within the scene, observe parallax and change viewing positions to look behind objects.

An in-depth analysis of this entire process is presented in multiple sources [68, 73, 13, 163].

Holograms are frequently used as security/authenticity marks, most notably on bank cards (such as the Visa Dove, shown in figure 2.6), due to the complexity of the process required to replicate them - with the 3D element of the image being unable to be replicated via a simple photocopy. There are also many one-off holograms viewable in certain galleries, in the UK, a collection is viewable on permanent display at the Camera Obscura in Edinburgh.

There have been previous attempts to utilise classical holography to deliver media to consumers before, most notably by RCA in the 1970s, with the HoloTape system, “A Low-Cost Prerecorded Television System Using Holographic Storage” [79]. Here, a full system of recording, storing and playing back holograms with a home player was proposed, where holograms embossed into plastic tape allowed greater redundancy than other storage mediums such as film reels or magnetic tape. Though ultimately the system was never commercially available, and was abandoned with VHS tape becoming the dominant home video format.

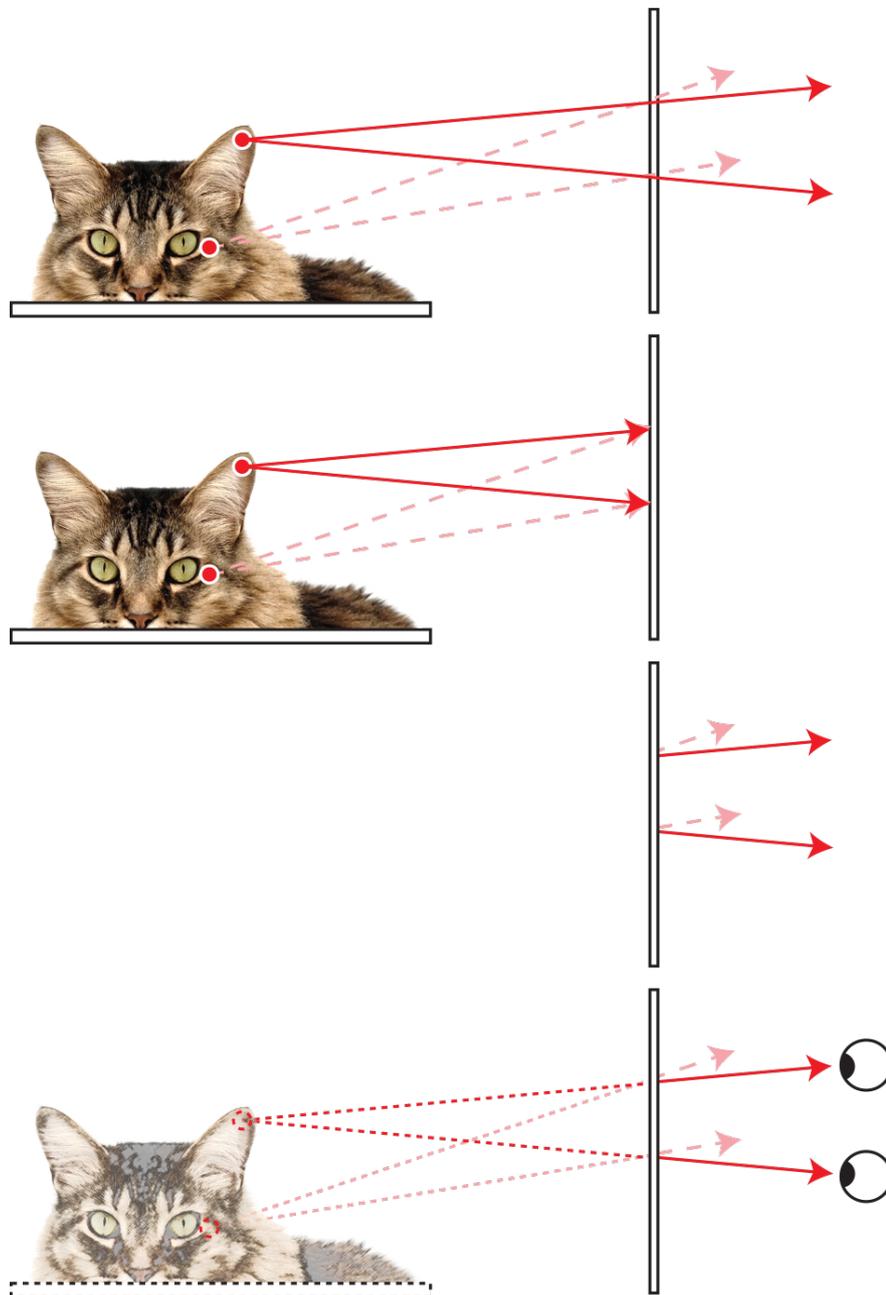


Fig. 2.5 A simple breakdown of what is observed when looking at a hologram. Top, viewing a cat through a window or film. Second and third show the incident and transmitted (refraction result) rays. In a hologram, we capture the light information hitting the window (or film). Last, we see the observer viewing the hologram result via the captured light information, still observing the object as if it were present behind the window whilst preserving the correct perceptual cues including convergence and accommodation. Diagram adapted from [13].



Fig. 2.6 VISA card hologram of a dove. The diffraction pattern is embossed into a metallic sticker, and in this image is illuminated by daylight.

### 2.1.3 Computer-Generated Holography

With the advancements in computer processing and other hardware in recent decades, trading a complex physical recording and reconstruction process for a convenient digital process is possible. The first Computer-generated Hologram was made by Lohmann and Paris in 1967 [134], where the calculation of light diffraction and interference occurs digitally. This resulting hologram can then be printed and recorded onto a photographic plate to be illuminated classically, or displayed on a Spatial Light Modulator (SLM), which is an electronically controlled device which can change the diffraction pattern it displays. Imagery from digital holograms may also be reconstructed in simulation.

As we are most interested in how CGH forms observable images, here diffraction is initially explained from the perspective of forming a replay field from a hologram. The computation of this diffraction and interference may be carried out digitally via the use of scalar diffraction theory [73, 106, 13]. The goal of the computation being to determine the resulting electric field of the observer/replay plane  $(x, y)$  after travelling some distance,  $z$ , from the hologram plane  $(u, v)$ . The Rayleigh-Sommerfeld diffraction formula provides a solution to this if the intended field is treated as a combination of points in 3D space.

$$\psi_z(x, y) = \frac{1}{i\lambda} \iint H(u, v) \frac{e^{ikr}}{r} \cos\theta \, dudv \quad (2.1)$$

Where:

- $\psi_z$  = Electric field in the replay plane at distance  $z$
- $H$  = Complex representation of the electric field at the Hologram plane
- $(x, y)$  = Replay plane coordinate space

- $(u, v)$  = 2D hologram coordinate space
- $r$  = Distance between the points  $(x, y)$  and  $(u, v, 0)$
- $\lambda$  = Wavelength
- $k = \text{Wavenumber } (\frac{2\pi}{\lambda})$
- $\theta$  = Angle of the wave-vector in relation to the normal of the hologram plane

This is shown in figure 2.7, where the hologram is a complex electric field in both the  $(u, v)$  plane and  $(x, y)$  plane after propagating distance  $z$ . [106] The double integral signifies that every point within the holographic replay field  $(x, y)$  is influenced by the contribution from every individual point across the two-dimensional hologram  $(u, v)$ . There may be more than one 2D plane in the replay field  $\psi(x, y, z)$  represented by a single hologram plane  $H(u, v)$ .

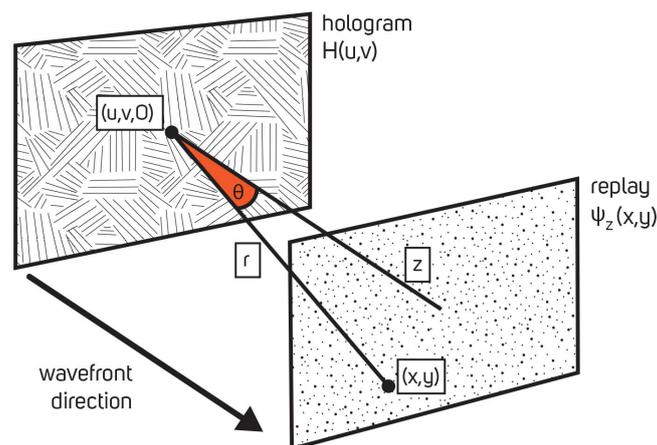


Fig. 2.7 The propagation of a wavefront through  $z$  from the hologram plane to a point  $(x, y)$  in the replay field.

This form of diffraction alone may be used to calculate the complex field required of a computer-generated hologram, but requires too much computation to be practical in most cases currently (excluding specialist hardware [220]), so further approximations are typically used. These approximations make use of the fact that as the waves propagate from the aperture, the wavefronts become more planar, of course relying on the viewer being at an appropriate distance for this to hold true, and consequently will not be possible for all applications. Figure 2.8 shows the Fresnel and far-field regions, where the Fresnel approximation and Fraunhofer approximations may be used respectively.

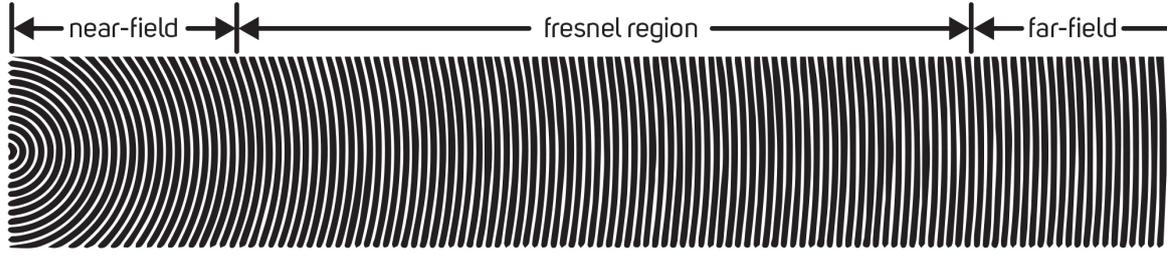


Fig. 2.8 Shows how as the wavefronts propagate across space, the sampling within a given space allows differing approximations to be made as to the wavefront. The near-field is represented by spherical waves, parabolas in the fresnel region, and planes in the far-field.

The Fresnel approximation assumes that the distance between the Hologram plane and the Replay plane is large enough that the spherical wavefronts may be represented by parabolas. This assumption can be made where,  $z^2 > u^2 + v^2$ , where  $\cos(\theta)$ , the obliquity factor, is equal to one, and therefore that  $r = z$ .

$$\psi(x, y, z) = \frac{e^{ikz}}{i\lambda z} e^{i\frac{\pi}{\lambda z}(x^2+y^2)} \mathfrak{F} \left\{ H(u, v) e^{i\frac{\pi}{\lambda z}(u^2+v^2)} \right\} \quad (2.2)$$

In the far-field, where  $z^2 \gg u^2 + v^2$ , the wavefront can be approximated as a plane using the Fraunhofer approximation. This implies that the image observed at an infinite distance from the viewer is essentially the Fourier Transform of the input complex hologram with a phase change:

$$\psi(x, y, z) = \frac{\exp \left[ ik \left( z + \frac{x^2+y^2}{2z} \right) \right]}{i\lambda z} \mathfrak{F} \{ H(u, v) \} \quad (2.3)$$

The reasonable distance ( $z$ ) at which such a hologram is observable is related to the width of the aperture (hologram),  $D$ , and is given by the ‘‘Antenna designer’s formula’’ [73].

$$z > \frac{2D^2}{\lambda} \quad (2.4)$$

Consequently, if suitably illuminating a hologram  $H(u, v)$  which encodes a single point, this point is visible in the replay field at the intended position  $\psi(x, y, z)$ . This reconstruction process, projecting the image and allowing wave interference to occur, is in itself a Fourier transform - the projection transforms from the Fourier/frequency domain into the image/spatial domain. When viewed by eye, this wave propagation in combination with the eye lens performs this transformation [73], transforming the field of light in front of the

viewer into the image domain, which is focused on the retina and consequently perceived by the viewer.

## 2.2 CGH in Practice

Since a Fourier transform provides an appropriate approximation of diffraction which allows the reconstruction of the replay field from the complex hologram field, the reverse can be applied - and the inverse Fourier transform is used to create the initial complex hologram.

$$H(u, v) = \mathfrak{F}^{-1} \left\{ \sqrt{I(x, y)} \right\} \quad (2.5)$$

Where the square root provides the amplitude from the intensity value of the point being transformed (such as a pixel value from a renderer, in linear). Then, the z position of the reconstruction point is determined by the "focus aberration" Zernike Polynomial, which adds a field-independent aberration correction to the hologram, which may be thought of as a digital lens, via a phase factor/profile  $\varphi(x, y)$  [106].

$$H(u, v) = \mathfrak{F}^{-1} \left\{ \sqrt{I(x, y)} \cdot e^{i2\pi\varphi(u, v)} \right\} \quad (2.6)$$

This allows the fast creation of a hologram which encodes pixels at a specific depth position - many individual point position holograms then are summed together in order to merge their respective complex fields and form a single hologram which reconstructs multiple points. But there are still several issues to contend with:

1. Reducing the number of Fourier Transforms that must occur
2. While the amplitude of the input image is taken from the pixel values, the phase information of the input image is unknown
3. SLM devices that can display complex values, and modulate both amplitude and phase, do not exist, and their bit-depth is limited

### 2.2.1 Modern CGH methods - Compute Reduction

The methods described in section 2.1.3 and above in section 2.2 show how diffraction can be approximated for the calculation of propagation for a single point, and consequently form a 3D scene by summing the contributions of potentially millions of points. When considering an entire scene for holographic reconstruction, every point in the resulting replay field needs to interact with each point in the hologram plane. This point-based method is often referred

to as a holographic ray-tracer [91], though some methods use a wave-based model [229]. The benefit with point-based methods is the most accuracy in the replay field, at the cost of the most compute. A dense sampling of points, and consequently a large number of Fourier Transforms, are required in order to form the appearance of continuous surfaces - otherwise the individual sparse points become visible.

There are many methods which reduce the compute requirements of holographic ray-tracing [33, 177, 206], but most gains are made by changing the model by which imagery is represented in the scene, with alternatives being polygon and layer based methods.

Polygon based methods will require about the same calculation time per polygon as per point [28], and given that a polygon (or micro-polygon) will cover a larger region of the image than a point, will require less calculation in total to achieve an improved resulting replay field and the appearance of a continuous surface. There are various methods that have been proposed to increase the speed of such techniques [161, 95].

Given that the compute requirements are tied to the number (and size) of the Fourier transforms performed, layer-based holography (also referred to as plane/image-based) can provide a significant reduction in computation required. This is where a single image, or a number of images composed from planar slices of a 3D scene [33, 30], are generated into holograms - each layer requires only one 2D Fourier transform. The number of depth layers specified dictates the compute requirements, and can provide an efficient trade-off between compute time and accuracy where many points can be placed at the same depth together.

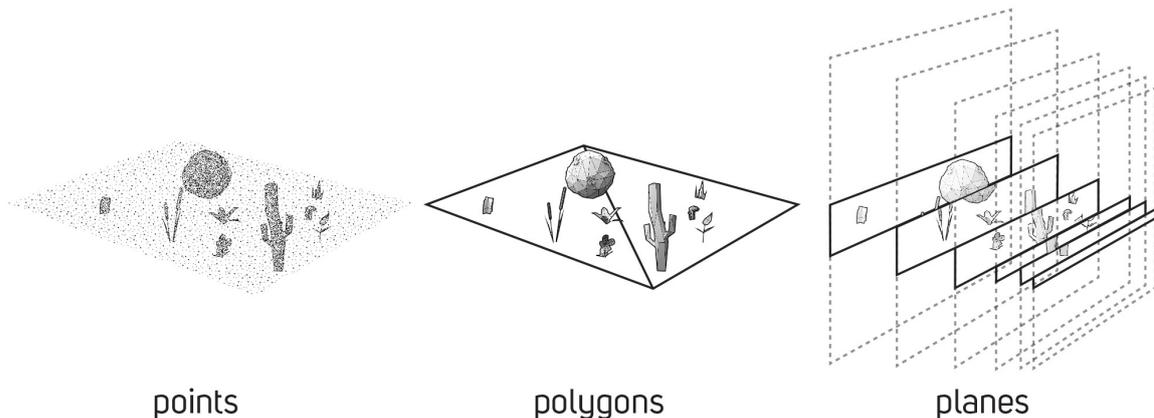


Fig. 2.9 An illustrative comparison between the side view of varying scene representations for holographic display. Each point, polygon and plane requires a Fourier Transform to be visible in the replay field. Points are the most dense representation, sampling every vertex in the scene, with planes the least, splitting a scene into 2D images at discrete depth slices.

Each of these methods, compared in figure 2.9 have their own considerations for scene shading and occlusion. Given that the scene will correctly appear to move in parallax

and the viewer may change the perspective from which they view the image, by moving their head/eye in the replay field, ensuring that specular highlights and objects moving in front of other objects behave correctly is important for the perception of the scene. In ray-tracing and polygon methods, lambertian, specular and other physically-based graphics models must be explicitly implemented to create a realistic image, whereas in layer-based holography, diffuse lambertian imagery is the default in part by the diffusing nature of adding random phase, explained in section 2.2.2. These reflection types are visualised in figure 2.10. Occlusions may be implemented by culling rays [228, 183] or by calculating additional planar/orthographic light propagation through space via Fourier transforms [73] and masking at planes [93, 71] but visual artefacts are frequently still present [24].

An extensive review of CGH for 3D Imaging is found here [177].

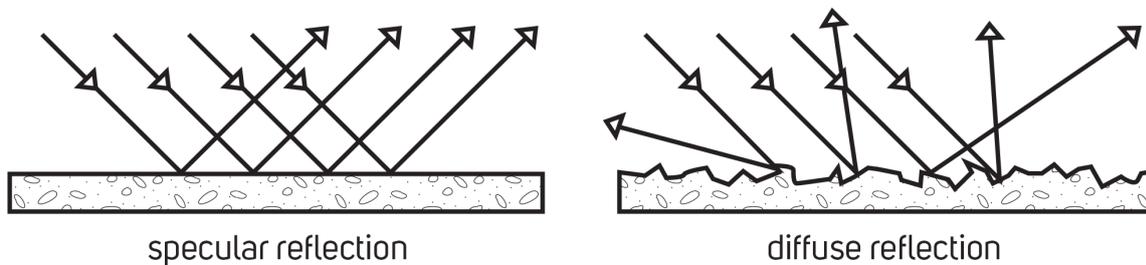


Fig. 2.10 Specular reflections are compared with diffuse reflections. Flat/smooth surfaces such as mirrors provide specular reflections, where the incoming light reflects at the same angle as the surface hit. Diffuse reflections are caused by rough surfaces, and scatter incoming light in many directions. Surfaces typically have both diffuse and specular components.

### 2.2.2 Random Phase

While the amplitude of the input points to the hologram is known - the phase information of the input is not. Fortunately, the eye is insensitive to phase so random phase information can be used in its place. Random phase in a hologram provides a wide diffusion of the object light [185, 134], and this diffuse resultant light helps ensure successful image reconstruction from the coherent source and consequent interference.

The randomisation of phase occurs via the phase factor/profiles introduced in equation 2.6, and is added uniformly across the field. This also allows a level of image redundancy, preventing ringing artefacts from dust, or even allows larger portions of the hologram to be blocked in the optical system [79, 173] while the full image remains visible at a reduced resolution. For display applications, the phase of the object field is far less important than the amplitude - hence the acceptable use of random values of phase to equalise the amplitude

across the replay field [106]. However, it also introduces significant noise into the image unless mitigated for.

The investigation of methods for phase retrieval, image optimisation and general noise reduction is a significant area of research in CGH, and includes techniques in the areas of:

- Higher resolution holograms, to produce more variation in noise in a given angular view[209], requires a higher resolution display device.
- Time-multiplexed imagery with varying random phase [22, 19], to average out the noise over time for it to not be perceivable [72], requires enough compute for a high frame rate.
- Iterative, trained or alternative algorithms to retrieve an optimised phase. Such as the Gerchberg Saxton algorithm [69], or recent learned/deep learning/neural network based hologram generation methods [113, 183, 165, 166]. A review of such alternatives is given by [222].

### 2.2.3 Display Hardware

#### Standard 2D Display

Modern conventional displays, such as Liquid-Crystal Display (LCD) panels, are generally made up of individual pixels that emit light via some means to produce a colour, and are directly observed. Here, an incoherent backlight (such as by a Light Emitting Diode (LED) array) produces a moderately wide spectrum illumination. This light then passes through a glass substrate (GS), and is blocked to varying levels by the Liquid Crystal (LC) layer, which controls the brightness of the sub-pixel. This then passes through colour filters (RGB), with 3 sub-pixels making up a single colour pixel and forming the intended RGB pixel value.

In order to successfully control the output colour values on the display via the pixel array, and to send the correct values to the LC from a device such as a graphics card, the following must be considered:

- The display quantisation - how many different values can the LC represent? Today, most conventional displays are 8-bit, and can display 256 values per colour.
- The Electro-Optical Transfer Function (EOTF) of the LC - how do the electrical signals correspond to an optical output? Conventionally this is called the "gamma" of the display and is defined by a curve, and determines how the display brightness appears across the range of pixel values.

- The Colour Space - determined by a combination of the illumination and the RGB "tristimulus" colour filters. Knowledge of what range of colours can be displayed at the output device allows the input graphics pipeline/device to transform imagery to look as close to as intended for any display capability.

### SLM Devices

In comparison, an SLM such as a Liquid Crystal on Silicon Device (LCoS)[127] will use its liquid-crystal layer to vary the refractive index of a pixel, and consequently alter the speed of the light being reflected from it, resulting in a potential phase change in the resultant light. Where many of these pixels are providing varying refractive indices, it causes an input coherent wavefront to output with modulations - and form the electric field according to the hologram displayed. The resulting wavefront then propagates, and interferes, to produce the replay field. A comparison diagram between an LCD and an SLM is shown in figure 2.11 - this shows a reflective SLM, with a coherent wavefront entering the device through the glass substrate, being modulated by the LC, and then are reflected out of the device with a new wavefront.

An SLM will display a hologram for only one wavelength (and therefore colour) at any given time, multiple wavelengths must be time-multiplexed. An SLM will have a specification wavelength operating range, and also a "gamma" which here results in being a phase modulation correction curve, rather than one changing brightness. Here, each LC cell is a pixel and displays the phase values only from a quantised complex hologram.

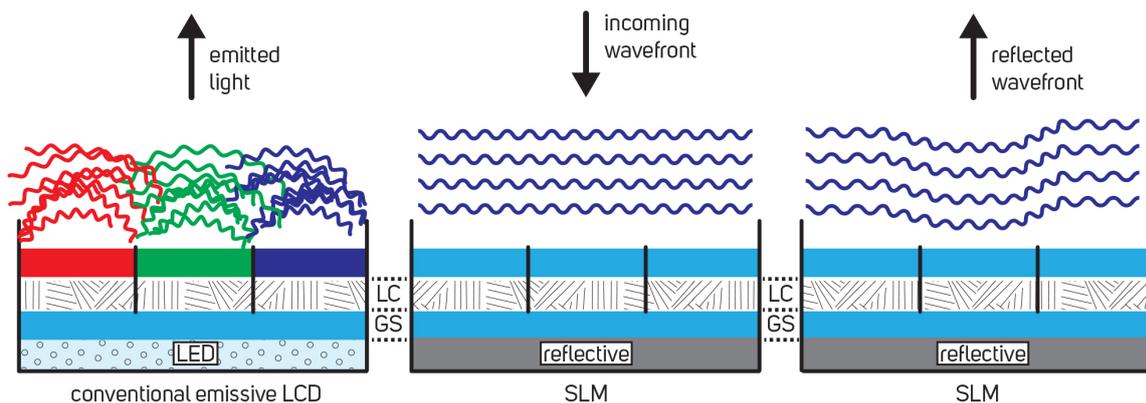


Fig. 2.11 An overview of the differences between a typical LCD and a phase modulating SLM. Light emitted by the incoherent source LED behind a glass substrate (GS) is blocked via a liquid crystal (LC) on a colour sub-pixel basis to form the resultant pixel colour emission via filters on the display. The SLM instead modulates the incoming wavefront using the LC, and reflects an altered wavefront according to the quantised hologram phase values.

### Hologram Quantisation

Whilst the calculated hologram is a complex-valued function, none of the existing SLM devices are able to display such values. An SLM device, which can display the hologram needed to produce the replay field - can only modulate phase or amplitude, not both. An ideal projection system would modulate both in a single device to display a full complex hologram, but in place, a pair of SLMs may be used [181] where one modulates phase and the other amplitude, or methods such as "Double-phase" may be used, which attempt to simulate complex values using two phase pixels [89]. Otherwise, in order to then display a hologram on a single SLM, it must be quantised to match the SLM architecture - this process introduces errors to the hologram and leads to additional noise [22]. The effects of noise introduced by quantisation can be reduced via similar methods to the reduction of noise in the field introduced by random phase shown in section 2.2.2.

The exact quantisation scheme required is dependant on the SLM capability. "Binary Phase" devices provide two possible phase states and can update quickly (usually kHz), and "Continuous Phase" devices allow a range of values, typically defined by a bit depth, such as 256 phase states at 8 bit, but are slower and can refresh on the order of milliseconds.

#### 2.2.4 Example Display System

An example near-eye projection system containing an SLM is shown in figure 2.12. The diagram shows the optical path from the laser illumination to the SLM, where a hologram pattern is displayed, causing the light to interfere to form the intended image, which is consequently relayed to the viewer. This also shows a semi-reflective beamsplitter, allowing imagery to be overlaid with a view of the real-world as per Augmented Reality Glasses - but can be viewed directly at  $H$ . Bottom right shows such a configuration, where the optical path is directly viewed - the green reflection of the replay can be seen in the viewer's eye. This projection system design is per-eye, and must be duplicated to form a stereo image.

The intended imagery which has undergone computation and displayed as a hologram on the SLM device is viewable on such display systems, with figure 2.13 showing an input image and the resulting output replay. As a hologram is dependent on the wavelength on which it is calculated, each colour, in this case red (632nm), green (520nm) and blue (450nm), is displayed and illuminated time sequentially, utilising persistence of vision [80] to form a single image. This must be above the flicker fusion threshold [139] in order to ensure that images displayed appear to be persistent.

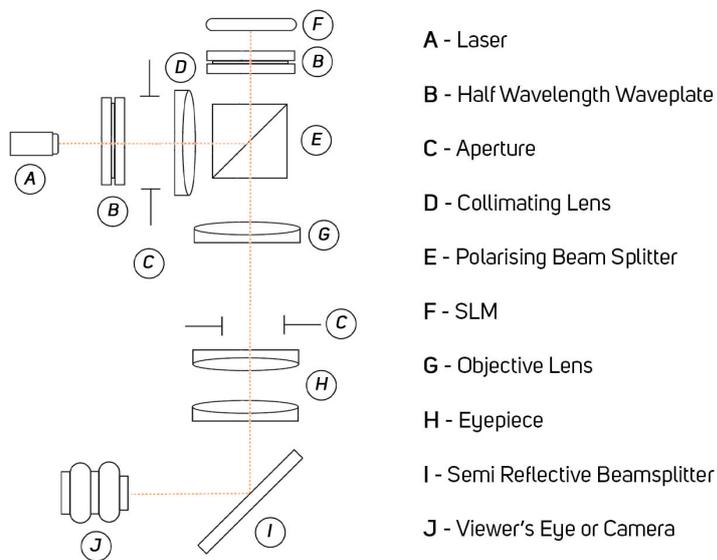


Fig. 2.12 Left: A diagram of the optical components required for a holographic display. The SLM is connected to the graphics output of a computer. Top Right: The resulting replay of holograms from a stereo SLM headset device, with hologram replay images seen projected onto the table from each eyepiece. Bottom Right: Directly viewing a holographic display, by looking directly into the eyepiece *H* in the diagram.

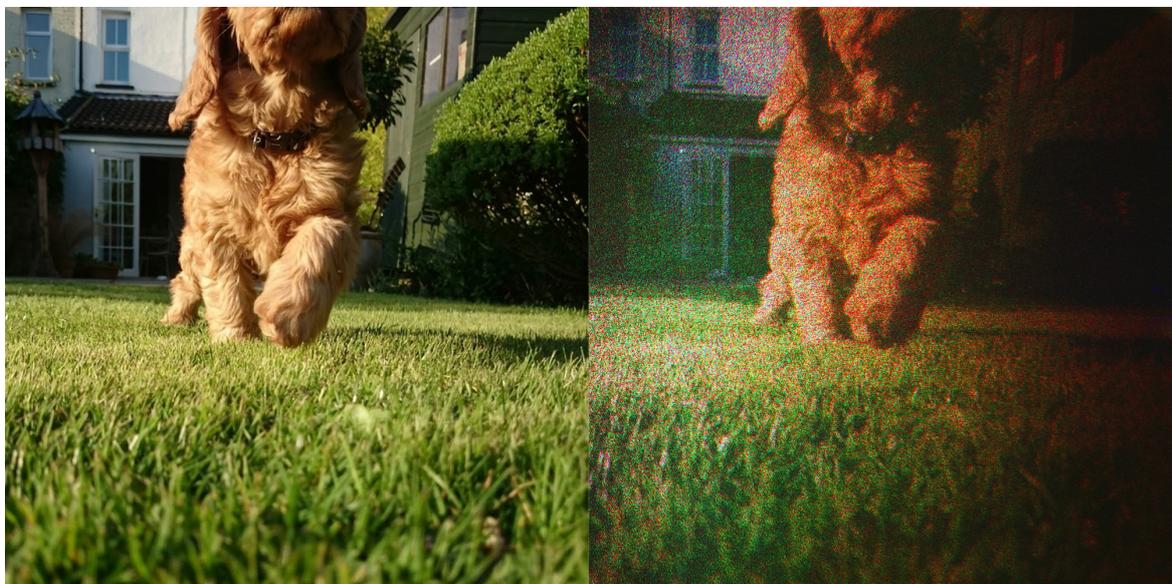


Fig. 2.13 Left: The input image to the hologram generation. A hologram is then generated via Fourier Transform. Right: The resulting replay of the hologram on a projector such as that of the direct viewing case in figure 2.12. The replay field is noisy as a result of random phase and quantisation.

# Chapter 3

## Production Workflow

### 3.1 Introduction

There is no established workflow for delivering any widely used production file formats to a hologram generation software suite in practice, and with such a high potential of realism from the display and so many interested parties, strong standards and robust practices are needed to ensure the highest-quality material reaches the hologram generation algorithms and ultimately the viewer's eyes.

#### 3.1.1 Technical Integration with Media Production

Each of the three techniques outlined for generating holograms, point, polygon and image based methods require different input data types, and therefore must be integrated into the media workflow at differing stages of the production process. Figure 3.1 shows an overview of a typical media production workflow for 2D display - this is easily adaptable to other forms, such as a fully animated production, which will not have on-set footage or similar, or a video game, which will effectively be entirely contained within specific elements of the "VFX" step and instead provide texture buffers for display.

Point-based methods primarily require the full 3D scene to be available, effectively replacing the existing ray-tracing or rasterisation based renderer. As previously mentioned, in order to replicate the realism of an existing production renderer, this holographic raytracer must provide all of the suitable shaders for object and materials representations - currently a feature complete holographic raytracer does not exist, with most only implementing a diffuse

---

Parts of this chapter were published in Demolder, A, 2022 [47] and Demolder, A, et al. 2020 [49]

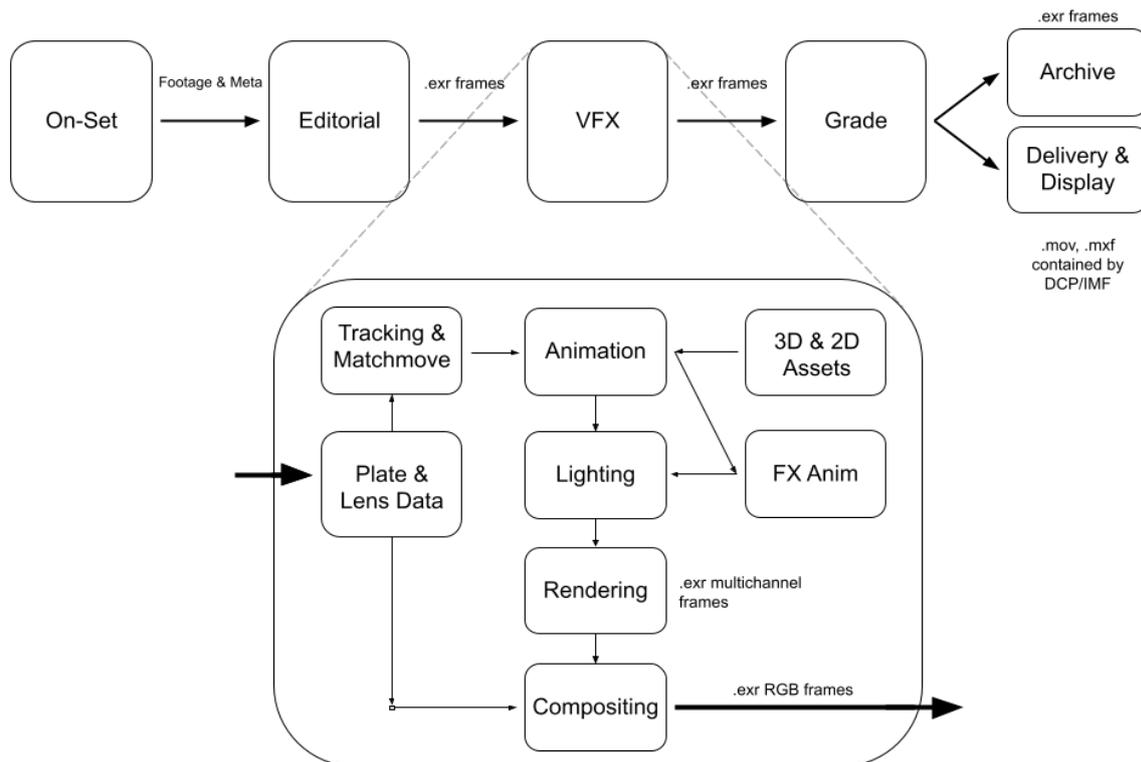


Fig. 3.1 An overview of a generalised media production pipeline, alterations may be made on a per-production basis to suit different needs. This overview is presented in order to determine where various hologram generation techniques may acquire their input data. Point-based and polygon-based methods are effectively required to replace an existing renderer in the pipeline, whereas an image-based method may access data outputs of existing renderers or compositing applications. It is unlikely that suitable data for hologram generation, such as depth, is stored in 2D delivery or archive files.

and specular shader. For such a raytracer to be successful, it must have feature parity with existing production renderers, and be suitably adaptable to enable sensible interactive look development by artists - but given the compute requirements, this is not yet feasible.

Polygon based methods also require the full 3D scene to be available, and suffer from the same shader/feature based issues as point-based methods as it is also required to be a renderer replacement, despite providing a speed increase. This method would source polygons directly from the DCC software (or directly from 3D graphics environments, such as OpenGL) and essentially take the form of a holographic fragment shader. Smooth shading is possible [152].

It is not yet clear how, or if, point-based and polygonal methods could interact with the compositing step and beyond in a production pipeline - but it is clear that image-based methods allow instantaneous compatibility with existing production practice. Image-based methods typically utilise a 2.5D input (RGBZ), or a multi-layered image representation [184]. Despite only providing a lambertian representation of the scene, and a selection of other issues due to such a high level of scene approximation as outlined in section 1.1 - it is the ideal candidate for integration with media production given the lowest compute requirements. A layer-based hologram with enough “layers”, is effectively a point-based method hologram (i.e. each point in a 2.5D image is treated as a layer), it provides a good candidate for the foundation of holographic integration into the production pipeline and allows control over quality vs speed. Image-based methods also allow the simplest integration into the media production workflow, and as shown in figure 3.1, imagery may be sourced at multiple stages, and includes conventional compositing/grade step support. Given that this imagery may be sourced from existing production renderers, production tools need not change, and the vast wealth of development in computer-graphics rendering can be utilised instantly.

Utilising layer-based methods to enable input from existing production sources, whether this is a path-tracer or a compositing application, will allow a sensible integration with CGH. This chapter establishes how this workflow should operate, and provides the means by which to immediately realise image quality improvement in CGH - by providing image data quality that exceeds display quality, as well as providing the framework by which to enable more advanced image techniques in later chapters.

This includes sufficient metadata for both Standard Dynamic Range (SDR) and High Dynamic Range (HDR) content, as well as the bit-depth required for the high luminance possible from a laser display, and accurate colour handling for any number of input colour spaces. This kind of data is commonplace for typical productions but is so far absent from CGH. This chapter solves these issues, and proposes a standardised container and theoretical framework to enable close integration to existing standards in use - the “High-Quality Hologram” standard, with the focus on supporting a close relationship between the render

output data and the display. This standard then enables additional creative controls, and enables later sections in this thesis to develop the range of image data available for hologram generation.

### 3.1.2 Lack of Artistic Consideration in CGH

From an artistic perspective, holographic display provides an extremely exciting opportunity for expression. CGH, once sufficiently advanced, enables realism, rather than photorealism, to be utilised within digital media. It provides a literal new dimension to other types of productions via full natural 3D viewing. Thus enabling immersive displays for many applications, whether it is AR via the use of near-eye displays for real-world virtual object placement at correct physical depths, or large holographic displays that provide a window into an entirely virtual environment indistinguishable from reality. This is a technology far from stereo 3D, and begins instead to provide light to our eyes in very much the same way that real-world objects are observed, rather than 2D approximations.

Much work has been done in delivering complete display hardware examples [118, 189, 132], and the algorithms to enable such displays, but as previously mentioned in section 1.1, no publications approach this from the perspective of integration with the professional media production pipeline that already exists to enable artists to deliver excellent visual work.

This chapter will focus on the process from the viewpoint of production for cinema, but the techniques will be readily adaptable for other applications such as real-time/game-engine based productions.

## 3.2 Background

Work in advancing display techniques from a production and perceptual perspective must be looked at in order to determine what must be adopted to enable a high quality holographic display experience.

The “production workflow” is not a fixed process, it is expansive, adaptable, and does not exist inside a single application. Media production pipelines will be formed of a complex array of processes and tools that may be off-the-shelf or custom made - but must all interact correctly together. An in-depth guide providing strategies to successfully handle such a workflow is provided by “The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures” [155].

Here we are particularly concerned with the stages during and after the VFX stage (such as shown in figure 3.1), where digital scene data exists. Rendering this scene-data to a 2D

output is performed by renderers such as ray-tracers [216] or rasterisers [154] - each of which have greatly matured since their first development, where light transport is now performed in productions with physically-based methods [61].

### 3.2.1 Digital Representations of Light

The resulting image data from a render will not immediately be ready for display, and this is by design. As the representation of light in these render systems is physically based (though represented as RGB values), image values will represent real-world light values with a linear relationship. This is important to preserve when passing image data to other applications in the pipeline - for many rendered elements to be composited together effectively, real-world light values are used, therefore this data must be kept in a format that is suitably non-destructive and preserves physical values for the compositing step. The form that this RGB tristimulus data takes is called a “working space” and is typically a scene-linear (or “scene-referred”) image [180], meaning that values uniformly map to real light values, and no perceptual quantisation is applied. Such a quantised encoding is applied for viewing (such as via an Electro-Optical Transfer Function (EOTF)) due to the limited bit depth of the display and is visibly lossy [75] - an example being the sRGB transfer function in an 8-bit container.

Such a quantised encoding is acceptable (and is specifically designed) for low-luminance display, which is referred to as a “Standard Dynamic Range” (SDR), but at higher luminances more bits are required to ensure no banding/artefacts are visible - and this is especially true when attempting to edit virtual or real-imagery during the production process, more data available in the source means more options and accuracy during the composite. Other perceptual quantisations exist for HDR images at higher bit depths [123], and aim to trade quality for memory, storage or transmission requirements.

A number of works in CGH correctly identify the linear nature of light [166, 222, 128], but still use input material recorded with an sRGB transfer function, and only approximately invert this by raising to the power of 2.2. By incorrectly working with this quantised imagery, the results of such techniques introduced by the research may be limited. This is a common error, and perceptual quantisation/uniformity "is either poorly understood or not recognized at all by a surprisingly large number of image scientists and engineers working today" [172].

### 3.2.2 Digital Representations of Colour

As well as an EOTF, the colourspace in which material is rendered or captured must be considered for display. The commonly known “sRGB colourspace”, which is actually formed of primaries from ITU-R BT.709 [99] (so will be referred to as BT.709), has become the

universal standard for modern displays. It is fairly limited, and does not cover much of the CIE diagram, which represents a colour system that identifies an area in which colours are visible to the human visual system. Larger colourspaces such as ITU-R BT.2020 [100] are specifically designed for much larger display gamuts, and defines physical wavelengths as the display tristimulus primaries (i.e. narrow-band R, G and B on the spectral locus of the CIE diagram) rather than wide-band phosphor based primaries, which can typically represent only a smaller range of colours.

During the production process, artists will utilise a “view transform” which ensures imagery is transformed correctly from both linear light and wide colour spaces into something compatible for the monitor they are working on. Ensuring that colours are presented correctly throughout the production process, and then viewed as intended after delivery, is critical in preserving a good viewing experience of a wide colour gamut display, as well as the creative intent of the material itself. No research in CGH appears to either correctly render imagery directly for the primaries by which it will be displayed, nor does any research appear to correctly apply a colour transformation to transform BT.709, which desktop monitors tend to use, to the correct holographic display gamut.

### 3.2.3 Delivery of Imagery

To review: an existing workflow overview is presented by figure 3.1. For a fully featured theatrical production, the general path to display is as follows: two-dimensional (2D) backplate footage is shot, trimmed to the edit, and passed to a Visual Effects (VFX) vendor along with any additional metadata. The vendor will recreate a lot of the environments and characters in 3D, push them through the VFX pipeline, and usually output 2D frames with numerous render-passes from their Digital Content Creation (DCC) packages (such as Autodesk Maya). This CGI imagery can then be composited with the plate footage in a DCC package such as Nuke. Imagery will for the most part in this pipeline exist in an OpenEXR [157] container, where colour (and other) information is stored with bit depths of either 16 or 32. This is then graded and mastered for deliveries.

The mastering and delivery specifications will vary. Recently, trends have pushed towards content being first mastered for HDR (such as a Dolby Vision [53]), and trim passes are produced for SDR. Each delivery file will have its own requirements, and may be based on standardised frameworks such as the Digital Cinema Package (DCP) [198] or the Interoperable Master Format (IMF) [195]. Versions delivered can include variations for cinema projection, Over The-Top (OTT) streaming services or physical media such as Blu-Ray disc. Then the content can be displayed on a 2D display device (or in 3D given a stereo pair) and also be synced with audio.

Once a production is correctly output and final creative touches made, it is delivered to a file for storage and onward distribution using “trim” passes, which quantise for specific display types.

Correctly utilising this process within CGH would enable light information to be preserved from the renderer to the display, enabling correct defocus behaviour, real luminances, and wide colour gamuts - as well as a wider array of information from the renderer (including depth related data) to enable advanced display effects.

## **3.3 Developing a Holographic Production Workflow**

### **3.3.1 Enabling High Quality Holographic Display**

Layer-based holography is often overlooked due to its approximations and associated artefacts, but most of these issues come from insufficient source material. A difference for holographic display vs conventional 2D display is that significantly more data is required from the VFX process for the best display experience. Though, much of the techniques can remain identical to the 2D output process – especially for content that is desired to be displayed in 2D, where the image may be projected into the replay field at a single fixed depth. The elements required for successful display are proposed to be split into 3 categories:

- 3D Data Representations
- HDR Specifications
- Colour Specifications

For the purpose of simplicity, the focus of the work in this chapter will be on enabling the imagery produced from the VFX stage to be produced into holograms. It can be assumed any captured or rendered plates have sufficient dynamic range, with an accurate Input Display Transform (IDT), or other colour management, and have had a successful 3D conversion or capture. It will be sufficient for this full scene data to exist in files via the DCC package, or via the compositing application.

As in conventional display, the largest quality drop occurs during the display process itself. In holographic display, this image quality drop occurs due to the current imperfections of the hardware and the need to avoid high computational loads - via approximations, such as depth layer reductions. However, with the delivery of scene data that is at the maximum possible production quality (as per 3.2), which is far more data dense than ever before, display technique improvements can be instantly realised.

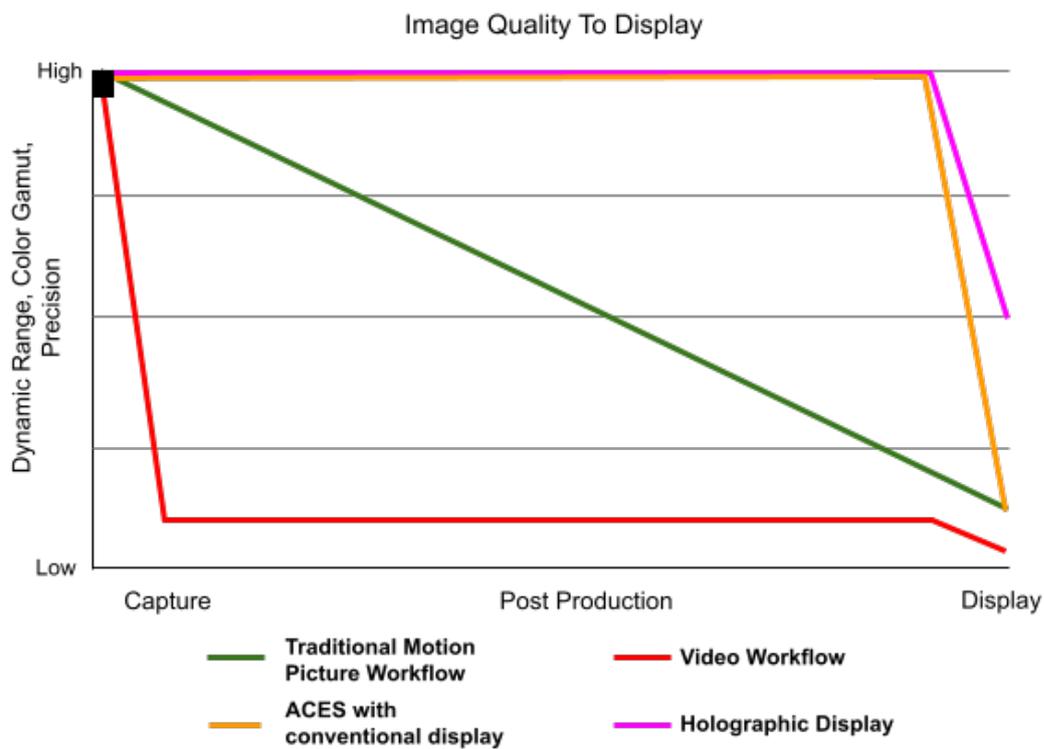


Fig. 3.2 Adaptation of a diagram taken from “Academy Color Encoding System (ACES) Case Study: Screen Gems”, presented by Alex Forsythe [1]. It provides a good visualization of how holographic display may fit in comparison to existing workflow and display techniques and provide improvements to the image being delivered to the viewer - on the path to delivering the highest possible quality.

### 3.3.2 3D Data Representation

For displaying imagery on existing 2D displays, pixels only contain the Red, Green, and Blue channel data (RGB). In holographic display, depth data is required for each pixel in order to correctly place it in the 3D space, as shown in section 2.1.3 – this is often referred to as Z-Buffer or Z-Depth. This data is produced during rendering and is a data channel that stores the distance from the camera to the specific pixel on the Computer-Generated Imagery (CGI) object, illustrated by 3.3. This data can also be produced from depth sensors attached to real-world cameras, or may be generated from stereo/mono video with tracking-based post-processing or neural networks. This is typically stored alongside the colour channels in an image to form RGBZ or RGBAZ (with alpha, representing transparency) images. More advanced depth data may be generated where depth data points are available behind other depth points, this is often referred to as "Deep" depth data.

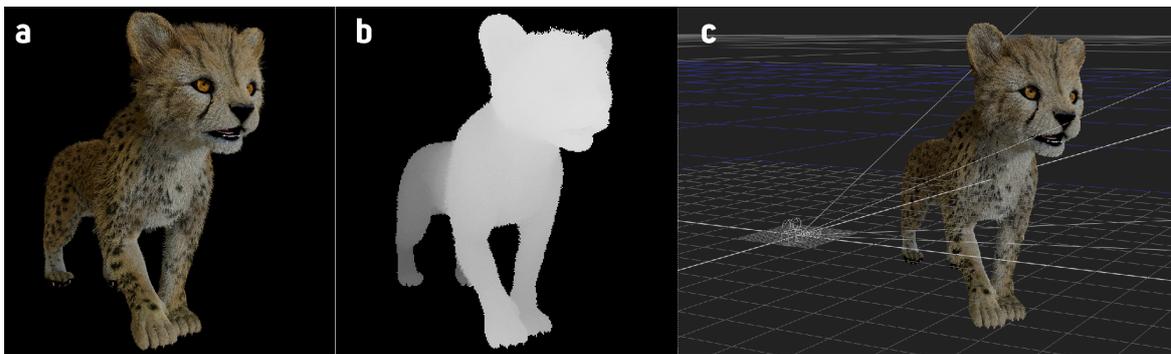


Fig. 3.3 (a) RGB pixel data, (b) A Visualisation of the depth (Z) data, and (c) RGB data re-projected in 3D space according to the depth.

### 3.3.3 Defocus Control

Depth data corresponding to colour pixels is handled by the VFX pipeline already, as it is useful for placing objects in 3D space in the compositor and for simulating depth of field late in the production process - but it is discarded for delivery as it is not required (even when delivering with a stereo pair) to display the final image. Pushing this depth data further into the pipeline does not mean artists or directors lose creative control over how depth of field is presented to the viewer, however, as the computational nature of the display allows for techniques that can adjust the aperture shape and f-stop, this is explored in section 3.6. This aperture information is stored as a mask image alongside the colour and depth data in a frame, and can vary frame by frame.

### 3.3.4 Multiple Depths

Many objects may be made up of semi-transparent, refractive, or reflective materials which cannot be represented by a single pixel of RGBZ depth data. Such additional depth data can be presented to the hologram generation algorithm via the use of additional Arbitrary Output Variables (AOVs) or render passes to be composited during hologram generation, this is explored further in chapter 4 or with multiple depth values stored per pixel such as with the OpenEXR Deep format [87, 107].

This allows for more complicated viewing situations to be supported, where a scene might contain fog, mirrors, or windows. In the case of a dirty mirror in a foggy room, where the subject of the scene is reflected in the mirror – depth information will be needed for the fog, for the mirror surface and for the subject. A viewer can then focus on the subject of the scene without hindrance. In the case of providing additional layers (in the context of scene depth slices) to the holographic generation algorithm, these also support object based occlusion masks to be stored. When calculated correctly, results in the viewer being able to look behind objects with correct front occlusion, without switching to a new rendered view as previously described – this provides the correct psychological cues between each “view” and is explored in section 4.4.1.

### 3.3.5 Multiple Views

Not all of the viewing situations are covered by multiple depths alone, and there is also the need for support for multiple views. If the viewing region is sufficiently large, holographic display naturally presents multiple views of the scene to the viewer. This results in the possibility for a viewer to look around/behind objects from many perspectives, or for extremely large viewing regions - multiple viewers to observe the same display and see different perspective views of the same scene. It is not possible to represent this information entirely correctly with only a single camera perspective, so it is necessary to ensure that many camera views are rendered and suitably organised in the output container. This can be seen as an extension of a stereo production pipeline and builds on OpenEXR’s support of views and multi-part image files. Rendering this imagery and enabling successful multi-view holographic display is explored in Chapter 5.

Though it must be noted that such a display feature may not be desirable in all contexts, as it may be distracting for the viewer and reduce creative control over the shot - literally diverting from the director’s creative vision/perspective. In appropriate contexts however, a multi-view display can be deeply immersive for the audience, as it places them in a given environment as if they were physically looking at it through a window.

### 3.3.6 Data Required

The new data supported for holographic display alongside RGB pixel data is summarised by:

- Multiple views per frame
- Aperture data per view
- Multiple colour and depth samples per view
- Multiple layers extracted from colour and depth samples

This fits into the workflow diagram as it leaves the VFX stage as per figure 3.4. As content produced is stored in a 16/32bit OpenEXR container after render and compositing - the additional data required as listed can exist in additional channels and layers in an OpenEXR.

As the holographic generation algorithm utilises floating point arithmetic on the Graphic Processing Unit (GPU), it is logical to continue this hand-off in float exclusively. The 16-bit implementation in OpenEXR is based on Nvidia's "half" half-precision floating point data type [108]), making it ideal for GPU processing. This datatype supports 30 stops of dynamic range (where a "stop" refers to the doubling (or halving) of the amount of light encoded), with an additional 10 stops at reduced precision at the low end, so is also more than sufficient for representing any scene. As SMPTE standard ST 2067-50:2018 [196] facilitates OpenEXR (specifically Academy Color Encoding System (ACES) container; described in section 3.5.2) in the IMF, it is also ideal to use an IMF during feature productions to deliver EXR frames to a hologram generation algorithm.

## 3.4 High Dynamic Range Handling

High Dynamic Range (HDR) is an important element of lifelike display. In holographic display, the majority of real-world luminance can be reached with laser illumination of the display. The luminance of such a display can reach peaks exceeding  $300,000\text{cd}/\text{m}^2$  [151]. As the viewer will be effectively looking directly into the laser light, eye safety, as well as potential viewer discomfort, must be considered. The luminance required will vary greatly per display application and viewing context. This distinction between viewing contexts is made by Kunkel et al. 2016 [123], and specifies a focus on the Entertainment Dynamic Range (EDR), which rules out a display system for full viewer light adaptation based on cost and the requirement for it to occur in existing viewing environments. The upper boundary of this EDR also considers phototoxicity ( $160,000\text{cd}/\text{m}^2$ ) and discomfort (around  $30,000\text{cd}/\text{m}^2$ ), and after a study of viewer preference, proposed a maximum of  $4000\text{cd}/\text{m}^2$  for diffuse

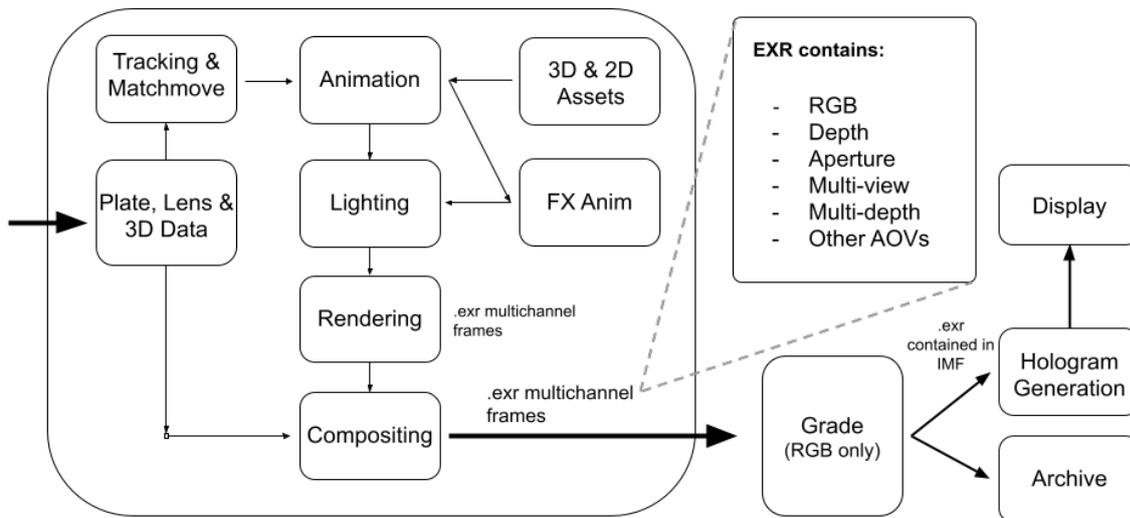


Fig. 3.4 A view of what data will be output from VFX to generate holograms. Input data may include 3D data captured on set. The grade/mastering process can still happen on only RGB data, but the additional depth, layer and view data should remain available. This can then be archived or generated into a hologram for display.

white, and  $20,000\text{cd}/\text{m}^2$  for highlights. In comparison, existing luminance standards exist for cinema theatrical projection at  $50\text{cd}/\text{m}^2$ , sRGB officially at  $80\text{cd}/\text{m}^2$  (though in practice this is typically as high as  $200\text{-}300\text{cd}/\text{m}^2$  in desktop displays/phones) and other EOTFs select  $1000\text{cd}/\text{m}^2$  under the BT.2100 standard for broadcast or home streaming. As well as supporting these existing proposals, proposed is a display system that can recreate higher luminances for better representation of real-world highlights if desired (remaining within safe levels, but allowing some creative discomfort), including overhead to allow for viewing even in bright environments. Acting more as a window into a real scene, rather than solely a representation of one.

HDR in holographic display is achieved with a combination of elements:

- High luminance laser illumination with a carefully designed optical system to prevent stray light or unwanted diffraction effects
- Equally, illumination which can achieve low luminance - as at low powers lasers may not emit light
- An SLM that can successfully replicate the diffraction pattern provided by the algorithm
- An effective algorithm that provides sufficient black levels (and therefore a higher contrast ratio, crucial for HDR)

- High bit-depth input image data.

A laser-based holographic display will also consume much less power than a backlit or emissive pixel display (such as in figure 2.11) to produce the same luminance, and will not have to dim the peak brightness over time for energy efficiency purposes or to mitigate risk of burn-in. Holographic display is efficient in that a significant portion of all of the light emitted by the source is redistributed into visible imagery.

### 3.4.1 Existing HDR

In conventional display, a HDR display standard system (such as Dolby Vision [53]) will have HDR specific metadata, which, if correctly utilised by the display, will control the backlight luminance of specific shots, or the entire production, such as:

- Display Min, Max Luminance ( $cd/m^2$ )
- MaxFALL (Maximum Frame Average-Light Level in  $cd/m^2$ )
- MaxCLL (Maximum Content Light Level in  $cd/m^2$ )
- An EOTF such as SMPTE ST 2084 Perceptual Quantizer (PQ) [197] or Hybrid-Log Gamma (HLG) [17]

Delivery image data for HDR is typically 10 or 12 bit, and when combined with an EOTF, viewing a 12 bit HDR image on a suitable display provides unnoticeable steps between data values, but 10 bit falls slightly short. Any unnoticeable difference is commonly defined as having a ramp that exists underneath the Barten Threshold [8, 9], as per figure 3.5. Using a quantising EOTF with a lower bit-depth is seen as a compromise in order to reduce storage/bandwidth required for content – but content output from cameras, rendering or compositing software packages will be stored as OpenEXR 16 bit as a minimum, typically without a quantising EOTF - instead being linear. This ensures the most data is retained for post-processing, rather than quantised for viewing.

### 3.4.2 Holographic HDR

For a proposed Holographic HDR display, the metadata required to be presented to a holographic display software suite will be very much the same for purposes of supporting legacy content - but differ in that a linear output (i.e. an EOTF of 1:1) is an option. Forming a linear display system results in image data values which are uncompressed/unaltered directly

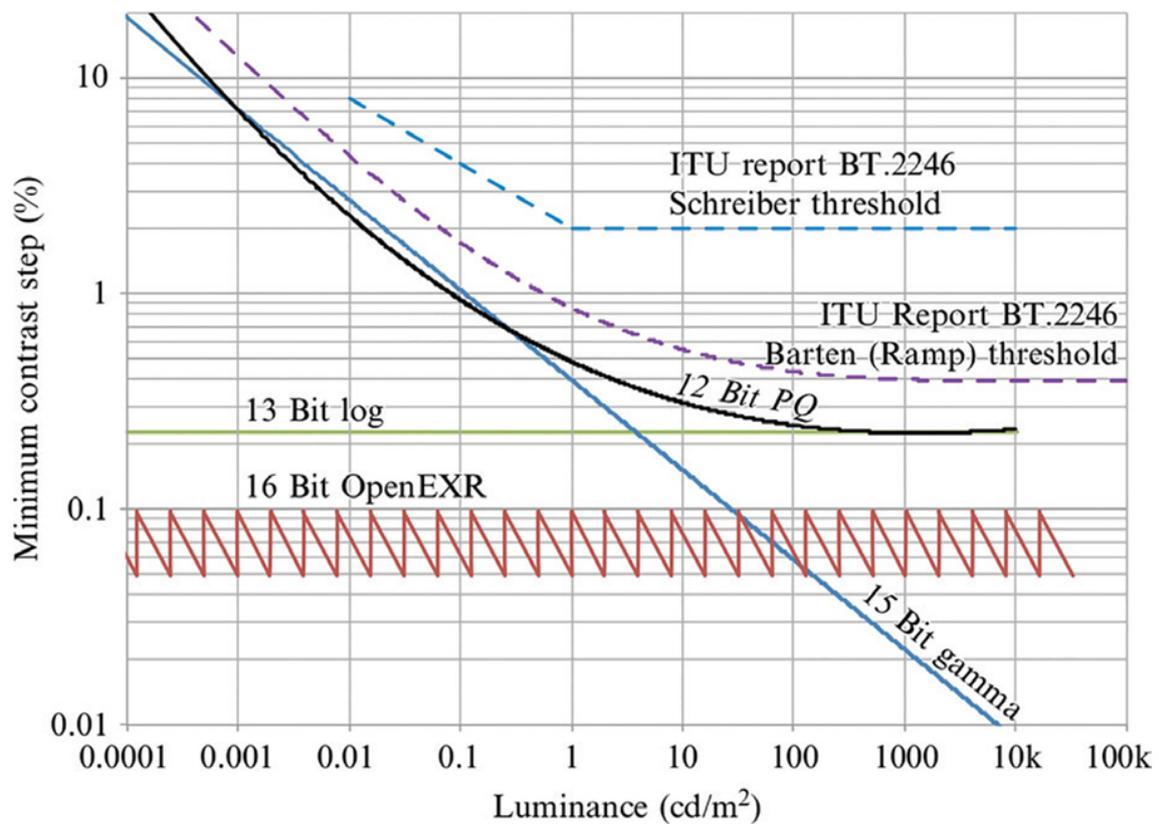


Fig. 3.5 Diagram taken from Kunkel et al. (2016) [123] demonstrating various quantisation approaches. EXR is stepped as it uses exponents but contains linear values; it also contains more information at lower luminances.

Table 3.1 An outline of Holographic HDR benefits and considerations.

Holo-HDR	Considerations to Address
High Ranges of Luminance	<ul style="list-style-type: none"> <li>• Suitably handling/clipping the brightest highlights (such as direct views of the sun/sunlight reflections).</li> <li>• Ensuring safe/comfortable luminance levels.</li> <li>• Similar challenges to existing HDR content, with backlit shots or strong highlights outside of subject interest being distracting.</li> <li>• Input source must be of sufficient bit-depth and fidelity.</li> <li>• The holographic display must have a low noise-floor.</li> </ul>
Increase in Range of Artistic Control of Final Image	<ul style="list-style-type: none"> <li>• End-to-End holographic display necessary for best content creation process.</li> <li>• Both real-world sets and Computer-Generated scenes will need to be carefully lit.</li> </ul>
No Perceptual EOTF for Holo-Gen	<ul style="list-style-type: none"> <li>• Extremely large file sizes, but ensures accurate hologram calculation. (Over-the-top services will likely prefer a more efficient use of bit-depth and an EOTF such as PQ if holograms are not streamed directly to the display.)</li> </ul>
Little-to-No Change to Content Creation Pipeline Required	<ul style="list-style-type: none"> <li>• The Grading process will need to preserve additional data-channels in the EXR such as depth, and where multi-view is used, consider luminance changes per frame per viewing position.</li> <li>• Creatives will need to familiarise themselves with viewing, and working on, a holographic display.</li> </ul>

representing real light values and overall aims to more consistently represent the intended luminance values of the mastered media than the existing HDR panels available. With this, a holographic HDR display with suitable optical and algorithmic design, is able to present almost a full scene-referred content experience to the viewer for the first time. This will present additional benefits and challenges, outlined in table 3.1.

A key part of offering a successful display system is that any existing content can be faithfully reproduced. By using limits on peak brightness, displaying imagery for the luminance at which it was created for, the creative intent of the original master can be preserved, whilst allowing for surround compensation (ambient luminance adjustment) and visible banding prevention by not exceeding the intended luminances of the media bitdepth. The reverse of such is also crucial – with suitable tone-mapping, or rather “display-mapping”

[86] built into the display pipeline. While quantisation is used on existing displays to efficiently allocate bits for human viewing, holographic calculation benefits from receiving higher bit-depth inputs, just as VFX post-processing utilises a linear workflow to preserve highlight details, holograms are ideally generated from linear images.

In defining a peak brightness limit in the interests of eye safety – it is important to define how any extremely bright pixel values should be handled. Existing HDR uses the MaxCLL value to define the total maximum luminance per shot in the preferred dynamic standards such as Dolby Vision, and similarly for holography a simple approach will be to define this threshold luminance – and above which, then any pixel values will not be 1:1 with the photons emitted from the display. This threshold may be equal to the value of an “scene white” or other reference diffuse in a known lighting environment. And the values above this threshold can either be linearly rescaled, or compressed via a curve. Assuming the display can reproduce a large enough luminance to ensure this threshold is sufficiently above usual scene light levels, there will be minimal mapping changes to the output, as shown in figure 3.6. This is similar to highlight compression approaches discussed in the work to determine the Perceptual Quantizer [123] and currently being explored in the ACES Output Transform Working Group [2] - it also replicates the behaviour of film, where high code values in a channel move towards white.

While this behaviour may be required for a physical holographic display, where hardware or design limitations warrant a luminance compression - in simulation, a full linear 1:1 mapping can be preserved. While recent work has shown that algorithmic changes can lead to an increase in peak luminance where hardware limits otherwise exist [114], it may still be the case that this new peak is not sufficient to match a 1:1 representation. Imagery may be re-exposed and viewed with an appropriate view-transform to quantise for the display that the artist or consumer is viewing, just as would occur when viewing imagery from a renderer or other high dynamic range source on a conventional 2D display. Another algorithmic and optical consideration is that of the noise floor of the holographic projector - an effective HDR display must not only have a high peak luminance, but also be able to have acceptable black levels, and noise introduced into the image during hologram generation, intentionally with random phase or not (see section 2.2.2), prevents a good contrast ratio on the display. Reducing the noise in CGH is beyond the scope of this work, but is a significant area of research in CGH, as despite the fine control over the laser illumination and light steering of the display allowing good black levels, if the image contains a region of high luminance, unwanted noise across the image will also be illuminated and consequently lower the contrast. A display cannot be HDR with high luminance alone, so is an important consideration to address when building a holographic display system.

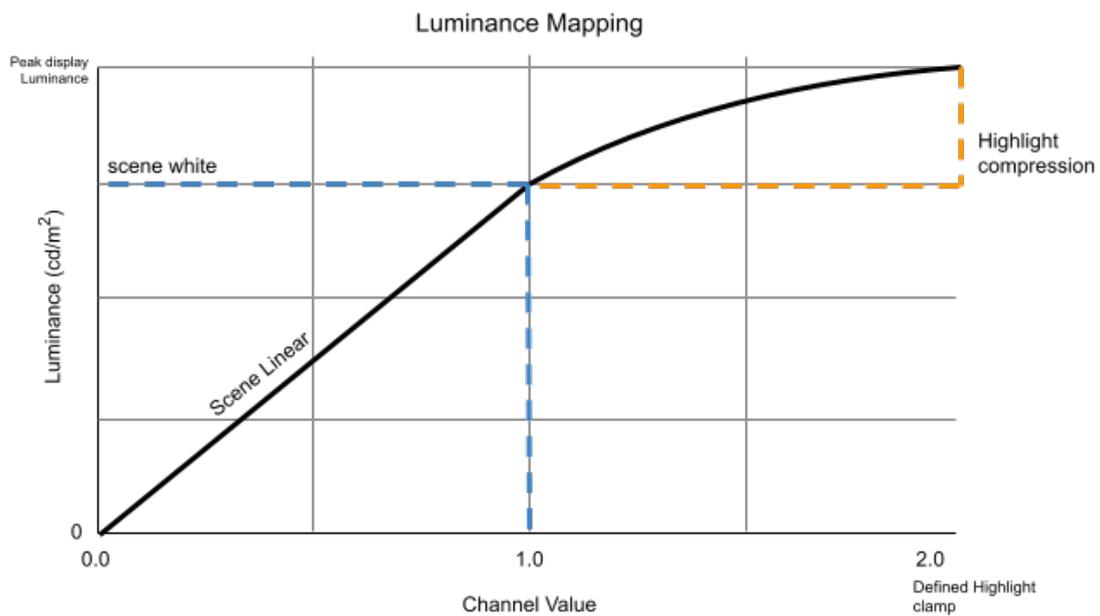


Fig. 3.6 With a high luminance display it is preferred to define real object luminance values rather than what is defined as “white”. This “scene white” can be defined against a real diffuse scene luminance. Values below this can be displayed as scene-referred (linear) and highlights above can be compressed against the peak display luminance, for example.

### 3.4.3 Luminance Reserved for Highlights

One of the benefits of a holographic display is that a large portion of the laser light is available to be distributed amongst the points in space that present the pixels of the input scene. Rather than a fixed uniform peak luminance, extra energy is available to be put into highlights if the display is not fully white. This fill factor of the display is often referred to as "coverage", and describes the percentage of non-zero colour values across the input image, illustrated in figure 3.7. A caveat of this is that where a projector is lacking the prerequisites for suitable laser driving design, the luminance of the display could change frame-to-frame, which is something seen in existing HDR panels described as "window size", this is often unpleasant to watch as the luminance is not consistent across a shot/scene as is intended by the HDR metadata. Equally, in a HDR holographic display, if a highlight headroom is not implemented and it is of a fixed luminance limit, the display will overall be darker than would be possible most of the time. A limiting factor in these cases is often the optical efficiencies, the class of the laser device in use and the power limits in place, but the safety and optical design of these projectors is not covered here.

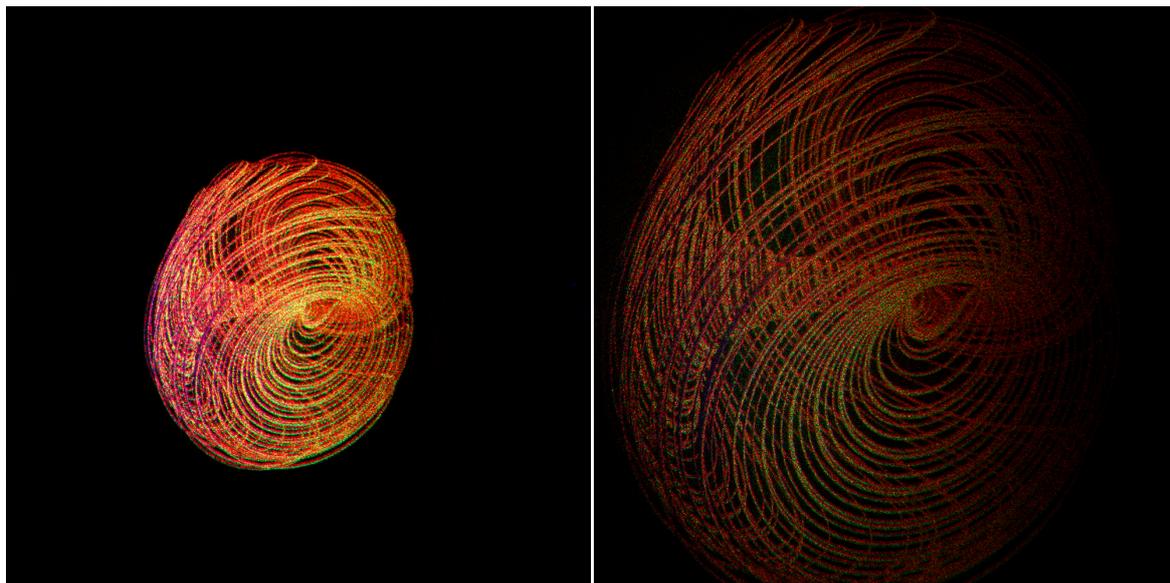


Fig. 3.7 Both images are displayed with the same input energy (total display luminance) to demonstrate coverage. As the left image has a lower coverage, the total light output is denser, and is displayed with higher luminance. If the total luminance available to the display is  $100\text{cd}/\text{m}^2$  (if the total display was "white"), then a coverage of 10% on a frame would result in the pixels displayed being  $1000\text{cd}/\text{m}^2$ . This setup and process is covered in detail by [151].

Figure 3.6 assumes that the display laser primaries are increasing together towards a peak luminance value where the full display is white – in practical applications this is much less likely to be a common case. Figure 3.7 shows this relationship on a physical holographic display, with manual energy clamping on the laser power to exaggerate the effect. Ensuring this total system works in unity is achieved by:

- Adjusting the total laser energy per frame, as a function of coverage. This ensures consistent brightness of objects – e.g. a code value of 1.0 in a pixel should be the same luminance in every displayed pixel on the display, regardless of coverage.
- Defining the uniform white luminance, or scene white, of the setup - to match a real diffuse white luminance (whether this is measured on set or creatively defined).
- Defining a highlight peak luminance value. This ensures that code values above display white do not exceed a defined luminance by clamping the input.
- Applying highlight scaling as preferred to linear ramp from display white to display highlight peak.

In this system, by utilising excess coverage energy where available, fully saturated primaries can be viewed with high brightness – which is an important distinction between a scene referred representation and what viewers are used to in a conventional display referred representation. In cases where excess energy is not available, they can be transformed towards white, but it is likely that a scene such as this is vastly over-lit and uncomfortable to view. Examples of image coverage values are given in figure 3.8.

Other than display transforms, very little will be required to be adapted in the HDR handling workflow in VFX or other post-processing operations. It is of course advisable to adjust displays used by artists – where holographic reference displays with the intended final algorithm and brightness capability are utilised.

Grading will be uniquely different in that in a multi-view hologram, the total brightness of one view may vary from that of another view on the same frame (and this display energy may be shared across views), but with sufficient view-analysis tools in grading software this should not present much of an issue when finalising the look of content.

The way artists think about digital representations of light will need to be more informed, and they will need to use real lighting values in scenes to have a direct understanding of what luminance a code value represents. This should hopefully be automatically intuitive when doing lighting or look-development on a holographic display directly.

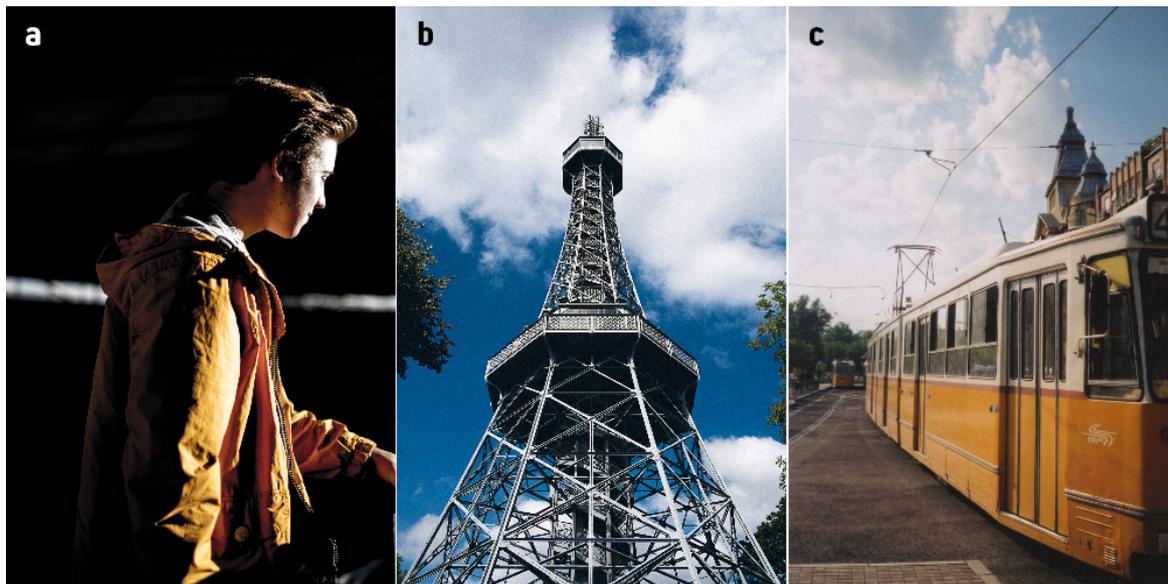


Fig. 3.8 Examples here are calculated using 8-bit images, for ease of visualization. a) Utilizes only 10% coverage. b) Utilizes 55% coverage, and c) 44% coverage. Even the images that appear bright leave a lot of headroom for highlight luminance.

### 3.5 Flexible Wide Colour Handling

Handling colours accurately should be one of the critical goals of any display system – and in a laser based holographic display, as the primaries are single-wavelength sources, a wide display gamut is automatic. For an accurate holographic colour handling system, three desirable features are presented:

- Preserving creative intent by respecting existing colour spaces for content, to ensure faithful reproduction where possible. This includes both smaller spaces such as ITU-R BT.709 [99] and wider spaces such as ITU-R BT.2020 [100]
- Enabling Academy Color Encoding System (ACES)(SMPTE ST 2065-1) [193] Output Display Transform (ODT) support for any selection of laser primaries, for content stored with AP0 imaginary primaries.
- Display-referred space support, for content made specifically with certain laser primaries in mind - which could make use of varying Colour Matching Functions (CMF) such as the more accurate “psychologically-relevant” CIE 2012 10-degree standard observer [149], or allow for more direct display of content from spectral renderers.

These cases should be supported by the holographic display suite, and the colour space information should be delivered via embedded metadata. Any colour transforms or chromatic

adaptations (for white point) will then be carried out before hologram generation, either by the generation suite itself or in advance. An example is given by figure 3.9, transforming from BT.709 to “laserspace” by the matrix shown in equation 3.1. This matrix is formed by:

- A CMF to determine the CIE XYZ space values for given wavelengths (here, 643nm (red), 530nm (green) and 450nm (blue) are chosen as they are commonly available laser sources). CMF data is readily available in sources such as the Colour Science library [37] (alongside many other useful utilities).
- A transformation calculation which inverts this XYZ matrix, and is then multiplied with the matrix that defines BT.709 in XYZ space - this may then be normalised.

$$\text{Laserspace} = \begin{bmatrix} 0.582 & 0.432 & -0.005 \\ 0.079 & 0.845 & 0.075 \\ 0.013 & 0.052 & 0.9345 \end{bmatrix} * \text{BT.709} \quad (3.1)$$

This is easily implemented in a workflow with tools such as OpenColorIO (OCIO) and enables correct transforms for a given holographic display from within any OCIO supported application, where a configuration is written to transform from a given input space to the display colorspace; this also applies to an ACES colour workflow.

Replays of holograms generated using such a proposed format are given in figure 3.10.

### 3.5.1 Laser Primary Choices

One of the challenges that may be faced by holographic display manufactures is standardisation of laser primaries. In theory there is a solution to this issue with the RGB colorspace in the ITU-R BT.2020 standard which would also allow compatability with existing 2D material. This compatability with specifically BT.709 was not taken into account when producing the "Laserspace" colourspace laser modules and consequently a small portion of the blue in BT.709 cannot be displayed in Laserspace, as seen in figure 3.9. The primary wavelengths outlined in BT.2020 (630nm, 532nm, 467nm) resolve this issue and is why it is the current gamut of choice for production work, and create a space that while encompassing almost all of pointer’s gamut [168], which is the gamut of real surface colours, is still only a percentage of the total possible visible colour. This is exactly why it is preferable to produce content using ACES, which is a colour system designed to encompass all of the visible gamut, as an appropriate ODT can be used to provide an accurate colour transform to any display gamut that may appear in production in future. Though, this presents some imperfections when using ACEScg as a working render colour space, as it is designed to match the primaries

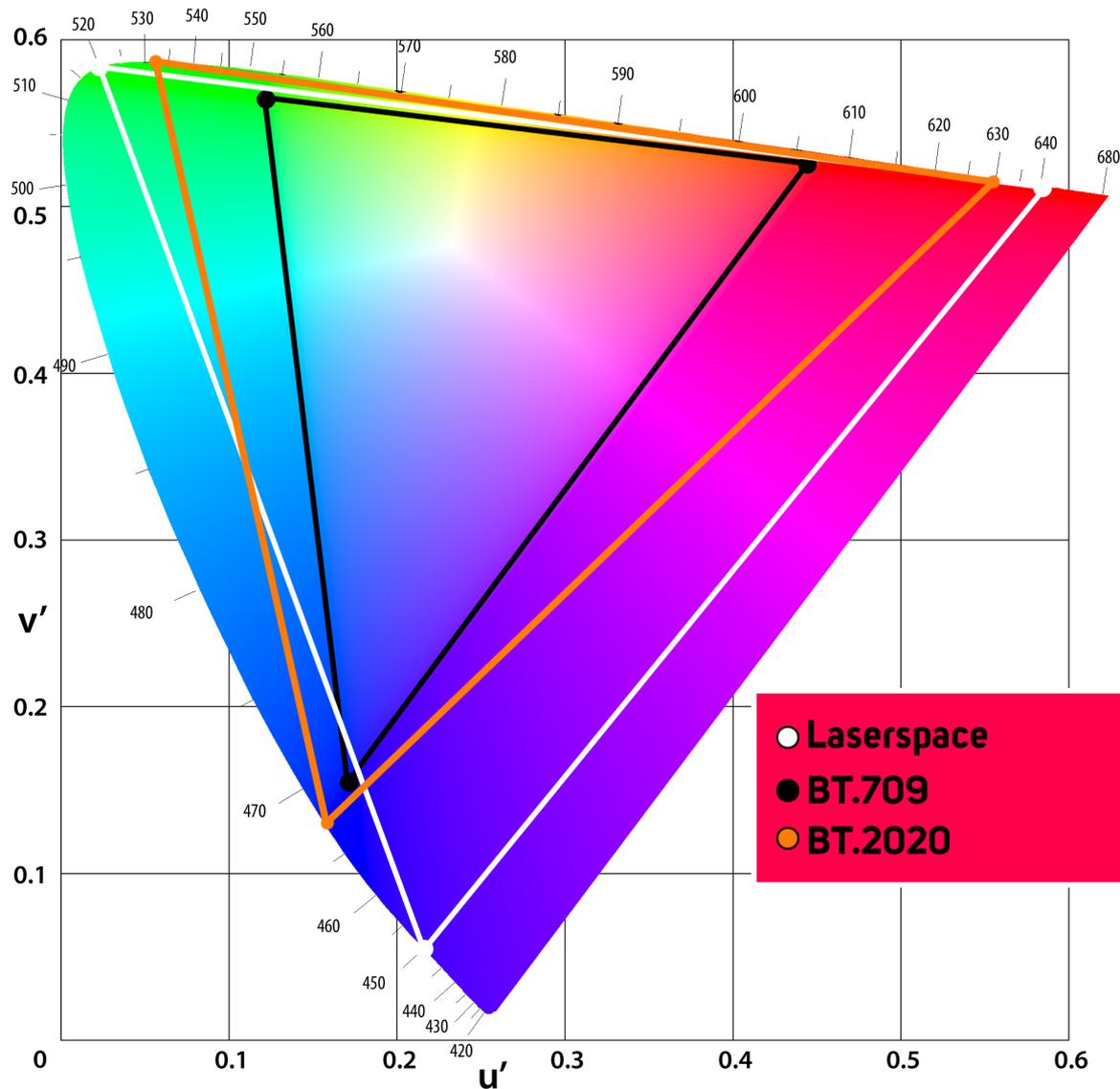


Fig. 3.9 The BT.709 and BT.2020 colorspace, compared against the primaries used in an example holographic projector to form “Laserspace”. The primaries at 643nm, 520nm and 450nm. Plotted on a CIE 1976 UCS  $u'v'$  chromaticity diagram, with the 1931 standard observer.

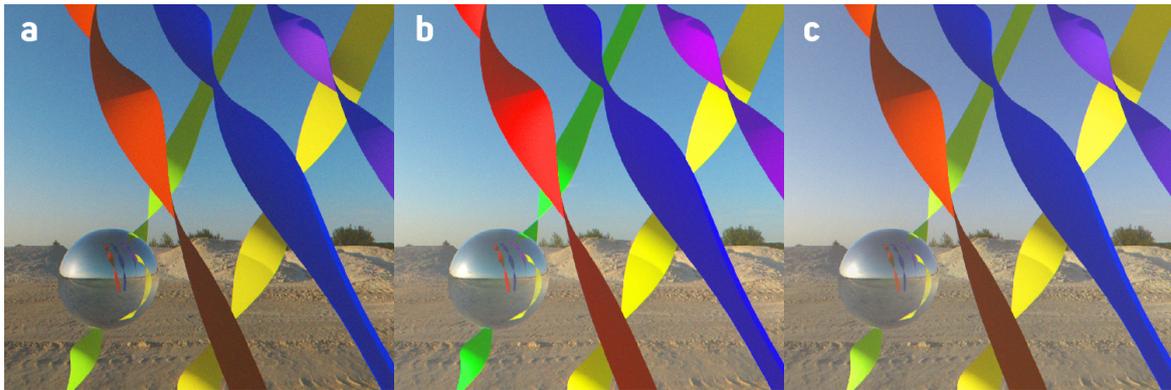


Fig. 3.10 Various holographic replays (simulation) from rendered content, using, a) BT.709 primaries transformed to Laserspace figure 3.9 and associated matrix, b) ITU-R BT.2020 primaries for both render and display, and c) ITU-R BT.2020 rendering with transform to Laserspace. For viewing purposes each is transformed again to BT.709. The most vibrant (and true to source) image is where the rendering primaries and display gamut match closest. Blues turn almost violet in conversion from BT.2020 to Laserspace. This occurs regardless of container bit depth.

of BT.2020 rather than allow the use of varying primaries. The slight drawback of ACES being tied to the CIE 1931 2 Degree Standard Observer is loss of compatibility with a more modern CMF, but the flexibility and legacy support provided by ACES is ideally suited for holographic display.

### 3.5.2 Academy Color Encoding System (ACES)

As OpenEXR is the container of choice for all previously proposed stages in the workflow, the ACES OpenEXR container (ACES2065-4, specified by SMPTE ST.2065-4 [194]) presents no issues with colour management process proposed for holographic display, but an expansion to the standard may be ideal, as noted by SMPTE ST.2065-4's informative notes [191]. This suggests support of all OpenEXR features - as is already used in practice. These expanded features, such as additional layers and channels, as discussed previously, are required by the hologram display suite for the full experience.

It is important to note that work is progressing on ACES modifications toward a more neutral Reference Rendering Transform (RRT), and a parameterized ODT, amongst others [40]. Control over the tonescale and access to hue preserving features will be essential in combination with the additional luminance overhead available on the holographic display for representation of bright saturated objects such as neon signs. These will greatly benefit the ease of workflow and will be eagerly received for holographic content delivery.

### 3.5.3 Modern Color Matching Function

In moving towards a future display, it is valuable to consider what other gains can be made for better colour handling. Given that some of the potential display primaries (such as the one used in figure 3.9) can be of a much shorter wavelengths compared to existing displays, there is additional benefit in using a recent CMF. The CIE 2012 10 Degree Standard Observer, which has been standardised from the functions proposed by Stockman & Sharpe [200] represents a “psychologically-relevant” CMF. The difference between the CIE 1931 2 Degree Standard Observer and the more recent CMF is minimal at most wavelengths, with the 1931 standard being the basis for most production, including ACES. The difference is illustrated in figure 3.11, and is significant in the shorter blue/violet wavelengths below 450nm - as the 1931 2 Degree Standard Observer significantly underestimates luminance at short wavelengths [199]. New content produced against varying CMFs can be supported by the holographic display as the primaries always have known points in CIE XYZ.

### 3.5.4 Chroma Sub-Sampling Support

Colour encoding is also possible in holography, where holograms have been calculated in YCbCr space, and this serves to reduce the computation time required [187]. This still features the same drawbacks as in conventional display, with visible colour detail being reduced in resolution. Displaying in ICtCp (defined by ITU-BT.2100 [101]) for holography under this scheme may also be possible, and allow HDR chroma subsampling in holographic display - but either encoding method can always be converted back to RGB for display. There is little benefit in colour encoding for the purpose of a high-quality hologram as data size is not a primary concern - but should be noted that it can be supported to display existing content.

### 3.5.5 Nearly Ultimate Gamut

It is certainly likely that more than three primaries (referred to as Multi-Primary display [85]) could be used in holographic displays to make use of a wider area of the visible spectrum. As shown in figure 3.12, using an optimised selection of lasers, Song et al. demonstrate a practical design of a "nearly ultimate gamut" of 6 primaries which cover 96.6% of the visible gamut [192]. This can be accommodated in content production for at least CGI content by rendering using a spectral renderer [102, 60]. In such a system, geometry and other scene data is sampled with real light wavelengths. Each wavelength's pixel data required by the display can simply be stored in a multichannel EXR and passed through the production process – but

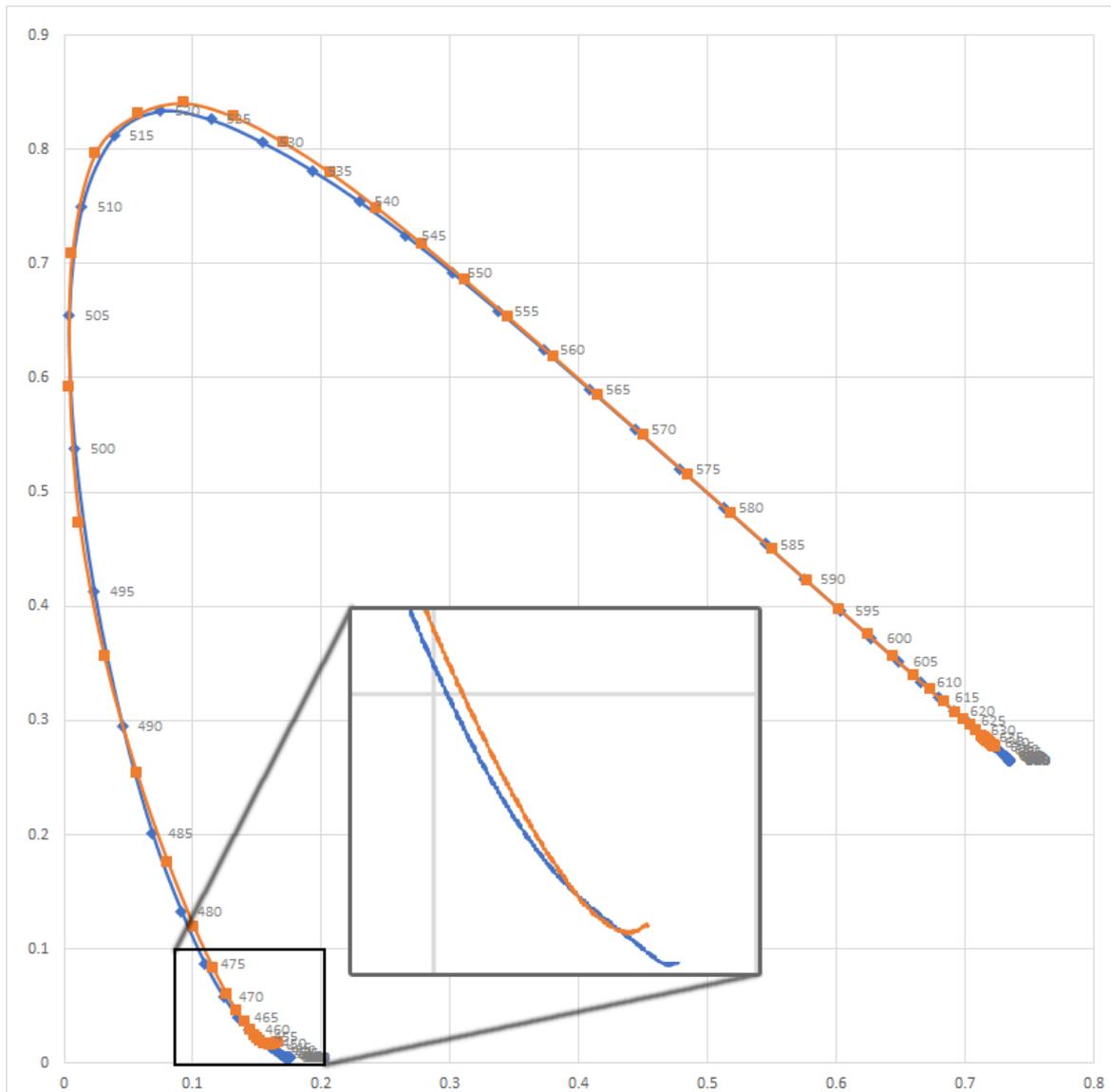


Fig. 3.11 1931 2 Deg (blue) vs 2012 2 Deg (Orange) spectral loci on displayed on a CIE 1931 chromaticity diagram. The short wavelength difference is well illustrated here. There is also a significant difference around 520nm-530nm green. These variations will cause significant visual differences when using monochromatic primaries.

this would require suitable support across many software vendors, alongside mass-adoption of the multi-primaries to become practical in widespread production. Converting from 3 primary RGB data to multi-primaries is also possible [176].

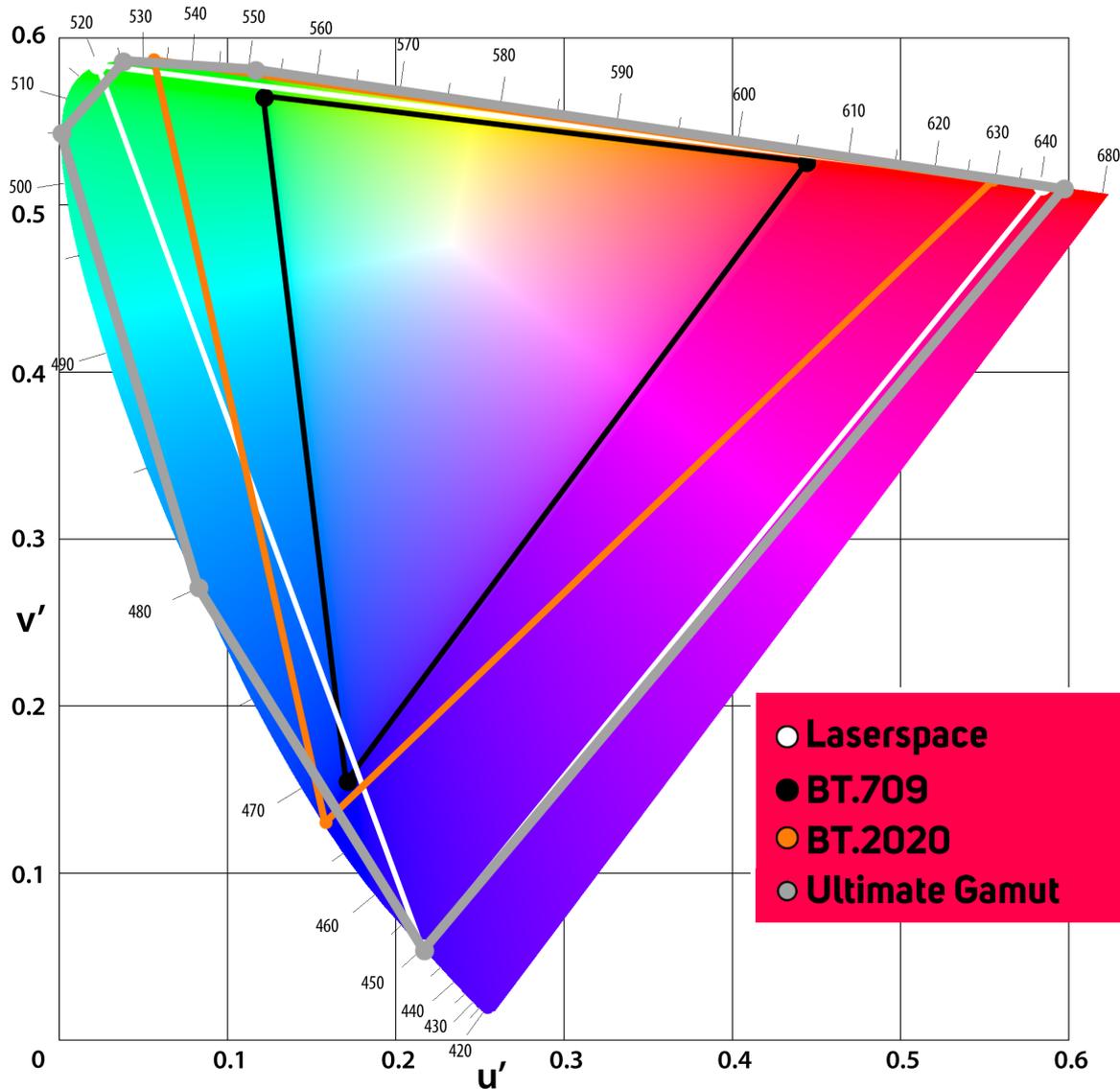


Fig. 3.12 Comparing sRGB with “laserspace”, the example holographic projector gamut, alongside the “nearly-ultimate” gamut proposed by Song et al. [192] Plotted on a CIE 1976 UCS  $u'v'$  chromaticity diagram, with the 1931 standard observer.

Outside of expanding the number of primaries used for pushing the boundaries of display quality, the ideal route to widespread lifelike colour reproduction on holographic display is likely to be most successful with ACES support. Not only from the perspective of supporting

varying output transforms, but also as it ensures that other display types with much lower peak luminance, and smaller display gamuts can also make use of media produced.

### 3.5.6 Proposed HQH Delivery Format

The High-Quality Hologram (HQH) delivery format aims to contain all the data required to deliver a rich holographic display experience, whilst ensuring compatibility with existing production practice. The standard must also contain specific hologram generation parameters as metadata that will affect image characteristics, as different CGH algorithms will have varying grain characteristics – as these are important considerations in the creative intent of the media. Given the previous sections, we can summarise the identified requirements as:

- A fully-featured EXR container: for additional channels, layers and metadata.
- Full scene-referred workflow: for real light values delivered directly to the holographic generation suite, and for near full scene-referred display.
- Luminance related metadata: for containing existing SDR/HDR standards and their EOTFs and for defining diffuse white and peak highlight luminance mastering values for near scene-referred display. Moving any tone-mapping to the stage of display. These must be respected by the display.
- Colour metadata: including input colorspace, intended display gamut and white point. This must also specify the CMF. APO input support, with metadata listing any ACES rendering models that are required should be specified for transformation at point of hologram generation.
- Hologram Algorithm Standard metadata: for ensuring any image characteristics introduced are chosen intentionally or revert to the highest quality algorithm available.
- All packaged as per the IMF framework.

Of course, a development system in the hands of many studios and other creators will best reveal what needs to be adjusted and consequently standardised. It is likely the ideal route to widespread lifelike image reproduction on holographic display is with an ACES implementation. Not only from the perspective of supporting varying wide gamut output transforms, but also as it ensures that other display types with much lower peak luminance, and smaller display gamuts can make use of media produced, whether it contains additional depth or hologram specific data or not. The key desire is that any holographic standards will be built on work that has been proven in production, and that they leave flexibility to

be refined, but provide a solid foundation now. Any imagery and associated holograms produced under this proposed standard may be archived together to form a “golden” copy of the production, whilst allowing for source material, complex holograms and/or quantised holograms to exist for later re-use or even re-generation as the display format progresses.

## 3.6 Artistic Optical System Controls

All existing media is recorded through a lens, whether this is physical or digital. Being able to control various aspects of the lens system has become critical in delivering a specific creative look to a given production. This section presents several methods of delivering such characteristics in holographic display systems.

### 3.6.1 Virtual Aperture Control

Lens characteristics such as bokeh make up an important part of the feel of an image. Many lenses are designed to produce pleasing effects, and many lens imperfections are emulated in CGI renders or added in post-production. Naturally, with our eye being a lens, we as viewers of the world experience and perceive lens effects such as bokeh; where in out-of-focus areas, points of light assume the image of the limiting aperture – our iris. This is an important part of human vision, not only in providing the ability resolve fine detail, but also to provide one of the several cues of object depth [36, 42, 140].

Given that the natural perception of depth with defocus is presented by CGH - controlling the optical characteristics such as bokeh shape and size would be a valuable artistic control. This is already controllable in 2D traditional media via the choice of lens, or via the use of kernels in the renderer/post-process, such as shown in “The Lego Batman Movie” in figure 3.13, where the bokeh assumes a heart shape for a particular shot.

This section proposes a novel method for artistic control of bokeh and depth of field in computer-generated holography, entirely in software with a minimal processing requirement.

The behaviour of depth of field resulting from a circle of confusion in holography is the same as in conventional imaging [144]. In a typical optical system, a physical limiting aperture would be required to provide a given circle of confusion (defocus) shape/f-stop, but in CGH as we are operating in the frequency domain, the resulting hologram may simply be masked to produce the desired bokeh and lens speed; provided the eyebox of the hologram is smaller than that of the eye of the viewer or smaller than the lens aperture of a camera viewing the hologram. The eyebox (or exit-pupil) refers to the area in which the replay field is projected in Z, and is the region that is intended to be viewed. In current near-eye



Fig. 3.13 The Lego Batman Movie [202], created at Animal Logic, makes use of a variety of lens effects - while most of this occurs in the renderer "Glimpse", this defocus kernel was applied in the composite [96]. using the "bokeh" tool from Peregrine Labs, now built into Nuke.

holographic display, the eyebox may be smaller than the eye aperture, as shown in figure 3.14.

Without intervention, given that the SLM usually is the limiting aperture in a holographic display, the defocus of holograms will typically match the shape of the SLM - a square eyebox will also assume a square shape, but with the presented method, the aperture shape, depth of field and chromatic aberration of the defocus can be controlled. All of which can be animated over frames and applied to pre-generated holograms or masked in real-time; both are visible on a holographic laser projector by eye.

Given a hologram generated from an input RGBZ image, a simulated replay with the square shape of the unmasked hologram is seen in figure 3.15, matching the eyebox shape. Here a mask may be digitally applied to the hologram pattern before it is displayed on the SLM, this is done simply by removing (multiplying) the pixels outside of the given desired shape area, for each colour channel. As per convolution theorem, this process is equivalent to the convolution of the input scene with the intended bokeh shape - but considering the reconstruction system, where the viewer's eye lens is performing the fourier transform back to the spatial domain as the SLM is imaged at the pupil, the consequent point spread function of the total optical system is formed by the convolution of this aperture distribution with the point spread function of the viewer's eye lens [43, 144]. Typically masks such as this are used to provide filters such as a lowpass, to remove frequencies within the image, such as seen in a 4f optical system. Indeed as a side effect of this masking, spatial resolution is lost alongside increasing the depth of field in the resulting image where the aperture is smaller,

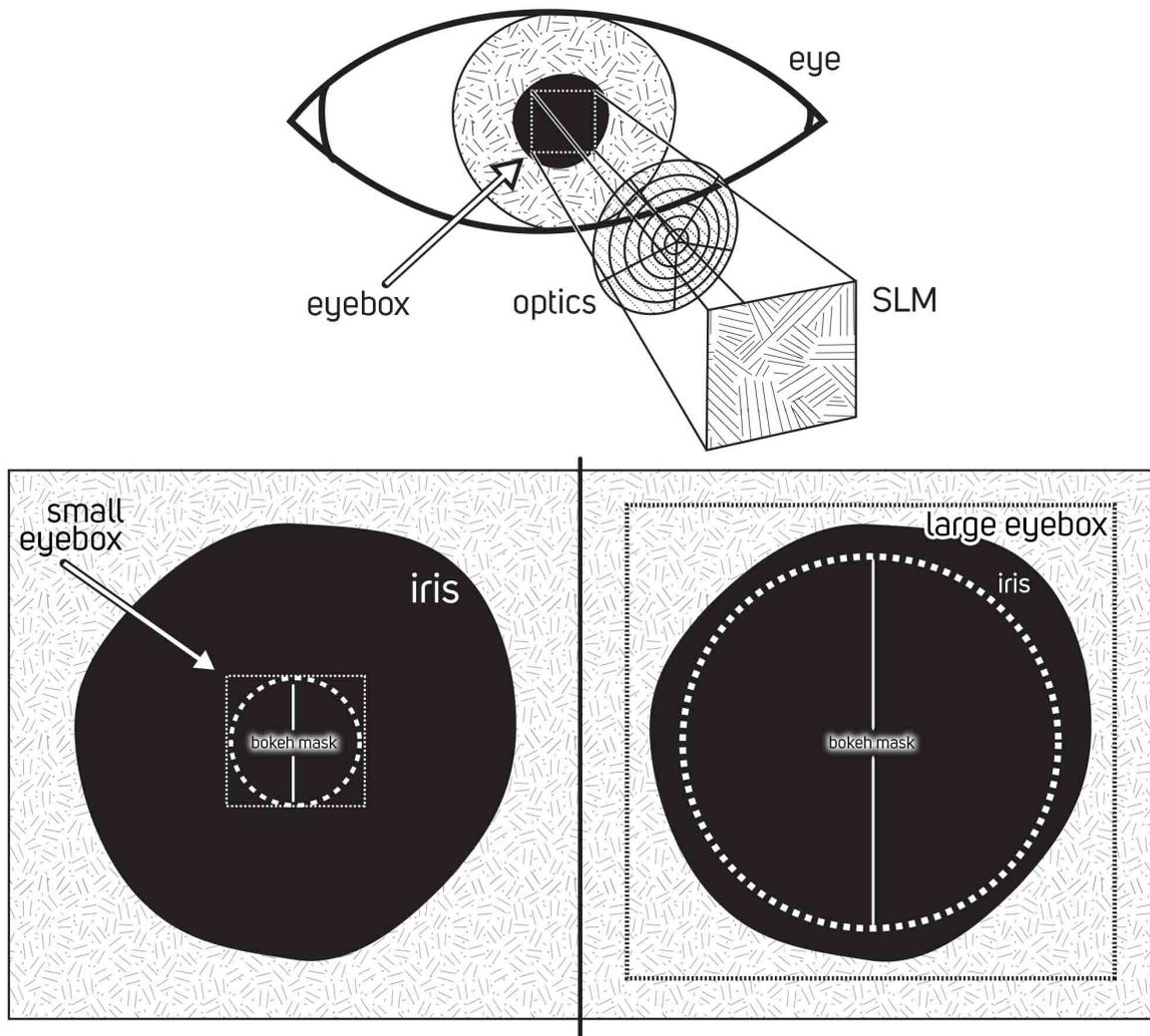


Fig. 3.14 For the bokeh controlling mask to be effective, it must be the limiting aperture in the optical projector, including the viewer's iris. A simplification of a projector (top) shows how the SLM may be imaged to form the eyebox at the pupil. Left: A small eyebox (indicated by the dotted square) means masking may occur without knowledge of the pupil size. Right: An eyebox larger than the pupil means for the bokeh control to be visible to the viewer, it must be smaller than the viewer's iris diameter.

just as a viewer observing a hologram will see a degraded image with a constricted pupil size. The scene is still visible in its entirety as the random phase used here (described in section 2.2.2) has ensured that the full scene is encoded across the entire hologram. Figure 3.15 compares the unmasked, masked and pinhole masked hologram patterns (though they appear grey, they are holograms) with their respective replays.

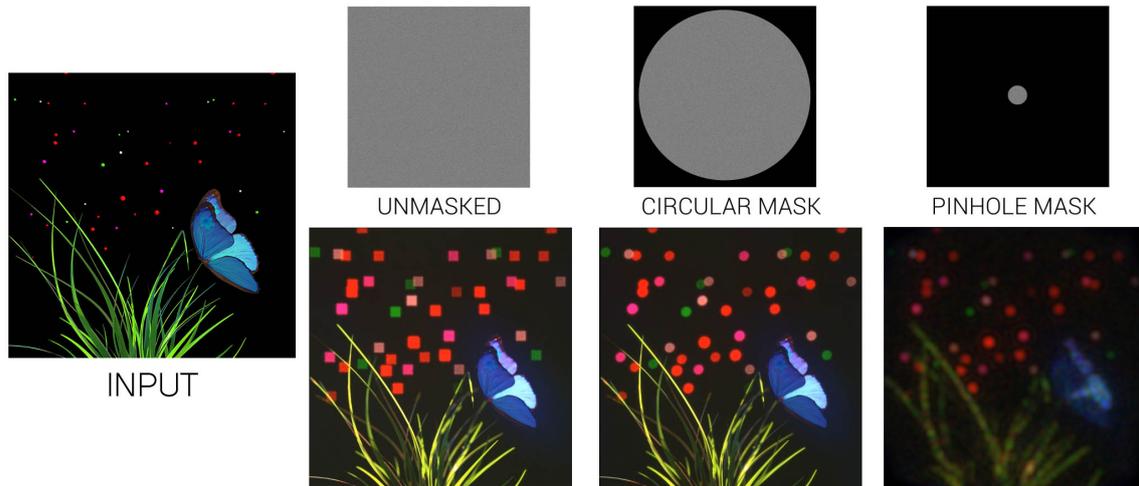


Fig. 3.15 Masking the hologram digitally will allow the shape applied to become the limiting aperture in the optical system, which controls the behaviour of the defocus. This is shown here by the hologram and respective replays - the square unmasked hologram results in a square defocus kernel matching the SLM shape, the circle mask creates a circular defocus. This also causes a loss of resolution, as shown by the pinhole mask.

The simulations in this section utilise a high number of time-averaged frames per capture in order to minimise the influence of noise in the field to ensure the effects of the defocus are visible - this would also be required in a physical display. This noise is especially visible in out of focus areas, in both complex and quantised hologram cases. Such artefacts in defocus, including edge fringes, may be resolved with optimisation that targets correct defocus-blur [112].

In order to provide the most natural defocus and to replicate eye characteristics, accurate human-acquired iris aperture data is used. Whilst this is not strictly necessary for general use, and the strength of effects will depend on the irregularity of an individual's iris, a visible difference should be visible in heavily out of focus areas, particularly with spots of light. Data kindly provided by Disney Research, from the publication "High-Quality Capture of Eyes" [14], allowed high resolution iris geometry captures across a range of pupil dilations to be input to the masking system. Here this iris geometry is rendered to 2D matte images, which is then used to provide the aperture mask - matching the defocus shape of a given eye,

with a result almost indistinguishable from real-world eye defocus. The resulting masks and replays of a realistic eye with two varying dilations are shown in figure 3.16. Such masks may be applied on a per-eye basis - whilst pupil dilation is linked between the eyes, the pupil shapes may vary.

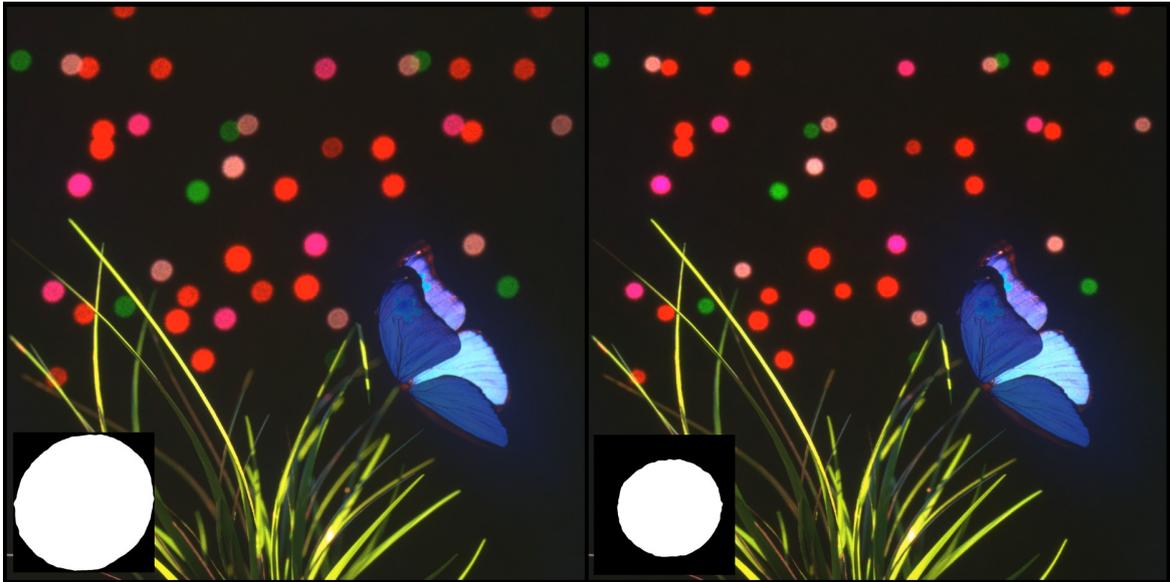


Fig. 3.16 Eye data can be used to create realistic iris masks which vary in dilation. In this simulation, the eyebox size is 2.048mm square, at a 2048x2048 resolution, with many sub-frames averaged to reduce noise. The size of the background dot defocus changes in relation to the aperture size.

These masks can also be more creative, as shown in figure 3.17, where the input mask is formed of a butterfly shape, and of an animated spinning VividQ Logo. As well as animation of the shape, the size may be animated in order to control the amount of defocus in the scene, and assist the attention of the viewer to focus on a specific subject in the scene. The resulting hologram may be projected as is, and still focus at the correct depths with sharp results; only the defocused areas inherit the effects as intended. The following effects are also possible with this method:

1. Each channel can have independent masks applied to provide chromatic aberration effects.
2. The masking effect can be used to reduce resolution in a specific axis, in a similar fashion to animal pupil shapes [5].

3. The look of an anamorphic lens can be introduced to the hologram without the requirement for any physical anamorphic lens elements in the display - any existing defocus kernel image used within a renderer may be applied.



Fig. 3.17 Input masks may be formed of any image, allowing the same effect as shown in figure 3.13, here a butterfly mask is used. These may also be animated, as shown on the right hand side with a spinning VividQ Logo. The aperture mask is stored in the same EXR as the RGBZ data.

Measuring pupil dilation using eye-tracking in combination with the proposed bokeh control system allows pupil sizes to be matched - where an approximation to the nearest captured dilation size as per [14] can be correctly applied to the holographic projection with the associated aperture shape. Additionally, this ensures that the emulated aperture is always smaller than the viewer's pupil aperture, as per figure 3.14, so that a given artistic bokeh or chosen f-stop can always be visible, even as eyebox sizes increase with advances in SLM resolution.

As the eye aperture is not as wide as some physical lenses used in film production that can achieve a significantly shallow depth of field, where the depth of field is the region (along the Z axis) of the 3D scene in focus, such an eye-tracked system may also apply some pre-processing to exaggerate defocussed areas in real-time rendered imagery by applying a blur kernel in 2D. This is similar to "varifocal" displays which track the eye and apply a digital blur (whilst moving the image plane to the focus distance of the viewer) to regions of the image which are outside of the focus field. Where this is an especially necessary adjustment is in brighter scenes, where the amount of natural defocus would be reduced due to the viewer's

pupil constricting to reduce the amount of light into the eye, and consequently deepening the depth of field. Adjusting the z-depth data to exaggerate depth differences and force the appearance of shallow depth-of-field is also possible with real-time and offline rendered imagery in combination with this method, but the results would not be as pronounced given this is still limited by the optical system. This bokeh control technique may be applied to both real-time or pre-generated holograms with mask shapes stored as an extra layer in the HQH compliant input.

### 3.6.2 Lens Flare and Glare

Other lensing effects are commonplace in existing media, despite not being present in normal vision, and also form a significant contribution to the character and photo-realism of an image. Effects such as lens flares, glare and dirt or other material on the lens is frequently used to the aid of the storytelling and visual experience, therefore it is desirable to replicate this in holographic display. Examples of this are given by figure 3.18.



Fig. 3.18 A selection of lens based effects in media. Top Left: Lens flares in The Lego Movie [203], Top Right: Lens flares in Transformers Revenge of the Fallen [204]. Bottom: Blood splatter on the lens in Breaking Bad [18] (left) and District 9 [52] (right).

Lens flares are typically caused by internal reflections within the lens capturing the image, and elements like dirt or blood on the lens are at the length of the lens from the sensor. Though, when these effects are added in post production, they will likely not need to be accurate, and instead give only the visual appearance desired - there are physically-based examples available with varying approximations [90, 164]. In holographic display, these

elements may be placed at depths close to the eye, where a viewer cannot focus on them by design, or placed at varying locations through the scene. Where elements are placed close to the eye, the resulting image will differ from that of which is input, as it will defocus - this will be inconvenient for the intended look of the shot, so it will also be desirable to place pre-defocused elements as would be used for 2D productions into the replay field.

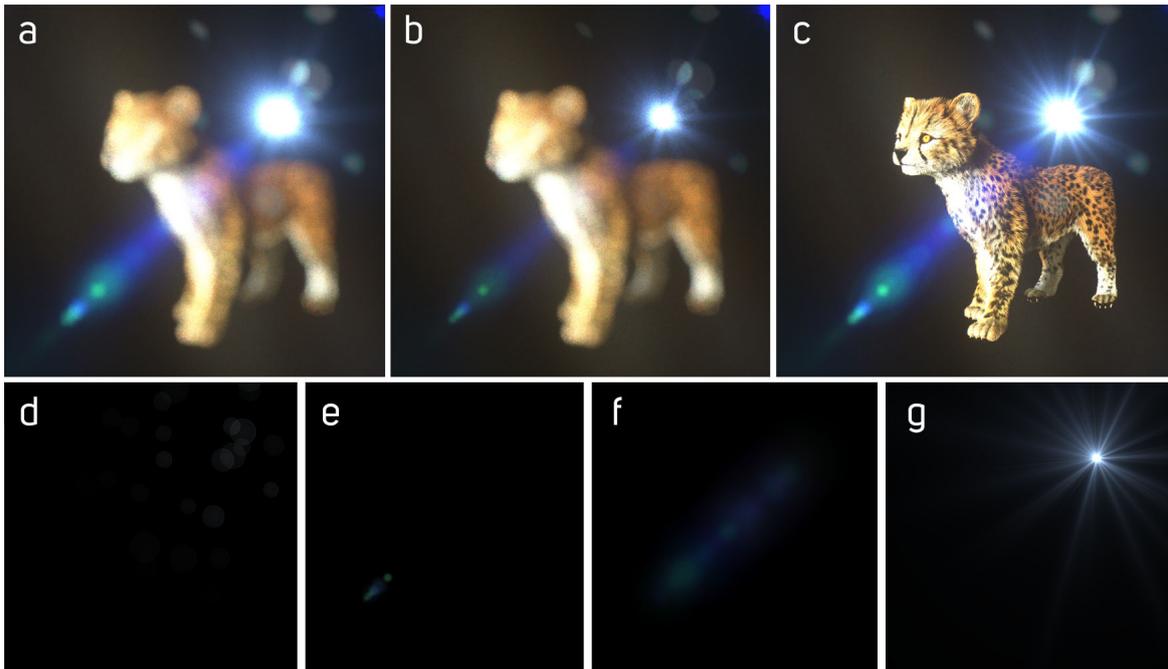


Fig. 3.19 A selection of flare elements (20cm from camera) are added to the hologram of a cheetah render (10m from camera). (a) The replay is focused at 10cm and significant defocus is shown in the flare. (b) By masking the aperture of the flare hologram with the method in the previous section, the defocus may be reduced, additional luminance was not added to compensate for the f-stop change. (c) The replay is focused on the cheetah (10m), the flare is in focus at 20cm because of the aperture control. (d-g): Various input flare elements can be used, each may be placed at individual depths as desired.

Using the method outlined in the previous section, the depth of field may be made deeper by using a smaller aperture shape on a specific element layer; the effects of this are shown in figure 3.19. With this method, flares or other elements will not defocus as much as the rest of the scene, allowing pre-blurred elements to be used more effectively.

So far in this section only photo-real lens effects or effects which are used in existing media formats have been explored, but adding effects visible to the eye itself can also be explored for potentially increasing realism, rather than photo-realism. Given that a glare can raise the perceived luminance by 20-35% [225], it is also a valuable tool where the luminance of the display is limited to a lower level to make a scene appear brighter. Glare layers may

be added as any other lens element, but some level of glow is naturally present in a replay given a high enough luminance given the viewing optics point-spread function. Though any manually curated lens effects may not match the characteristics of that of a real eye glare under all conditions, given luminance and eyebox constraints. In order to reproduce the appearance of such, the techniques of “Temporal Glare: Real Time Dynamic Simulation of the Scattering in the Human Eye” [175] may be applied using the aperture control method outlined in the previous section. This allows elements such as vitreous particles and eyelashes to be implemented in the holographic display, where they will not form naturally at lower luminances. These three methods are shown in figure 3.20.

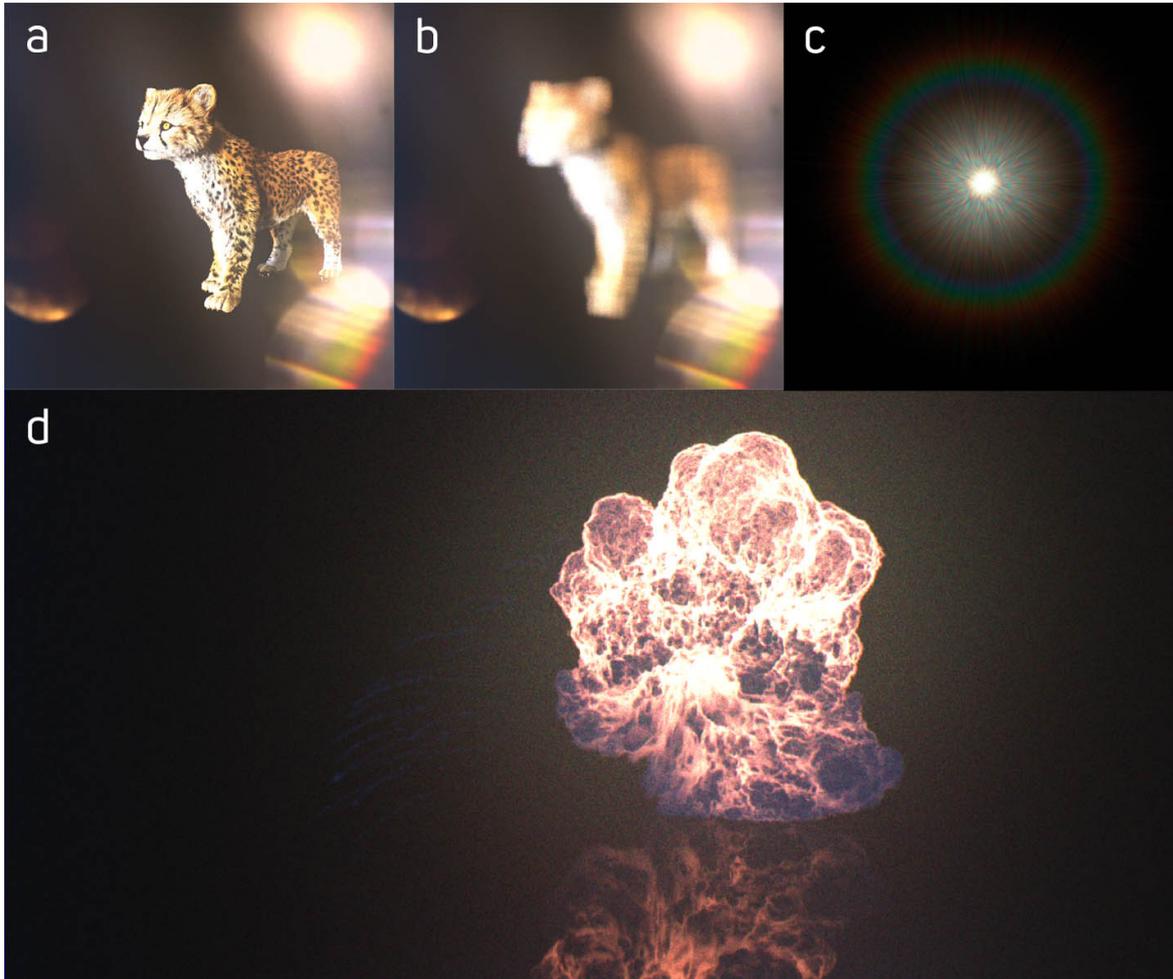


Fig. 3.20 (a) and (b) Show manually added lens glares added that do not vary in focus using a controlled aperture size. (c) Shows a Human Eye Scattering glare from [175], which may also be implemented in the imagery using the methods outlined. (d) Bottom shows how a natural glow/glare will form around high luminances in cases when correctly using a HDR pipeline. Here a simulated a fuel explosion (SideFX Houdini, Mantra) presents a diffuse glow when the image is replayed, and is not present in the original render.

### 3.6.3 Field Curvature

Field curvature may be visibly present in images where the centre is in focus but the edges are not. This is considered an optical error where the lens elements may cause the resulting image to be formed in a dome-like fashion rather than one that is a plane - but it also adds some pleasant effects to the image, making it appear dream-like and helps draw the attention of the viewer. Figure 3.21 illustrates how this may appear on an image of a grid, with the masking region visualised.

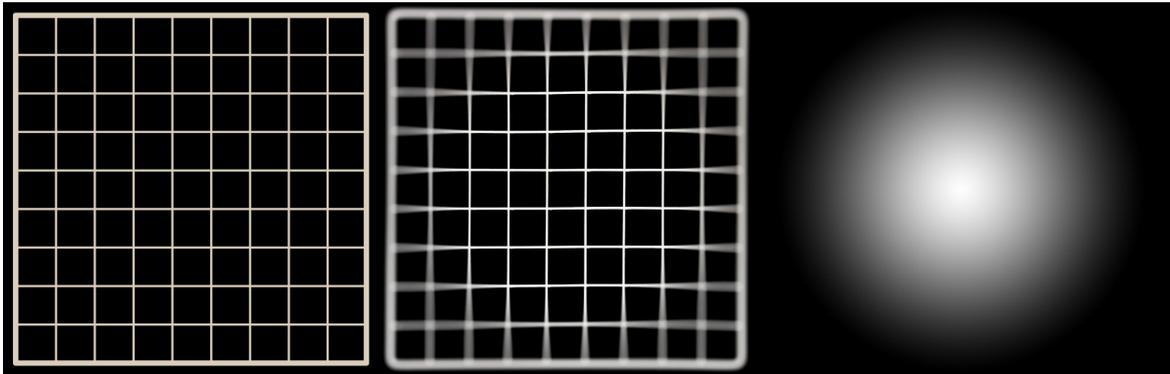


Fig. 3.21 Left: An input source image. Mid: The resulting replay when field curvature is applied. Right: A visualisation of a depth multiplier mask.

This is achieved in holographic display by using a vignette/radial mask to multiply the depth values in the input, this can be performed within a compositing DCC before passing images to hologram generation. Values in the centre will remain in the correct position as the input depths, whereas values at the edges will be multiplied in order to, in this case, be moved closer to the viewer to create a warped replay field. This technique can also be used to correct errors within the display system optics to ensure that imagery that is intended to be planar does indeed all focus at the same depth, as well as a wider range of other aberration corrections, with a wide exploration available in "Adaptive Aberration Correction for Holographic Projectors", Kaczorowski, 2017 [106].

## 3.7 Implementation in Production Environment

Given the previous sections, there are a few final considerations required in order to implement a full holographic production workflow. It is important to outline this early as to ensure any further work or research that occurs can be put into practice in a way that is compatible with the existing production environment, existing artist tools and the existing production process. Proprietary solutions can be of course developed in isolation for CGH, but it would

Table 3.2 Production Environment Processes

Process	Form	Interaction	Result
Production Configuration	Config File or UI Builder Application.	Director, VFX Supervisor, Technical Artist	Input Image Specifications with compatible Holographic Algorithm Specifications for HQH Format. Includes max luminance and wavelength choices.
DCC App Config	DCC Plugins or Setup Scripts	Automated per Config	Loads appropriate config into DCC packages for correct settings/ exports (incl. layer requirements), and general scene/render setup. Allow artist hologram previews.
Colour Management	OpenColorIO Config	VFX Supervisor, Pipeline TD, Deployed to Artists	Colour space locked across all artists working on production, ensures all content looks as intended on holographic projector.
Optical System Controls	Compositing Nodes and CGH Generation Step	Cinematographer, Compositor/ Automated	Consistent defocus look, and control where it is required in the composite for defocus size/ shape/animation. Includes control of field curvature.

be detrimental to ignore the wealth of tools already built for successful image creation in existing media production.

### 3.7.1 Tooling Configurations

Given that any combination of the tools and methods proposed by the previous sections could be selected to be integrated within a given production, it is by design that this workflow is not too dissimilar to existing practice. Specialist tools, listed below in table 3.2 can be dropped in within existing production steps, and most automated to some degree.

A production configuration can be built as a reference for any applications in the pipeline to apply the correct settings for scene building or file exports, as well as define the hologram generation algorithm settings as a baseline, some settings can be overridden on shots where required. A sample config is given below in figure 3.22.

This config file is not only loaded by the hologram generation package, but also any DCC packages that interact with the pipeline, via plugins or scripts. It is also the goal that the hologram pipeline is in line with the VFX Reference Platform [210] release in order to maximise compatibility across productions. DCC applications such as Maya or Houdini

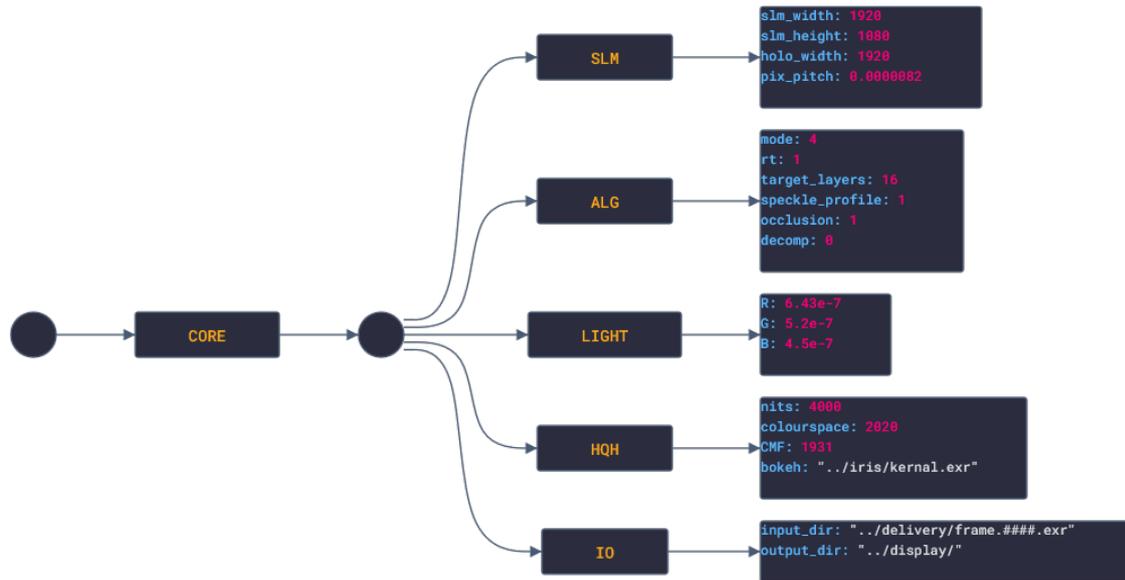


Fig. 3.22 An example configuration that specifies hologram generation properties and defines some HQH standard properties. This config file may also contain the path to the OCIO configuration, define specific views to generate, and to map specific image channels to additional render layers - such as AOVs, explored in section 4.3.2, or elements such as flares.

should apply a scene setup routine for hologram generation, though it is likely a studio's pipeline will already have a mechanism for this. This includes procedures such as:

- Colourspace, by loading an OCIO profile to apply working space, display transforms, and output colour management. The transform for a given holographic projector is calculated from wavelength primaries and a chosen CMF, where it is then added to the OCIO config. Here, shaders are exported for GPU acceleration of colour transforms and the application of the EOTF.
- Cameras, by generating a virtual camera rig with appropriate number of cameras and resolution settings for each. Camera rigs may wish to consider the field of view of the corresponding holographic projector depending on the display application.
- AOVs, for automatically setting up scenes to output any required passes for compositing and for hologram generation, such as various depth passes (see section 4.3.2)

And this can then be adjusted on a per shot basis, such as by restricting the range of views, having a requirement for Deep rendering (see section 4.5), or render layer specifications. It will certainly be the case that any DCC software viewport and renderer will require a form of real-time or interactive hologram generation to allow previews directly on holographic

displays. The same plugin that handles config loading can also handle previewing, in a similar fashion to how a renderer plugs into a DCC app. Given many studios are moving to Universal Scene Description (USD) [208] based workflows, any hologram generation plugin could exist in the form of integration with Hydra [59], an open-source framework to transport scene graph data to a renderer. In Hydra, renderers are abstracted, this allows both viewport and final frame rendering, with hologram generation, via a Hydra enabled application.

A compositing application will be setup similarly, either in 3D or with Multiview enabled in the node tree (both of which are supported by Nuke). In this case a shot may have specific view-based needs, such as colour adjustments or bokeh shapes, so there is a requirement for such a setup to easily enable a spatial level of editing, such as the “OneView” node in Nuke. It must also automatically re-merge all of the required views into a single output EXR.

Once the composite is assembled, the imagery is passed to the colourist. Many grading focused packages, such as Blackmagic Davinci Resolve [44] and FilmLight’s Baselight [10] support a stereo workflow, so out of the box a standard grading process is supported and can be applied to all views of a scene. Any enhanced editing is not within the scope of this thesis, but artists could benefit from features such eye tracking or virtual view simulation for view based spatial colour grading, grade interpolation for smooth view specific effects and a “global view” for a clear perspective of grading the entire 3D scene.

Hologram generation itself, either for final-frame or for mid-production previews can be easily bundled into render jobs. Render management systems such as Deadline [46] or OpenCue [156] or similar. Alongside offering large flexibility in deployment, these systems allow jobs submitted to contain any type of work, including initiating hologram generation on a frame sequence via command-line interface (CLI). Imagery for holo-gen is collected from render nodes and bundled into the HQH standard file and submitted to the hologram generation CLI as a complete final IMF package, or as preview dailies as frame sequences. Hologram generation is highly parallelisable and can be performed on the GPU in batches for speed, and with the increase in prevalence of GPU rendering being utilised; this should not require any additional setup for an existing pool of GPU render nodes. Any difference in FFT implementation for layer-based holography may differ based on CPU or GPU processing, but any variations will not be significant, if at all visible. This can all be specified and applied to the production via automated procedures according to the configuration, and any additional image channels required can be enabled in the appropriate DCC software.

Using the workflow and enhancements developed across the previous sections, which are described in table 3.2, examples of the end-to-end compute process for holographic display are described in the next section.

### 3.7.2 Compute Process

The holographic compute process of this workflow assumes an existing studio pipeline, along with render nodes that have GPU hardware available. Figure 3.23 offers a high-level overview of what this workflow might look like in practice - where holographic display offers improved 3D viewing of the scene, benefiting the animation or modelling process by displaying a more natural viewport, right through to the final grading process for delivery. Individual workstations or servers may render interactively and preview holograms via a holographic display, or the previews may be simulated. Given the parallelisation support, hologram generation may be sent to a local compute cluster made up of networked workstations or dedicated render blades [70].

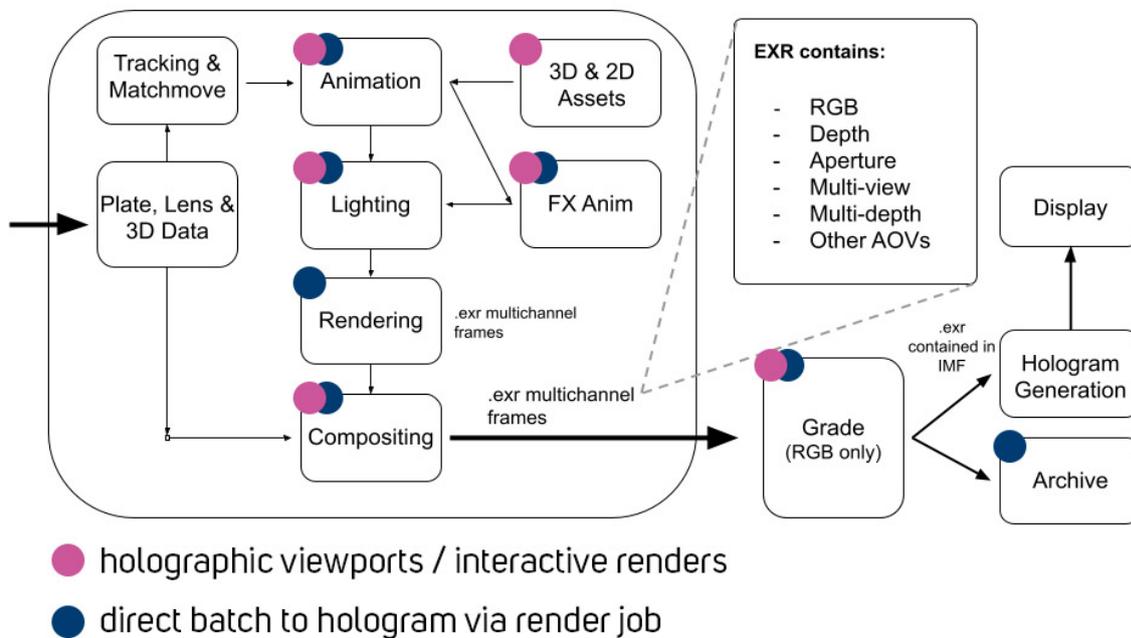


Fig. 3.23 An overview of stages of the VFX process with suitable interaction with hologram generation marked by the coloured dots. An interactive or batch render process will utilise the specified config to provide the production environment settings for submitting hologram generation tasks.

Compute can also be performed on remote machines - Though it's important to note that for any interactive viewport, the usual way remote desktops operate would not be able to support the streaming of holographic frames without an extremely capable internet connection or other newly implemented compression algorithms - traditional video encoding is not suited for holographic data due to the differences found in the signals compared to natural imagery. Compression of holograms and holographic video is an active area of

research [147, 179, 15, 213]. Any holographic processing for artist-side viewing, at this time, should be performed locally, but is increasingly possible using remote compute. Just as the rendering compute implementation itself can vary, the holographic generation compute can also make use of this provided, whether by generic compute tasks such as via Google Cloud Platform with OpenCue or Amazon Web Services with Deadline - an example overview is outlined in figure 3.24. There are many out-of-the-box solutions for render job scheduling which are not covered here, and neither covered is the case of full realtime holographic rendering in the cloud to be streamed directly to a viewing device. In any of these cases, GPU acceleration is preferred, but not strictly necessary.

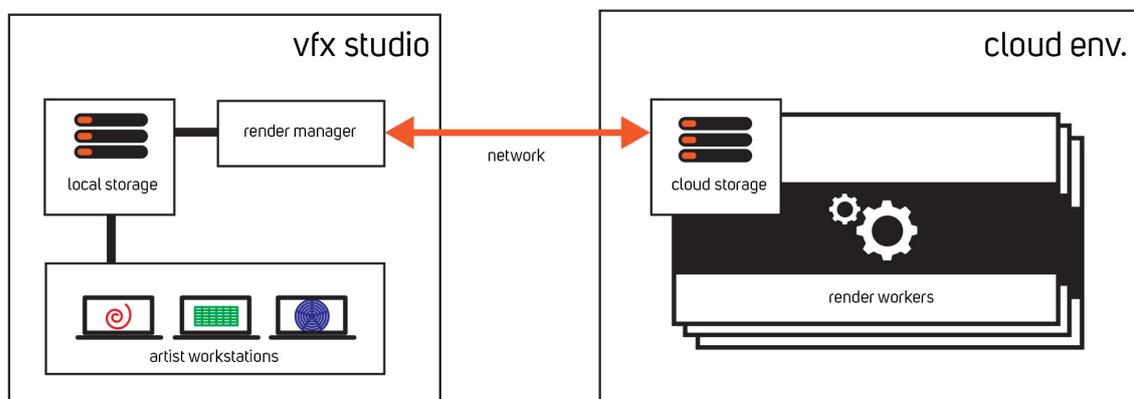


Fig. 3.24 A simple infrastructure overview showing how a VFX studio may utilise cloud rendering - the same render workers can also perform hologram generation, ideally on GPU enabled workers.

## 3.8 Holographic Projector Calibration

Ensuring a display is calibrated is essential for production work. Colours, luminances and whitepoints must meet a specific display standard to ensure that all artists working on the production will be seeing the same images - or so that a specific display can be used as the main reference display, this is all a normal part of the production process and crucial to ensure artists can see what they're working on accurately.

In order to calibrate a holographic display's colour and peak luminance, a similar technique may be used as a conventional laser projector, with Target Colour Gamut Data (TCGD) and Measured Colour Gamut Data (MCGD). Reaching a specific TCGD is achieved by adjusting the drive current of each laser primary, and adjusting for operating temperatures. The maximum drive current of a given laser, optical system efficiency and eye safety will

determine the luminance limit of the system. Such a system is described in detail in this Laser technical FAQ from Christie [126].

As the whitepoint (the chromaticity of a white object under an illuminant) is defined by the media colour space, this may be read from the incoming hologram metadata, such as that defined by OCIO, and used to load a given colourspace profile to the display. Colourspaces Rec.2020 and Rec.709 use the D65 illuminant, for example, which is defined by the CIE - and corresponds to average midday light in northern europe. D65 has a correlated colour temperature (CCT) of approximately 6500 K. This colour temperature is so named as it defines how cool or warm an image appears, in kelvin. The planckian locus (the curve) in the central region where "whites" appear on the CIE diagram in figure 3.25 defines a region of chromaticities that may be made up of black-body light sources.

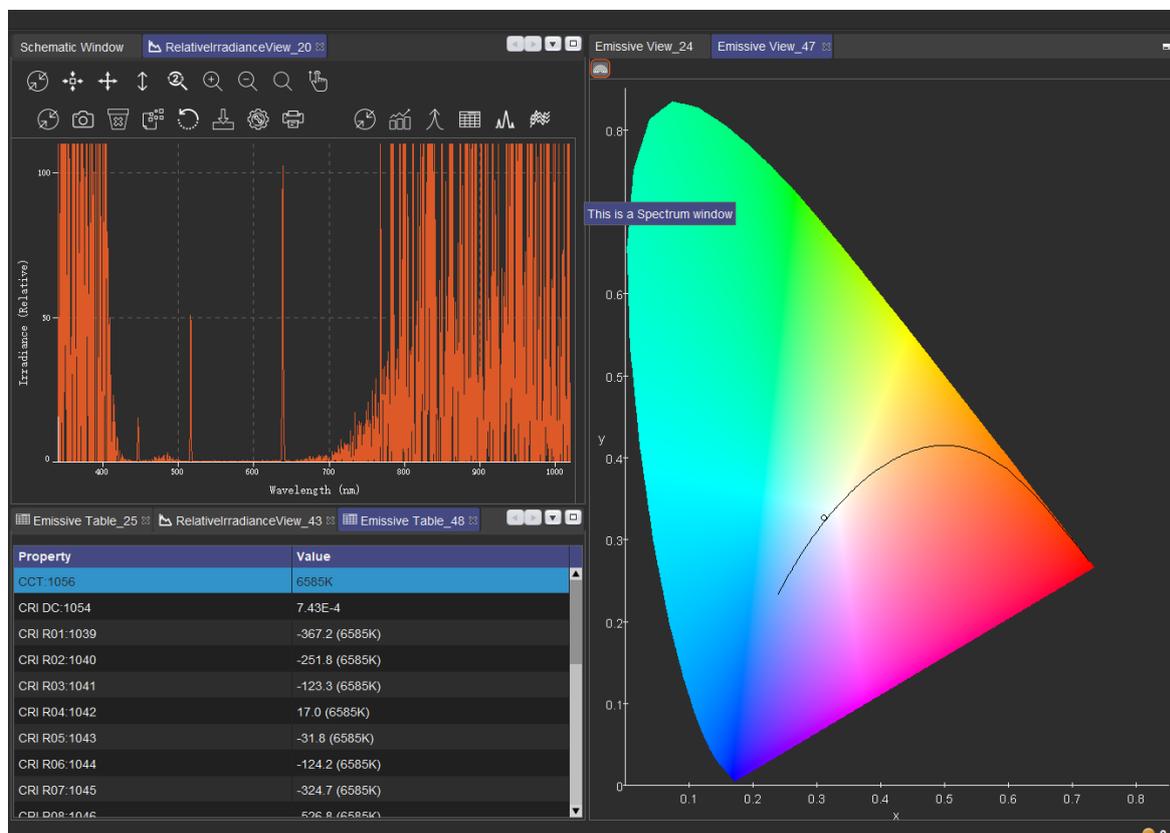


Fig. 3.25 A screenshot from "OceanView", spectrometer software measuring a real holographic projector displaying a white image. The spectrum of the hologram is shown on the orange graph, with 3 peaks at the R, G and B laser wavelengths. The circle on the CIE diagram, at the planckian locus, is the measured CCT of the whitepoint being displayed, read as a value of 6585k with an aim for 6500k (D65). The power being delivered to each laser can be adjusted to adjust this whitepoint.

This can be measured and thereby set in a holographic display by utilising a spectrometer, here an Ocean Insight Flame is used, and figure 3.25 shows a screenshot of the “OceanView” software during the calibration process. The projector must be allowed to reach operating temperature before measuring, and is recommended to be about 15 minutes (as defined in the International Committee for Display Metrology’s measurement guide, the Information Display Measurements Standard [97]). Here a white plane is displayed on the holographic display, and a lens is used to focus the replay field into a small point which is directed into the fibre cable of the spectrometer. By then adjusting the laser drivers, the correct laser powers can be found to reach the specified whitepoint, here D65 is also chosen. This same measurement data forms the MCGD, and can then be used to form the target display gamut for colourspace conversions to be performed correctly. As this specific projector is a prototype, this driver power adjustment did not have the resolution required to match the 6500 target precisely, with a result of 6585k shown in the CCT property.

Here, the luminance of the display may be measured, and as the relationship between the current provided to a laser and the power of the light exiting the laser is linear, the set whitepoint currents may be scaled to adjust the display luminance to one that is specified. When combined with a complete optical aberration correction system, such as in [106], this proposed calibration system provides a method of ensuring display accuracy and consistency.

### 3.9 Production Workflow Results

Prior to the work in this chapter, VividQ only enabled input to their hologram generation algorithms via the use of proprietary format “CDM” (Colour Depth Map), and had a script that could run in the game engine Unity to output the 8-bit RGBZ channels. This format was unhelpful as no image editing applications could load it, and it only supported SDR with limited metadata, so a full EXR loading implementation was instead written in its place. The OpenEXR library, a familiar staple in advanced media applications and studio pipelines, is used to access the data outlined in section 3.5.6. OCIO was then included in the generation software, ensuring correct transforms are read from the workflow config, and applied appropriately for the defined output display wavelengths. This allows hologram generation to be compatible with the proposed workflow and the workflow of likely every major production studio.

### 3.9.1 Improved Image Quality & Control

By improving the quality of the information being delivered to the hologram generation algorithm, by passing through unedited “raw” render output data, it removes the limits imposed by any historic SDR workflows. The increase in bit-depth not only unlocks the possibility of large luminances and wide colourspaces, but also serves as a valuable tool to debug holographic replays when working in an SDR environment, as shown by adjusting the exposure in a replay after its simulation in figure 3.26. This is simply not possible when working with typically quantised input data formats. This presented full scene-linear workflow is required to properly conserve light values, and other research in CGH would greatly benefit from such. By ensuring that the scene data is not the limiting factor, for example, research in lowering the noise floor of holographic replays, whether in algorithm or hardware, would instantly enable an improved holographic HDR experience in practice.



Fig. 3.26 A range of exposures across a holographic replay of the same input scene, of -10, -5, -2, 0, 1, 2 (stops), where the top row has an 8 bit input with no consideration given to EOTF or colourspace, and assumes the sRGB standard. The bottom row is produced after following the methods proposed in this chapter, a full HDR linear workflow with colour-management (rec.709). This full scene-linear workflow is required to properly conserve light values.

Multi-focal productions will be a change for how artists and directors work, but this work in section 3.6 also proposes methods to make use of this creatively whilst also replicating effects that are familiar, such as via lens effects which are already commonly used.

### 3.9.2 Conclusion

Providing a system of hologram generation compatible with existing workflows is not only convenient but also critical to accelerating the acceptance of holographic display by lowering the barrier of entry. Using the workflow proposed, scenes such as that of figure 3.27 may be created in a DCC package without technical hindrance. This example is made of several

elements composited together in the application Nuke, where depths were manually assigned to elements and then all written out into a single EXR file for processing into a hologram for display on hardware. Providing a process that can integrate with existing software and workflows has been a key element of this chapter.

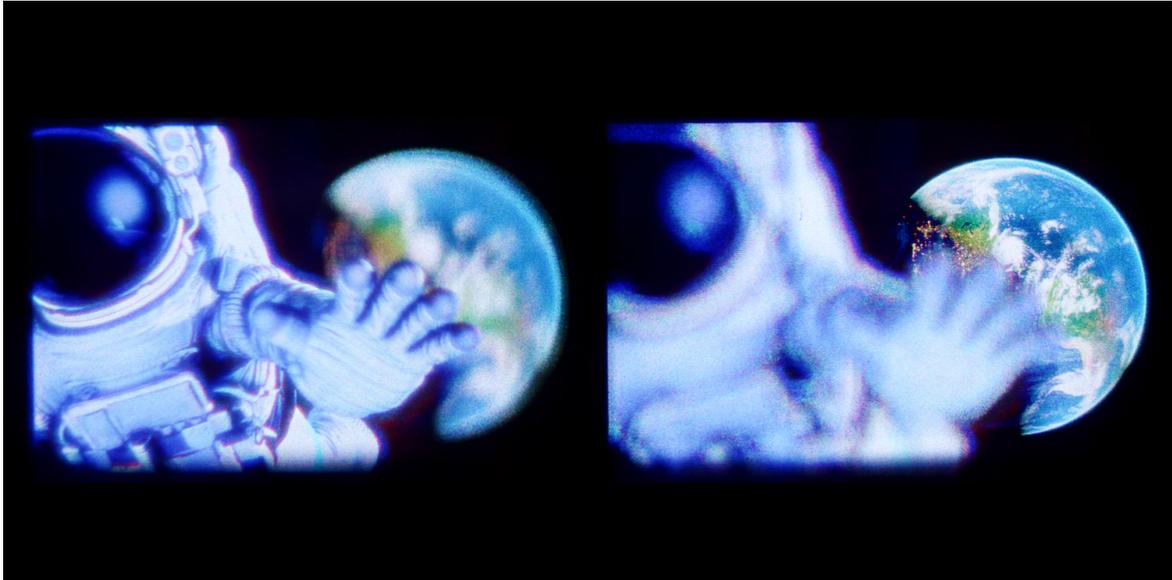


Fig. 3.27 A holographic replay captured (Using a Blackmagic Pocket Cinema Camera 4K) from a real projector of a scene built in The Foundry's Nuke. Left shows focus on the spaceman, right shows focus on the earth. The scene source has been fully processed using the workflow laid out in this chapter, illustrated here by media created specifically for holographic display.

This unlimiting architecture ensures that, as in 3.27, new display hardware can be instantly taken advantage of - but also means that existing media may be instantly processed or “remastered” for holographic display using data that likely already exists in production archives. The creative intent of previous productions can be reproduced precisely by enabling the luminance and colourspace standards set by the original, as shown in high quality holographic replays in figure 3.28 for both rendered and captured imagery. Equally, new productions may wish to make use of advances in colour science by adopting new colour matching functions, and multi-spectral display - all of which are possible under this proposed workflow. No matter what display standard or hardware developments arrive beyond that which is proposed here, this workflow will be able to adapt to support it.

The opportunity to define a next generation display pipeline is most importantly a chance to learn from the issues facing previous and current display technologies. By defining the ideal theoretical workflow early, it is hoped that a discussion can be started on how to work together towards an immersive future with holographic display. To that end, a complete



Fig. 3.28 Top: A holographic replay simulation of "Cosmos Laundromat", rendered in Blender and delivered to Netflix in PQ HDR with a colourspace of DCI-P3 (sourced from [opencontent.netflix.com](https://opencontent.netflix.com) [148]). Bottom: A holographic replay simulation of an image previously shot using the Blackmagic Pocket Cinema Camera 4K using the BRAW format, in HDR with a colourspace of BT.709. The scene source has been fully processed using the workflow laid out in this chapter, illustrated here by media not originally intended for holographic display.

---

workflow framework for holographic media has been proposed, alongside dynamic range handling, wide colour gamut handling, and legacy content support - as well as a selection of unique controls such as optical effects. This culminates in a selection of techniques that allows creatives to reproduce their creative intent in existing media on holographic display, and to create new media with broader technical and artistic possibilities in 3D. This chapter also enables the chapters proceeding it, by providing a foundation on which to utilise existing production pipelines to greatly increase the quality of the holographic display image - in a way which is practical and readily integrated via existing software packages.



# Chapter 4

## Multi-Sampled Depth

### 4.1 Limits of the Z Channel

As described in section 2.2.1, layer-based holography makes a number of approximations in order to prioritise speed. These approximations include representing a scene from a single render-perspective, and using a 2.5D source that contains colour data with a single depth channel, which places each point in 3D space. This RGBZ source for hologram generation is sufficient to represent solid diffuse surfaces from only this given perspective. When trying to represent more complex scenes, or when viewing a scene from a different position to which it was rendered from, a single depth value is no longer suitable in a number of cases.

By incorrectly making the assumption that every object in a scene can be represented by a single depth value per pixel, many objects will exist wholly or partially at incorrect distances from the viewer, and exhibit incorrect parallax. When such a replay field is viewed from a position other than the precise virtual camera position, or when the perceived image has sufficiently defocused elements, the occlusion that has occurred in the source render is visible, meaning that the viewer can see behind objects - but nothing has been rendered for these regions, gaps appear, and the immersion of the scene is broken.

In order to represent a scene realistically, these complex situations must be considered and consequently supported. Figure 4.1 shows an image which has a number of complex scene features which cannot be represented by single depth values. Here, the image is of someone taking a photo of a landscape out of a bus window - and the viewer looks over her shoulder. There are few solid items in this scene that may be represented by a single depth

---

Parts of this chapter were published in Demolder, A, et al. 2021 [50] and Patent GB2214017.2 [48]



Fig. 4.1 This image, a still from a raw format video, shows the large amount of complexity in a real scene that must be replicated by any holographic display system. Features include refractions, reflections, see-through material, motion blurs, volumes and anti-aliasing in thin objects such as hair, as well as occlusions around regions of blur and defocus.

value with hard edges, such as a window frame and her head, face and phone - but the rest of the scene requires special attention. Her strands of hair are thin, requiring anti-aliasing, and her glasses refract the scene - each of which defocus to reveal the background. The curtain over the window at the right side of the scene is not fully opaque, and behind it the window both refracts and reflects, showing the landscape and a reflection of the interior of the bus. The foreground outside of the window is in motion and blurred, the water in the distance reflects the sky, atmospheric effects reduce visibility towards the horizon, and clouds in the sky attenuate the light. So far this also only assumes that the viewer is fixed at the render/capture position, and in holographic display the viewer may move to a new viewing position within the eyebox, and imagery occluded (behind objects) will need to be revealed, and other parts of the image consequently occluded.

Where objects can be viewed in reflections (such as in mirrors) and through refractions (such as through glass or water); the depth at which these objects come into focus is defined by the focal power of the material in question. When rendering convincing imagery in computer graphics, realistic representation of these materials is essential to the realism of the final image. These material representations do not pose an issue for rendering for 2D display, as only the resultant pixel intensities are required. Occlusions and semi-transparencies are calculated - and focus is pre-determined by a depth of field routine in the renderer, or by

post processing. Doing this post processing requires the same depth data of the scene as in image/layer-based CGH, where z-depth values are used that correspond to the focus distance of objects. While this technique works well enough for the first object hit by the raytracer, it is unable to accurately encode the depth at which virtual objects appear after reflection, refraction or simply where there are any semi-transparencies. It is hence necessary that in order to create perceptually correct scenes, the renderer must pass multiple depth values into the pipeline which allows the hologram generation engine to ensure realistic representation of such complexities.

Figure 4.2 shows how this same scene may be represented by single depth values, compared to how it ought to be represented by multi-sampled depths.

This chapter outlines and contributes novel methods to support such complex situations in imagery in holographic display, including:

- Materials that are semi-transparent
- Optical properties of materials - reflections and refractions
- Motion blur
- Volumetric effects such as smoke, fire, mist/fog
- Anti-aliased edges (edges of objects, thin objects such as fur/hair)
- Occluded regions revealed by defocus
- Occluded regions revealed by parallax
- Occlusions caused by parallax

How each of these features is achieved in conventional computer graphics is described in section 4.2. These elements are then enabled holographically by proposed methods, and they may then successfully be viewed in the replay field - where the image correctly forms a representation akin to the bottom image of figure 4.2, rather than the incorrect approach of the top image.

This presents a significant increase in image fidelity to the viewer, whilst still remaining compatible with the layer-based method. The production workflow outlined in the previous chapter enables more than one depth value per pixel to be delivered to the hologram generation algorithm, and this chapter enables the algorithm to successfully utilise this data to deliver an improved viewing experience.

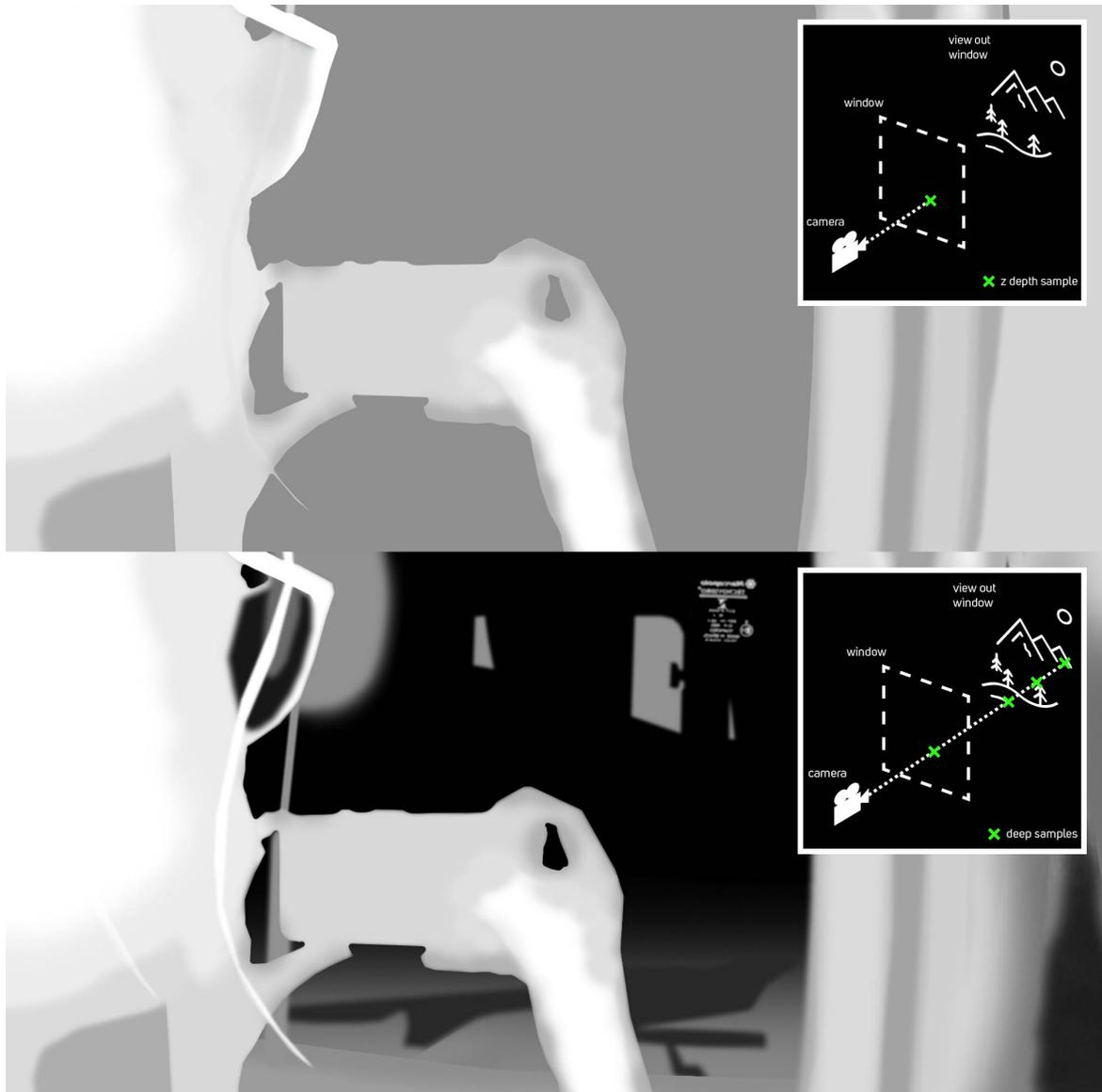


Fig. 4.2 An illustration from a single perspective of how the image in figure 4.1 may be represented by a single depth value (top) without depth data from through the window, compared to multiple depth values per pixel (bottom) which allows the accumulation of multiple depth samples through Z and a wider range of image effects to be represented. Note semi-transparencies through the curtain, reflections in the window and water, refractions through the window and glasses, motion blur in the road sign and improved edges in elements such as the hair.

## 4.2 Background

In order to explore how layer-based holograms may overcome the lack of multi-depth representation, a number of features in computer graphics which have already resolved these issues for conventional 2D output displays must first be explored. These are then compared to what is state-of-the-art in CGH.

### 4.2.1 The Z-Buffer in Computer Graphics

Ray-Tracing algorithms, “hidden-surface algorithms”, or “visible-surface algorithms”, allow the display of visible surfaces from a given perspective. These algorithms allow for the realistic rendering of complex scenes for displays that are raster-based, where pixels are used to form an image [23, 16, 216].

For an input 3D scene, typically made up of geometry such as polygons, surfaces visible to the camera must be identified. Various techniques exist to process such regions of a scene - one of the earliest examples being the “Painter’s Algorithm” [150], which orders polygons in the 3D scene by their depth, painting polygons over others in the scene. Other developments come from occlusion culling, which prevents geometry from being drawn to a screen texture in the first place, whilst the Z-Buffer algorithm stores only the frontmost point value in image-space [23]. These algorithms then present pixel values only for surfaces which are visible. Surfaces which are occluded by other surfaces are discarded in the image buffer, as they will not be displayed. GPUs will accelerate this Z-buffering by performing these checks in hardware implementations.

These methods do not necessarily have to use a realistic light transport model, and may have various models applied to vertices or faces. Models such as Lambertian reflectance (for diffuse surfaces) and the Blinn/Phong shading model [16] (for surfaces with a specular component) may be used alone. Other techniques, such as Whitted’s “Improved Illumination Model for Shaded Display” [216], allow light to be calculated in such a way that includes both refractive and reflective surfaces - which increases the complexity of the surface removal, but also greatly advances the image fidelity. Whitted’s model introduces the concept of the ray-tree, and primary and secondary rays, where light paths from various surface interactions contribute to final pixel values through multiple bounces - where a ray does not terminate at the first intersection with an object, unlike with the z-buffer. This ray-tree model is the basis on which modern renderers work. The method by which a ray interacts with a surface is called a “shader”, and there are many shader models (that evaluate how a scene should be sampled, into pixels) that apply varying illumination models (which are continuous) in modern renderers which span between physical correctness and artistic control, one such

example being Disney's Principled BRDF (bidirectional reflectance distribution function) [21]. A number of parameters and approximations will also be made in order to reduce computational complexity - such as defining a "trace-depth" limit, which defines how many bounces a ray may make, for example specifically between reflective materials.

With producing imagery on a fixed grid of pixels, issues such as staircasing or jagged edges, known as "aliasing" are apparent [23, 216]. Aliasing may occur at the edges of objects, where objects are smaller than the relative pixel size, and where textures are projected onto surfaces. Ultimately, aliasing occurs as a result of undersampling, and may be resolved via methods of low-pass filtering, or utilising a "warnock-type" divide-and-conquer technique [214] to subdivide pixels for intensity sampling. Other methods will average a number of subdivided screen-space samples within each pixel, where foreground and background pixels are averaged [188]. In recent real-time applications, temporal based anti-aliasing is common [223].

These anti-aliased edges, as well as objects or shaders that are semi-transparent, may introduce an alpha channel to the framebuffer [169] if elements are not rendered all together at once. For a given object, these pixel values represent the "matte", which defines the occlusion the object has. This allows for a new image to be formed of several separately rendered elements - this is referred to as compositing, as mentioned in the workflow chapter previously.

All of these techniques may form imagery whilst utilising a pin-hole camera model, where a perfect lens is simulated - and all of the scene be in-focus. But real lenses have a non-zero width, which results in a "circle of confusion", and therefore a defocus effect. This is shown in figure 4.3. Adjusting this lens aperture shape and size has been explored in section 3.6. The larger the aperture, the larger the circle of confusion and consequently the size of the visible defocus. In computer graphics, use of such a camera model allows more realistic rendering to occur [170] for 2D display, typically via a thin-lens model [83]. As defocused objects are blurred, they introduce a region of semi-transparent occlusion.

In a similar fashion, semi-transparent regions may also be introduced by motion-blur. This is where the camera shutter-time is simulated in the image renderer. In a raytracer, this may be calculated and produced in the raster output by utilising the path and velocity of objects in the scene, alongside the exposure time of the camera. This may occur via various methods [171, 39, 143].

As these techniques are all applied in order to determine the raster output, the Z-buffer is often either discarded, or retained as a depth raster image for only the first-most ray-surface intersection. This depth value, typically being a single buffer, does not contain any awareness of any semi-transparencies or secondary rays, such as anti-aliasing, motion blur, defocus,

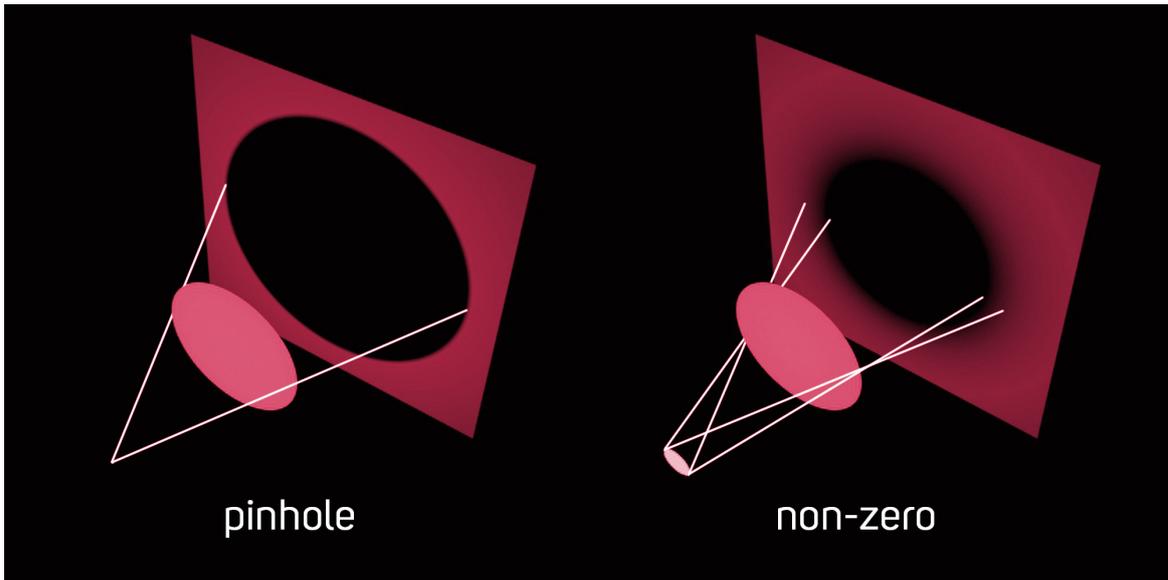


Fig. 4.3 A pinhole aperture is compared with one that is non-zero. In this illustration in the pinhole model, light is projected from a single point and results in a shadow that has sharp edges - it is not defocused. In the non-zero model, the light source has a physical size and the resulting shadow has a blurred edge. While here this is shown via light sources, the same applies to imaging apertures. A pinhole camera will always be in focus and will not exhibit blur, while one that has a physically sized aperture will.

reflections or refractions or for volumetric shaders. The z-buffer was never intended to be used outside of the renderer and is limited. Yet where this depth data is utilised in a post-processing fashion at the compositing stage, such as for depth of field emulation or depth compositing, it is problematic even with a high bit-depth value. As the edges of an object's depth buffer are aliased, but the colour may be anti-aliased, the colour information will not properly represent the object's occlusions defined by the depth buffer. Equally if the depth-buffer is anti-aliased, the depth values no longer represent the actual depth of a pixel. This is especially apparent on edges between foreground and background objects, where depth values may end up somewhere in-between the front and back depths. Similar representation issues occur on objects that have a thinner width than 1 pixel, such as with hair, or where rays are sampled across space/time (defocus/motion) - a single depth value cannot represent these pixels correctly.

Better information may be formed in the renderer to improve the accuracy of depth information - this was introduced first as an extension to REYES [38], which stored multiple hits through Z in the buffer to allow transparency and constructive solid geometry, and later by "Deep Shadow Maps" [135], which allows a representation of fractional visibility for a given pixel through all possible depths. Rather than recording a first-surface intersection for

a ray at e.g. the front surface of an object with semi-transparency, this depth representation records the light values through many positions in depth, sampling through the object and the objects behind it, resulting in a deeper depth buffer. This technique then enables “deep compositing”, which allows renders with deep data to be brought into compositing DCC packages via the OpenEXR format [87, 107], and ensures that the problematic cases are minimised. Now, anti-aliased edges, fur/hair, motion blur and complex semi-transparent volumetric objects may be represented, and combined in a compositing application with other elements, correctly.

#### 4.2.2 The Z-Buffer in Layer-Based CGH

While a hologram may be formed where each hologram pixel individually contributes to the replay field, as outlined in section 2.1.3, and may be integrated with techniques such as the blinn/phong model [153] for view-dependant shading, the scenes in layer-based holograms are instead sliced from the output 2D image of a conventional renderer. This image is sliced into depth regions/near-far clipping planes which allows a number of image planes to be used to represent segments of the scene and reduce the number of Fourier Transforms required. This slicing occurs by using the Z data (which stores only the front-most ray/surface intersection) to determine a number of fractional images, and these may be sliced evenly, or with a weighting to the front, as an object closer to the viewer will have more parallax - so will need more granular depth information for the correct phase factors to be applied to each layer, here with the "defocus" Zernike polynomial aberration being applied to change the focus of the layer wavefront - an overview of this is available here[124]. This ensures that the layers appear to have the correct focus and parallax behaviour. Such front-weighted slicing is often performed using constant dioptric spacing [7]. There is also a suggestion of a minimum layer density [71], which ultimately recommends a rather dense layered image.

Each of these layers in the stack is calculated in batches on the GPU, with individual holograms being calculated per colour channel, as each is wavelength dependent. As holograms are additive in the same way light is, once each layer with its associated depth is generated into a hologram, they may be added together. The final hologram will contain all of the information for all of the depth layers encoded.

A key issue identified in holographic display/CGH is how such a hologram performs occlusions between each layer. As layer-based CGH utilises the RGBZ channels associated with a raster image that has already been rendered, it has already been occluded for 2D output with a pinhole aperture. Yet, the replay field formed by the hologram is an image formed by a non-zero aperture, and is 3D. This causes the following viewing artefacts:

- When a viewer is not at the precise position of the render perspective within the replay field eyebox, it is possible to see behind objects and layer slices, as shown by figure 4.4 - in these regions behind layers there is no colour (or other) information as it has not been rendered to the scene-input, so areas behind layers appear black. Layers may also overlap to cause bright artefacts. This presents only a literal representation of a single perspective of the scene to the viewer.
- When a viewer focuses on objects in the distance, layers closer to the viewer will defocus. At the edges of this defocused object, the black gaps of missing information again appear, as shown in figure 4.5. If the viewer focuses on other layers closer to them, a black halo will form around the object layer as the background defocuses.

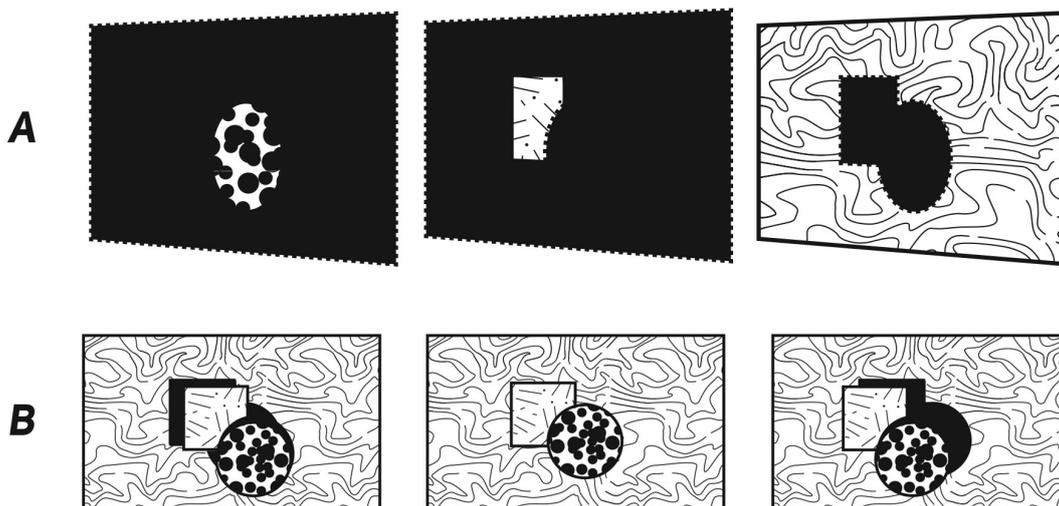


Fig. 4.4 Occlusion issues in viewing position when utilising pre-occluded inputs for layer-based CGH. A) Shows the input layers, a circle, a partially occluded square, and a partially occluded background. B) Shows this scene as viewed from left, centre and right positions respectively.

There are two key ways to resolve the occlusion issues in CGH:

### Geometric Occlusion

Geometric Occlusion is not too dissimilar from how occlusion works for the purpose of outputting a 2D image from a 3D scene as described in the previous section when employing



Fig. 4.5 Occlusion issues in defocus when utilising pre-occluded inputs for layer-based CGH. Left shows the input layers, a blue swirl layer which is placed close to the viewer and a background layer which is partially occluded. Right shows this scene when viewed with focus set on the background - here the black void of missing information appears as the blur swirl itself defocuses and overlaps the background, this is highlighted by the enlarged segment.

a circle of confusion, shown in figure 4.3. In a holographic ray tracing renderer, where a hologram is generated that represents every possible view of a scene, a method of geometric occlusion (such as Chen, Wilkinson, 2009 [33]) is used to effectively block objects in a 3D scene from being viewed where an object exists in front of it, in relation to the viewpoint. In this scenario, the full scene data is required, and every vertex or fragment of an object, from every view, contributes to the hologram.

### Wave/Diffraction-based Occlusion

Wave based occlusion refers to using diffraction to handle occlusion by wave (or ray) propagation. It is applicable to layer/point based holography, as a given layer in depth is propagated towards the next depth layer, and is masked by any object that exists there. Methods such as the fourier based “ping-pong” algorithm [93] describe this process. An overview of wave based occlusion methods and extensions (as well as geometric occlusion) is described in Chen’s thesis [32]. Other examples of this include Shi et al. [183], which introduces an “occlusion-aware point-based-method”, shown in figure 4.6, which uses the cone of propagation and a boundary check condition to perform occlusion on any

wavefronts. Other methods of wave based occlusion are explored in the following publications [54, 104, 71].

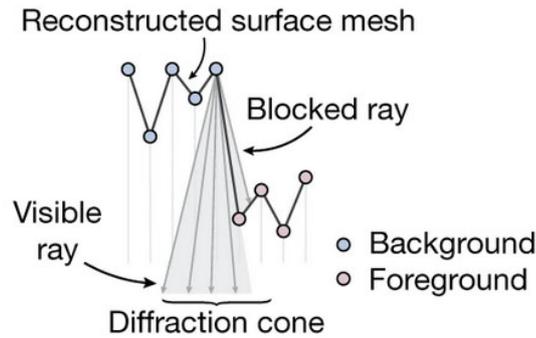


Fig. 4.6 The point-based method of occlusion from Shi et al. 2021 [183]

The critical difference between each is that geometric occlusion is crucial to making acceptable ray-tracing based holograms, which inherently have multiple views encoded in them - while wave based occlusion is for layer-based holograms, which are used to create holograms for planes parallel to the hologram/viewpoint. When viewing a raytraced hologram, generated from a full 3D scene, from an “off-axis” view, referring to a view not at the centre of the image, the viewer just sees a new view, so naturally can see behind an object. In layer-based holography (or any method using a 2.5D input), off-axis viewing of a generated hologram causes the perspective of a scene to disintegrate - the viewer sees nothing behind the object, even if the object can correctly block light from behind itself from this new view. This issue is also visible in Shi et al. [183], where foreground layers, such as the bunny (left) in figure 4.7, exhibit a dark line/halo where the viewing perspective is not exactly on-axis with the render perspective. Other examples in the literature, such as Gilles’ thesis [71] even demonstrate this where using a value of layer separation to make depth appear continuous, shown right in figure 4.7.

Recent methods suggest instead the use of a “Layered Depth Image” (LDI) [184], which is essentially identical to clipping the scene into separate renders by using camera clipping planes in order to retrieve some occluded scene information. Another suggested alternative is using lightfield inputs [24], in order to provide integral imaging of the scene, which also resolves a number of issues with scene discretisation and provides some depth information to elements that have reflections. Most recently a “defocus-blur-considered depth map” is proposed [221], which sees scenes rendered multiple times with a varying focus distance (varifocal images) as training for the neural network which generates the holograms - where the zdepth map itself is also blurred.



Fig. 4.7 Occlusion artefacts visible between layers in the literature, right from Shi et al. 2021 [183], between the bunny/vine and the background, and left from Gilles, 2016 [71] between many layers sliced from a car model.

These existing methods, despite providing improvement to the data input in the form of expanding 2.5D data for the benefit of occlusions, both require rendering the source scene multiple times - with numerous RGBZ images required as input to generate the holograms. These methods also do not propose solutions that resolve the fundamental issues with using Z-buffer data for image formation as outlined in the previous section. Cases of Z data severely mismatching colour data are shown in figure 4.8, with the top layer from the unity game engine, and the bottom from the “defocus-blur-considered depth map” method [221]. This means that semi-transparencies cannot yet be correctly represented as they can be in conventional computer graphics, aliasing is visible and there is not yet support for multi-sampled depths in any form.

### 4.3 Holographic Compositing & AOVs

Given the similarity between the issues that have previously faced computer graphics rendering and the issues that holographic display faces, it would be unwise to not take inspiration from the solutions previously presented. In the media production workflow, as per figure 3.1, image data is passed from the renderer to compositing before display. The compositing application, when provided with the output of techniques outlined in section 4.2.1, allows a complex 3D scene (even if composed of a selection of 2D layers) to be formed successfully. In this process, computer graphics developments have minimised the resistance and data loss between renderer and the compositing workspace to ensure error free imagery when being edited. The parallel that must be drawn here is that the requirements for compositing are extremely similar to those of holographic display - and by leveraging this fact, the quality of imagery delivered by an adapted holographic generation algorithm may be vastly improved.

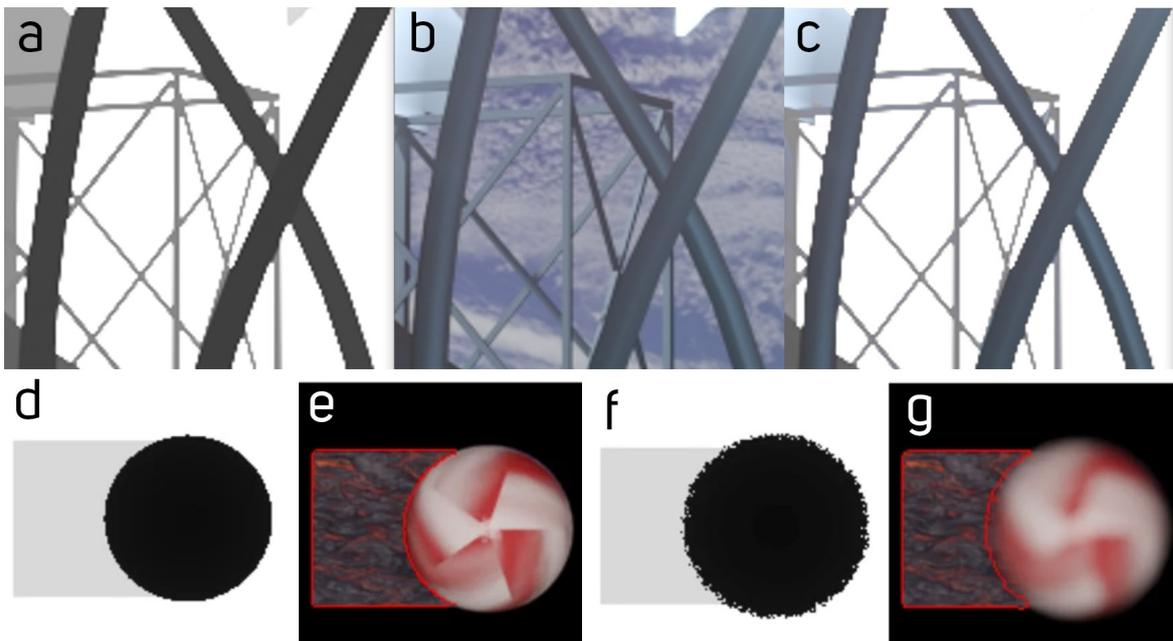


Fig. 4.8 a) Shows the Z render pass from game engine Unity, note the lack of anti-aliasing to prevent errors with depth on edges. b) Shows the colour pass for the same scene, which is anti-aliased. c) When forming a layer input for CGH from the depth and colour, the aliasing errors get applied to the colour pass, resulting in this case in artefacts in the front layer, and colour bleeding on the back layer. d and e) Similar depth aliasing issues and colour pass respectively from [221] and; f and g) The suggested method of a “defocus-blur-considered depth map” with unacceptable levels of edge artefacts and depth/colour mismatch - where the back layer is outlined in red.

This section introduces the concept of “Holographic Compositing”, where data gathered from the renderer which is typically used in the production process may also be utilised for display - in such a way that compositing does not “flatten” the 3D data into a 2D output, but rather preserves the 3D data of elements so that the holographic display itself may directly composite the image data in the replay field. Such a technique enables the advances made in computer graphics to be directly utilised in holographic display.

A typical relationship between renderer and compositor will be formed of multi-channel image data (usually in the floating point EXR container via the HQH standard defined in the previous chapter), and will provide a render of the scene in components, which will allow easy adjustment without the need to re-render. As mentioned in the previous chapter, these are referred to as Arbitrary Output Variables (AOVs). Such passes may include diffuse, lighting, global illumination, reflections, specific objects or even passes where the scene is lit by specific lights, as well as various data channels such as alpha, object distances/positions and motion vectors. As these image channel components are ultimately rendered separately and combined by the renderer output routine, as described by the ray-tree [216], they may instead be output individually to retain flexible control over elements of the image. This is also the same mechanism by which Porter and Duff composite different image elements [169]. Image channels will then be appropriately adjusted and recombined to form the final image, an overview of this is shown in figure 4.9.

Being able to perform such compositing of holograms allows the anti-aliased and semi-transparent accumulation of both separate render elements (such as objects), and image channels (such as reflections), which can also be associated with correct depth data, to provide correct holographic occlusion and representation of secondary rays.

### 4.3.1 Holographic Compositing Operations

"Holographic Compositing" requires that the conventional 2D compositing operations are recreated in fourier (frequency) space. For the purpose of re-assembling a selection of AOVs and layers to form the intended image, the primary interest is on the Plus and Over operations. Multiply and Minus, as well as a selection of other operations related to the Over are also described. These presented methods are applicable to any kind of CGH.

#### Holographic Plus

The holographic plus is exactly as it sounds, one hologram added to another. This is the routine by which layers are combined in layer-based holography. After each layer is fourier

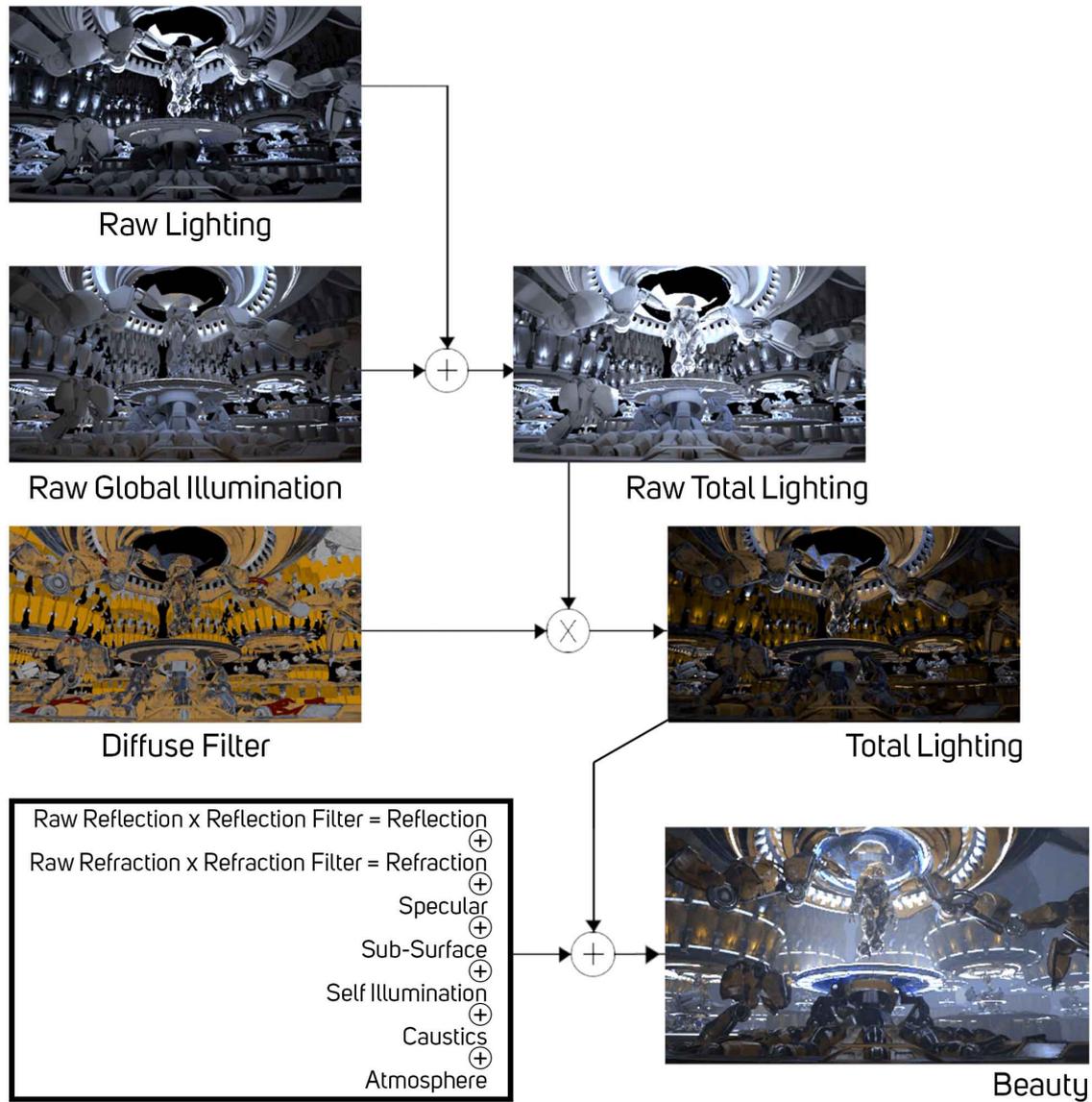


Fig. 4.9 The process in which different render components are re-assembled in order to form the final intended image. Adapted from Chaos Group's VRAY Renderer Documentation [26], it is shown how the final image, often referred to as the "Beauty" image, is formed.

transformed, the complex values are summed. These layers may be at any depth or already be formed of many layers themselves.

$$A + B \tag{4.1}$$

Layer-based holography assumes that the layers are pre-occluded by the input renderer, and as the light values in the replay field that the hologram will form will be additive, images will overlap without clipping (assuming a floating point input and sufficient physical hardware). If the layers are not pre-occluded, both the total of layer A and B will be visible regardless of their depth, and these overlapping regions will again appear brighter. Figure 4.10 (left) shows two layers, a cylinder and a background, from a 2.5D source, which have been added as per existing layer-based hologram methods.

### Holographic Over

An Over operation is the most common compositing operation, and allows one layer to overlap and occlude another layer.

$$A + B(1 - a) \tag{4.2}$$

A Holographic Over takes the form of one layer being occluded by the alpha mask (indicated by the lower case in the equation) of another. Here, as the occlusion must occur in 3D space, the wavefront of layer *B* is propagated through space to the depth of layer *A*, it is masked by the inverse alpha of layer *A* in the image domain, and then propagated to the hologram plane. Such an algorithm is referred to as the "Silhouette mask method", an occlusion algorithm as outlined in the previous section. As per Section 2.2.1, this propagation can occur by any diffraction method, but an FFT is used with the Angular Spectrum method for propagation between parallel planes with an orthogonal projection - and while some literature suggests that this mask must be binary [71, 227], this is not the case, and any alpha value may be used in order to introduce semi-transparencies, supporting anti-aliased edges and layers with semi-transparent materials. If a full image layer (i.e. one that has not been already geometrically occluded, or had a 2D holdout (stencil) applied) is used as layer *B*, it will be correctly occluded via a wave/diffraction-based method, so both layer *A* and *B* will appear to defocus correctly, without artefacts, even in parallax.

A simple overview of forming the Holographic Over, between two layers, is as follows:

1. Two input layers of RGBA data, a background *B* and foreground *A* with defined depths ( $z_b$  and  $z_a$  respectively) for each image plane.

2. FFT layers  $B$  and  $A$ , each layer is composed of RGB components. These layers will be at optical infinity.
3. The background layer  $B$  is propagated (via Angular Spectrum Method / FFT with focus phase factors) to  $z_a$ .
4. The alpha of  $A$  is resized to scale with the wavelength of each RGB component. This is defined by  $\text{ImageSize} * \text{ShortestLambda} / \text{SelectedLambda}$  and assumes colours are scaled around the smallest wavelength. Therefore the stencil for the shortest wavelength (blue) does not require resizing.
5. Each scaled alpha of  $A$  is then inverted, and used to multiply the selected hologram, forming a stencil of layer  $A$  on layer  $B$ , in the image/spatial domain.
6. The layer  $B$  is now propagated to  $z_b$ , its intended position in space.
7. The layers  $A$  and  $B$  will now be added to form a single hologram.
8. The total hologram is shifted to the pupil - propagated to depth 0.

The layers may also only be propagated forward, and accumulate layers as they travel towards the viewer. The background layer, occluded by a foreground object, but not with the plus required to display the foreground object as per step 7, may look as per figure 4.11. Figure 4.10 (right) shows this with a cylinder against a full image background, each rendered on individual layers, compared to a 2.5D input which has been pre-occluded by the renderer and split into layers after the fact.

The Over, via use of holographic occlusion, may be adapted to form the following operations in table 4.1, whilst this expands the operations define in Porter and Duff [169], it encompasses a larger selection of operations which are available in modern compositing applications, such as The Foundry's Nuke [65].

The operations Conjoint-Over and Disjoint-Over may be useful when combining render-elements which may be partially pre-occluded. These edge cases are useful where a background layer may be already stencilled with the occlusion of a foreground element, but where that layer also wishes to occlude other layers, so a plus is not sufficient by itself. Given that layers must be occluded holographically, it may not be immediately obvious why this is still necessary, but it will be important to cover edge-cases when utilising the occlusion methods outlined in section 4.2.1. Though this unintended doubling of occlusions can be avoided by ensuring that layers behind foreground elements are not pre-occluded (such as from a 2.5D source).

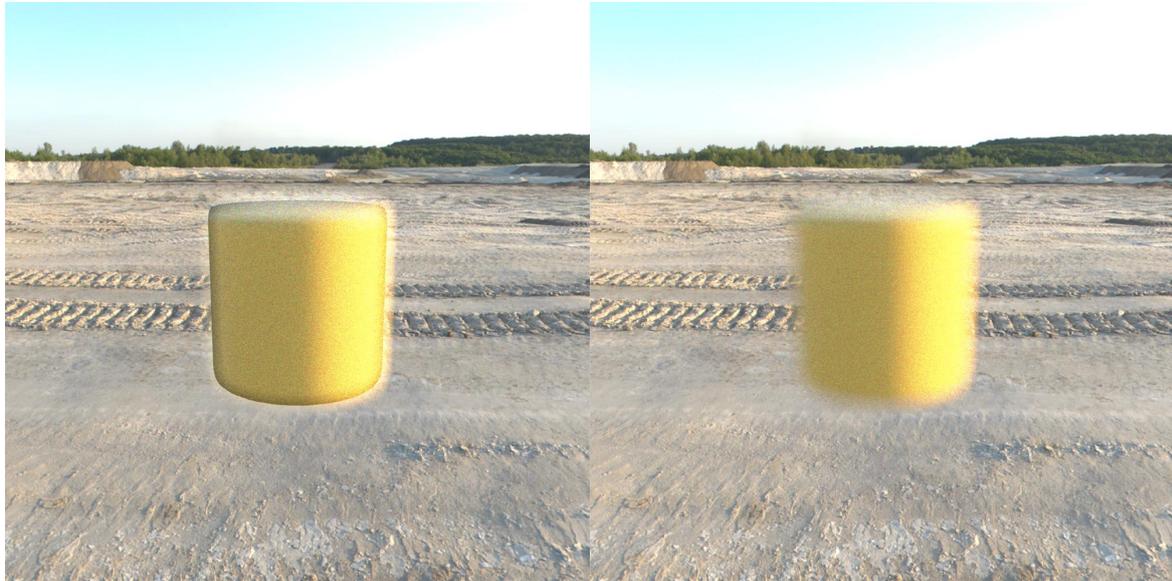


Fig. 4.10 A yellow cylinder is composited with a background layer, left shows this process with a pre-occluded background (such as from a 2.5D source) with a holographic plus operation, and right shows the same composite with a holographic over. The focus is set on the background.

Table 4.1 Holographic Compositing Operations Enabled via Occlusion.

<b>Operation</b>	<b>Algorithm</b>
Atop	$Ab + B(1 - a)$
Conjoint-Over	$A + B(1 - a/b)$ , $A$ if $a > b$
Disjoint-Over	$A + B(1 - a)/b$ , $A + B$ if $a + b < 1$
In	$Ab$
Out	$A(1 - b)$
Stencil	$B(1 - a)$
Under	$A(1 - b) + B$
xor	$A(1 - b) + B(1 - a)$



Fig. 4.11 How a single layer appears when it has been occluded - here a background layer is occluded by the foreground layer of a cube. The focus is set on the background, and the region which has been occluded represents the defocused foreground cube.

### Holographic Multiply

The holographic multiply, as shown, can perform an occlusion, as it is the same mask as from a conventional "Over" operation. This does not have to be across depths, and may be at the same depth. This allows the multiply operation shown in figure 4.9 to be carried out during the holographic composite, rather than in source layers - this is where a total lighting pass is produced by the multiplication of raw total lighting and the diffuse colour, this is carried out via the occlusion method detailed, and shown in figure 4.12. This does not have to be a single channel mask, and can mask any colour channel.

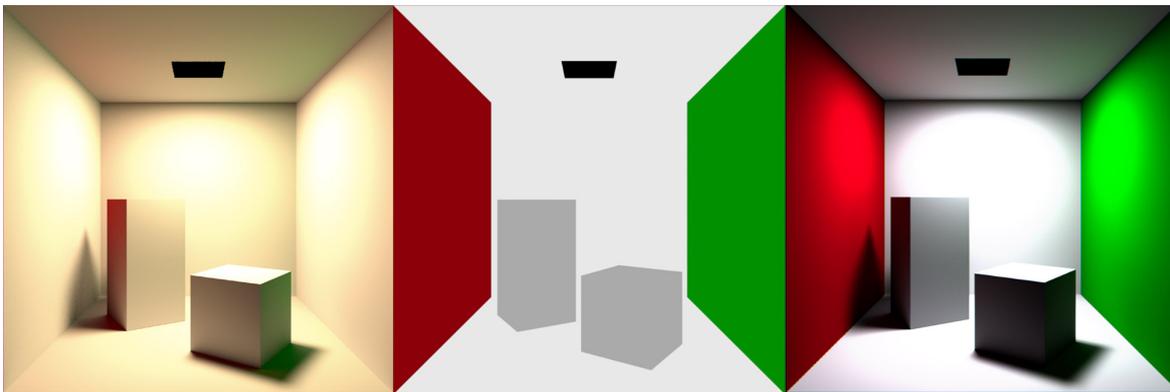


Fig. 4.12 The resulting holographic image (right) is produced by multiplying the total lighting pass (left) with the diffuse pass (centre). This is a single 2D layer and the masking of the diffuse layer by the total lighting is performed at the same depth.

### Holographic Minus

A Holographic Minus operation only allows the removal of a known layer that exists in the total hologram.

$$A + B = \text{Total} \quad (4.3)$$

$$\text{Total} - A = B \quad (4.4)$$

The layer being removed from the total hologram must be identical to that of the original, including the same random phase. If the original information for this layer is not known, a viable alternative is to use an occlusion mask (via a multiply) at the appropriate depth to remove the layer's contribution to the total hologram between a specific zFront and zBack (either side of the depth of the given layer) in depth. However, anything already holographically occluded cannot be recovered. After the layer or layers have been removed,

they may then be replaced with new adjusted layers or otherwise have the removed regions filled in.

### 4.3.2 Ray Length Based ZDepth

#### **An aside: ZDepth vs Distance**

It's important to note that zdepth usually is not actually the distance from the camera to a pixel in world-space, but instead is the distance along the z-axis from the centre of the camera to the plane along which the pixel in question is situated. This can be converted to the real distance by knowing the angle of the pixel within the camera fov. In fourier plane/layer based holography the zdepth is technically correct, as it is also planar - but it must be remembered that this is purely an approximation, as the actual projection should be spherical and with perspective.

#### **Using Real Distances: Optical Ray Lengths**

Now that render AOVs and layers may be successfully holographically composited, additional data is required to ensure that AOVs which are formed by secondary rays [216], such as reflection and refraction passes, have correct depth data.

When using a 2.5D input, or where 2D layers have fixed depths, any of these lighting effects achieved via secondary rays in the ray-tree will not be represented correctly, this is illustrated in figure 4.13 with a defocus post-process. Here, as the z-buffer is being formed of the first-hit geometry in the scene, rather than the optical length of the ray, the depth data is incorrect - the focus is on the window, yet the view through the window, and the reflections in the window, are in focus.

The 'depth' of a point refers to the distance from which multiple rays from that point appear to diverge. In the case of an object being viewed in the reflection of a flat mirror, the total depth is the length from the object to the mirror and then to the viewer. Figure 4.14A shows the primary ray from viewer to mirror, and a secondary ray from mirror to object. As previously described, a z-depth pass will provide the length of only the primary ray, so must be added to the length of the reflection secondary ray for the correct depth of the object in camera space.

A refraction depth pass is performed in a similar fashion, as seen in Figure 4.14B. In the case of looking through glass, the primary ray is formed by the depth from viewer to the surface of the glass, but the length of the secondary ray is determined by the shape of the surface as well as the refractive index. The resulting image may be real or virtual, and have magnification. The focal length of the lens is given by the lensmaker's equation:

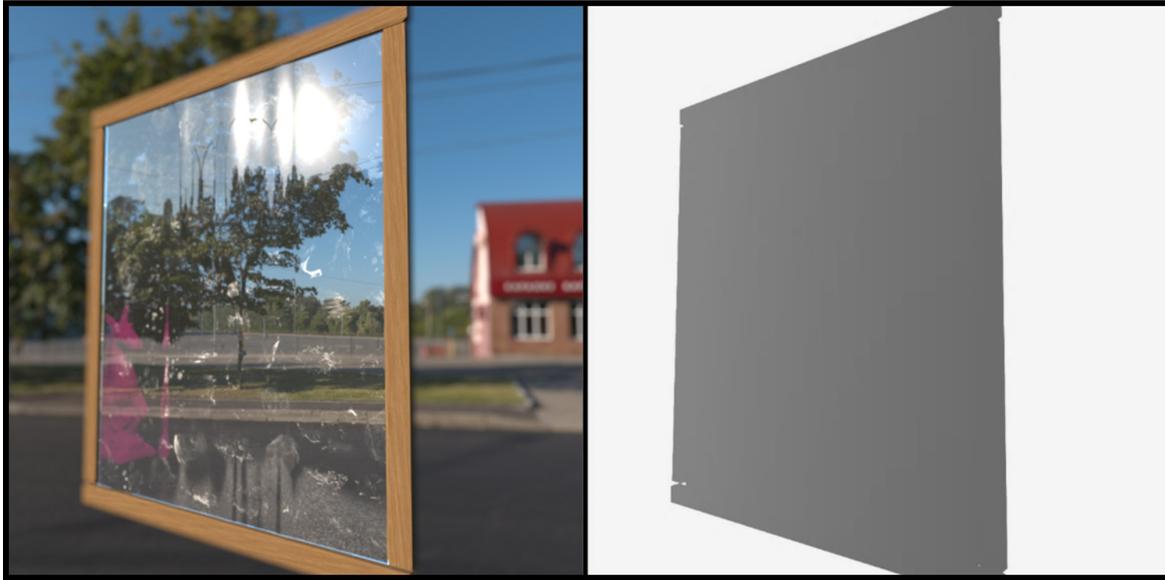


Fig. 4.13 The depth pass (right) represents only the first-hit geometry of the scene, rather than the optical length of the ray. Left, a defocus is simulated and applied as a post-process, with focus set on the glass pane - the same outcome would occur if this was via a hologram using the same source. The background and reflection seen through the pane are incorrectly in focus.

$$\frac{1}{f} = (n - 1) \left( \frac{1}{R_1} - \frac{1}{R_2} \right) \quad (4.5)$$

Where  $f$  is focal length,  $n$  is Index of Refraction (IOR) and  $R_1$ ,  $R_2$  are the radii of curvature of the surface. The new secondary ray depth of the object image visible through the lens is given by:

$$\frac{1}{S_2} = \frac{1}{f} - \frac{1}{S_1} \quad (4.6)$$

Where  $S_1$  and  $S_2$  are the depths to the lens and the object respectively. When added,  $S_1$  and  $S_2$  provide the total correct depth value. Most renderers will already store a value for focus depth per pixel per render-pass per trace-depth as part of the ray tree; this may be retrieved via custom AOVs.

Once these passes are formed, each AOV is sliced into a suitable number of layers, generated into holograms using the associated depth information. The distances are in meters for ease of use in normal compositing operations, but converted later during hologram generation to wavelengths, as each colour channel will have a different wavelength and consequently have a different number of oscillations in the same physical distance - requiring different holograms to be generated for each wavelength. These are then holographically

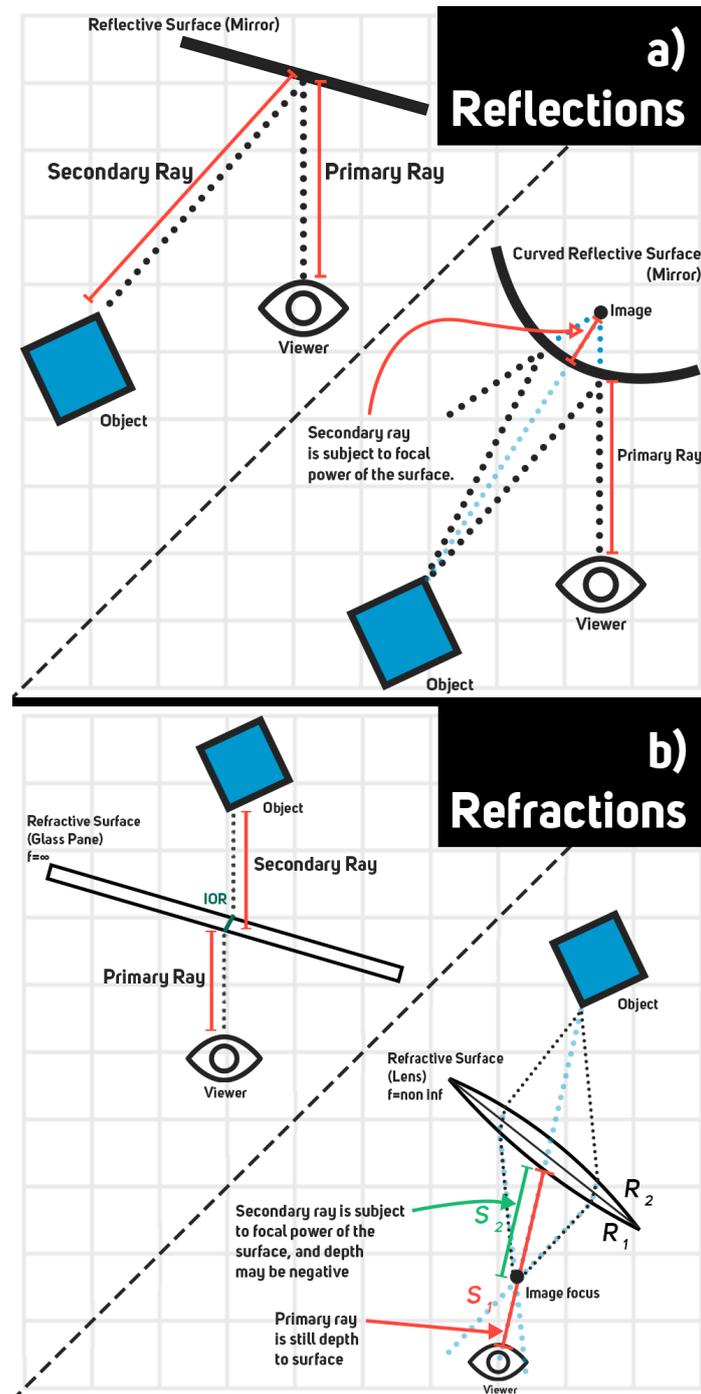


Fig. 4.14 (a) The total depth required for reflections is the summation of primary and secondary rays, following the law of reflection and any focal power. (b) The focal length is calculated with the Index of Refraction (IOR) given by Snell's Law, and the thickness of the material. Distances of secondary rays may be negative. A material may exhibit both reflective and refractive properties simultaneously.

composited. The reflection and refraction passes require a plus operation, as they combine with the total lighting to form the intended image, this is shown in figure 4.15.

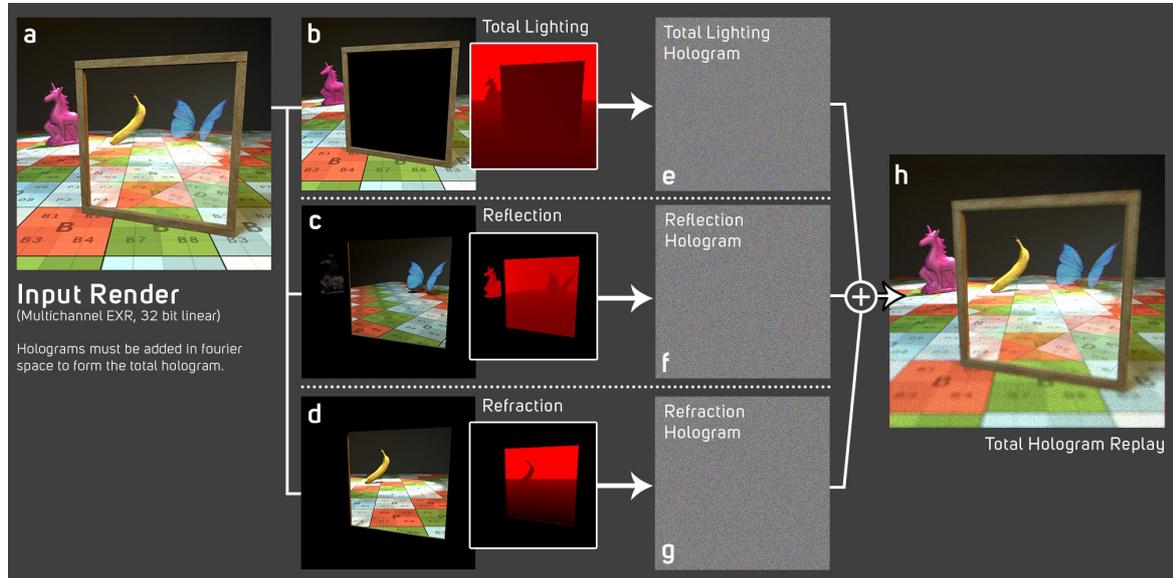


Fig. 4.15 (a) The input multichannel file. (b) Total Lighting render, z-depth. (c) Reflection render, associated depth. (d) Refraction render, associated depth. (e) Hologram of Total Lighting only. (f) Hologram of Reflection only. (g) Hologram of Refraction Only. (h) The total composite hologram replay, via a plus operation of the holograms for each render pass. This is the final image that is viewable on the holographic display.

As the image formed in the replay field is still diffuse, the imagery formed will not be view-dependant, such as how a specular reflection behaves, but given the reflections and refractions will have the correct depth information, they will defocus correctly and move in parallax according to their assigned depth layer. Where there are multiple light bounces between reflections, it will be important to increase the trace depth of the ray in order to represent each additional bounce - each bounce will require a new AOV+depth pass in order to be composited into the total hologram correctly. Distances further from the viewer will require fewer layer allocations to remain perceptually correct.

This method then allows the physically correct representation of reflections and refractions, so materials such as mirrors, glass, and water may now be correctly viewed - as well as improved material representation for materials that exhibit any reflective properties, as most materials are not purely diffuse. Table 4.2 shows the AOV RGB data, and associated depth pass required for correct representation, also see 4.9 for the usual 2D process. Figure 4.16 shows the result of the compositing process for both reflections and refractions, and figure 4.17 shows the result of the process compared to figure 4.13 shown at the start of this section.

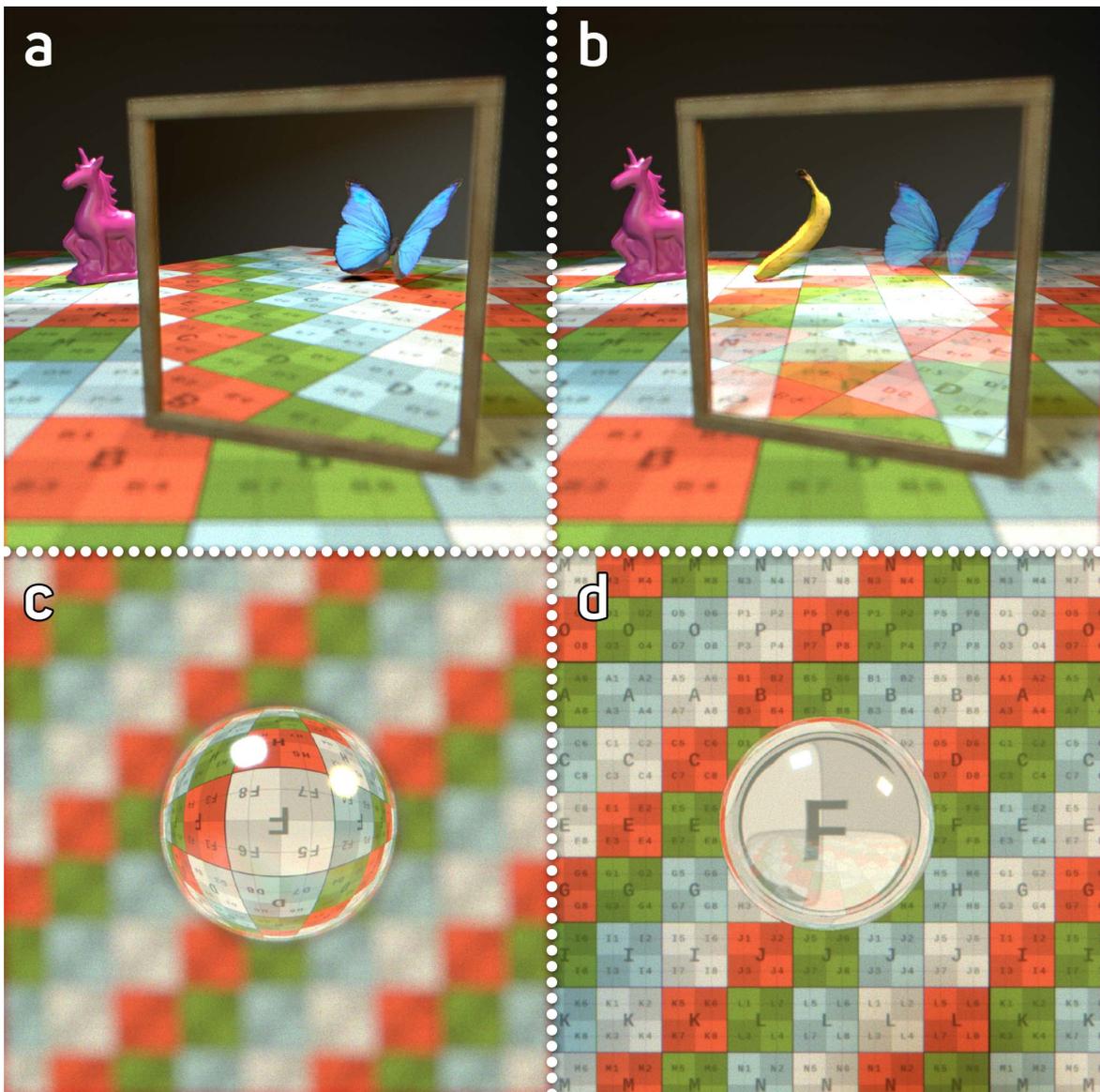


Fig. 4.16 (a) Replay of butterfly reflection and unicorn (both 9m) (b) Changing the mirror to glass reveals a banana (9m). (c) Grid at 1m from camera focused by a lens ( $f=100\text{mm}$ ) placed 850mm from the camera, the grid focuses at 550mm (real). Reflection focus is 950mm. (d) Grid, 1m from camera focused by lens ( $f=100\text{mm}$ ) placed 950mm from camera, the grid focuses at 1050mm (virtual). Reflection focus is 1050mm.

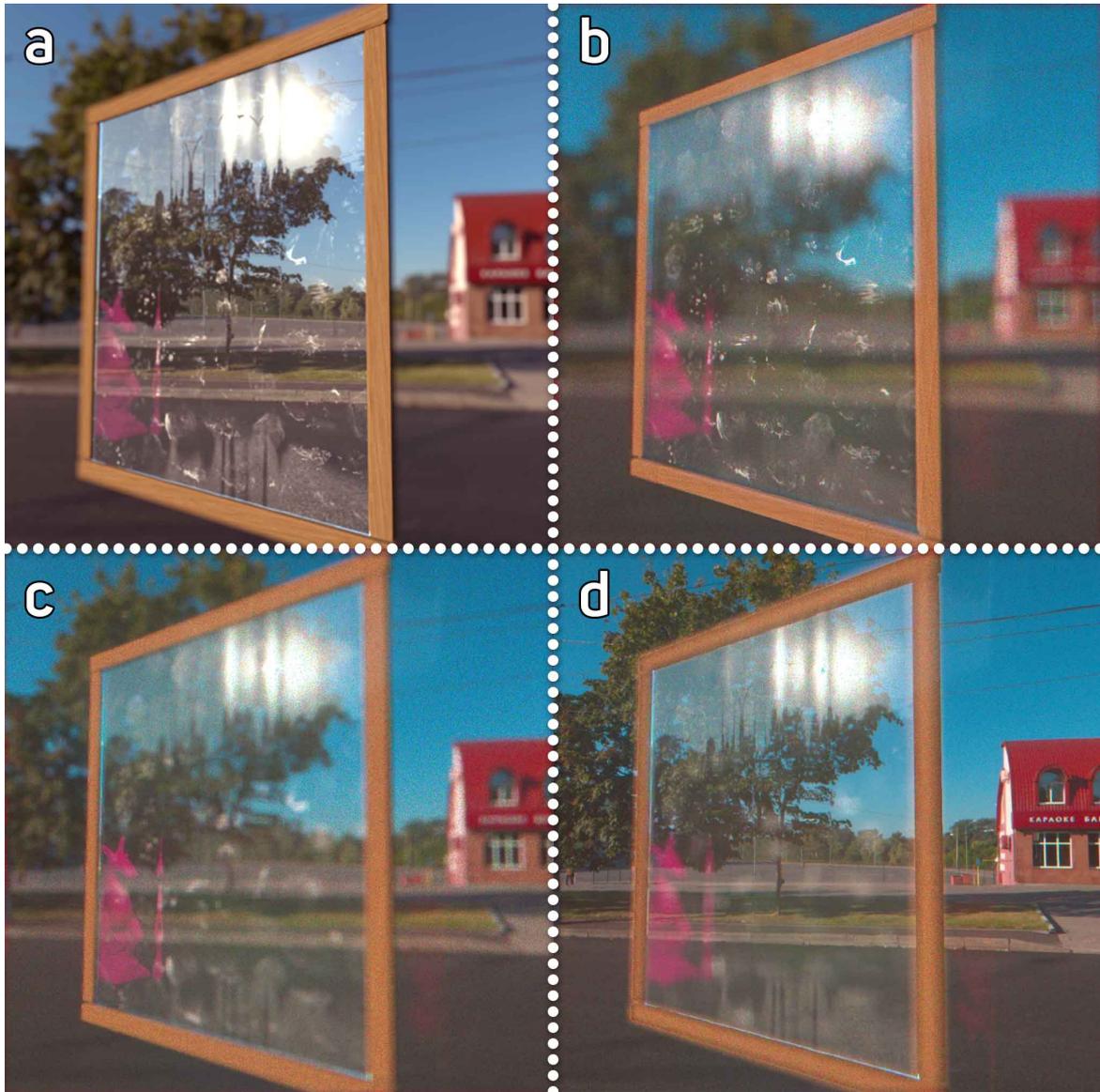


Fig. 4.17 (a) The original simulation from figure 4.13, shows incorrect defocus. (b) A holographic replay of the same scene but with multi-depth holographic compositing, focus is set at the pane surface and shows correct defocus through the window. (c) Focus set on the reflection of the unicorn toy in the window. (d) Focus set on the background.

Table 4.2 Render Passes for Composite

<b>RGB Channels</b>	<b>Depth Source</b>
Total Lighting ((Raw Lighting + Raw Global Illumination) * Diffuse Filter)	Z-depth (primary ray depth)
Reflection (Raw Reflection * Reflection Filter)	Reflection secondary ray depth + primary ray depth
Specular Highlights	Specular secondary ray depth + primary ray depth
Refraction (Raw Refraction * Refraction Filter)	Refraction Secondary ray depth + primary ray depth
Others (such as SSS, Emission, Caustics)	Z-depth (primary ray depth)

### 4.3.3 Limitations of AOVs

In order to make the most of the holographic compositing process, scenes must be appropriately divided into layers. For the most success in the result of a holographic Over, a full unoccluded background layer is required, and in order for a layer to be successfully represented with semi-transparencies, it must have alpha information.

It is possible to represent everything with enough render layers in this fashion where objects are not split across layers. Having these isolated layers means that motion-blur, hair/fur and other objects that require semi-transparencies such as smoke will have the appropriate alpha information to receive the benefits of the holographic over - and occlude other layers successfully.

This kind of scene slicing must occur in the renderer, and these render layers may be all contained by a single EXR output. In order to replicate this behaviour in a 2.5D render, without additional render layers, Object IDs [88] or cryptomatte [67] may be used, which provides a unique colour value for each object to indicate the region of an image which is inhabited by a given object. This allows the slicing algorithm to ensure that objects remain on individual layers, and this data pass may then be anti-aliased to encode semi-transparencies, rather than have the appearance of aliased edges when slicing a scene from a 2.5D source.

This is in agreement with the alternative methods discussed in section 4.2.2, which explore using layered-depth images, as a single 2.5D pass in the typical RGBZ format is not sufficient for a good viewing experience. While section 4.3.2 has proposed a method to enable reflection and refraction representation in 2.5D rendered imagery, only the colour information of the scene that is visible may be utilised, which results in artefacts at the boundaries of layers as shown in figure 4.7. This boundary layer issue is also present when

Table 4.3 In-painting Methods and Quality of Output

<b>Method</b>	<b>Detail in Output</b>
Edge Blurring (i.e. Gaussian) [7]	Low
Image Pyramid [3]	Mid
Content Aware Fill [6]	High
Neural Network [121]	High

using a layered-depth image and an object is sliced across multiple layers - objects are composed of hollow meshes, and when this is rendered via scene slices, the inside of the mesh becomes visible even when occlusion is utilised correctly.

In order to resolve these issues, information behind layers must be gathered in such a way that may allow 2.5D inputs, as well as information for where objects are sliced across layers.

## 4.4 Occlusion Synthesis

### 4.4.1 Insufficient Scene Data

Missing scene information causes occlusion artefacts, and the nature of layer-based holography means that there will always be missing scene information where objects or scenes are sliced from a particular source. This must be mitigated in order to reduce the visibility of errors in the replay field during a holographic composite.

This section proposes a method of “occlusion synthesis” whereby in order to fill gaps behind layers, whether it is one object occluding another or one object occluding itself, scene information is in-filled, in-painted or otherwise generated. Ideally, the original scene data should be used to fill occlusions, but this is not always possible.

This is particularly important for creating holographic versions of existing material, which may have RGBZ renders, or RGBZ via a depth conversion/generation process, but not individual render elements for every object. Synthesising data behind objects using depth would allow artefact free viewing without having to re-render source material, or go through significant re-mastering or re-compositing. Such synthesis would also resolve artefacts where objects are split across multiple layers, and allow objects to span a greater number of layers. Both of these artefact cases are shown in figure 4.7.

Synthesising this data occurs by utilising one of a selection of algorithms, each of which can deliver various levels of details. Some example methods are provided in table 4.3.

### 4.4.2 Process Overview

In order to synthesise any scene data for occluded regions of a scene, these areas must be identified. This can be done using a combination of Z and alpha data (which may be derived from an Object ID pass) to enable semi-transparencies, or Z alone.

Layer slicing can be performed to give a stack of layers through depth, and mask regions may be identified where a front layer contains information. This mask region will then be provided to the in-filling algorithm. The identification of the region eligible for synthesis is shown in figure 4.18.

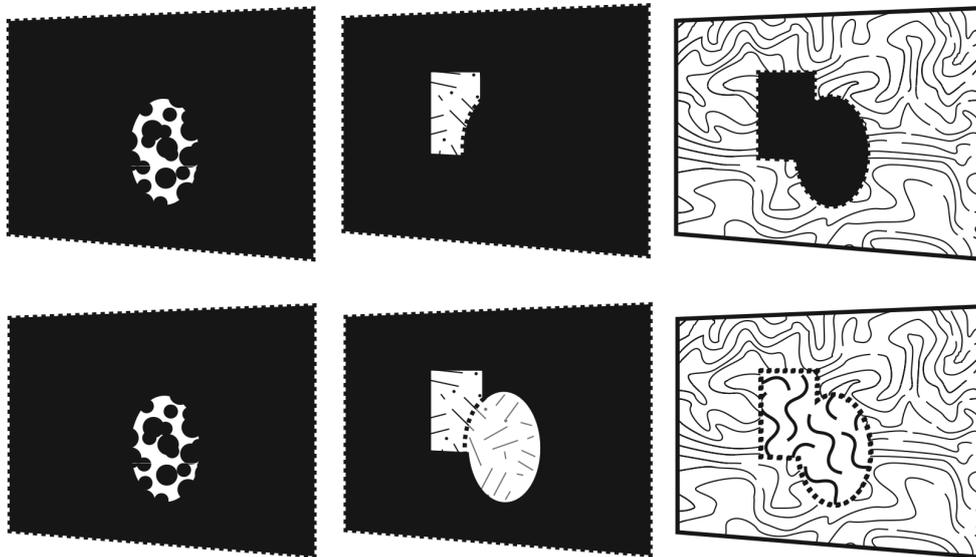


Fig. 4.18 Each input layer is shown in the top row, a circle at the top of the stack, which occludes a square, and both occlude the background. This occlusion corresponds to regions that are eligible to be in-filled in the bottom row. Specifically, the regions on layers behind the circle and square on the forward layers.

The indicated regions are then filled with new image data, whether this is colour, depth or otherwise, and replace the original layers - this new layer stack is then generated into holograms and composited for display. A comprehensive overview diagram of this process is given in figure 4.19 and corresponds to the following description of the occlusion synthesis stages:

1. A scene-input is presented to the holographic display system. This could include any number of views of the same scene (as per chapter 5) - as each viewpoint is processed individually and follows the same process as if it were the only view provided. This

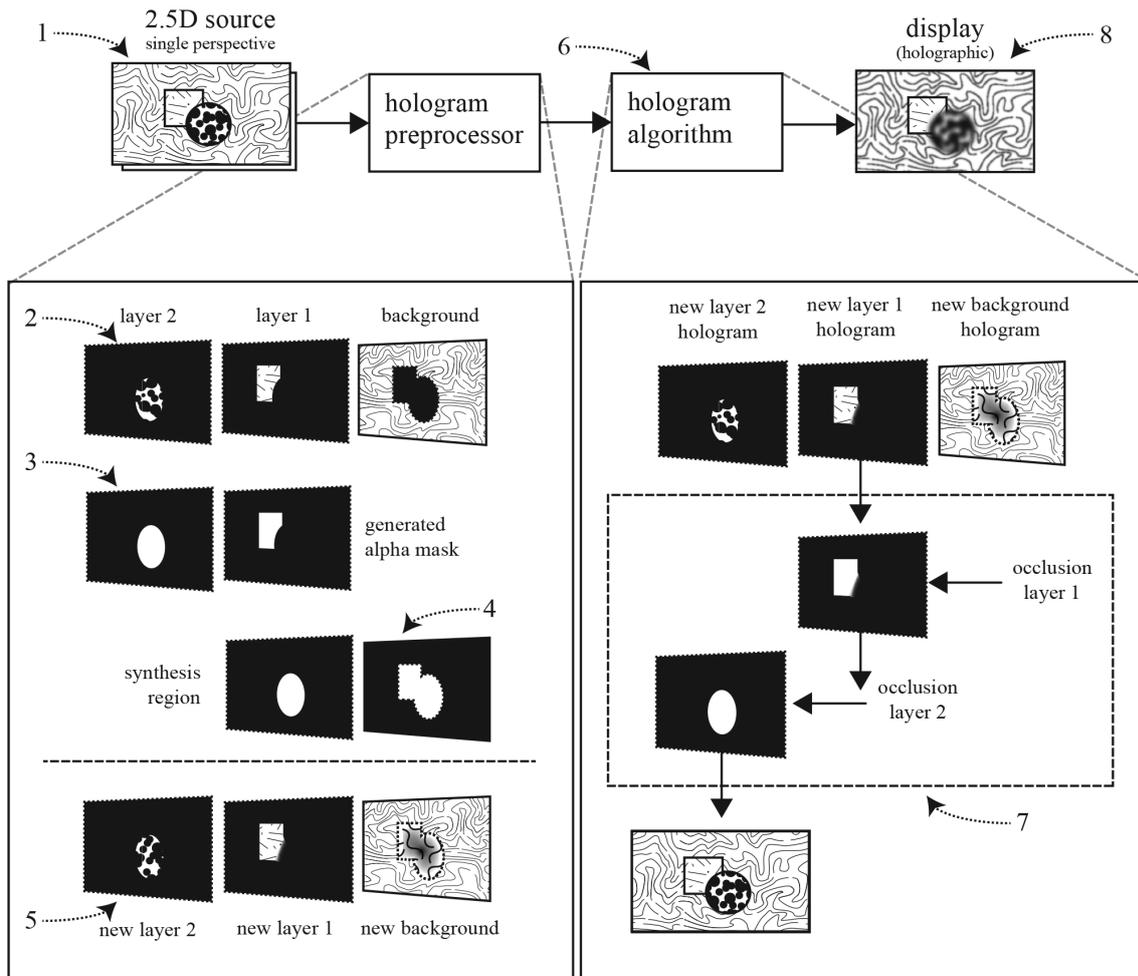


Fig. 4.19 An overview of occlusion synthesis and where it fits into the holographic image generation and compositing process.

- example uses a 2.5D input, but may also include whole layers, volumes, or other 3D representations of objects.
2. The display system is instructed, or determines, the number of layers that the source will be displayed with. This may be limited only by the scene-input.
  3. Each layer, of any number of layers, is used to generate an associated mask, which describes the region of itself in the shadow of forward layers, using the Z channel data provided for a mask of forward layers, or an additional channel from the image renderer (such as an alpha, or object ID).
  4. The mask is then used to highlight the areas of each layer that require additional synthesised image data to remove shadow areas.
  5. The highlighted areas go through a method of image fill synthesis (potentially using methods such as, but not limited to, image blurring, colour stretching, Image Pyramid algorithms, Content-Aware Fill, or neural network generated imagery) and the new layers replace their original, resulting in a stack of layers with synthesised sections.
  6. This stack of layers is presented to the holographic display algorithm, where holograms are generated for each layer at the correct associated depth.
  7. The layers are holographically composited with overs - each layer goes through a holographic occlusion algorithm where required, this requires an alpha mask area, such as the one generated in step 3, with the synthesis adjustment, to determine the regions that can be holographically occluded. The layers may or may not be combined during this process, but should be in some fashion (via holographic compositing, or displaying them individually across-time (time-multiplexing), for example) in order to be displayed.
  8. The resulting hologram is viewed on a holographic display system without artefacts.

### 4.4.3 Worked Examples

The above process is followed here with a worst case scenario, where only a 2.5D RGBZ source has been provided to be displayed. Figure 4.20 (left) is the input to the system that generates a hologram, Figure 4.20 (right) shows the background layer independently. Both the background and foreground layers will be generated into holograms and displayed at the correct depth to preserve the perceptually correct display.



Fig. 4.20 Left shows the input to the system as a 2.5D source, right shows the background layer extracted from the input.

A viewer expects to be able to focus on a distant object and have closer objects defocus in a physically correct way. In a hologram generated from the above input, the defocus is physically correct, but the way that the foreground occludes background (occludes) within its defocus is incorrect. This incorrect appearance of defocus is due to the missing information from the background image, and is pictured in figure 4.21 (left). The foreground object appears to remain incorrectly in focus due to image data from the background becoming visible when it defocuses - the cut-out of this foreground object from the back becoming visible confuses the viewer.

The method described in this section solves this issue by incorporating synthesised image information within the gap in order to prevent the artefact becoming visible. A range of methods (listed in table 4.3) could be used to generate such image information to fill the gap. One such result is shown in figure 4.21 (right). Here, an “image pyramid” method is used via OpenImageIO’s “fillholes\_pushpull” function [158], which delivers results similar to Nuke’s inpaint node. Any method used must be temporally consistent when used for moving imagery. Any choice of synthesis method will need to balance between quality vs time for computation available, which will vary per use case.

The synthesised background image (figure 4.21 (right)), generated by using the foreground layer as a mask to highlight the region of interest, can be used as a drop-in replacement for the source background image shown (figure 4.20 (right)), a hologram can be generated of both layers which then appears to be perceptually correct to the viewer, and is correctly

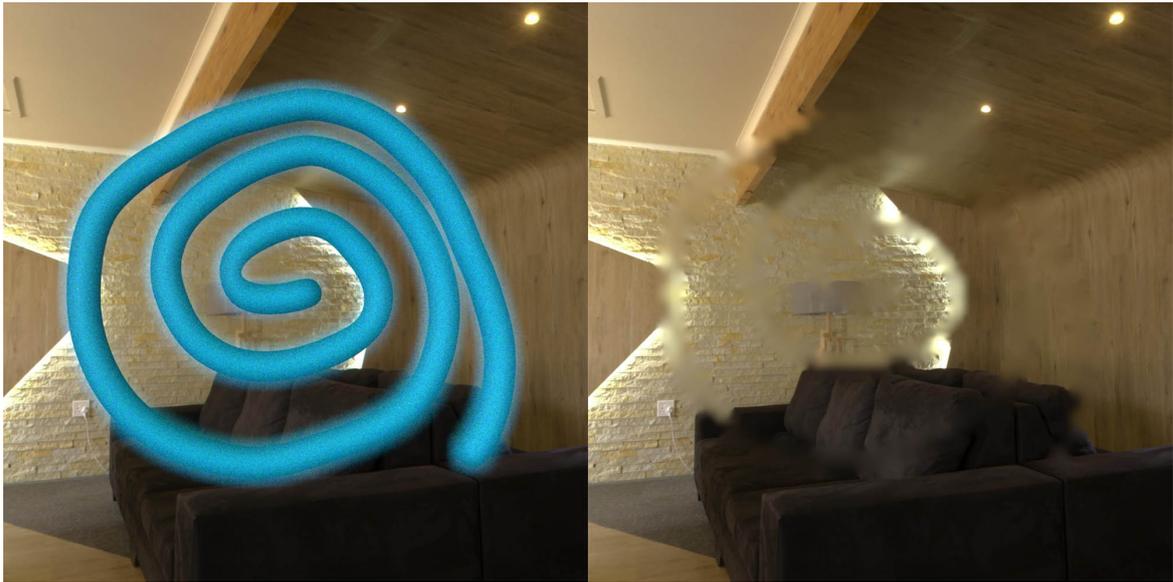


Fig. 4.21 Left shows the hologram replay result with no occlusion (composed with a plus) and no occlusion synthesis - focus on background. Right shows the synthesised layer result via an Image Pyramid algorithm of the background layer.

presented with occlusion (using a Holographic Over) that utilises the additional image data. Figure 4.22 demonstrates the results when viewed.

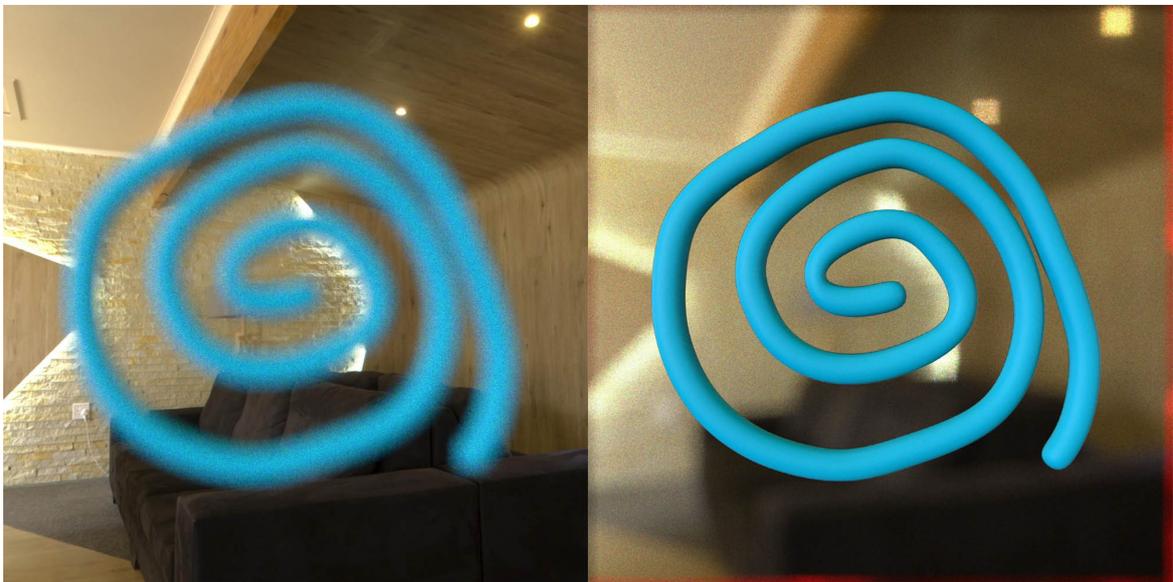


Fig. 4.22 Replays at both background and foreground focuses respectively. The background layer used is the result of the in-filling process and the the layers are composited with a holographic over, and therefore occlude correctly and appear error free.

This method of course will extend to any number of layers, and with an appropriately selected level of detail provided by an image synthesis algorithm (table 4.3), will be suitably error free if compared to a render which contains the actual scene data. This is seen further in figure 4.23, which shows a comparison between using a Holographic Over alone, compared to an over with occlusion synthesis at various focal depths.

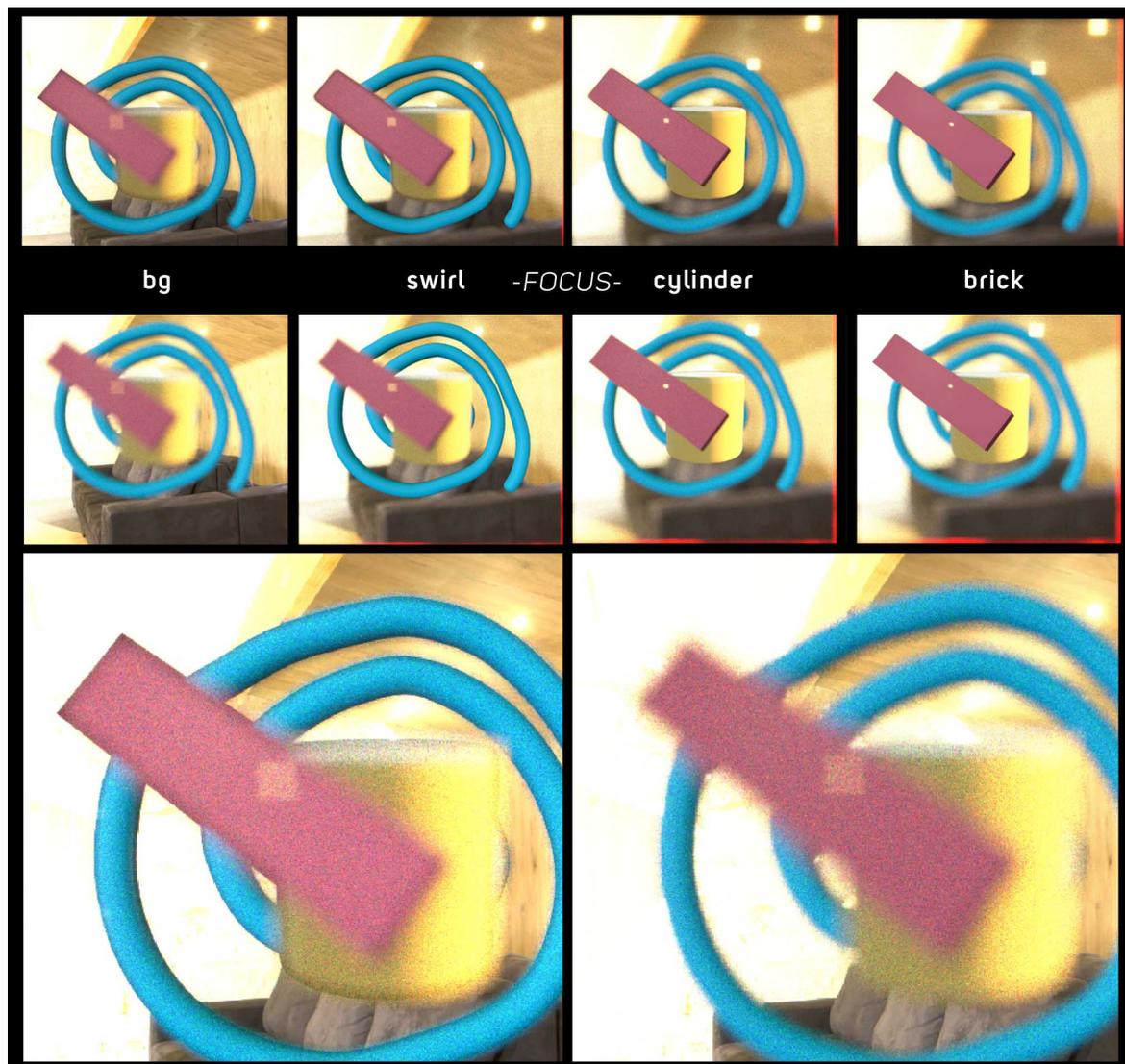


Fig. 4.23 A scene comparing occlusion techniques with varying focal points. Top Row: Occlusion (Holographic Over) method from an RGBZ 2.5D source. Mid Row: Occlusion (Holographic Over) method from an RGBZ 2.5D source with synthesised image data via the introduced method. Large zooms: Comparison of over from RGBZ 2.5D (left) which shows artefacts vs over with synthesised data (right) which is artefact free - with the focus on background in both images.

The results of the method are then shown from a viewing position on a real projector in figure 4.24 - where occlusion synthesis is used with a Holographic Over, artefacts are no longer visible from any pupil position in the eyebox. Artefact free viewing occurs both between different objects, the dog cube and the partial sphere, and between different layers of the same object, the partial sphere. The front layer contains the dog cube and front portion of the sphere, and the back layer contains the remainder of the sphere, which is occluded by the dog cube and itself.

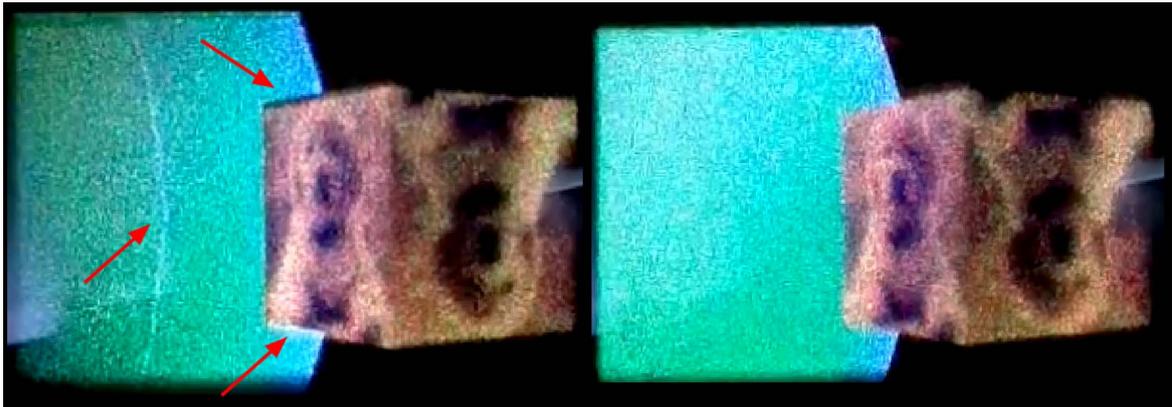


Fig. 4.24 Images captured from a physical projector before and after occlusion synthesis. Left: Red arrows indicate the artefacts visible when viewing a holographic image from a 2.5D source. Right: The artefacts are no longer visible when occlusion synthesis is utilised.

#### 4.4.4 Limitations of Occlusion Synthesis

When presented with limited scene data to generate holograms with, the occlusion synthesis method provides a powerful solution. It allows layers of the hologram to appear without artefacts when combined with the Holographic Over, and is a necessity where inputs for display are only 2.5D RGBZ and of a selection of explicit 2D layers where objects are split across layers. It may be used with Object IDs in order to support anti-aliased edges and semi-transparencies where additional AOV data is provided.

However, the quality of results provided by the occlusion synthesis method rely heavily on the quality of the in-painting/in-filling/synthesis algorithm used, and the context. While it may be sufficient for a display with a small eyebox, and little parallax, to use only an edge blurring infill, larger eyebox displays, where there is a larger chance the viewer may purposefully look behind objects, will require much higher detail synthesis.

A set of occlusion synthesis level of detail (LOD) distances may be set, where small depths between layers may utilise a low fidelity synthesis in order to just remove artefacts,

such as where an object is split across close layers, and a more comprehensive synthesis method with detailed outputs can fill larger depth differences and ensure this is temporally consistent.

Whilst occlusion synthesis is a good fallback where sufficient scene data is not available, and for low-power and real-time devices that could utilise a synthesis method instead of a larger rendering compute burden, it is not the optimal solution to successful multi-depth display in CGH. It would instead be ideal, that where a scene is being rendered for the purpose of holographic display, that the fully or partially occluded areas could be presented to the hologram generation algorithms - given that the original scene data is available, it should be utilised.

## 4.5 Deep Holograms

As introduced in section 4.2.1, Deep imagery is a multi-sampled rendered representation of a 3D scene. When a conventional image is rendered, a single sample is produced which may contain information in the pixel such as RGB, Alpha, Depth, Object ID, Reflection, Refraction and so on. Whereas a deep image will provide this information as well as additional alpha information for a number of depth sub-samples. Rather than giving depth data at a pixel for the first ray intersection as in z-depth, Deep contains both the first ray intersection, and any additional intersections through a number of additional samples through depth at a pixel in (x,y). This is demonstrated in figure 4.25. The top left "z volume" pass demonstrates how typical z-depth would represent a volume (such as low-density fog/smoke), and will vary in implementation - where the best case will have a full volume rendered but prevent the correct depth of objects seen through the semi-transparent volume. The bottom left "deep volume" image shows the volume of the fog now represented in a way that samples can occur through this volume and out to the other side, collecting depth information for anything in, or behind, the fog. For motion blur, this also means correctly sampling objects behind the path of an object in motion.

In computer graphics, deep rendering and compositing enables the use of semi-transparent objects, motion-blurred regions, anti-aliased edges, and in-render depth of field - whilst preserving depth data in such a way that additional objects may be composited into such a render correctly. An example of this, with anti-aliased edges and fine edges of fur, is shown in figure 4.26. A full scene rendered in deep, with a VDB volume, is then shown composited with the cheetah, where edges and depths are accurately preserved in figure 4.27, in the 3D view it is also seen that the motion blur on the ball also retains accurate depth data.

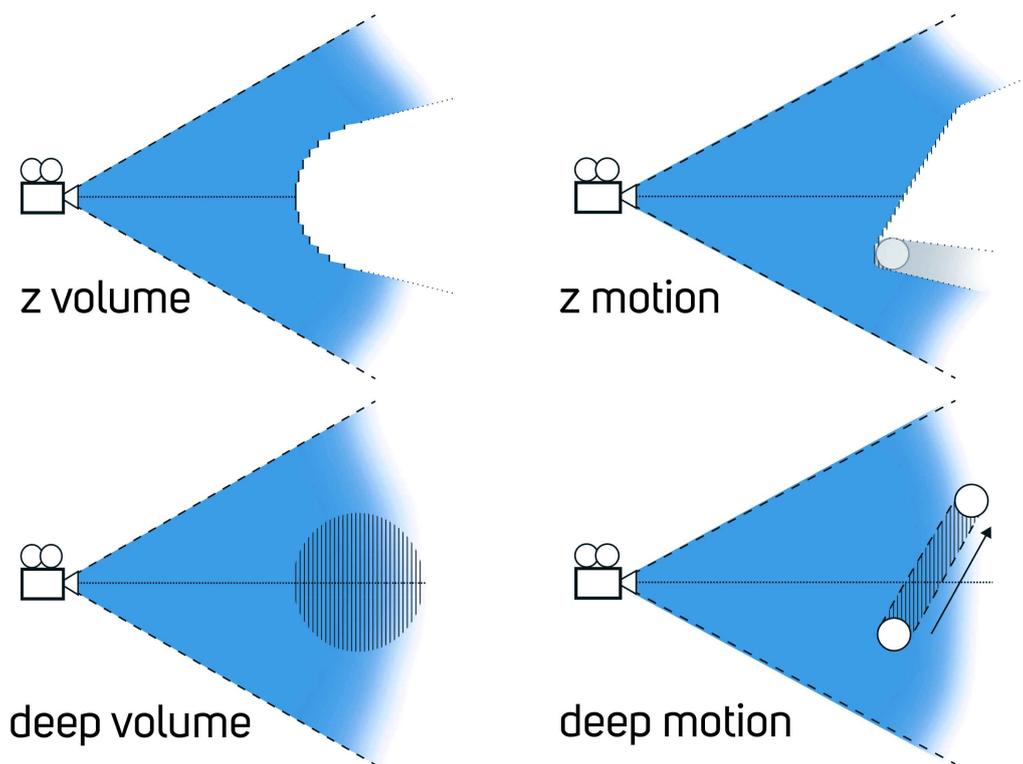


Fig. 4.25 Comparisons of both a volumetric semi-transparent sphere and of a solid object with motion-blur between Z-depth passes and a Deep render pass. Z depth forms a solid stop at the first intersection, and has trailing edges as fine details are lost to sub-sampling. Deep, by comparison, is able to sample accurately across semi-transparent, volumetric and motion-blurred objects.

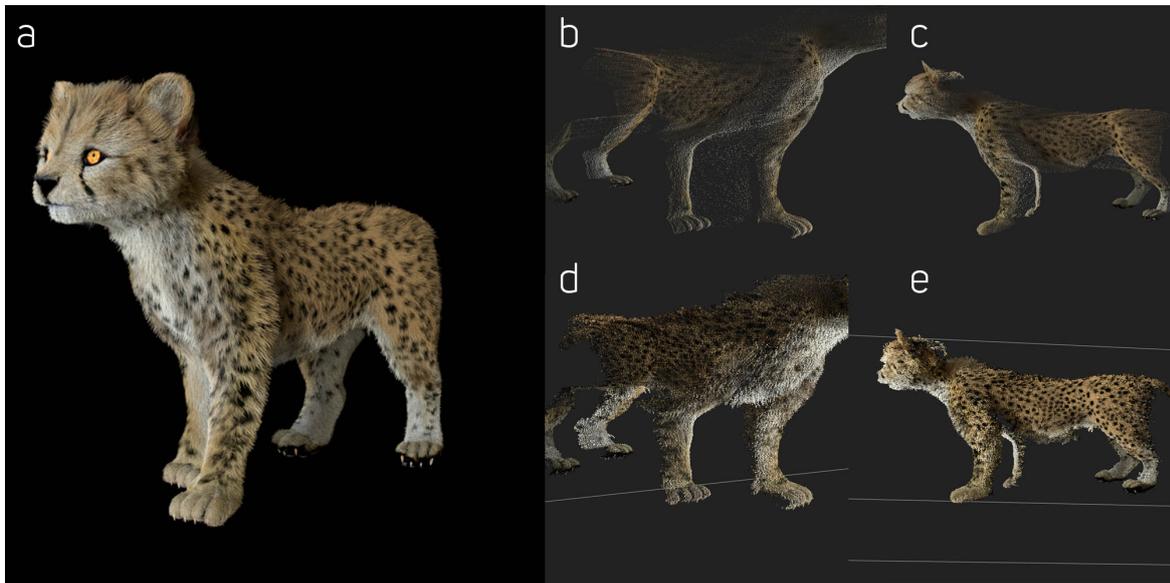


Fig. 4.26 (a) The input render image. (b) Reprojected view of the front legs of the cheetah using zdepth values. (c) Side view of the same zdepth image. (d) The deep representation equivalent of *b*. (e) The deep representation equivalent of *c*. Notice no stray points and significantly more detail in the deep representations.

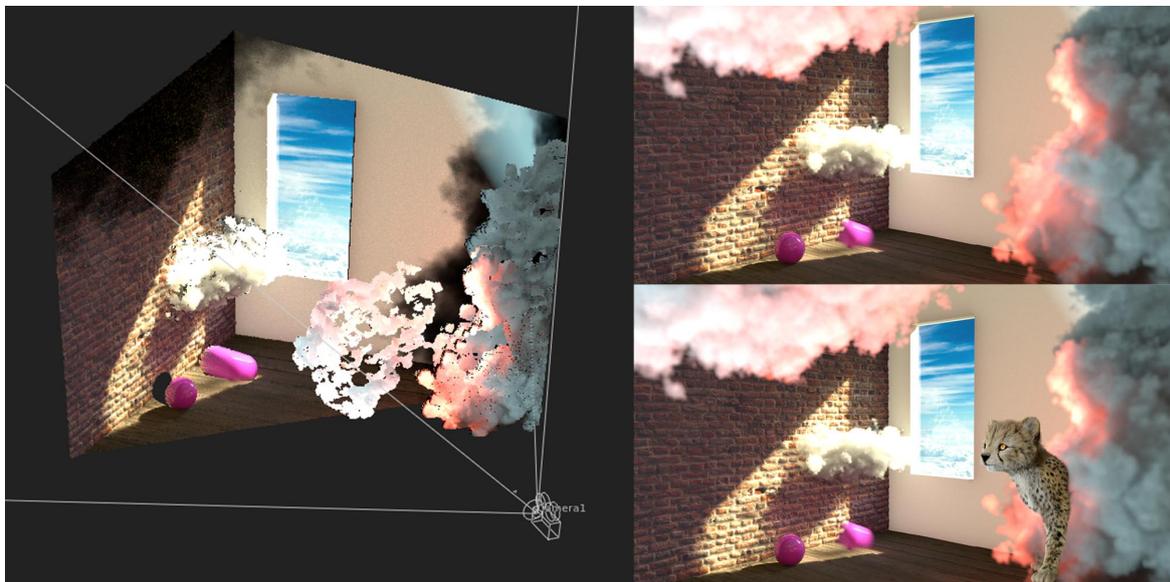


Fig. 4.27 Left shows the deep render scene points visible in right top. Right bottom shows how deep scenes allow the compositing of complex objects - here the cheetah shown in figure 4.26, with fur, is composited into a cloud in the room without artefacts.

As outlined in previous sections, it is in the interest of CGH to support such semi-transparent data in order to form more realistic imagery of 3D scenes for display, and expanding the holographic compositing operations introduced to support Deep compositing would be of great benefit.

Rendering a scene for holographic display using deep resolves all of the issues outlined at the start of this chapter when combined with ray-length z-depth for reflective and refractive effects, and a sufficient expansion of deep rendering techniques to include fully occluded regions. These improvements are suggested in order to provide the blueprint of a one-shot render system for successful holographic display of any kind. This section will provide an overview of how one may utilise deep compositing with the available techniques for layer-based display.

### 4.5.1 Layer based CGH from a Deep source

Given that deep images contain many samples throughout depths, an algorithm must be proposed in order to appropriately allocate such samples onto discrete depth layers which may be used for hologram generation.

As a deep sample represents a volume in space, rather than a particular distance value at a pixel, and deep samples require splitting and subsequently merging where many volume samples occur at the same layer. A number of scenarios occur when performing this operation, some of which are shown in figure 4.28, and are otherwise described in this section. In a deep image, an integer value is provided per pixel to indicate the number of samples available in that pixel, and this is used to iterate between the front and back depth samples. Each of these depth values is associated with an alpha value, which is used to calculate the sample colour - the full theory of which is available in [87, 107], where the following equations are taken from.

As per figure 4.28, for a given deep image, the depth and alpha samples stored in a pixel determine the colour and occlusion contributions for a region of the image. Where volume samples (those which represent semi-transparent regions) are present, they are not intended to represent that of a point in space like a point cloud - but rather that of the contribution across depth of an absorbent material between two points. In order to approximate the scene by discrete layers, such volumes will require splitting at a given depth, which will create two samples. Sample splitting, or extracting a subsample from a sample of length  $l'$  from a sample of length  $l$ , which provides a new sample colour,  $C'$  and a new alpha value,  $\alpha'$  is given by:

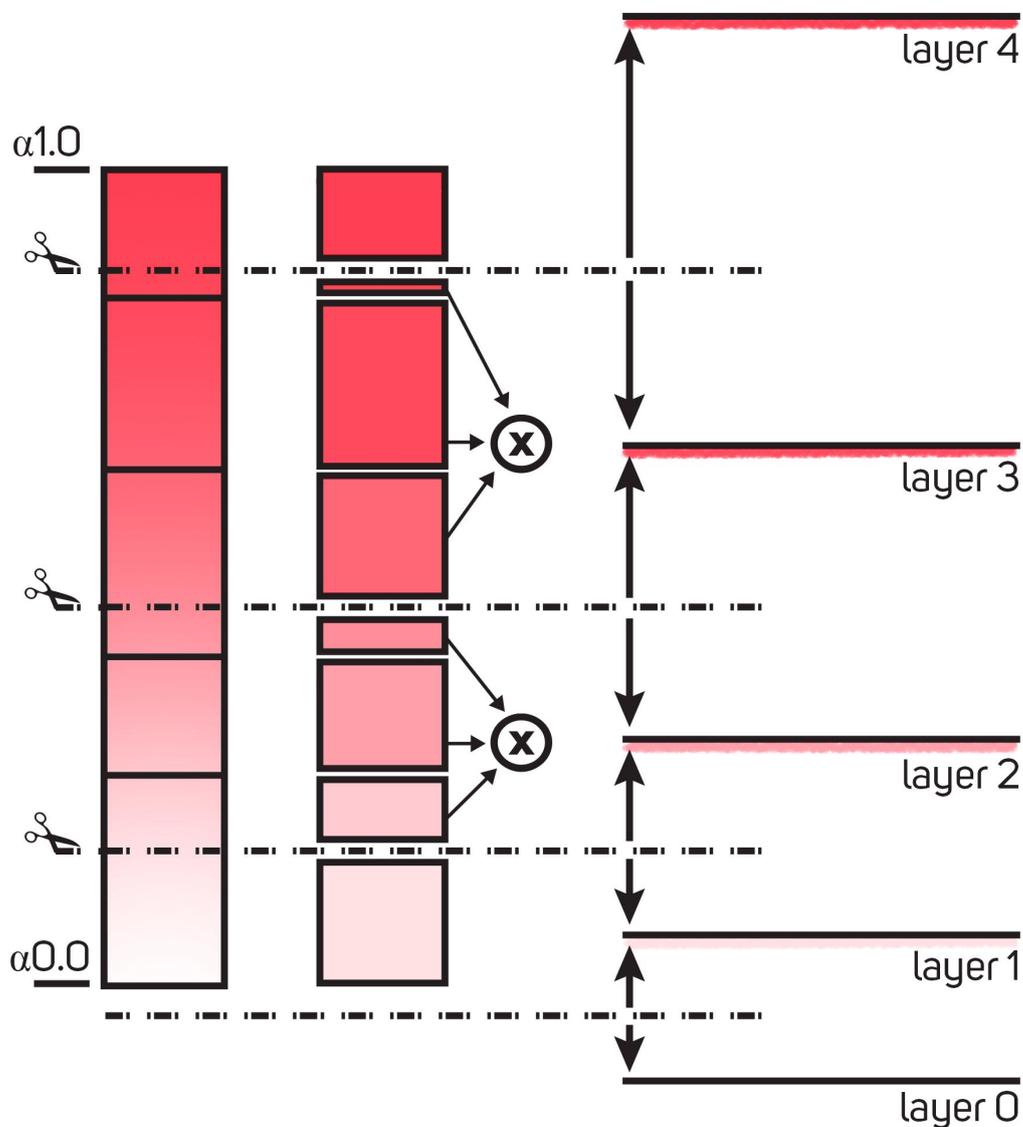


Fig. 4.28 A volume sample from a single pixel (left) is allocated to hologram layers.  $z_{\text{front}}$  is placed at some depth represented by  $a$  acquiring a 0.0 value, and  $z_{\text{back}}$  where  $a$  is 1.0. The complete deep sample, consisting of many subsamples, is consequently sliced halfway between given layer spacings, in this case in dioptres (central). These resulting subsamples are then merged in order to determine the correct colour and alpha information for the requested layer (right).

$$r = \frac{z - z_{\text{front}}}{z_{\text{back}} - z_{\text{front}}} = \frac{l'}{l} \quad (4.7)$$

$$\alpha' = 1 - (1 - \alpha)^r \quad (4.8)$$

$$C' = \begin{cases} rC, & \text{if } \alpha = 0 \\ C \frac{\alpha'}{\alpha} & \text{otherwise} \end{cases} \quad (4.9)$$

To demonstrate, a volume sample such as a fog, which absorbs a portion of light between depths  $z_{\text{front}} = 0$  and  $z_{\text{back}} = 10$  for example, with an alpha of  $\alpha = 0.5$  is split at the centre of the sample,  $Z = 5$ . The sample has a colour (in a single channel) of  $C = 0.6$ . When split, the resulting subsample will have resulting values of:

$$r = \frac{5 - 0}{10 - 0} = 0.5 \quad (4.10)$$

$$\alpha' = 1 - (1 - 0.5)^r = 0.29289... \quad (4.11)$$

$$C' = C \frac{\alpha'}{\alpha} = 0.35147... \quad (4.12)$$

Here the depth at which samples are split,  $Z$ , is defined as the halfway point between two allocating depth layers, and as samples have a constant optical density and colour, the two resulting sub-samples contribute to their closest respective layer. Where a sample will be sliced by many depth layers, the same operation occurs per depth layer, and a larger sample may progressively be sliced into smaller sub-samples. Any sample splitting is non destructive, and the original sample may be retrieved by performing an over operation using the sub-samples. Given that the original sample was split in half, each resulting subsample colour,  $C'_a$  and  $C'_b$ , and alpha,  $\alpha'_a$  and  $\alpha'_b$  each have the same values respectively:

$$C = C'_a + C'_b(1 - \alpha_a) = 0.6 \quad (4.13)$$

$$\alpha = \alpha'_a + \alpha'_b(1 - \alpha_a) = 0.5 \quad (4.14)$$

Sub-samples allocated to a given layer require merging to form a new single total sample. These may be a combination of fully transparent (where a point is emissive, but does not occlude), semi-transparent, or solid samples - these cases must be considered for when calculating the new merged sample values:

$$\begin{aligned}
\alpha_c &= 1 - (1 - \alpha_a)(1 - \alpha_b) \\
\mathbf{C}_c &= \begin{cases} \mathbf{C}_a + \mathbf{C}_b, & \text{if } \alpha_a = 0 \text{ and } \alpha_b = 0 \\ \mathbf{C}_a, & \text{if } \alpha_a = 1 \text{ and } \alpha_b < 1 \\ \mathbf{C}_b, & \text{if } \alpha_b = 1 \text{ and } \alpha_a < 1 \\ (\mathbf{C}_a + \mathbf{C}_b)/2, & \text{if } \alpha_a = 1 \text{ and } \alpha_b = 1 \\ \alpha_a \left( \frac{\mathbf{C}_a \log(1 - \alpha_a)}{\alpha_a} - \mathbf{C}_b \right) / \log(1 - \alpha_a), & \text{if } 0 < \alpha_a < 1 \text{ and } \alpha_b = 0 \\ \alpha_b \left( \frac{\mathbf{C}_b \log(1 - \alpha_b)}{\alpha_b} - \mathbf{C}_a \right) / \log(1 - \alpha_b), & \text{if } 0 < \alpha_b < 1 \text{ and } \alpha_a = 0 \\ \alpha_c \left( \frac{\mathbf{C}_a \log(1 - \alpha_a)}{\alpha_a} + \frac{\mathbf{C}_b \log(1 - \alpha_b)}{\alpha_b} \right) / \log(1 - \alpha_c), & \text{otherwise} \end{cases} \quad (4.15)
\end{aligned}$$

A large number of sub-samples allocated to the same layer requires the same merging process to occur to form the resulting display layers, as per the central section of figure 4.28. This allows the formed display layers to then consequently correctly holographically composite together to form a correct representation of the 3D scene. Figure 4.29 shows the resulting display layer stack from the layer slicing method, and figure 4.30 shows the total result of this proposed method - where a selection of solid objects are suspended in a VDB volume.

The resulting scene shows that semi-transparencies can successfully be captured from a renderer and holographically composited for display. Minor visible differences seen between the input and resulting hologram replay are explained by an accumulation of error across many layers via an approximation in calculating the layer-energy for simulation, seen in the pink cloud turning white (as per the holographic HDR highlight compression). Additionally, colour transforms alter the holographic result. The image exhibits the glowing and dark border artefacts at edges, related to occlusion, due to the lack of information behind solid objects, but can successfully allow accurate display of volumes such as fog and fire, motion blurred objects, thin details such as hair and fur and also accurate anti-aliased edge information.

## 4.5.2 Deep Occlusion

As the resulting display-layers output from the previous section have RGBA data, and an associated depth, each layer can successfully be composited holographically in 3D space using over operations. This allows the correct display of numerous partially occluding layers, such as motion blurred objects, slices from a volume, or anti-aliased edges. However, this



Fig. 4.29 The resulting layers extracted from a Deep EXR file. The layers are extracted according to dioptre weighted spacing and the rules which dictate how samples are split and merged when compositing such renders.

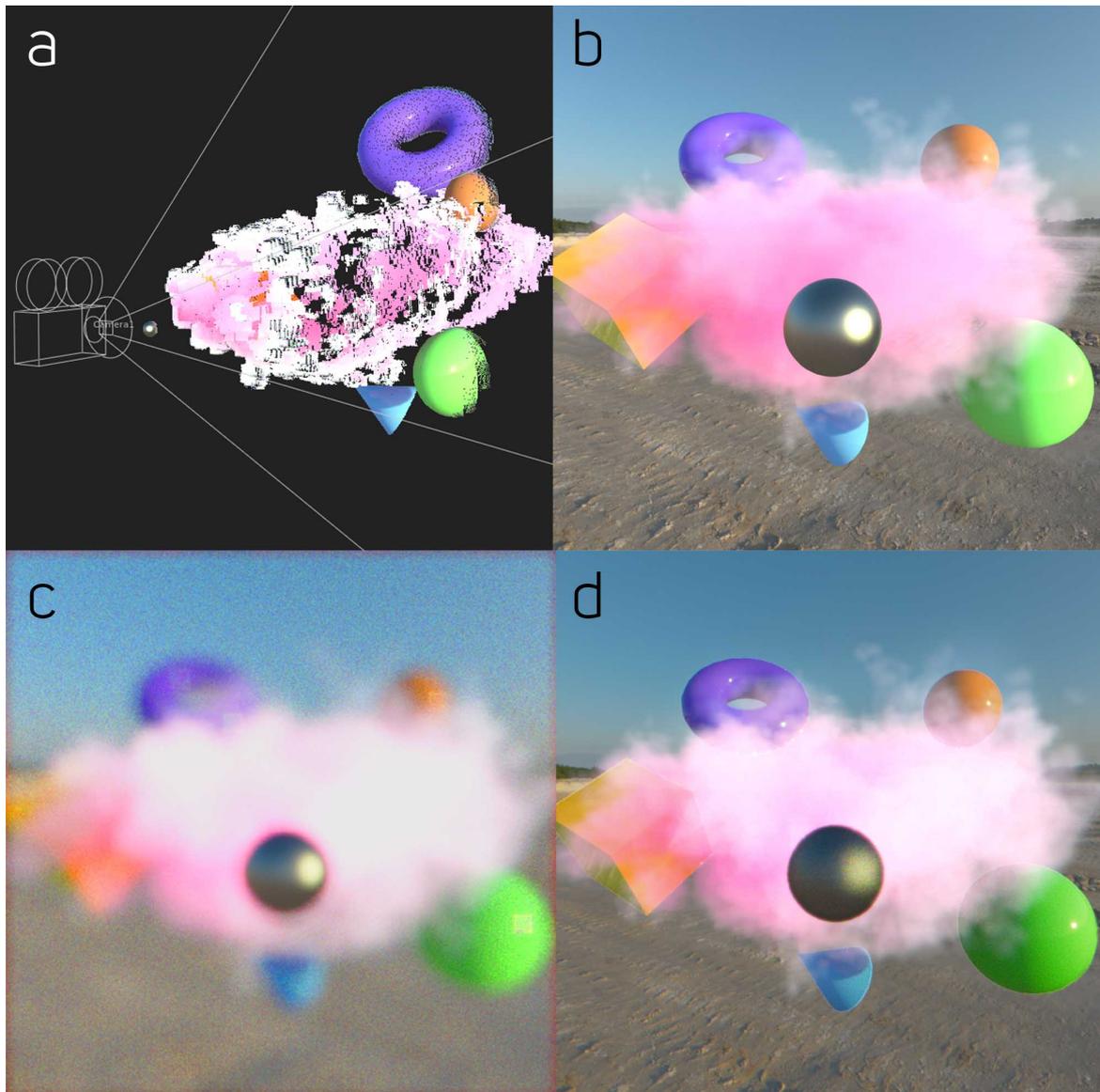


Fig. 4.30 (a) The input scene deep-to-points projection. (b) The input scene from the render perspective. (c) Holographic replay with focus on the very front of the central ball, which is very close to camera. (d) Holographic replay of the resulting layer sliced and composited scene, with focus at the centre of the cloud.

does not allow for the case where solid objects may wish to be composited together, or where a solid object is rendered embedded within a volume- as sampling through space will stop when a solid object is hit, or where only one sample in a pixel is presented. As explored in the previous section, occlusion synthesis may be applied to these cases to ensure that solid objects do not leave artefacts in the holographic image, and this may easily be integrated alongside the deep layer slicing algorithm presented here.

Deep occlusion synthesis masks may be created where layers beyond a sample which has an accumulated alpha of 1.0 exist, where a layer contains the total volume sample between  $z_{\text{front}}$  and  $z_{\text{back}}$ , or where  $z_{\text{front}}$  and  $z_{\text{back}}$  are at the same depth, indicating a point sample, this is shown in figure 4.31.

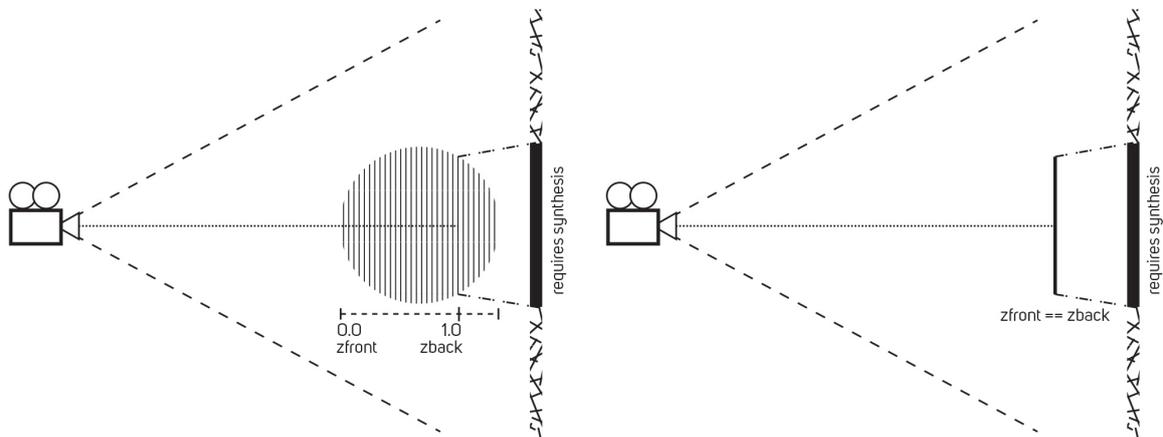


Fig. 4.31 Left shows a volume which has an accumulated alpha of 1.0, resulting in a region in shadow from the camera perspective which is not rendered, and consequently requires occlusion synthesis. Right shows the same scenario but where the front and back samples are at the same depth, representing either a volume sample flattened into a discretised layer, or a point sample.

It is also important to note that while the documentation for deep EXR files indicates use by only sampling through volumes or other semi-transparencies until a solid point is hit, the sub-sampling storage interface in a deep file may be adjusted for additional means. Similarly, a note in the documentation of Pixar’s renderer, Renderman, indicates such on the topic of sampling until a solid surface hit [167] - “This limits the ability to perform operations such as removing opaque objects by their identifier, or perform defocus on deep compositing images”. So while deep imagery, in its base form, will assist with the composite, and by extension the holographic display of semi-transparencies, more can be offered to render and store occluded objects.

If integrating such a consideration to a renderer, and complete freedom in the composite is valued even at the cost of render times, all occlusions should be rendered - including entire regions behind any solid objects. Where a deep image is being rendered for holographic display, or output for display from an existing compositing application, it becomes unnecessary to render all occluded regions, as there will be a limited area which will be visible by the viewer. In order to reduce additional rendering, or reduce storage requirements, this render region may be defined by the aperture of the holographic display system. As shown in figure 4.3, at a given viewing position, the circle of confusion defines the visible regions of the image which would otherwise be occluded when a scene is projected under the pinhole model. Here, the field of view of the hologram may be used to define the render regions required, as shown in figure 4.32. This rendering may simply occur by a similar mechanism to which defocus is calculated in raytracing, but without utilising the defocus result of occluding objects for the final framebuffer.

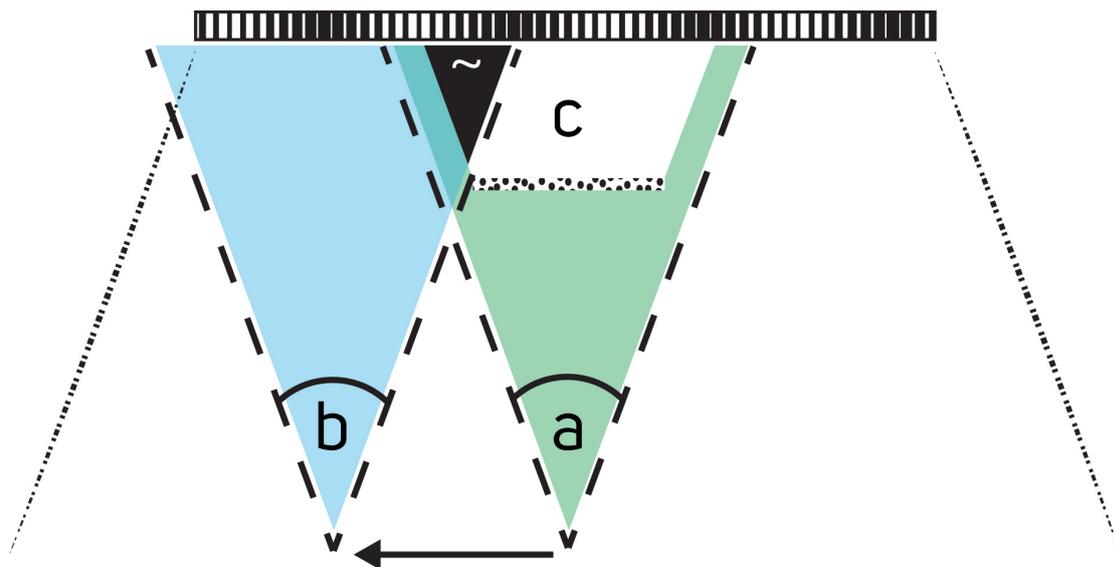


Fig. 4.32 Viewing position *a* shows a viewing frustum towards the virtual image projected by the SLM, the region *c* is in the shadow of the virtual object. When *a* moves position, indicated by *b*, the frustum is unimpeded by the virtual object, causing the region marked by  $\sim$  to become visible. This region however, is occluded from viewing position *a* when rendered with a pinhole model.

In storing such samples where solid point samples may occlude other solid point or volume samples, a scheme may be used per pixel where rather than inferring colour data across depth via the alpha, sample colour may be explicitly stored. Additional adaptations to deep, such as OpenDCX (Deep Compositing Extended), provides support for per-sample metadata and allows sub-pixel masks for improved hard-surface compositing [58]. Whilst

such alterations are suitable for internal use within a studio, wider adoption for holographic display would require end-to-end support across a range of DCC packages.

## 4.6 Renderer Integration Streamlining

While each of the techniques proposed in the previous sections allows various aspects of the holographic image to be improved, in order for a complete solution to be realised, they must be combined in such a way that allows the most original scene data to be preserved.

### 4.6.1 Complete Scene Representation

A renderer which delivers a complete “hologram-ready” scene representation, to the proposed HQH standard proposed in section 3.5.6, must include:

- Sub-sampled depths
- Colour and depth data for occluded regions
- Colour and depth data for specific optical properties which allow imagery to be formed through reflection and refraction.

All of this data may be stored in an EXR using the deep format scheme, and this also enables the multi-sampling of depths via reflections and refractions. Such a scene which requires this detail in order to successfully be displayed is presented in figure 4.1.

Delivering such complete scene representations is not only beneficial for holographic display, but would also bring benefit to conventional 2D compositing workflows - allowing a single render file to be used to contain a complete scene which is editable using existing deep based tools. Such a technique, for example, when combined with the use of an object ID pass, allows for objects to be removed and replaced by new versions in the composite, and this is also true of the holographic composite. In cases where occlusions do not change, such a method also will mean only the replaced object will require to have a new hologram generated, rather than regenerating the entire scene. For example, given the scene in figure 4.30, an object can be updated as follows:

$$\text{Total Scene Holo} - \text{Original Object Layer Holo} = \text{Scene With Object Occlusion Holo} \quad (4.16)$$

$$\text{Scene With Object Occlusion Holo} + \text{New Object Layer Holo} = \text{Fixed Scene Holo} \quad (4.17)$$

Such a scene representation provides a sufficiently lossless capture of the scene from a given perspective, and is suitable for accurately providing the scene data to the hologram generation algorithms, in order to successfully display the scene. Even where scenes are approximated for the purpose of reducing compute, an image may be produced that is far more faithful to that which exists in the authoring DCC application than from solely an RGBZ output image.

#### 4.6.2 Context Based Layer Slicing

Given that holographic display offers a 3D viewing experience, it is usually assumed that the input should be fully 3D - but it is often the case that a lot of media is created by compositing various 2D layers, even when creating 3D scenes. Examples of this include the use of 2D matte paintings for backgrounds or other objects at various distances, 2D sprites for integrating fog into scenes in 3D space, or simply 2D animations with multiple layers. Layer-based CGH can holographically composite and display such elements directly, and does not require that scenes be represented entirely by a 3D source. Ensuring that the context of the media intended for display is taken into account by a display algorithm is essential for both conveying the artistic intention whilst ensuring computational efficiency.

While this thesis has so far concentrated on delivery of imagery from 3D renderer sources, slicing a scene into representative layers is one that is actually intrinsic to 2D animation, and parallels can be drawn to the construction of layer-based CGH.

Techniques in traditional animation such as the Multi-Plane system, made popular and patented by Disney [51] in the 1930s, places stacked animation celluloids (cels) on glass panes, placed at different depths in space to then be captured by camera. This allowed layers to be moved in parallax and have varying focal depths, and is seen in use in traditionally animated films such as *Snow White and the Seven Dwarfs*. Later this process would become fully digital, and is still how 2D animation occurs today - these are rarely formed of more than a handful of depth layers. The original process is shown in figure 4.33.

If one wished to utilise animation with traditionally painted cels for a CGH production, cels take the form of layers, and may be scanned and holographically composited. The same method applies to digital layer techniques, and any otherwise digitally animated frames. Here, it is the decision of the artist as to how scenes must be divided across layers, and this is still very much a control available in CGH. Other similarly suited media includes side-scrolling or other 2D games.



Fig. 4.33 Traditionally, a number of layers from a single painting (left) would have been separated so that they can be placed at some distance apart from each other (centre) in order for physical camera effects such as defocus and parallax to be seen in the final animation (right). This example taken from the episode "Tricks of Our Trade" (1957) from the series "The Magical World of Disney" [205]

Allocating layers when extending into other sources, including complex 2D footage or material from 3D sources is where such decisions become more complex. For a production where all elements are viewed in full 3D, as discussed in the prior sections in this chapter, various depth and object representations assist in placing layers and consequently reducing artefacts where imagery is discretised. This is suitable for full 3D animation, and is suitable for productions which are willing to have entire 3D scans of sets, or otherwise have 3D conversions of any footage shot on a typical cinema suited camera. This requires significantly more work, both on set and in post, and significantly more data storage for shots - and practically speaking is likely far more information than is required for the artistic intent in most shots.

Rather than fully 3D shoot, render and composite shots, the fact that layers are being allocated does allow a reduction in data requirements, and assists in composing shots that may be formed from mixed sources. Modern examples of this include any 3D rendered material that is composited with 2D plate footage, which is a feature of any blockbuster film. Older examples could include such productions as "King Kong" (1933) [119] or "Jason and the Argonauts" (1963) [103], which composites regular 2D plates with physical model stop-motion animation, or "Who Framed Roger Rabbit?" (1988) [217], which uses painted cels. Here in both cases, the footage which is filmed in 2D is sliced up into appropriate layers (rotoscoped or keyed) in order to facilitate the compositing of new scene components. These resulting layers may then be used as input layers for the holographic compositing process - and this is applicable even where no new objects are to be composited or other VFX work to take place. The layers may then be placed manually in depth as an artistic decision. A similar process of manual depth placement already happens when compositing 2D layers such as fog into 3D scenes, or when converting a production into stereo for stereoscopic 3D

release, including the painting/addition or synthesis of new depth information to assist the composite (and disparity) - except here all of this depth information can then be passed into the hologram generation algorithm.

Ultimately, when not working with complete 3D representations of scenes, artistic decisions will be required to allocate layers best to the context, otherwise layers may be allocated automatically, either to meet a computational requirement or a required layer density for particular complex shots. Automatic allocation, as discussed previously, may occur by various schemes, utilising the techniques proposed for multi-sampled depth enablement proposed in this chapter - including utilisation of Object ID, deep data or occlusion synthesis. This automatic allocation may not be perfect in representing the scene as intended. A summary of the source considerations are provided in table 4.4.

As each layer is an additional compute burden, it is in the interest of any automatic allocation algorithm to try to reduce the number of layers required. For some content this produces unacceptable results, and more layers cannot be avoided - but other scene sources are well suited to a very small number of layers. The type of display, and the required level of detail required to view the hologram will also need to be taken into account when calculating the compute budget. Any near eye display will require a display per eye, which doubles the number of layers required to be computed if the content needs to be viewed in stereo 3D, which of course must also be taken into account when rendering or otherwise animating source content.

Any limiting of layer counts will increase parallax issues between layers - continuous surfaces (surfaces that span many depth layers) require more layers to reduce this error. For interactive rates during production, or real-time rates, layer discretisation is a requirement, and this is the trade-off that layer-based CGH knowingly makes - so it is hoped that the methods presented in this chapter will sufficiently increase visual fidelity using a limited number of layers in these situations, and provide suitable controls for artists.

### 4.6.3 Other Scene Data

Other scene data which is often used in compositing may also be utilised by a given hologram generation algorithm. Such data could be used for performing motion blur in fourier space via FFT [171] as shown in figure 4.34. This also will partially reduce the overall computational requirement compared to doing the same operation for an image-domain result, where the original technique is compared with the same method adjusted for holographic display, as shown by figure 4.35. Here, motion vector data, which may be stored as 2D or 3D vectors, which is commonly used in compositing for performing motion blur, may be utilised to perform the same blur, but instead at the display stage. Rendered frames will also typically

Table 4.4 Summary of Scene Source Considerations for Layer Slicing

Scene Source	Slicing Method	Requires Synthesis	Semi-Transparency Support	Resulting Display
2D Full Layers (multi-plane)	Manual allocation of depth layers	N	Y	Exact recreation of 2D layers in 3D space
2D Composite Source (e.g. mixed 2D plates + explicit layers)	Manual allocation of depth layers	Partially, may be required for plates only	Partially if built of explicitly composited semi-transparent layers	Exact recreation of 2D layers in 3D space
2D + 2.5D	Mixed automatic + Manual allocation of depth layers	Partially, may be required for plate. Required for 2.5D	Partially if built of explicitly composited semi-transparent layers	Exact 2D layer recreation with layers automatically in the appropriate region
2.5D Source	Automatic allocation of depth layers	Y	N	Layers automatically allocated in the appropriate region
2.5D Source with AOV's including Object ID	Automatic allocation of depth layers, assisted by object ID data.	Y	N. May be enabled by multiple render layers	Layers automatically allocated in the appropriate region. Enables reflections and refractions
Limited 3D Deep Source	Full automatic allocation, assisted by object ID	Partially, required for regions behind point samples	Y	Enables semi-transparencies from a rendered source. Missing precise image data for occluded regions
Full 3D Deep Source	Full automatic allocation, assisted by object ID	N	Y	With a sufficient number of layers, an indistinguishable reproduction of the input source imagery

include data such as exposure time in their metadata, and indeed that is a requirement for this method where the motion blur is not performed in-render.

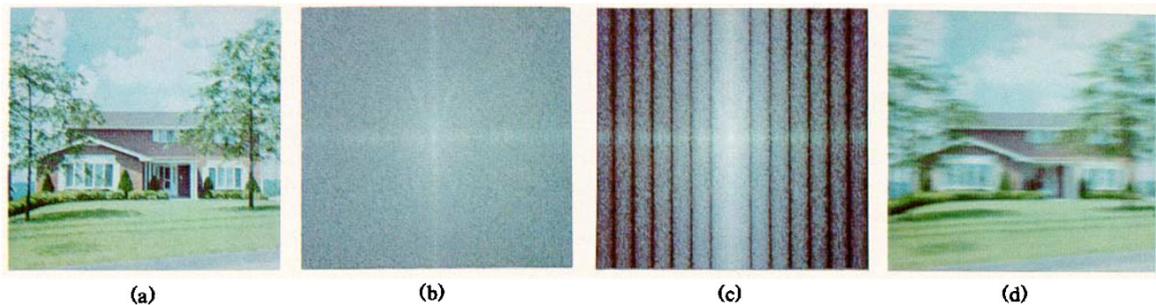


Fig. 4.34 From Potmesil and Chakravarty 1983 [171] - Uniform motion-blur with frequency domain computations: (a) Input image, (b) FFT of original image, (c) FFT convolved with point spread function (PSF), (d) Blurred image.

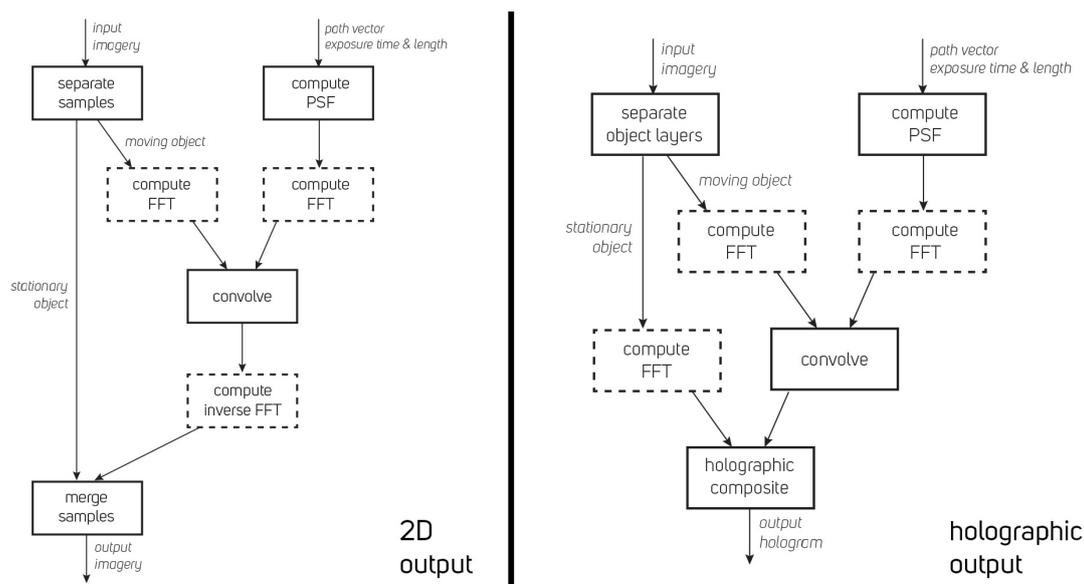


Fig. 4.35 Left shows a block diagram of the motion-blur process as described by Potmesil and Chakravarty [171], which is intended for the output of 2D imagery. By adapting the process to keep imagery in the frequency domain, shown right, holographic objects/layers may be displayed directly after the point spread function (PSF) is applied.

Samples can be thought of as layers in the same way that a layer can exist with only 1 pixel in it. For compute reductions sake, many samples should ideally be bundled into a discrete number of layers for the motion-blur to be calculated, and this would allow for non-uniform blurs across the image. As the samples that make up the total image are composited holographically, and the inverse FFT step is effectively performed by the lens in the viewer's

eye, the computation is not required - and the sample merging may occur instead after the specific region to blur is convolved. Such a specific region may be indicated by the motion vector data pass alone, or also be assisted by Object ID data, to ensure that regions may correctly be blurred when occurring in regions which are partially obscured. Additionally, shadows of objects which are to be motion blurred must be indicated, as must reflections or refractions of such objects - this may be indicated by additional AOVs with associated vector and ID data. Ideally, performing motion blur in this way will be used with a source where regions occluded by the object being blurred have been rendered and are available (such as in a complete scene deep format image), or this may otherwise have occlusion synthesis performed.

## 4.7 Multi-Sampled Depth Results

This chapter has presented the issues that hinder CGH imagery from presenting convincing imagery when using limited source material, either in limited depth or other lacking layer information. By comparing these issues to the solutions presented in conventional computer graphics and image rendering, this chapter has developed and presented techniques which allow holographic display to overcome such limitations. The methods presented enable successful representations of complex scene imagery in CGH using data which is readily available from existing production renderers.

The introduction of the concept of holographic compositing enables the artist to realise significant control over the final displayed holographic image, and also allow such a concept as full 3D display to be seamlessly integrated into existing production practice. Here, existing compositing practices have been adapted and employed with additional render data, for example to provide the additional data required for physically correct reflection and refraction passes in depth, which is presented by full scenes displayed in figure 4.36 in simulation and figure 4.37 captured from a real display. Holographic compositing also enables the proper composite of layers sliced from deep rendered imagery for 3D display. These methods are the first to consider proper anti-aliasing and related accurate depth information, the first to consider any form of volumetric representation in CGH, the first to enable motion blur, the first to enable reflective and refraction representations in layer-based CGH and the first to provide a viable one-shot complete-scene rendering proposal for any form of CGH, which supports any kind of scene representation supported via existing renderers.

This chapter has also proposed a novel method of Occlusion Synthesis, which reduces, or fully removes, visible artefacts in imagery which contains insufficient data for holographic display, such as from a 2.5D source, or where objects have been otherwise sliced across



Fig. 4.36 Holographic replays (simulated) of a scene containing a lens of  $f=180\text{mm}$ ,  $0.5\text{m}$  from camera. Above: Focus is at  $0.46\text{m}$  on the smudge on the magnifying glass surface, the focus of the reflected lights is  $0.52\text{m}$ . Below: Focus ( $1\text{m}$ ) now on the text printed on the film box and bunny statue, the film box is physically placed  $150\text{mm}$  behind the lens. Rendered in Maxon Redshift, from a scene in Autodesk Maya.



Fig. 4.37 Holographic replays (real) of a scene containing water droplets on a pane several centimeters from the camera, with a city scene behind it. The drops both refract and occlude the background. Captured on a BMPCC4K, from a physical projector, with source scene combining a video source with droplets simulated and rendered in Houdini with Mantra (SideFX), and adjusted in Nuke (Foundry).

depth layers. By utilising various methods of image synthesis, the quality, and compute, can be carefully controlled - enabling high quality offline-based image formation, as well as fast approximations for real-time applications. This contribution alone resolves the key visible issue with layer-based and RGBZ holography identified in the literature [24]. The methods developed here are also not limited to holographic planes which are perpendicular to the viewer - and layer allocation on tilted planes [142], which could include individual polygons, may make use of the methods provided, but is outside the scope of this work.

As well as the developed methods that underpin successful holographic display, recommendations have also been developed as to which methods are suitable given a particular input scene source, all with the goal of enabling unhindered specific artistic intent.

# Chapter 5

## Multi-View Holographic Display

### 5.1 Multi-Perspective CGH for Large Format Displays

So far the works presented in this thesis have not been explicitly concerned with the end device in which the holograms generated have been displayed, but have mostly assumed a near-eye configuration. A hologram generated by current layer-based CGH methods will represent a 3D scene from a single fixed perspective. This assumption simplifies not only the technical requirements of a display (i.e. to a display with a small eyebox which is possible with today's SLM devices), but also the rendering of any imagery - as only one camera perspective is required.

However, an ideal holographic display device should allow the viewer to adjust their viewing position without the requirement of wearing the display itself, and allow many viewers of the same display at once. Such a hologram may be in the physical form of a computer monitor or TV. Though, miniature versions could be applicable for VR headsets. Such a large-format display may be thought of as identical to looking through an open window, where the viewer can observe a scene on the other side of the frame, and objects can reach or fly in.

Equally, in the case of the single perspective near-eye displays, and as explored in Section 4.2.2, it is frequently the case that the viewer may go “off-axis” from this fixed perspective. Such a scenario requires occlusion synthesis in order to prevent the perception of the image from falling apart. This is sufficient where the viewer is within a relatively small eyebox, but when the size of this eyebox increases, even to the extent where the viewer may move their head by several millimetres, the single-perspective method can no longer sufficiently represent the scene in 3D - and is instead perceived exactly as the 2D layers that the scene is built with. This of course is appropriate for a certain selection of content, such as conversions of 2D animation that makes use of techniques such as the Disney

Multiplane Camera [51], or in 2D side-scrolling video games. But where one wishes to display 3D objects in a “continuous” fashion, new perspective views are required to maintain a convincing experience. These must be catered for both in the rendered source material, and in the hologram generated and consequently displayed on the SLM.

A comparison of near-eye with small eyebox, near-eye with larger eyeboxes and large scale display devices is given in figure 5.1. And imagery captured from a real SLM in figure 5.2 shows how different images may be presented to the viewer in different parts of the eyebox. These sub-holograms viewed in select parts of the eyebox are also referred to as hogels, and each sub-hologram can present a different image to the viewer.

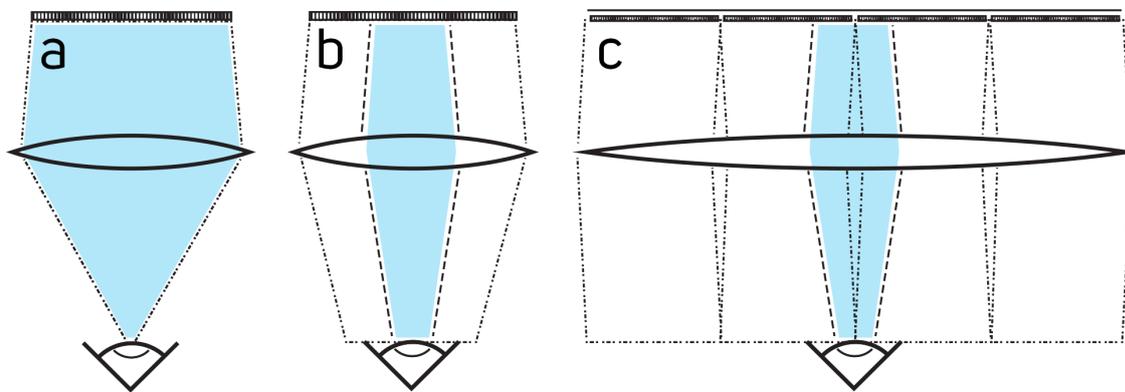


Fig. 5.1 A comparison of various eyebox scenarios: a) A small eyebox which fits entirely in the pupil ensures the entire SLM is used for the image. b) A larger eyebox results in the pupil becoming the limiting aperture, and only a segment of the SLM is viewed at once. c) An extremely large eyebox, formed by combining multiple SLM devices, still results in a pupil-limited aperture - but there is a larger extent to which the viewer can move through the image.

This chapter proposes a suitable production system for generation of holograms which allow the user to move from the central optical-axis of the display, primarily concerned with large scale displays. This chapter outlines methods of capturing and rendering scenes in practice to generate a hologram with a large eyebox, and proposes an efficient CGH display method which allows the creation of holograms that can encode 3D scenes. Whilst the ideal physical hardware for this is yet to be developed, assumptions can be made for suitable simulation - and the techniques developed here are applicable to any future large format holographic display system.

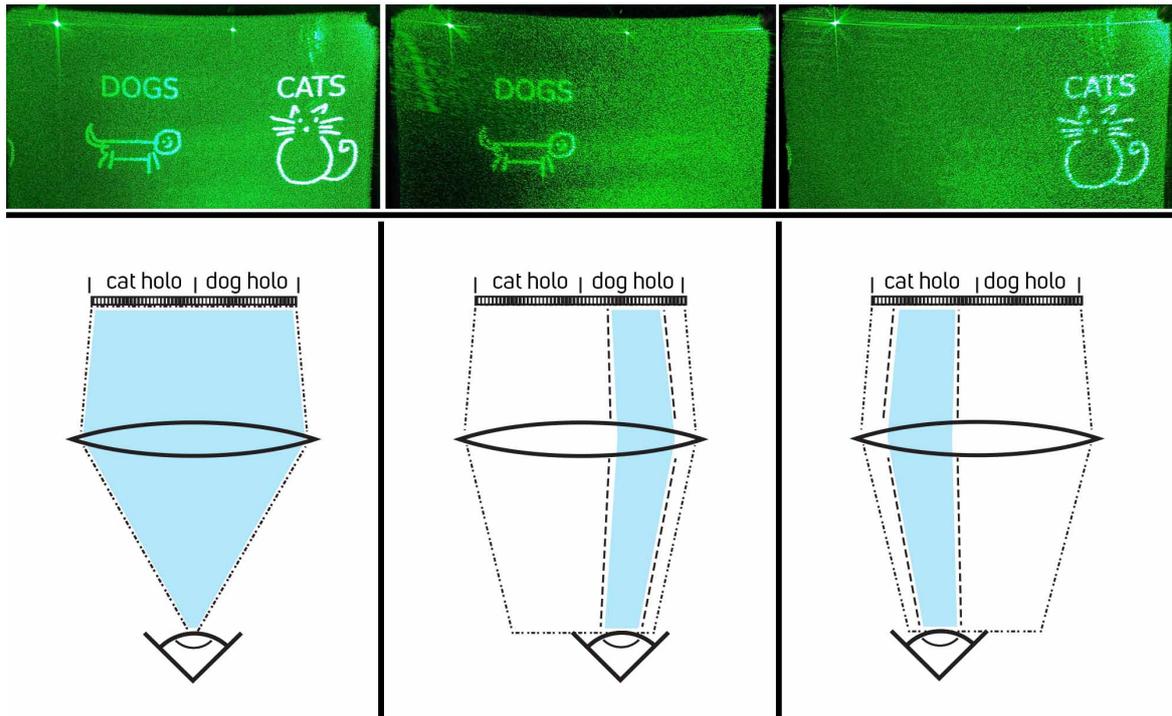


Fig. 5.2 Here three viewing scenarios are presented with a single wavelength hologram made up of two sub-holograms: Left, the entire SLM is viewed at once, and both the cat and dog image can be seen. Centre: The eye is presented with a portion of the eyebox which contains only the dog sub-hologram - so only the dog is visible, despite in image-space the dog appearing on the other side of the device. Right: The eye moves to the other part of the eyebox in which only the cat sub-hologram is visible. Note the reduction in resolution where the viewed hologram area is reduced. Each sub-hologram can present an entirely different image to the viewer.

## 5.2 Related Work

Research on displays that can represent multiple perspectives is generally now divided into two general tracks, that of lightfield display, and that of holography.

Here, the relationship between lightfield display and holographic methods is explained. Followed by an examination of the nuances and relationships between the various holographic techniques, what they each offer, their tradeoffs, and how they might form a large format 3D display.

### 5.2.1 Lightfield Display

Lightfield, as it is known today, was first outlined by Levoy and Hanrahan in 1996 [130] as a rendering method which allows the generation of new synthetic views from a collection of input views.

It is this “field”, where radiance is represented as a function of position and direction, that allows this interpolation of unique views to take place. This representation of a given 3D scene also allows the display of aforementioned unique views, with the visual medium being first explored in 1908 by Lippmann [133, 12], who proposed a fly’s eye lens array for image capture to a film plate and subsequent re-projection as 3D display. It is the multiple views provided to the viewer which fundamentally provides the depth cues.

A modern lightfield display system provides a three-dimensional experience to the viewer via the use of these many views of a scene, which are typically displayed on a conventional 2D panel with an additional optical layer. Figure 5.3 shows how such a field of light may be captured, and then consequently displayed.

Each view is aligned with and viewed through cylindrical lenslets (often referred to as lenticular, shown in Figure 5.3 C) providing one dimension of parallax, or spherical lenslets (often referred to as integral photography) for two dimensions. Both of these types of display sacrifice spatial resolution in order to gain full parallax (angular resolution) [78], and have a fundamental limit [12] in the spatio-angular tradeoff due to the diffraction limit of individual lenses [125, 25], these limits mean that lightfield display will not suitably match human visual acuity.

### 5.2.2 The Holographic Stereogram - Comparison with Lightfield

The comparison between lightfields and holography is most easily made by the “Holographic Stereogram”, both of which use “epipolar volumes” to represent 3D imagery formed from an array of input images. These imaging methods are making use of a hybrid of photographic

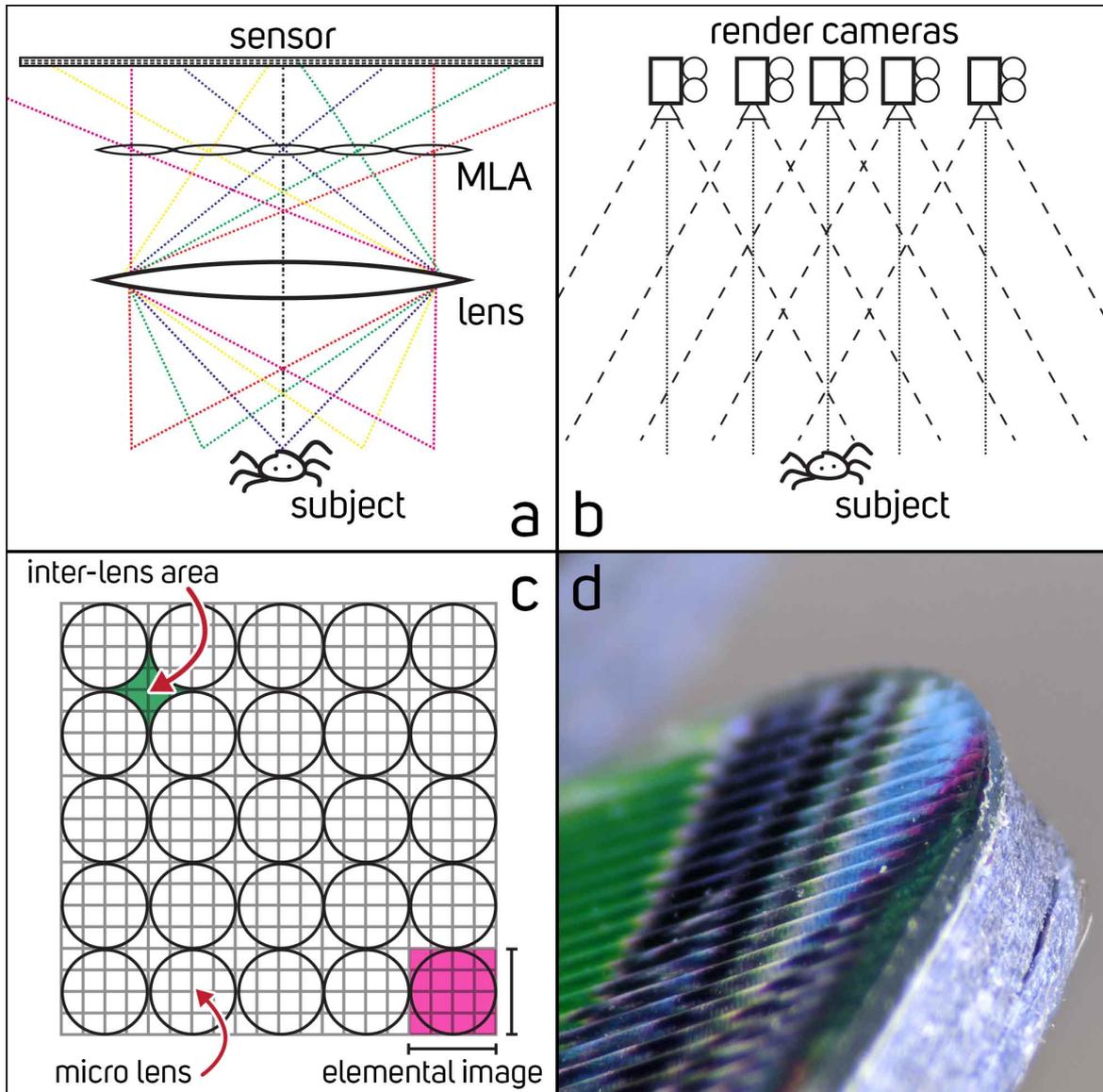


Fig. 5.3 (a) Provides a diagram of how a lightfield may be captured on an image sensor using a micro lens array (MLA). The subject is focused on the array and consequently split into elemental images on the sensor. This setup may also be comprised of multiple unique cameras, as per (b) Where multiple virtual render cameras capture the scene in an array. (c) Shows a lenslet display, where each elemental image is shown on a conventional raster-based display behind an accompanying micro lens. (d) Shows an image of lenticular ridges fused onto a puzzle piece, under a microscope.

and holographic techniques, designed to simplify the capture and display process [12]. This was originally intended to apply to analogue holography, where a range of discrete viewing positions could be captured in a less complex holographic recording process to allow wider applications of the imaging technique - but still applies to CGH, where a reduction in compute is desirable.

Using a number of 2D images, a scene's 3D qualities is approximated in a holographic stereogram - this may be viewed through a number of slits, or viewed via a transfer hologram [78], but in either case, the hologram itself does not provide a sufficient focus cue and has a limited depth of field [77, 137]. Some stereograms may be only equivalent to a 2D binocular stereo display with parallax [111], and some with only horizontal parallax - the resolution and depth of focus of which is limited to its component photographs. A comprehensive comparison of analogue holographic stereogram techniques is given in "Survey of holographic stereograms" by Stephen Benton [12], where lightfields and holographic stereograms are described together, because they are both the same representation of a 3D scene.

### 5.2.3 Holographic Field Sampling

A fully computed hologram (such as one from a "holographic renderer", as discussed in Section 2.2.1) is not dissimilar to a sufficiently dense lightfield display of spherical lenslets or similar holographic stereogram. Here, each pixel on the SLM is effectively a lenslet, and represents a unique perspective into the scene, but is far denser than the diffraction limitations of lightfields allow. While this in theory provides a superior experience in quality, the information encoded in a fully raytraced hologram (from a 3D scene of points or polygons) is far denser than is required to represent an apparently continuous scene to the viewer.

It is also important to note that the field of view of the SLM device is coupled with the pixel's diffraction cone. Devices with smaller pixels will diffract the incoming wave with a wider angle, and consequently provide a wider field of view for the resulting image perspective. Where larger pixels are used, a smaller field of view is achieved because of a smaller diffraction angle. In addition to this, as planar waves are used in fourier holograms, the correct projection, and propagation between layers (as per e.g. the angular spectrum method [73], also applicable for occlusions), is instead modelled as orthogonal rather than perspective, where other optics are not added to the system or the illumination and diffraction approximation are not changed to spherical waves [182, 183]. However, as the approximation to planar waves is primarily a convenience, we do not have to produce resulting imagery with an orthogonal projection [45] even where the propagation itself is orthogonal. Assuming a complete orthogonal system can itself be beneficial to the hologram calculation [94, 71].

Of course, in practice it is not yet commonly viable to produce digital diffraction grating devices of such dense resolutions and reducing the computation required is still desirable. This is where sample-reduction display techniques, such as lightfield and other methods of holography benefit - which both can also make use of the existing featuresets in production renderers.

### 5.2.4 Holographic Field Sampling Reduction

Given the significant density required for the display to be effective via such lightfield and holographic stereogram methods, a large body of research has focused on reducing the sampling requirements of holography. One of the outcomes of this is of course image/plane based holograms - but where the viewer is concerned with multiple views, which naturally corresponds to displays of a larger area and consequently a significant range of input views, there are a selection of techniques available.

#### Phase Added Stereograms

Adding a focus cue and extending the depth of field of a holographic stereogram can be achieved by ensuring each view (often referred to as a sub-hologram or hogel) has phase information. By adding phase factors (see section 2.2) to the 2D parallax images used in a stereogram, depth is achieved without introducing significant additional calculations [219], and has the benefit of increasing the resolution of the image. Accurate Phase Added Stereograms (APAS) [110] and Overlap Add Stereograms [160] improve PAS by improving image quality of the resulting holograms. Other research introduces methods which improve speed with Fast Phase Added Stereograms (FPAS) [109], improves field of view with spherical wavefronts for use in near-eye displays [182], and allows methods of arbitrary viewpoints to be used to make up the elemental hologram [120, 82].

#### Hogel Free Holograms

Hogel Free Holography (HFH), Chakravarthula et al. [24], introduces a technique that by instead of displaying hogels or tiles which each represent samples of a wavefront, directly synthesises holograms with a full wavefront from the input of a lightfield render. It has the benefit of being able to accurately represent the 3D input scene in a much faster fashion than ray-tracing. This is yet to be shown in a much larger viewing panel, but such a technique may be combined with multi-SLM holography as is.

As the views packed into both HS, PAS and HFH are so dense, they utilise geometric occlusion (i.e. occlusion performed in the computer graphics renderer), rather than holo-

graphic occlusion. While this reduces holographic compute requirements, it adds additional rendering requirements, this is extremely significant when rendering complex offline scenes.

### **Multi-SLM Holographic Display**

There are many works which develop multi-SLM holograms - Coarse Integral Holographic Display (CIHD) [190] provides a good summary of the area, and itself treats each hologram tile independently - combining many tiles into a single, large eyebox “super” holographic display, the array configuration of which can be customised. This method initially uses an Angularly Tiled Layer Hologram algorithm [31], where views of a scene are sampled around an arc, this increases the field of view rather than increasing the eyebox. This work is later followed up with “Holobricks” [132], which implements CIHD in a seamless modular projector system, using multiple SLMs, though only demonstrated in 2D. CIHD and similar techniques require a significantly less dense collection of viewpoints of a scene than other HS techniques.

### **5.2.5 Multi-View via Eye Tracking**

An extremely efficient way to perform large-format multiview holography is to track the viewer’s eyes. By replicating eye positions and angles in a viewing system, whether this is correlating to pre-rendered offline camera positions or updating positions in a real-time engine - it ensures that at any given time, only 2 exact viewing positions and associated holograms, are generated and/or displayed.

In a large-scale display system, such as a multi-SLM display, the viewer will view only a fraction of the total display area at any given time. Naively, a hologram for each SLM would be generated and illuminated, but a display which utilises eye-tracking would only be required to generate appropriate holograms for the two eyebox positions of the viewer - this is shown in figure 5.4.

In these larger display systems it is likely that multiple viewers may be present at once, and the eye tracking system requirements consequently begin to become more complex, but still deliver significant computational savings for the holographic panel.

To note, there is a distinct difference between tracking eye positions in a larger replay field and gaze-contingent varifocal displays which make use of near-eye gaze tracking to either provide focal changes to the display system [57] or to provide foveated rendering [211]. Further discussion on the use of near-eye eye tracking in CGH is provided in Section 6.2.1.

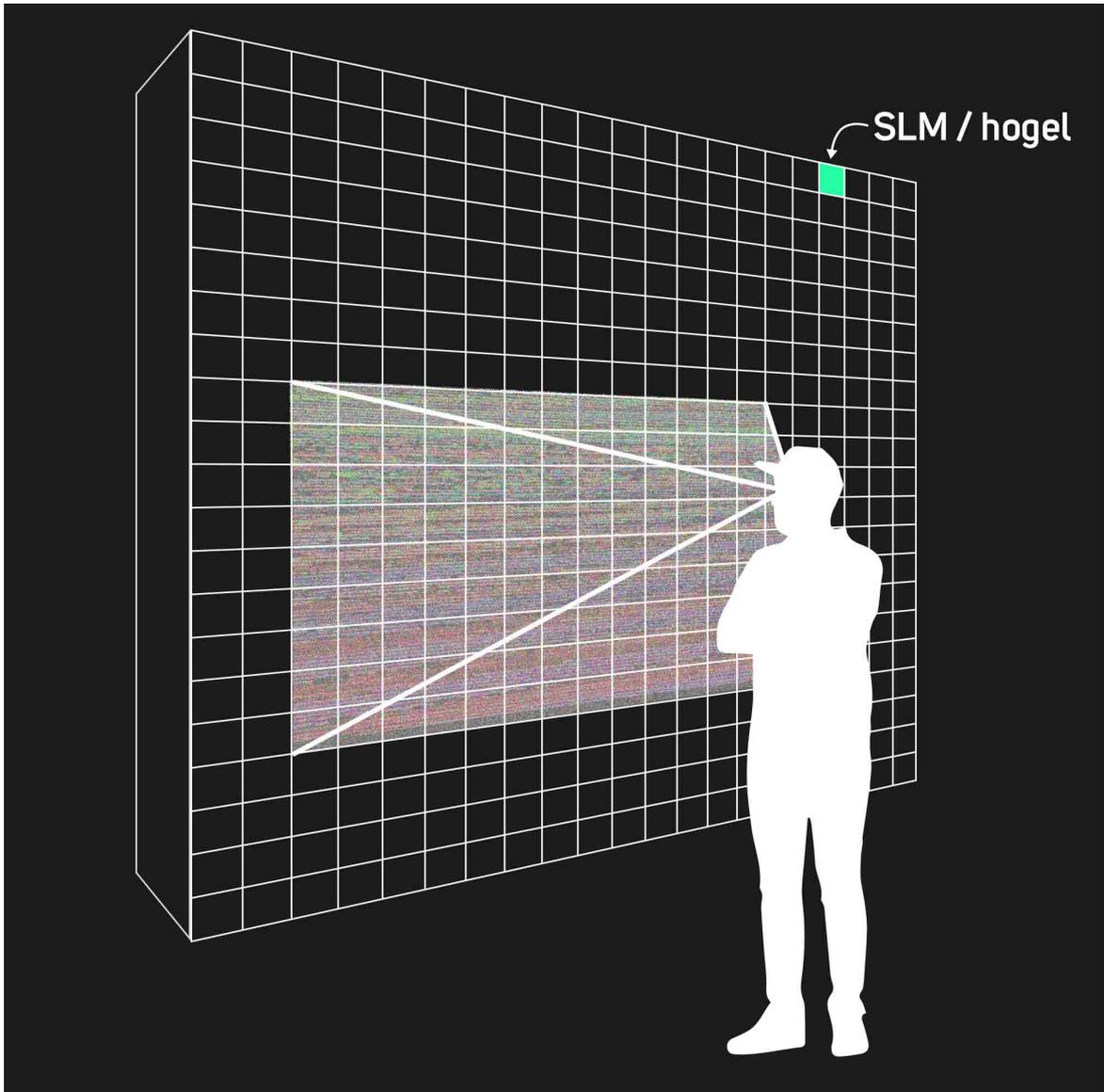


Fig. 5.4 A large-scale display may utilise eye tracking of the viewer in order to only require compute for a fraction of the given display area. Here, most of the slm devices can remain blank, whilst the active area displays imagery which propagates to the viewer eyeboxes. This depicts an enlarged version of the "Holobricks" system [132].

### 5.2.6 Multi-plane with Multi-view

Across much of the literature referenced in this chapter so far, multi-plane holography (where the scene is sliced into many depth layers, i.e. image/layer based) is often quickly dismissed as unsuitable as it does not allow continuous focus or accurate occlusion. As part of the sampling reduction of the hologram, depth discontinuities are introduced, and the wavefront produced by the discrete components is not as densely sampled as one produced by a full wavefield.

Even with these issues, there are benefits over the use of typical Holographic Stereograms, which suffer from the same spatio-angular resolution trade-off as in any lightfield display, where a larger hogel size leads to a decreased total spatial resolution. [138, 160]

Multi-plane holography is often utilised in multi-SLM display- as multi-plane holograms are an ideal candidate for large multi-view holographic panels due to their lower compute requirements. The hogels are often placed along an arc in order to reduce the artefacts introduced by occluded areas of layers that become visible when viewed off-axis [190] as well as increase the FOV where the pixel pitch of a given SLM is not yet small enough to provide this diffraction angle natively. Though, as explored in section 4.4.1, with the introduction of the occlusion synthesis technique, a viewer off-axis will usually observe the layered representation of a scene break apart, revealing gaps where objects have been split across layers, and gaps behind objects when they move in parallax. This issue is also highlighted by Chakravarthula et al.[24] when proposing hogel-free holography. However, with sufficient layers allocated for a given scene, and the occlusion synthesis technique (section 4.4.1) applied, this chapter proposes that the concerns raised may be far less perceptible.

Elemental holograms then, with sufficient occlusion data, may be tiled in a planar fashion to increase viewing area, as shown by Holobricks [132], and as described by Halle [78] without being plagued by the artefacts detailed in the literature. This will be developed by this chapter.

Even where there are not enough layers to represent a scene continuously, it is accepted that this approximation of depths is one of the controls by which to reduce the compute requirements of a scene. Where computation is less of a concern, the full RGBZ hologram may be computed in its place, with each RGBZ hologram forming a hogel. "Computer holography acceleration algorithms and hardware implementations" by Tomoyoshi Ito and Tomoyoshi Shimobaba [186] also suggest this combination of RGBZ and multiview as "promising in terms of computational cost, image quality, and viewing angle.

## 5.3 Discrete Hybrid Holography

As explored in the previous section, there are numerous techniques that allow the use of multi-view imagery to form a three dimensional image. Outside of utilising a holographic raytracer, there is always a need to sparsely sample a source scene, whether this is for multi-view acquisition from real world scenes, rendering for a lightfield display, or forming new interpolated synthetic views. These can be fed into a multitude of methods which generate holograms - the challenge for which comes with having the ability to successfully balance the rendering/compute burden with the output quality of the hologram in a flexible way. This is yet to be addressed.

A multi-view technique is required that can not only deliver a final experience to the viewer at production quality, but also provide incremental versions in preview quality to an artist or director working on media - as well as have a route to be suitable for real-time applications. A good multi-view experience should not only provide good quality imagery to the viewer across a large viewing area, potentially across a display the size of a wall, but also leverage a range of controls to appropriately balance compute with quality to prevent hindering the creative process.

This section proposes the “Discrete Hybrid Holography” method, which utilises traditional rendering techniques to create a multi-perspective view-sampled hologram for a fraction of the computation of a holographic renderer. The same high-quality viewing experience is delivered in each perspective view when presenting a fully featured multi-plane hologram as in a single perspective - but here is discretely sampled across a number of tiles/hogels and layers. Tiling in both X and Y via hogels and Z via depth planes results in the compute task being sensible to manage, parallelisable and scalable with desired quality and physical display size requirements. An outline of the proposed holographic generation system is presented in figure 5.5, where multiple holograms are tiled together from multiple renders from different positions - the viewer can then move between these tiles to observe an apparently continuous sampling of the scene. In comparison to a lightfield display, here the viewer does not need to be observing multiple tiles at once for defocus to occur.

Discrete Hybrid Holography avoids a number of the drawbacks of typical multi-plane holography by utilising the works presented in the earlier chapters, and as explained in section 5.2.6. These drawbacks most significantly include incorrect parallax effects owing to occlusion, solved by application of occlusion synthesis (section 4.4.1), and lack of parallax in any view-dependent effects, solved by providing depth information for reflections/refractions (section 4.3.2).

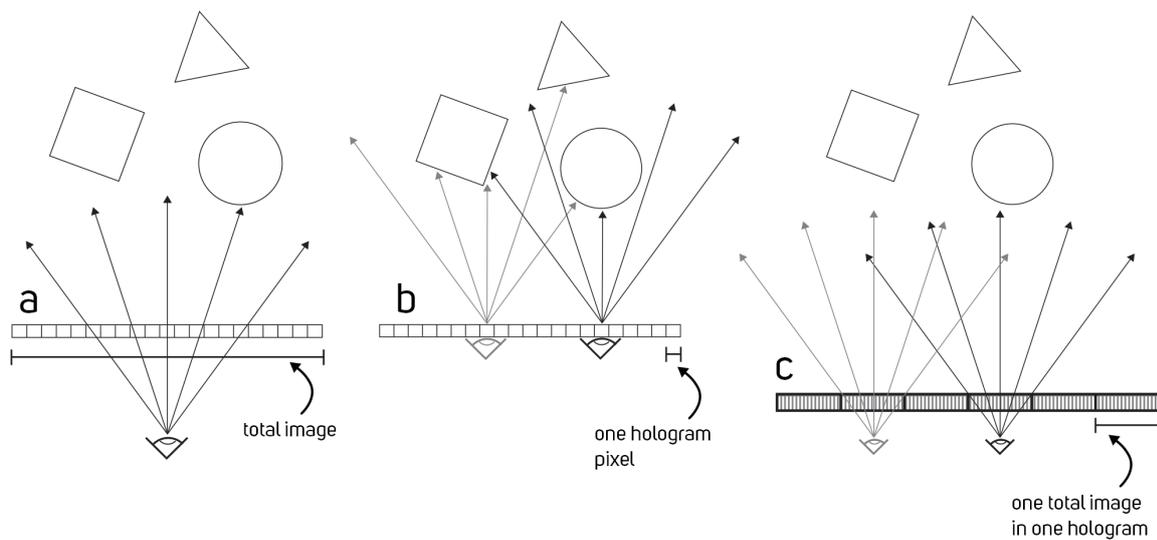


Fig. 5.5 a) Represents a traditional renderer, where pixels samples a segment of the scene, and (b) represents a holographic renderer, where each pixel samples the entire scene. The proposed hybrid approach (c) involves effectively placing a render perspective of (a) into each (enlarged) sample of (b) according to the appropriate X,Y position. This does not provide a truly continuous perspective, but can be controlled to define the number of views required for it to appear so as the viewer moves between tiles. Providing a discrete sampling of (b), whilst utilising the much faster, practical and advanced existing renderers.

### 5.3.1 Hogel Layout Design & Holographic Computation Algorithm

Forming a grid of independent holograms aims to replicate the process of tiling physical SLM devices - where each sub-hologram is significant enough in resolution in order to maximise the benefits of using image/layer based holography. Most existing approaches of forming a multi-view hologram of individual tiles, or hogels, requires a dense sampling of views as when the eye is in between views, artefacts appear in the image related to occlusion and missing image information -

"Results of research on CGH generated by multiview images [82] suggested that if the size is about 1 mm, observers do not feel a jerky motion of reconstructed images." [92]

However, even when formed from 2.5D source imagery, using the occlusion synthesis technique to form holograms as shown in section 4.4.1 can result in the successful viewing of holograms with a much larger hogel tile size using enough depth layers. The presented method is suitable for both dividing up a high resolution SLM into subsections, or for tiling multiple physical SLM devices in an optical setup.

#### General Algorithm Overview

Imagery for each layer in each tile is presented to the algorithm (as per the HQH standard proposed in section 3.5.6, with further information regarding views in 5.3.5), and occlusion synthesis can occur where required. From here, the layer stack can be brought into frequency space using the hologram generation algorithm of choice. In the example presented here, FFTs are used with random phase for each layer. An occlusion algorithm can then be used, here using the "ping-pong" algorithm [93]. An example of the input imagery and processing steps are given in figure 5.6.

In order to have the correct parallax within the tile, the hologram must be generated with the correct centre position. As a hologram with random phase contains all of the image information spread across the image (e.g. if half of the SLM was covered, the entire scene would still be visible, just at a lower resolution bandwidth), it can be cropped to represent that scene at a given viewing perspective. If a hologram of a given tile is generated to be the total size of the large format holographic display, this crop allows the parallax to be continuously viewed whilst ensuring that the tile shows the correct spatial information. However, generating large resolution holograms for them to only be cropped is a waste of computation, and instead each hologram may have its centre position adjusted via use of Zernike Polynomials. The second and third Zernike Polynomials, tilt in X and Y, will translate the optical axis (and therefore the camera position) of each holographic image and allow it to be positioned correctly in the total tile grid. Each layer hologram is generated at

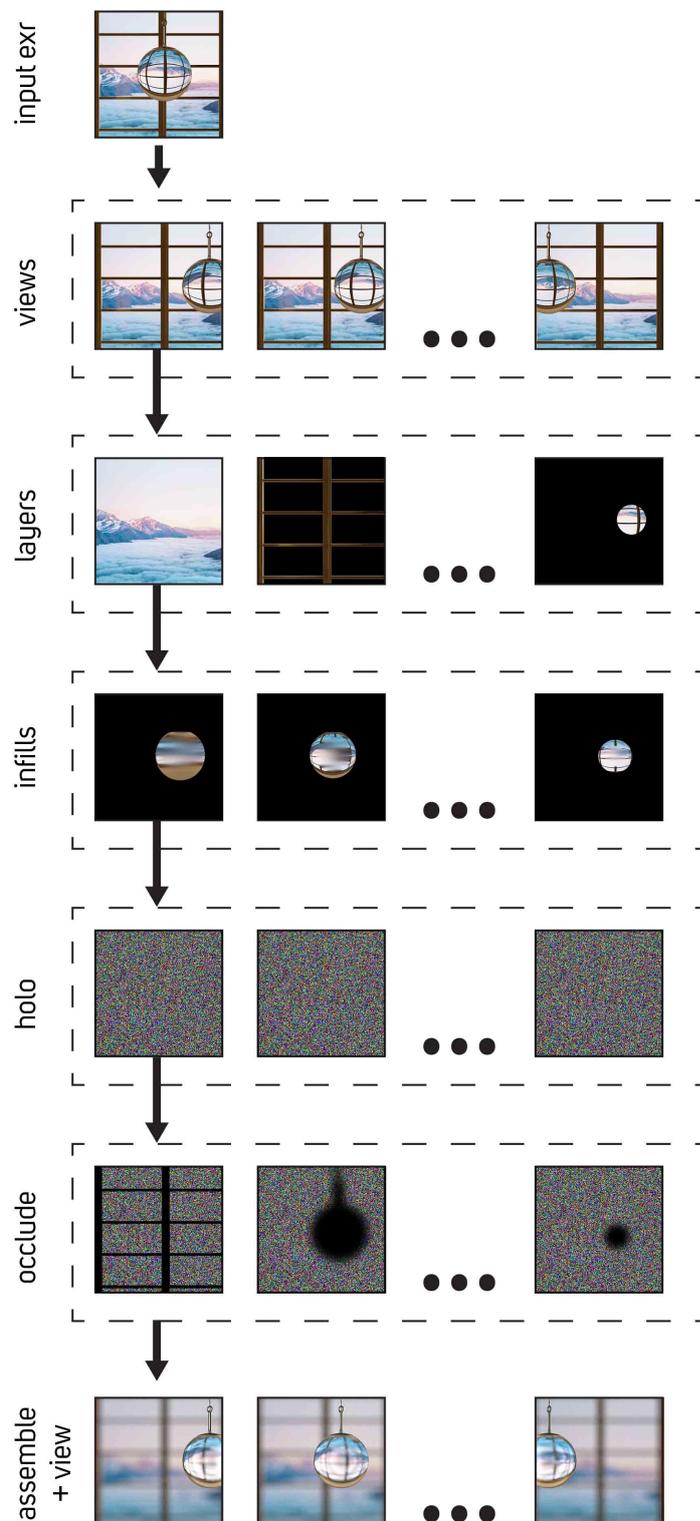


Fig. 5.6 The steps of processing an image stack undergo in order to be tiled as a complete hologram. An input image is broken into the constituent views and layers, occlusion synthesis occurs alongside hologram generation, and the resulting holograms can be composited and tiled for display.

optical infinity, and its defocus Zernike (see section 2.2) is applied at the centre position for the entire display relative to the tile position. This is shown in figure 5.7.

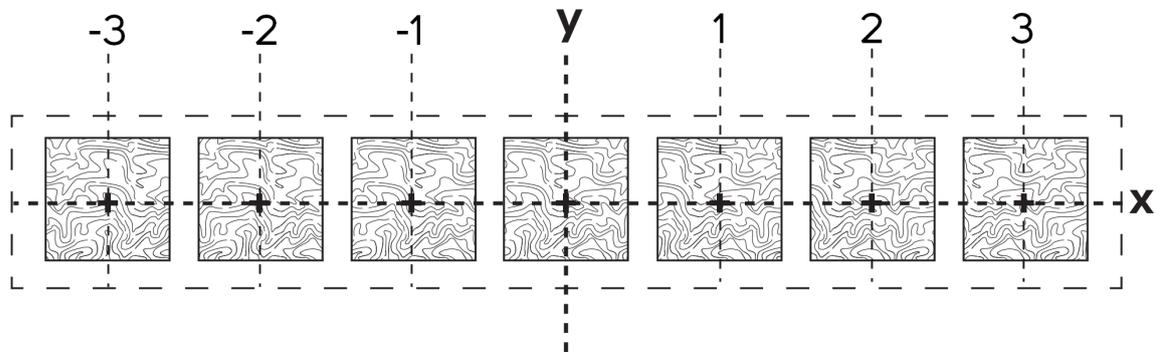


Fig. 5.7 Tiles are placed in a local co-ordinate space, each tile is 1 unit wide and is translated around the origin to their intended position, which ensures that parallax between tiles is correct. Altering Zernike Polynomials is much more efficient than generating a hologram of the total display resolution which will then be cropped to a sub-hologram/hogel tile.

With each tile generated with the correct centre positions, they may all be displayed - and the viewing experience will be correct.

### Design Factors

“Clearly, though, the stereogram designer has a wide latitude in deciding how accurately the hologram should approximate a continuous three-dimensional image.” Halle, 1991 [76]

There are a number of design decisions to make when forming such a hologram, each of which may vary the viewing experience, the variables are as follows:

- View Density
- Layer Density
- Rendering vs Hologram Resolutions
- Occlusion Requirements

The factor at the forefront of this is view density, as it is known that with enough views, a lightfield is created in which there is effectively continuous handover between provided render perspectives. These are geometrically occluded at the render source, meaning holographic occlusion, and any associated layer synthesis to make this acceptable, is not

required. However, this is burdensome on the compute process, and makes little sense to be in an image-based process, and may as well be processed by a fully native holographic-renderer. Making the most of a sparse view tiling ensures efficiency, but without considering occlusion when rendering or processing layers, the viewer sees additional occlusion artefacts for each new view - this is shown in figure 5.8.

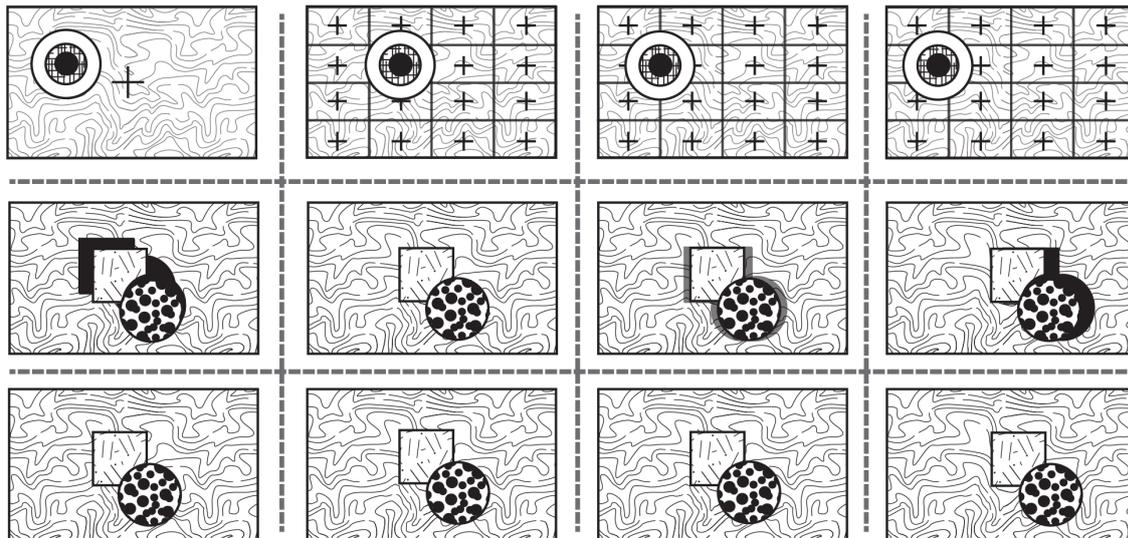


Fig. 5.8 Comparing eye positions within the eyebox across single-view and multi-view holograms - with the resulting scene perception for the viewer with and without occlusion synthesis applied to tiles

This is the only artefact that appears when viewing objects that are diffuse, and the benefit here being that with an increased eyebox size, objects can appear to move in parallax significantly further in relation to the viewer. The real bonus benefit of multi-view though of course is enabling view-dependent effects, meaning scene-global visual changes such as a perspective transform, or viewer angle dependent changes such as varying reflection or refraction via surfaces.

### 5.3.2 View Handover & View Dependant Effects

A sufficiently continuous perception of the 3D scene across a large viewing area is reliant on any “view-dependent” effects to be carefully considered. Given that parallax (also known as motion perspective) is significant for depth perception [42, 117], it is crucial that it is properly accounted for. Given that in layer-based holography the imagery formed is diffuse, new input views must be provided, with a sufficient amount of layers, to correctly allow

correct source scene perspective across views. Reflective and refractive passes may be added to cover additional cases.

### **Diffuse Perspective**

If the number of views and layers composited to form the hologram are not dense enough, the perspective of the scene will not appear to be continuous to the viewer, and will ghost/fade in as the viewer changes their viewing position. If layers are dense, there may be enough parallax across continuous surfaces or edges to provide enough of a correct handover between sparse views. If views are sufficiently dense, only 1 layer will be required for continuous viewing, but the layer will not provide the correct depth focal cues to the viewer without additional techniques (such as gaze-contingent blur simulation). If views and layers are dense, compute is higher and quality is higher, but it is likely preferable to increase layer count before increasing view count in most cases, remembering that parallax on layers ensures they continue to appear in the correct relative position. With a lower number of views, each view size should be sampled to fit into a pupil width - so to not sample the scene more than is required. A comparison is shown in figure 5.9.

### **Reflection Sampling**

As per Section 4.3.2 which explores depth passes for reflections, in order to accurately represent a reflection in a hologram, the pass must be placed correctly in space using depths gathered from the renderer. When this is correctly used, this has the additional benefit of ensuring that objects viewed in reflections also have the correct parallax. This means as the viewer observes a scene which contains for example a mirror, from two different perspective holograms, the parallax of objects viewed via the mirror will be correct, provided there are sufficient depth layers allocated. More complex surfaces which are reflective may require more depth layers to ensure that the parallax has improved perception. These more complex scenarios are outlined in section 4.3.2, but one such example of this is where a reflective surface contains another reflection within it - this is shown in figure 5.10 via two mirrors. The first set of views show a reflection of the bookshelf with the dinosaur figure, and the second set places a mirror in front of the bookshelf, showing a new reflection back to the octopus.

As reflections are viewed through an object (e.g. a reflection appears to be within the bounds of a surface), occlusion is an important consideration. Successful representation of reflections via a surface must be restricted to only that surface, even when reflections will have independent parallax to the surface. Reflections may also have to be occluded by other

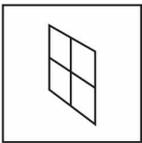
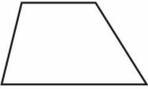
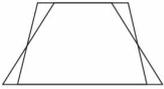
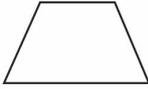
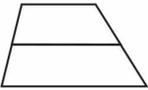
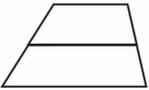
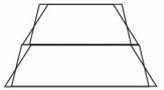
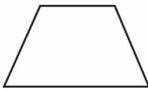
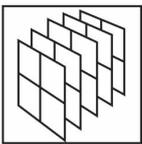
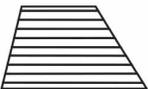
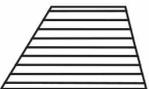
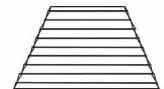
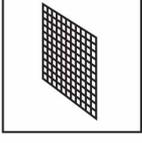
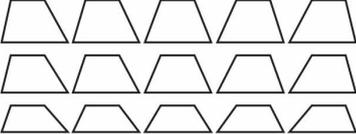
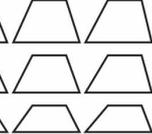
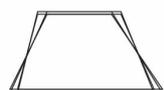
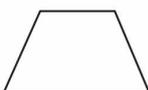
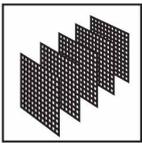
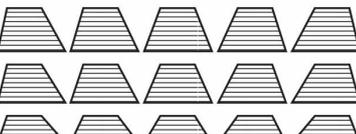
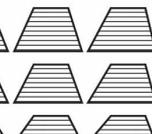
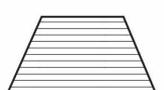
	density visualisation	left	views	right	perceived view	ideal view
1 layer low num views						
low num layers low num views						
dense layers low num views					a 	
low num layers dense views						
dense layers dense views					b 	

Fig. 5.9 A comparison of many view/layer combinations of a simple (flat) surface, which presents a challenge to perspective continuity. The perceived results which contain the least error are indicated by (a) dense layers and low number of views, and (b) dense layers and dense views.

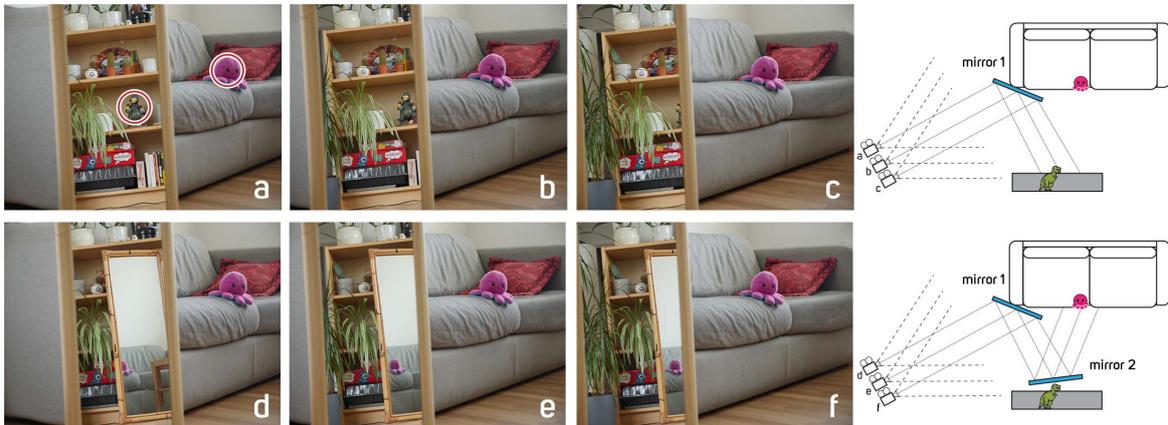


Fig. 5.10 Camera views a, b and c show the scene containing only one mirror, but the scene in camera views d, e and f now contains an additional reflection level with the second mirror. Each mirror acts as a new window into the scene which must have more layers assigned in order to ensure the correct parallax is preserved across the views. The dinosaur and octopus are circled in (a), and the diagrams on the right show the layout of the scene shown in the photographs.

objects visible in the reflection - and a reflective object will need to be treated as its own occlusion layer stack.

As shown in figure 5.11, when a viewer moves within the eyebox between multiple views, and objects move with parallax in each view, without knowledge of the parts of the scene that are revealed, the same issue with occlusion artefacts as discussed in section 4.4.2 and described above arises. The occlusion synthesis technique can be applied here, and of particular benefit may be an infilling algorithm that makes use of knowledge of hidden regions via other render views in order to construct a continuous view of the scene. When this is achieved, a viewer can correctly focus on an object via a reflection whilst moving through the eyebox with an artefact free experience.

### Refraction Sampling

In a similar fashion to reflections, refractions are also “surface bound”, and move with parallax according to the depth at which the object refracted is in focus when the viewer moves across views - locked to the confines of the refractive object, such as a glass of water or a lens. The same occlusion performed on the reflection pass can be performed on the refraction pass, and with sufficient depth layers, the perceived perspective change should be sufficiently continuous.

Without enough layers, there may be significant gaps which disrupt the parallax effect, and also disrupt any perspective changes when moving to a new discretely sampled render



Fig. 5.11 Occlusions within reflections must be taken into account - occluded data can be used from nearby views in order to infill missing information from a given view. Here, the regions marked in black (b and e) show sections of the scene which would be missing as the viewer moves between the two neighbouring sparse views for a single reflection (a and c) and for a reflection within a reflection (d and f).

view. Scenes that contain refractive and reflective objects (perhaps excluding planar surfaces if they are parallel with the hologram plane, such as mirrors or windows, in some cases) will require significantly more layers to accurately portray realistic lighting effects. Though, even outside of these cases, if objects with view-dependent shaders are present in the scene, many more layers will be required, but this may be balanced somewhat with a denser number of views. This denser number of views will also ensure that more accurate view specific effects are present, as they originate from the original scene renderer, rather than being approximated via parallax and depth layers. The diffuse holograms do not necessarily have to be as dense as the key view-dependent passes, and may be sampled less frequently than reflective and refractive render inputs. An example configuration of varying view densities is diagrammed in figure 5.12, in this case with an uniformly increased view density in the reflections, and a selective density increase in the refractions to provide better view sampling for view-dependant effects.

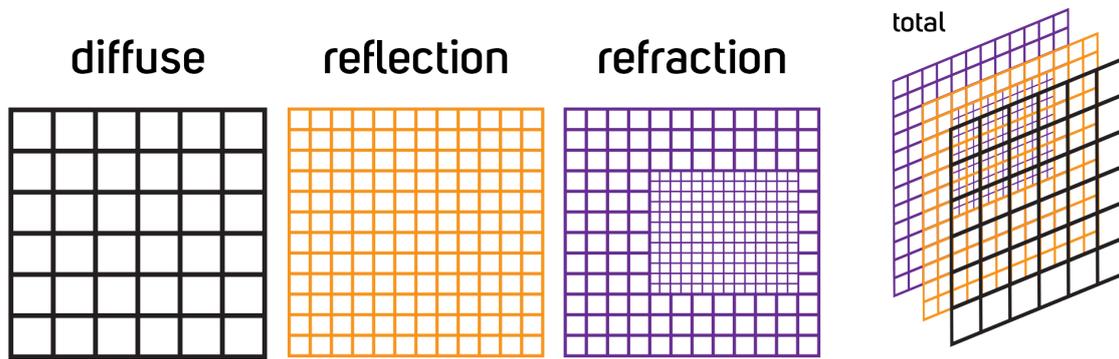


Fig. 5.12 Differing view density sampling is an excellent control for the allocation of computation where it is required. Example view tiling densities are shown here. A sparser diffuse set of views is used, with a denser, but uniform reflection set. The refraction set is denser yet, but only a select region where additional detail is required. These then can be composited into the total hologram for display.

### 5.3.3 Compute Balancing

After a certain threshold, there is sufficient information in the scene to be able to properly encode a continuous 3D view. The issue with this threshold is that it is continuously changing itself - whether on a shot-by-shot basis or even within the length of a given shot or camera move (or even continuously, such as in a first or third person view in a game). This means that choices need to be made on how to assign the number of views and layers fairly frequently - and whether continuous 3D perception is prioritised over a lower compute burden and a more discretised view of the scene.

#### Layer Density & View Density

If the eyebox is significantly large, objects are significantly close to the viewer, or if surfaces are continuous and parallel to the hologram plane (where they may cross many depth layers) - many more layers are required in order to create a realistic viewing experience within a specific view. When a viewer may transition between two views in this scenario, a continuous experience will be had if the depth layers are assigned at the same depths across these views, as shown in figure 5.13, though there is a tolerance at which a viewer may no longer notice changes in depth of an object as they move between views. This gap between neighbouring points is said to appear to be continuous if an angular separation of 3mrad or below is between them [136, 11], which approximately subtends 0.3m at a distance of 100m. Though this makes the assumption that gaps between layers have not had occlusion synthesis style

corrections, and the density of layers may also decrease as the viewer age increases, due to loss of accommodation amplitude in the eye [55].

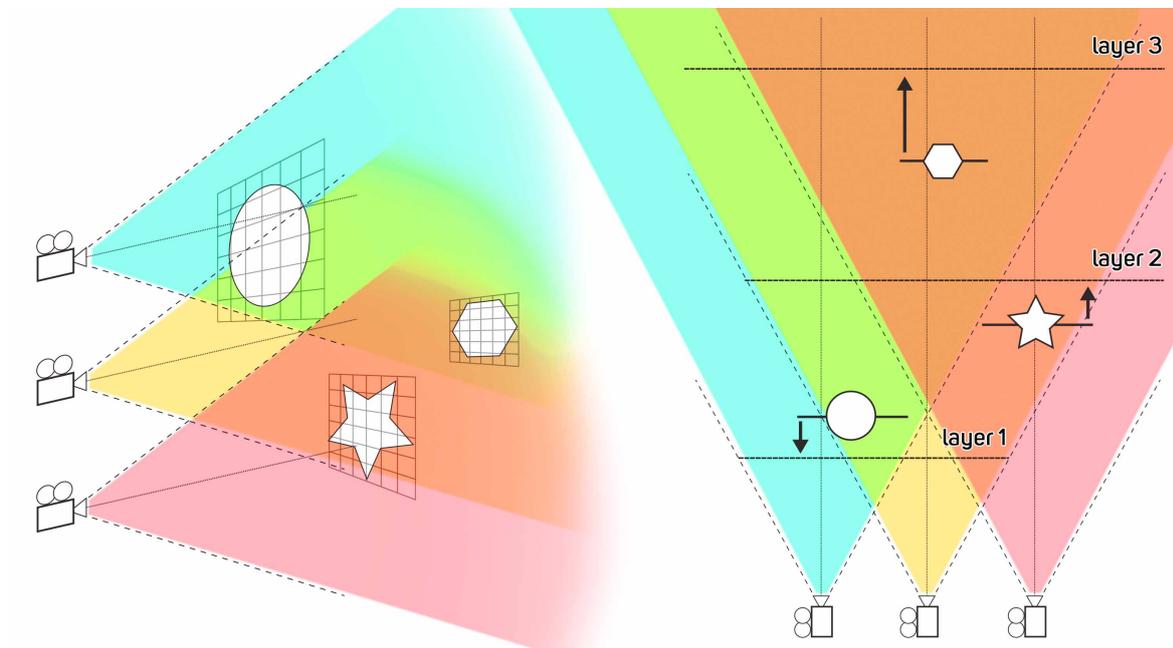


Fig. 5.13 Layers allocated must be equal across views to ensure a consistent and comfortable viewing experience. Here, Layer 1 is allocated in views 1 and 2, layer 2 is allocated across views 2 and 3, and layer 3 is allocated across all 3 views.

The success of an image-based multi-view hologram after being perceived as continuous is based on its compute saving, and as the density of these views increases, less is contributed to the overall experience while the compute time increases - when a sufficient perception of the scene can already be made even with one view. It is likely that in the region of one view per pupil width for a diffuse pass (the perceived view in figure 5.9 marked (a) gives a good approximation of this), and two views per pupil width for more heavily view-dependant features such as reflection/refraction passes of the scene is a sufficient range for a discrete multi-view display system, though optimising these variables for different quality requirements are outside of the scope of this work.

In practice, items like matte paintings (backgrounds to scenes) can be formed of a single layer and could be the same base hologram in each view, as it will be placed at optical infinity. But more complicated geometry, especially that which pushes across more depths, will have more spatially varying view-dependent effects, and require denser views/layers.

Fast processing of such sparser multi-view holograms benefits well from parallelism, where each view can be independently processed - whilst the layers within a given view may

not be generated multi-threaded to the same degree due to the inter-layer dependencies when performing occlusion.

### Render & Display Resolutions

Here a distinction must be made between the input image resolution, the SLM resolution and the end result resolution of the system. As the final image seen by the viewer is not made up of pixels but instead of connected “blobs” (where using random phase) which attempt to form the input imagery, compared in figure 5.14, the input resolution and output resolution are in different domains. Though it is certainly a requirement that the SLM resolution itself is high enough to display a sufficiently high resolution hologram, in order to minimise the effects a diminished effective resolution caused by large noise/speckle sizes. Reducing the size of the noise is done so with a higher resolution (in pixels per degree), but the input image resolution does not need to match this.

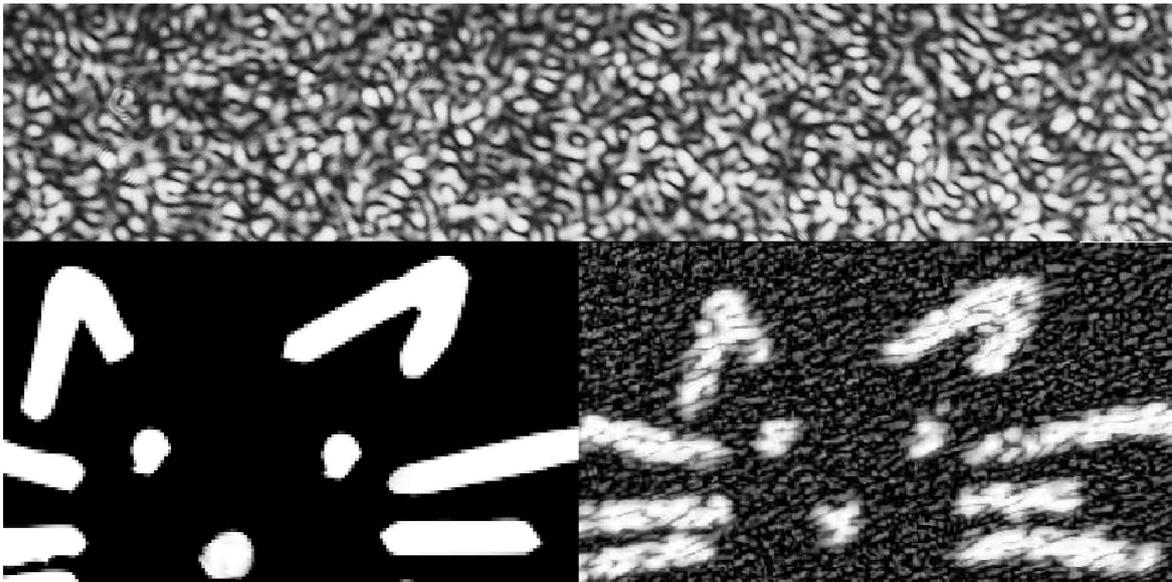


Fig. 5.14 A comparison of uniform (random phase) speckle or "blobs" (top), with how such speckle replaces the appearance of pixels in an image, shown by the low resolution cat image raster (left) and the associated hologram replay (right).

Using high resolution input imagery is also problematic for render times, so it is a viable option to render at a much lower resolution and upscale the result. This can be performed even with a basic sampling method such as nearest neighbour, and is done so in this chapter. Though the goal of an overall system is to ensure that at least 60 pixels per degree are achieved for matching human visual acuity for a given display [41, 207] (see section 5.3.6

for further details), and this is applicable to any individual view tile/hogel, whether this is a single SLM divided into multiple views, or a selection of multiple SLMs.

### 5.3.4 View Rendering & Capture

In order to gather all of the views required for a multi-view hologram, a camera rig must be created. This can be either physical or virtual, but here virtual cameras will be explored, and physical equivalents can be built, if appropriate hardware exists, according to the same specifications. From a production perspective, a combination of photogrammetry, volumetric capture, lightfield style capture systems or Neural Radiance Fields (NeRF) can be used to digitally represent real-world scenes, and this can then be rendered using the appropriate virtual cameras to match a given display specification. It is possible to somewhat convert holograms from one display system to another via scene reconstruction, but is simpler to just re-render the original source, and is not covered by this work.

#### Virtual Scene Rendering

When creating a virtual camera rig for multi-view holography, for a realistic view, it is important to know the physical attributes of the projector it will be calculated for and consequently displayed on so that the images displayed will be the correct scale and the virtual camera projection can match the display projection.

The attributes that must be considered are:

- Pixel Pitch
- Tile Resolution
- FoV (specifically angle of view, AoV)
- Number of Tiles

The number of tiles and tile resolution is content dependent and a balance of image qualities, as explored in the previous section. A physical SLM will have a resolution, and a pixel pitch - this, when combined with an optical magnification, then defines the physical size of the SLM, which matches the eyebox for the viewer, this should also be used as the camera back size. Magnifying the hologram will increase the FoV/AoV of the display, but at the cost of eyebox size, a concept referred to as *étendue* [27]. As this varies by wavelength, the smallest wavelength (blue, in an RGB projector) defines the focal length of the virtual camera across all colours, and ensures that larger wavelength image channels can be scaled in the algorithm to match the frame size of wavelengths that diffract less.

$$\text{SLM size} * \text{Magnification} = \text{Eyebox} = \text{Camera Back} \quad (5.1)$$

$$\text{FoV}_x = \frac{180\lambda \frac{\text{Resolution}}{\text{Eyebox}}}{\pi} \quad (5.2)$$

As the FoV (here shown in the  $x$  axis) and tile spacing will define the nearest depth at which object may appear to the camera between views, this value can also be calculated and used to set a near clipping plane [182], which prevents rendering of objects at a problematic distance to the camera, defining a safe rendering region shown in figure 5.15, in the viewport and in the renderer:

$$\text{Z Safety} = \tan\left(\frac{(180 - \text{FoV}_x \frac{\pi}{180})}{2}\right) * (\text{PixelPitch}_x * \text{Resolution}_x) \quad (5.3)$$

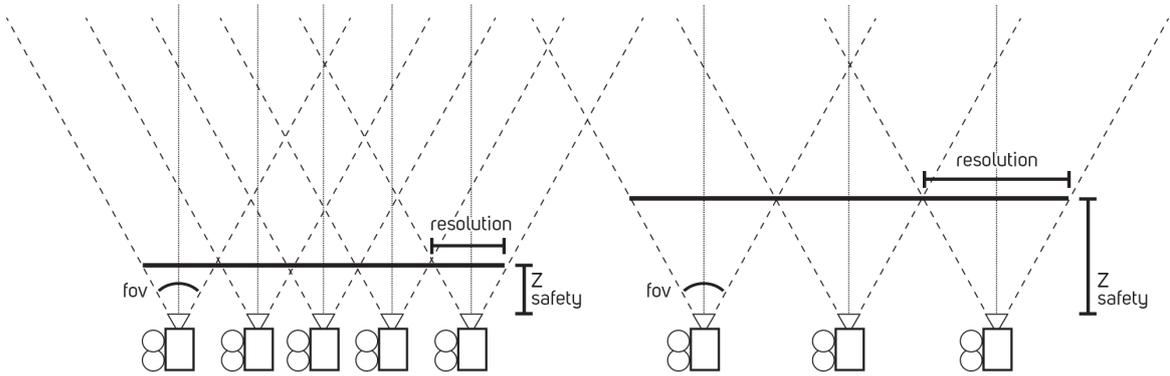


Fig. 5.15 The near clipping of each camera is defined by the Z Safety. Here, the same FoV is used, but the image resolution of each tile (whilst not increasing the pixel density) is increased on the right hand setup, which increases the Z Safety requirement - and consequently the near clipping distance that must be used in the renderer.

If the DCC in use does not have internal helper functions for calculating or setting focal length, or if physical cameras are being used, a conversion is required.

A diagram of the relationship between focal length and field of view (angle) is given in figure 5.16. The focal length can be calculated from the angle of view (AoV, though not to be confused with "Arbitrary Object Variable" used in rendering outputs) and camera back size using:

$$F = \frac{\frac{1}{2}h}{\tan\left(\frac{\text{AoV}}{2}\right)} \quad (5.4)$$

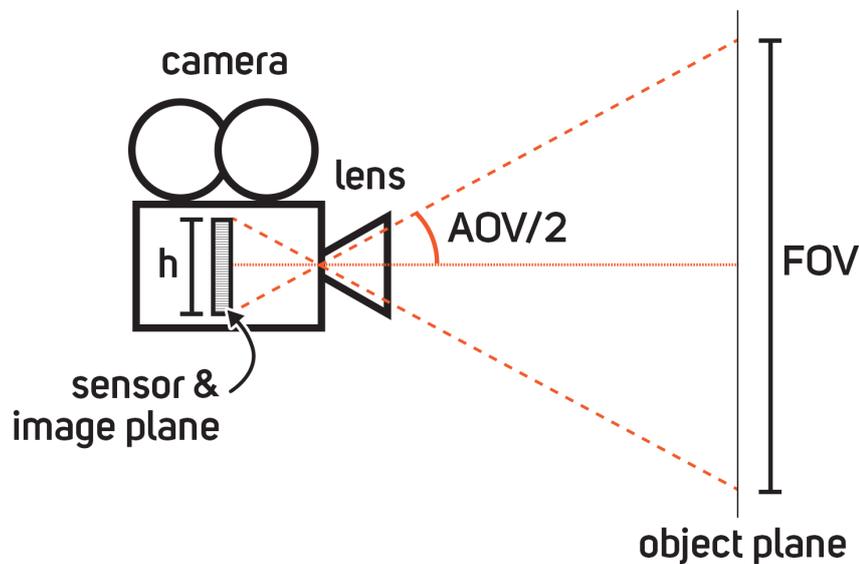


Fig. 5.16 The focal length of a lens is given by the relationship between the view of the field and the image plane, where the sensor is placed, or where the pixels are sampled in a render. This is the plane at which the hologram is also generated for a given display, and also illustrates the link between the "camera back" and the SLM in the discrete-hybrid method. Details such as film-gate are not included in this diagram.

In Autodesk Maya for example, the focal length (measured in mm) must be set, rather than the field of view - and additional unit conversions will be required as the camera back size parameter is in inches.

```
_f = (0.5 * _cameraBack) / (math.tan((pi / 360) * _AOV[1]) * 0.03937)
# .03937 = convert mm to inches, maya uses inches for camera back
```

Once these values have been defined by the physical display specification, a virtual camera rig can be spawned. In this example, a 3D scene is built in Maya, and a virtual camera rig is created using Maya's Python Scripting. The script creates a GUI which prompts the user to input the required multi-view render specifications, and will perform all of the calculations in order to make a virtual camera rig. The rig is able to be transformed and animated in 3D space as expected, whilst still being able to render a sequence of views that, when turned into holograms and tiled, will create a physically correct viewing experience. A witness camera is also provided to ensure that the artist working on the scene is able to see the full field of view provided by all independent cameras. This is shown in figure 5.17.

In this example implementation, a single camera traverses all the views across time, and renders an image sequence - each camera view position is set by the script as a keyframe. This is intended to be used with a static scene, but dynamic scenes can be animated that create

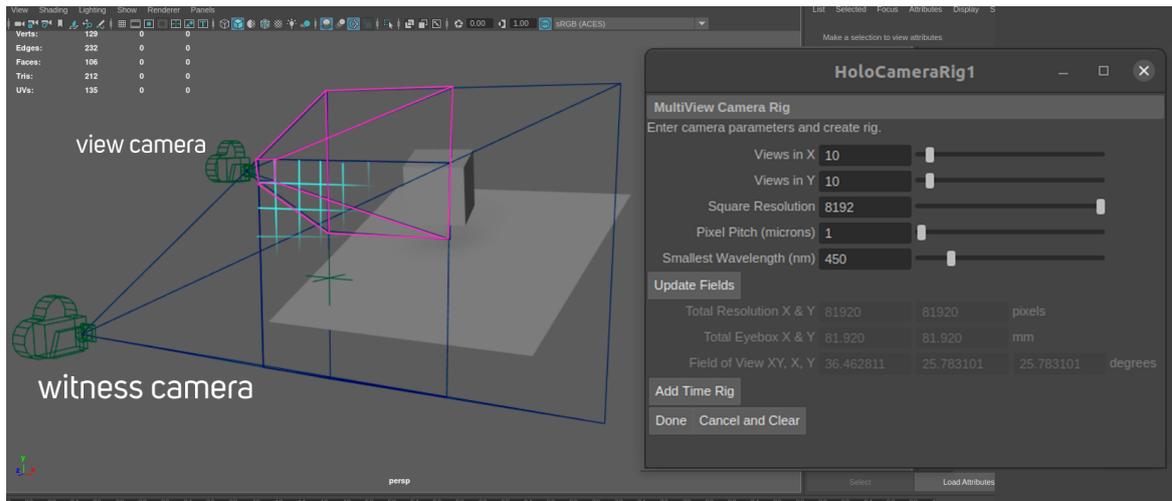


Fig. 5.17 A script generated tool in DCC package Autodesk Maya allows virtual camera rig parameters to be adjusted (or automatically set) for correct rendering for holographic generation per a given display system. The view camera captures the image (with one view frustum shown in pink) for each grid tile (shown in blue) within the total capture plane. The witness camera provides a useful frame of reference for 2D rendering of the scene.

a similar effect to a lenticular animation (like those novelty rulers that change depending on the viewer angle), where, as the viewer moves across the eyebox, the viewer sees new views that encode time into eye position, a comparison of which is given in figure 5.18. Rendering for animation across all views requires a camera that exists at each view position in the rig in any given frame - support for rendering multiple cameras per frame is supported by most production renderers.

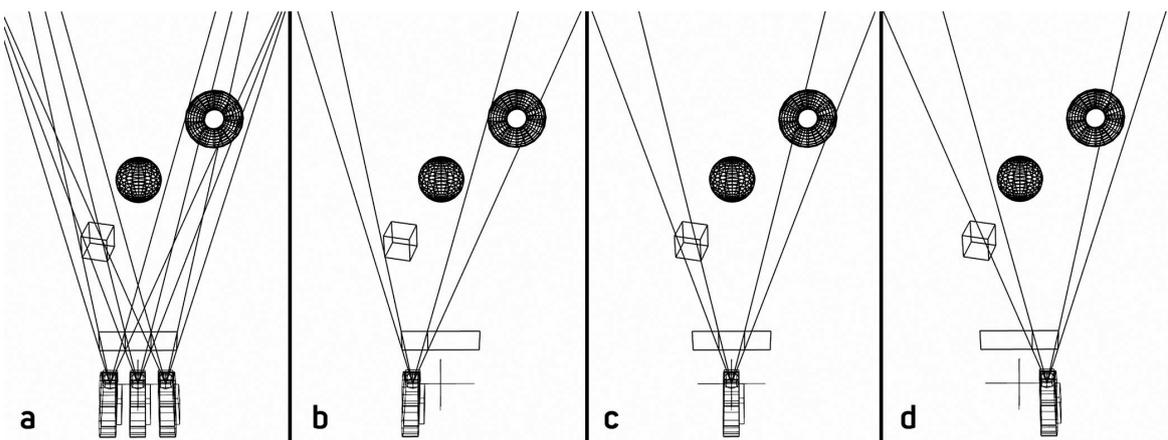


Fig. 5.18 (a) Shows all cameras present simultaneously. (b-d) Shows a single camera traversing across views per frame: positions 0, 1 and 2 respectively.

The resulting render views produced are akin to a lightfield render, but are more sparsely sampled. With the developments presented earlier in this work, a successful 3D image is perceived by only 1 render perspective, and with the addition of view tiling, a given hologram extends this to multiple perspectives - where this method only requires 2 views as a minimum, shown by figure 5.19. A multi-view hologram also benefits from the other advancements made in the previous chapters, and views are rendered using correct colour management, have real-world luminance support, contain alpha channels for use in occlusion, and reflection & refraction (plus depth) channels for physically correct defocus. The layer allocation of these holograms is content dependent and explored in sections 4.6.2 and 5.3.3, but any render view may also contain multiple layers for use with holographic occlusion algorithms.

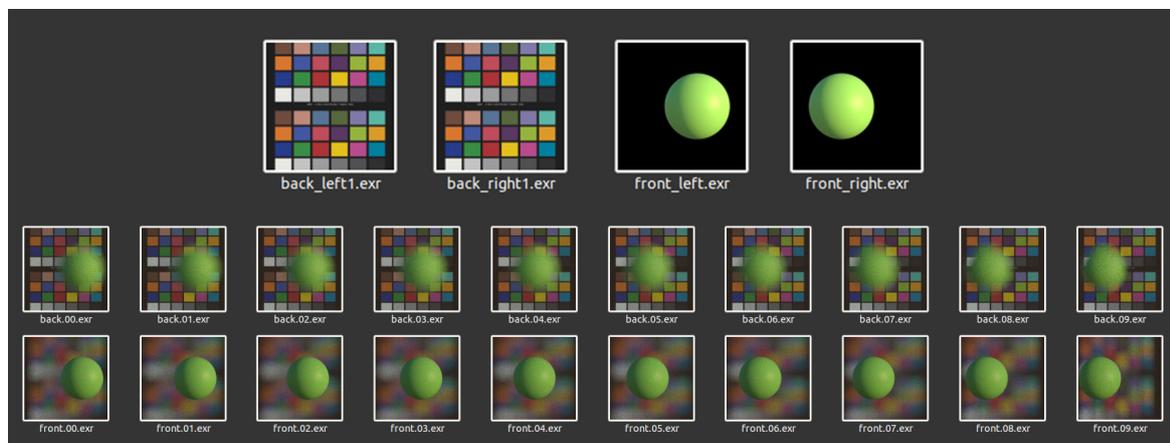


Fig. 5.19 The two input views (shown top: with background test chart and foreground green sphere) can be consequently viewed as any number of views across the eyebox - here demonstrated with bottom: 10 new viewing positions simulated, with focus back and front. (The resulting image views may be viewed as stereograms)

An issue that will occur in the real-world with multi-view holograms prepared from offline rendered materials is that these are rendered using a fixed virtual camera rig that may not necessarily match the display on which it will be viewed. This reiterates the importance of a rigid set of display standards to ensure the viewing experience is matched across devices, not just in the methods outlined here, but even whole classes of device construction - without this there will be significant issues in cross-compatibility of media. Where material is prepared with one particular display configuration in mind - it can either be re-rendered, or the views can instead be interpolated from those that already exist.

### View Interpolation

Ensuring that enough views (or the correct view positions) are present to represent the 3D scene is important to not disrupt the viewing experience. It may be the case that source content has already been rendered with a fixed number of views, or a real-world scene has been captured from a sparse sampling of camera positions. New views can be interpolated from these datasets through a selection of techniques in order to improve the spatial resolution of such a scene.

Recent work has highlighted the approaches of NeRF [146] and MultiPlane Images (MPI) [64] that allow novel view synthesis from sparse input views. These views come via a continuous representation of the scene in 3D, meaning that data such as Z depth, and view dependent effects can be preserved and acquired for display. Any method of capturing or synthesising these novel views can be utilised to create holograms, and indeed there are numerous methods that demonstrate hologram creation from lightfield style sources [34, 160, 24]. When working with inputs to the hologram generation that are only 2.5D, this is especially valuable. Full representations of the scene can also be built using only 2D inputs, similar to photogrammetry. An overview of the general process of new view synthesis is given by figure 5.20 - each new resulting view can be used as the source for a holographic view.

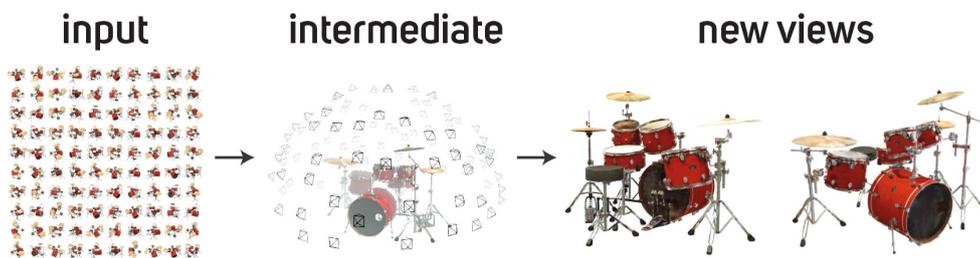


Fig. 5.20 A range of input images are used to seed the intermediate form of the scene representation. This intermediate stage may take the form of a radiance field (NeRF) or the more conventional form of generated geometry akin to photogrammetry, without view dependant effects included. The intermediate may then be used to generate any number of new views, which can then be used for hologram generation. Adapted from Mildenhall et al. 2021 [146].

Importantly, these methods can be used to create the discrete views required for the discrete hybrid hologram generation as described in this chapter, and will be more computationally effective than generating a full-size hologram from all of the available views. By providing control over the view and layer density sampling, the artist has full direction over the display result when weighed up against compute times for given shots.

Additionally, these same methods of view interpolation are beneficial for occlusion infilling synthesis, allowing a “local” multiview, as described in section 4.4.2.

### 5.3.5 View Storage

Given that a significant number of views may be produced for holographic material per frame in a given render sequence, it’s important to handle these appropriately. Fortunately, much of the computer graphics pipeline is set up to handle working with stereo renders from the previous popularity of 3D film production which saw use of polarising glasses in cinemas and in consumer TV sets, and this can be used as a foundation of storage and handling of multiple unique images per frame.

OpenEXR, much like its support of multi-channel images, also supports containing multiple views. These, when stereo, may be referred to as “SXR” (Stereo [e]Xtended Range) files, and when containing any number of views, may use the extension “MXR” for Multi-view eXR. The benefits of using this format for storing views for holographic display generation are that it supports all previously utilised features, but now within multiple views, and is exactly what is required for sensible view handling as was noted in the HQH standard outline. A production renderer could indeed merge all of these views into a single file written to disk, but where that may not be possible, this can easily be added using OpenEXR, either directly or via OpenImageIO (OIIO), using a dedicated tool which can be executed after render. There are many routes for implementation, using C, Python or even oiio tool via a bash script:

```
oiio tool "view1.exr" "view2.exr" --siappendall -o out.exr
```

Producing a single file for these multi-view renders ensures potentially thousands of views are simple to manage - and setting a “default view” within these EXR files means a representative view of the render can be edited in applications which do not support multi-view. In supported DCCs, these files can be opened and manipulated as expected if required as shown in figure 5.21.

Note that while multi-view is supported in EXR, users should utilise “Multi-part” instead, which is the same but allows a complete header to be stored for that image, and allows any “combination of tiles, scanlines, deep tiles and/or deep scanlines” [157]. This will be the more relevant feature for use in storing any combination of image elements for holographic display, and is required for storing data such as a camera transform in the header of each view - in place of having an assumed view ordering. There are also benefits in that it will allow deep volumes to be stored in a multi-part file independently of the rest of the scene, making

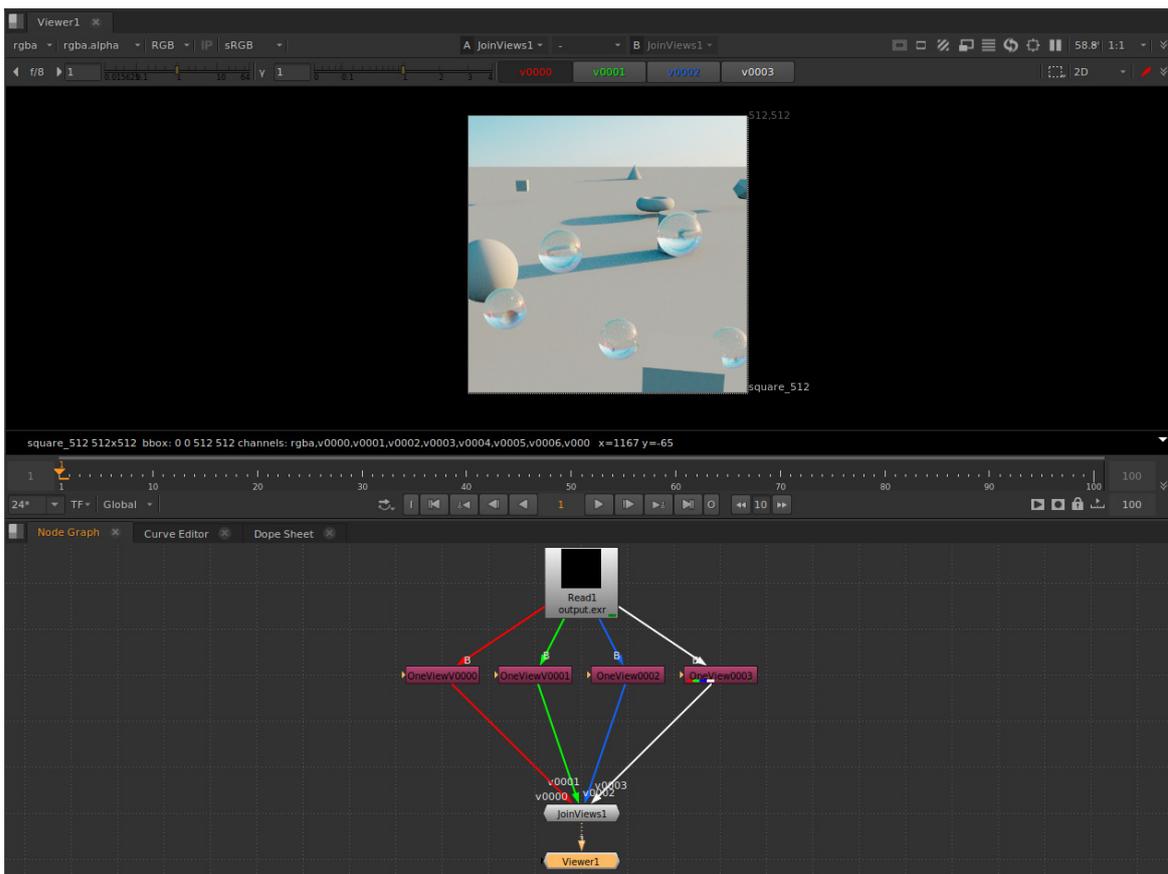


Fig. 5.21 The Foundry's Nuke providing support for editing multi-view imagery. Here the merged views from a single EXR can be accessed and edited, before being again merged for hologram generation.

any layer slicing easier to perform. Each image part may contain any number of image or data channels appropriate to that specific view, such as motion vectors or refraction/reflection passes, and can still be adjusted in the composite prior to being passed to the hologram generation engine.

### 5.3.6 Considerations for Physical Display

In a real display system built to provide a multi-view holographic experience, there are a number of considerations to balance. Whether it is determining the number of views themselves, or determining how each of those views are displayed or aligned - various additional factors need to be explored.

#### Optical System Configuration

A significant factor in the good performance of any display is of course the overall display resolution itself. In multi-view CGH, the overall experience is influenced by:

- The pixel count resolution of the input imagery
- The pixel count resolution of the hologram, and subsequent SLM(s) it is displayed on
- The pixel density (pixel pitch) of the SLM(s), in microns
- The pixels per degree (PPD) of the final image at the eye
- The noise in the field

Aspects such as the field noise (via the algorithm design or other calculation/display errors), the pixel density and the resolution of the SLM displaying the hologram will tend to be fixed in a given display system, and the PPD of the display will also be set by the optical design.

Given that human visual acuity is 60 pixels per degree (PPD) with 20/20 vision [41, 207], as long as this angular requirement is met, the spatial component may be varied to create a larger eyebox with additional SLM devices. Though it is important to note that it is common for those with corrected vision to have better visual acuity than 20/20, and human acuity ranges between 60-200PPD (0.3-1 arc-minutes) in practice.

### Trade-offs in Projector Designs

There are a number of choices to be made when configuring a design. Here constraints will typically be tied to the hardware choices, whether this is the SLM pixel pitch and resolution, or exist in designed trade-offs, between optical magnification for eyebox size vs fov size. Outside the scope of this work is the optical design itself, whether this is formed of a waveguide or pancakes lenses for compact designs, each will introduce further constraints on the system. A comparison of simple example display configurations which focuses on the projector before any combiner are detailed in table 5.1. For most of the theoretical display systems, an ideal 1micron pixel pitch is used to simplify comparisons, where smaller pixels allow larger diffraction angles (and therefore a higher field of view). Comparisons are also made between étendue tradeoffs, where an eyebox size may be increased instead of an increased field of view. The resolution per tile is indicative of SLM resolution, in a multi-SLM setup (owing to current SLM developments reaching maximum resolutions of 8K). Display *F* is an exception, where 10x10 views are given on a single 8K resolution SLM, and provides a larger pixel pitch more indicative of SLM devices today.

Display *A* and *B* compare a system in which 2500 tiles, each of 8K resolution achieve >60PPD with an eyebox large enough to cover both eyes. Display *B* offers a larger area of head movement and potentially reduced field noise (via 127PPD), but at the cost of a lower field of view. The same eyebox size as Display *B* is shown in Display *C*, but made up of 10x10 16K resolution tiles - here a dense number of pixels is delivered per degree, but at a fraction of the required field of view (where it is generally accepted that humans can see 200-220 degrees [201, 41]). Display *D* halves the resolution per tile, and consequently the Eyebox and PPD. Display *E* halves the per resolution tile again, but uses some magnification to increase eyebox size. Display *F* shows that a system using a single SLM device, similar to those available today (in 2023), even with a high resolution and an eyebox that covers a dilated pupil, are not yet suitable for successful large-scale multi-view displays.

In order for both eyes to fit in a single eyebox created by the display, the viewer's inter-pupillary distance (IPD) plus half the pupil dilation size must be smaller than the value shown in the table. This value also defines the range of head movement. Each tile in the display may be physical, via a unique SLM, or scanned, where a single SLM may multiplex to many tiles where it refreshes suitably fast. The 60PPD per view value does not take into account field noise and cannot be 1 : 1 compared to conventional display resolutions.

It is inevitable that for any suitable high-end image quality that display systems must utilise much higher resolution and pixel density SLMs, even where algorithm improvements reduce field noise. Lower resolution display devices used in a multi-view configuration will result in low perceived resolutions and a significant proportion of field noise, remembering

Table 5.1 Comparison of Display Configurations

Display	Pixel Pitch (microns)	Resolution per Tile	Views	Magnifier	Eyebox (cm)	FOV (degrees)	~PPD (per view)
A	1	8k	50x50	5	8.2	129	64
B	1	8k	50x50	2.5	16.4	64	127
C	1	16k	10x10	1	16.4	26	635
D	1	8k	10x10	1	8.2	26	318
E	1	4k	10x10	2.5	16.4	64	64
F	4	819	10x10	1	0.819	26	32

also that the quality of the final perception of the hologram suffers when the limiting aperture is the eyebox and is closed down, an ideal display will see results that are acceptable between the range of pupil sizes 2mm-8mm [84, 174]. Making sure enough resolution is available even where the eye is constricted is critical.

### Tile Multiplexing

Requiring a unique SLM in a large scale display per tile could be a factor that significantly increases panel cost. Part of a given large-format panel design could include time-multiplexed (spatiotemporal multiplexed) views, whereby a single SLM provides holograms to a number of view tiles across time, but where the viewer perceives such as effectively instantaneous. This can be achieved either via tiling super-resolution holograms utilising tip/tilt zernikes [106], or via an arrangement of mirrors [122].

Equally this could be applied to a system where a pair of SLMs are reflected through an optical system toward a tracked viewer with eyepiece [4]. Such a system could also provide imagery to multiple viewers using a time-multiplexed setup.

Examples of both of these scenarios are outlined in figure 5.22.

Where bordering tiles are not multiplexed, care must be taken to ensure that optical path lengths are suitable so as to ensure the phase of the light between each tile does not cause unwanted interference.

### Curved Display

Where an individual viewer is in front of the display, it may be ideal to present a curved display system, as shown in figure 5.23. When multiple holograms are presented in a curve (or a curved array), the field of view of the display may be enlarged. Note that the figure presents

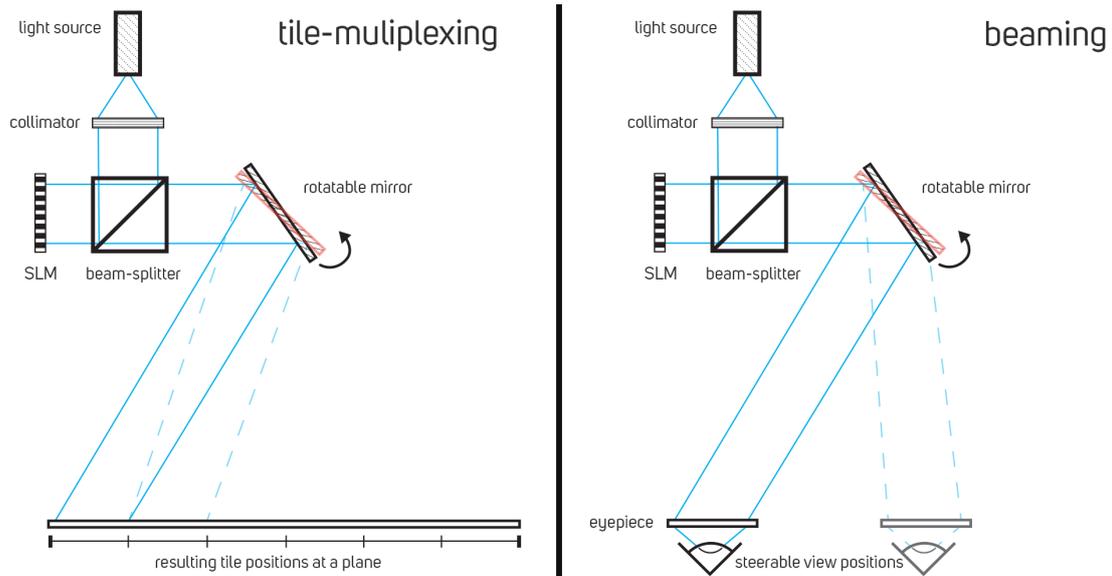


Fig. 5.22 Basic overview of the optical design of tile-steering large-format displays. Left: Tile-multiplexing, where using either a rotatable mirror or selection of mirrors, one SLM provides imagery to a range of positions in the field over time. Right: A beaming display, where the optical system steers the image toward the viewer(s). Note that a lens or other optical element is required to fill the field of view in a beaming display, but this does not necessarily have to be at the eye.

a simplified view of the tiled field of views, the diffraction angle, and consequently the field of view of a given SLM is dictated by the pixel pitch (smaller pixels diffract more). Hahn et al. [74] show how an array of SLMs may enlarge the field of view of a holographic display. This is applicable to both near-eye and large-format displays, and despite the hardware constraints required of multiple SLMs, another example is given in [224]. Rendering source imagery for a curved display could use existing 360 degree lens shaders, or be achieved simply by rotating the render camera per tile. For the approximation of a spherical arrangement, various combinations of square, pentagon and hexagon shaped hologram tiles can be used to form polyhedra.

### 5.3.7 Limitations

The ultimate goal of adding multi-view to computer generated holograms is providing superior realism for the viewing experience via continuous perspective changes that match the viewer position. Using a discrete sampling of the scene will always disrupt that continuity - and the viewing experience will be jarring for the viewer when the sampling of the scene is too limited in views and layers. However, given that the renderer required for a convincing

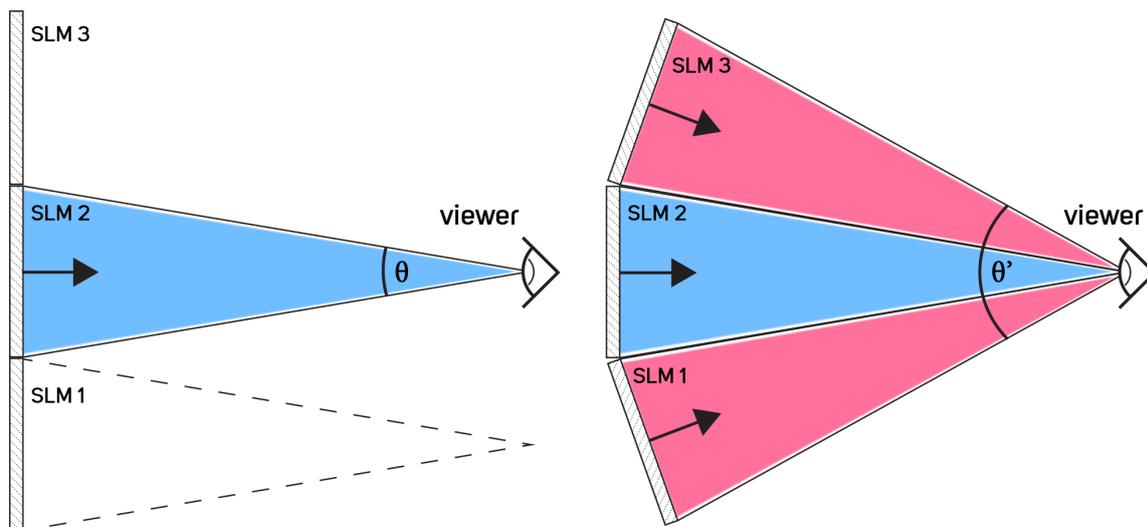


Fig. 5.23 A simplified illustration of increasing visible field of view. Left shows SLM devices with a limited diffraction angle, and consequently a small field of view,  $\theta$ , presented to the viewer, with the viewing frustum coloured blue. The viewer can only see a limited field of view of the total scene. Whereas on the right, with the same SLM devices, the circular arrangement ensures that the field of view is extended,  $\theta'$  for the viewer, with the extended frustum coloured pink.

fully continuous “native” raytraced holographic representation of a 3D scene is not yet sufficiently developed, requires extensive computation, and may even provide more detail to the viewer than is perceivable, a discrete hybrid tiling system is certainly a suitable candidate for holographic imagery production.

The complications in getting this right for a given shot lie in providing an efficient number of views and layers to the viewer given a pupil size and eyebox size. For an artist trying to view this work in (ideally) real-time, it is perhaps acceptable to track their eyes in order to very efficiently render and generate holograms for their particular viewing position in relation to the display panel. When there are many viewers, it becomes more challenging to track viewer pupils to adapt this for real-time generation. For full production quality, there is no need for this generation to happen in real-time, and it is then a matter of ensuring that there is not an over-assignment of computation regarding the number of unique views or layers - these scenes will still need (perhaps manual) optimisation to provide the experience required, and such a workflow could require lower quality cuts across varying devices, much like SDR grade passes are produced from a HDR master.

The overarching issue lies in how to successfully represent a perceptually correct scene with  $N$  number of layers, though this is not an issue of multi-view directly, but one of layer-based holography. This is especially apparent when dealing with changes in viewing

perspective of a scene, notably one with flat continuous surfaces or through complex view dependent materials such as refractions or reflections - more layers can be added to ensure parallax is correct across the surface, but this is still ultimately a discrete sampling. Replacing samplings which are perpendicular to the viewer of flat surfaces, for example, with a tilted plane [142] of a matching normal to the plane itself promises a far more efficient use of compute. This, in combination with each discrete hologram view could achieve superior effects within the overall display system, and similar could be achieved for refractive and reflective surfaces by increasing sampling rates whilst avoiding ray-tracing.

Given that the devices capable of such scale to show significant multi-view do not yet exist, it is difficult to predict what exactly will be required - but it is likely that a set of display standard brackets, like the resolution classifications for traditional panels, will be able to dictate the render and generation qualities required.

## 5.4 Multi-View Results

This chapter has introduced a method by which multi-view holograms can be generated, where a significant eyebox size can be created, with high tile resolutions and sufficient perceived quality. This presents a fully featured 3D viewing experience to the observer. The resulting holographic image contains significant view dependent cues, uses existing mature computer graphics renderers, and is integrated with well established production workflows - while using a fraction of the compute of a holographic ray-tracer.

### 5.4.1 Discussion & Results

The images presented in this section show this in practice by displaying a challenging scene built in Autodesk Maya. The scene is rendered using Maxon Redshift, with the presented multi-view virtual camera rig. Render outputs are saved into HQH standard EXR files. Each view file is then merged using the pipeline tool presented, and fed into the proposed hologram generation algorithm.

The input scene consists of a glass ball hanging close to the viewer (20cm), a window frame (3m) and a mountain view in the distance (2km). A selection of the image channels used are shown in figure 5.24, and an overview of input views vs output views is shown in figure 5.25.

The inputs are made up of only 6 layers, with 10 views. Each hologram tile is 8192 x 8192 in resolution and forms a total hologram resolution of 81,920 x 8192 pixels. Here it is simulated on a display system where each pixel on the SLM has a pitch of 1 micron, is not

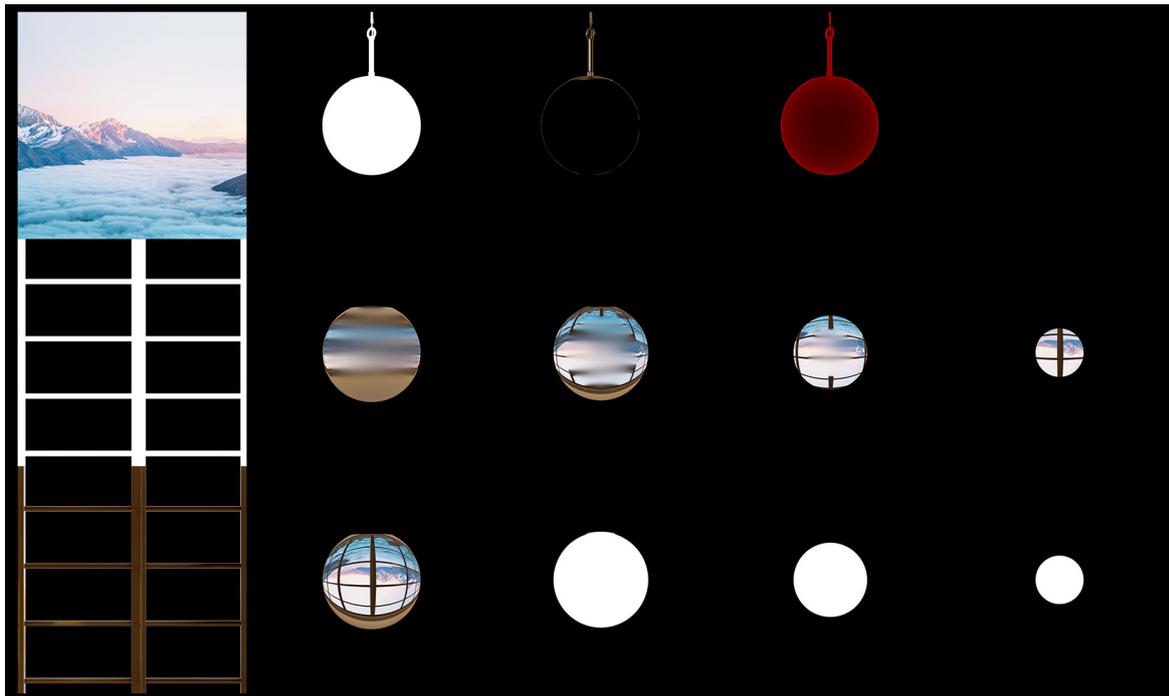


Fig. 5.24 The range of different input channels from the single render of the demonstration scene. This includes layers which have undergone infilling, and the refraction depth pass.

magnified, and forms an eyebox of (X) 81.92mm x (Y) 8.192mm - enough space in the Y axis for both of the viewer's eyes to view the scene in stereo with parallax. Having an  $>8$ mm X height provides a large enough eyebox for a fully dilated pupil to maximise defocus effects. The render view inputs themselves are 2048 x 2048 and upscaled to the 8k resolution of the SLM tile in order to super-sample for noise reduction. The resulting field of view is 26 degrees in X and Y (36 degrees diagonal) and provides 300 pixels per degree (PPD), 5 times larger than the 60 PPD required by human vision - in order to densely pack and average out noise. A diagram of the optical system is shown in figure 5.26.

Once these holograms are generated, they are immediately ready for display, and have the correct properties for artefact free viewing across the large eyebox with improved noise characteristics, especially in defocused areas. With the holographic defocus matching the experience presented in the in-render depth of field simulation, the viewer is then free to focus throughout the scene without caveats. A range of focal points and viewing positions are shown in figures 5.27, 5.28 and 5.29.

Where the eye varies in pupil size, a smaller region of the eyebox is seen, which results in a reduced perceptual image quality due to less image information reaching the retina. This bandwidth issue is one of the fundamental problems with holography, and is still the case

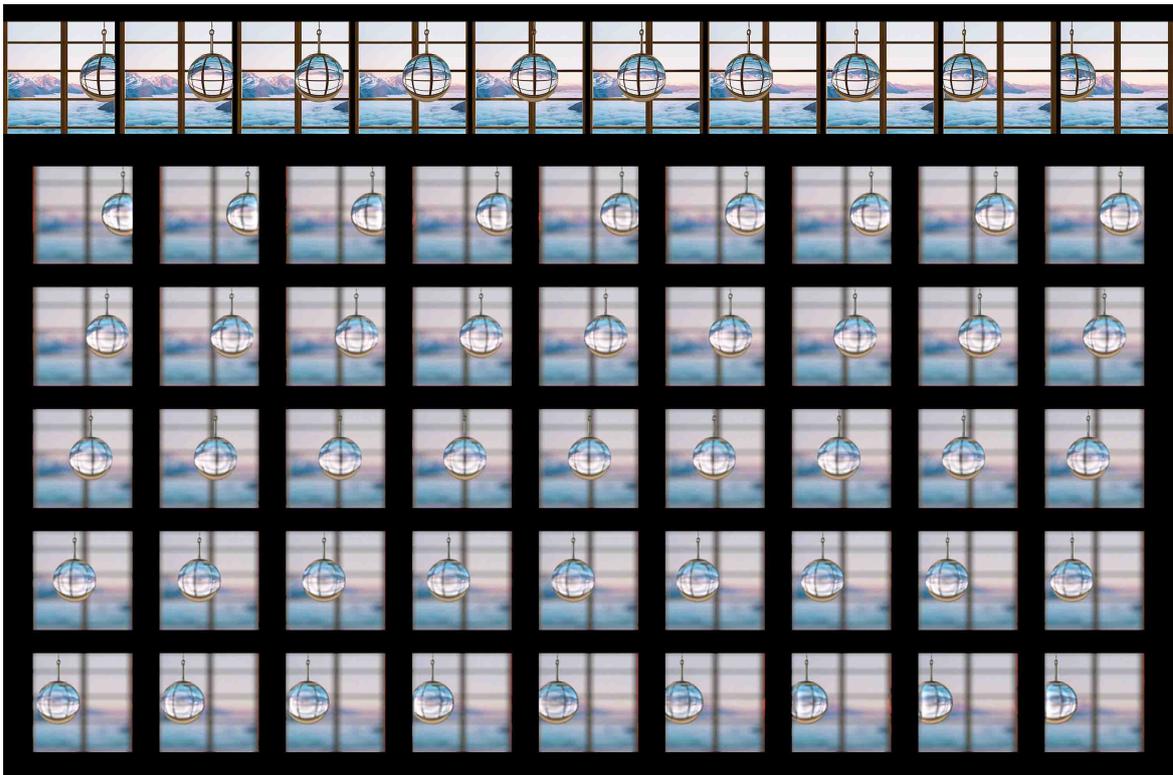


Fig. 5.25 The top row shows the 10 input render views from the source 3D scene. Underneath is 25 output views, which are replayed from new viewing positions from the resulting hologram - each viewing position is 1.8mm apart, and has a pupil size of 8mm.

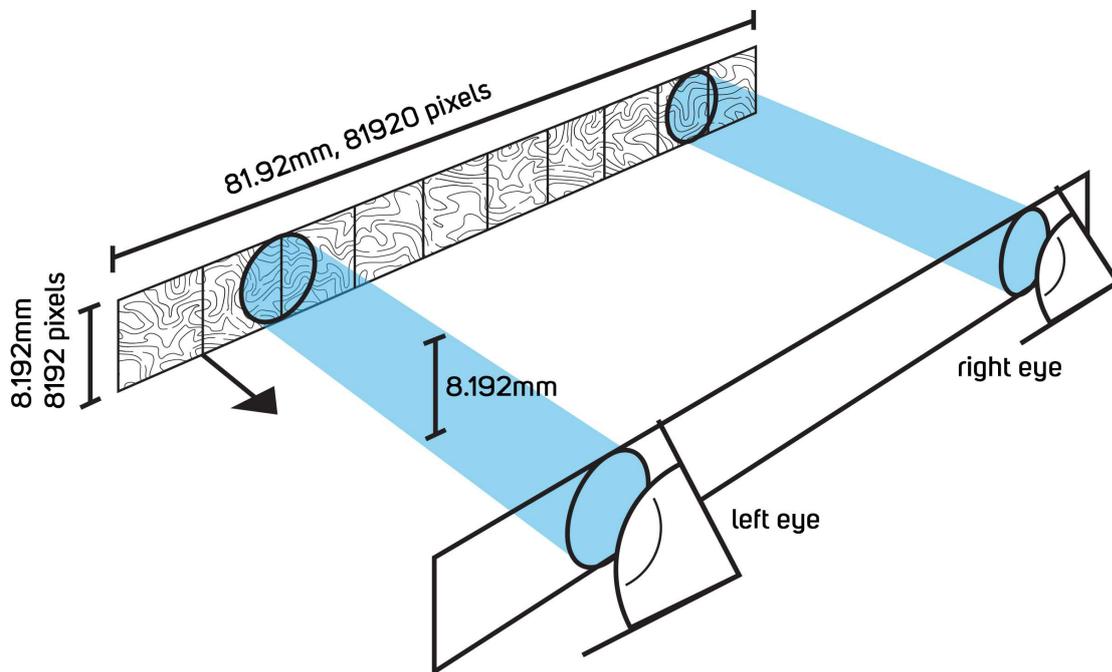


Fig. 5.26 An overview of the theoretical display system used in this section, one slice in Y of the display system D (10x10 views) from table 5.1. Here the viewer may see a combination of views at any given time.

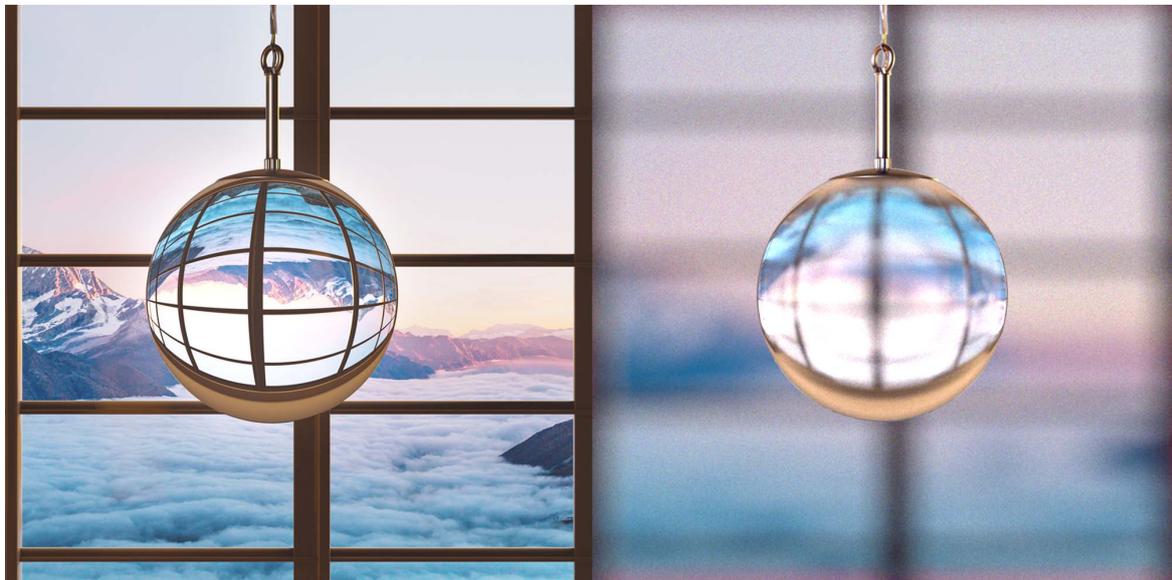


Fig. 5.27 Left: The source render scene with no defocus - a pinhole aperture. Right: The replay of the hologram including refraction depths and occlusion synthesis to enable correct defocus for the viewer.

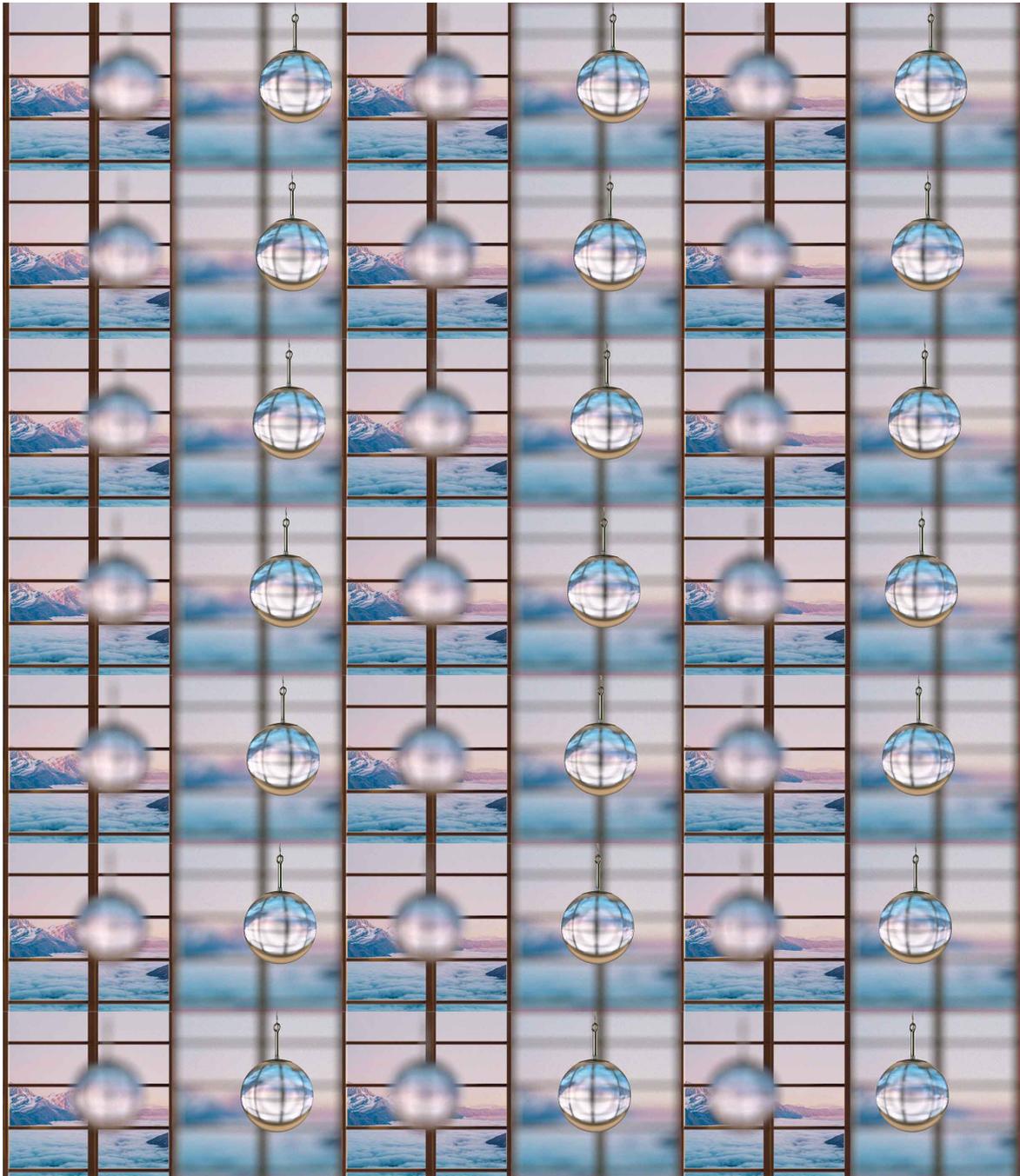


Fig. 5.28 A range of hologram replay positions with focus pairs at infinity and 20cm.

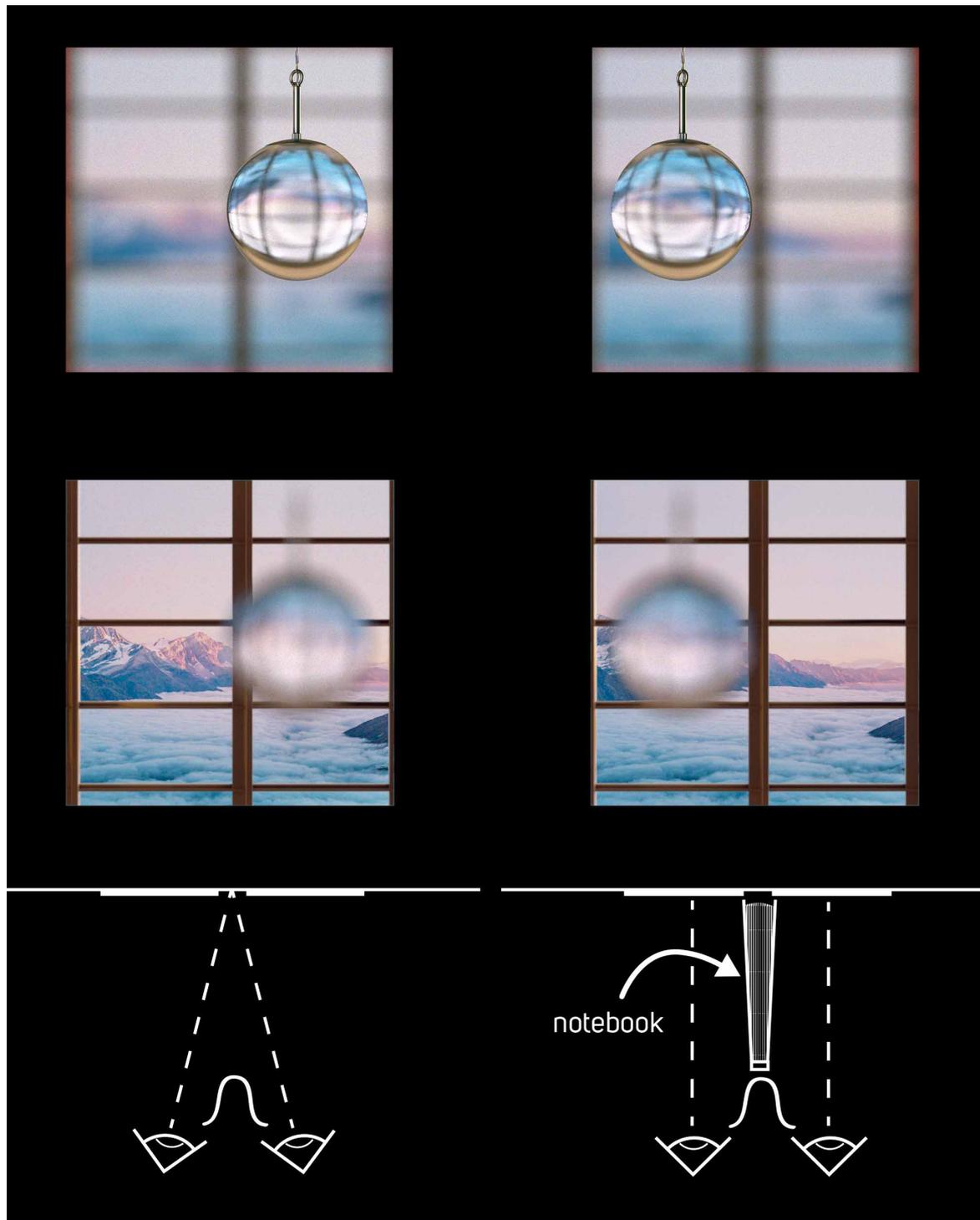


Fig. 5.29 Two stereo pairs of the scene with front and back focus. The viewing instruction diagram shows that usually, your eyes will verge in the center of the two images - if you bring your head quite close to the screen and "zone out" to verge in the distance, you can bring the two images together to form a stereo view. It may help to make the image smaller on screen, so that each image tile is about 2-3 finger widths wide, and use a notebook in order to see each image only in one eye. For the top image-pair, verge on the glass ball. For the bottom image-pair, verge the background. Viewing this may take some practice!

here - but figure 5.30 shows that despite any changes in pupil size, the multi-view image is able to persist. As the pupil decreases in size, so does the defocus.

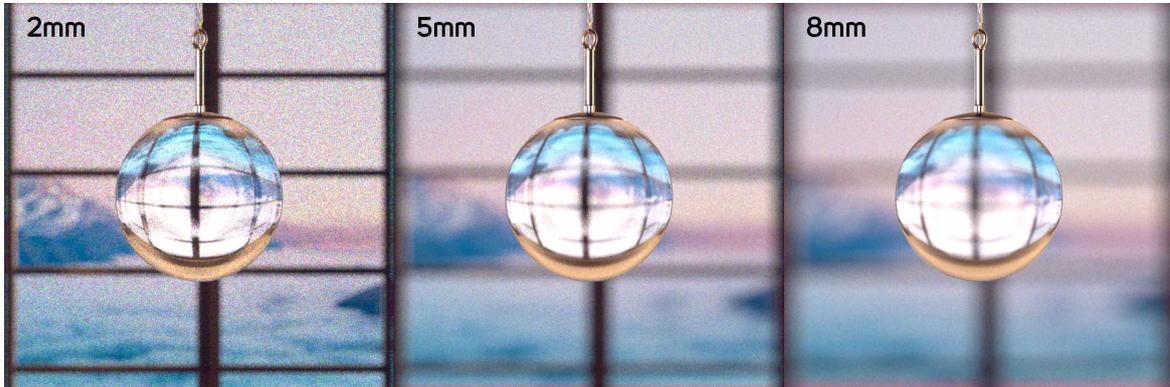


Fig. 5.30 Varying the iris size in the replay shows how image bandwidth is affected - smaller pupil size viewing a larger eyebox results in a decrease in image quality, but smaller eyebox sizes result in less defocus. In practice, scene brightness could be reduced to prevent pupil construction in order to preserve defocus effects if this is an issue in a given display system.

The glass ball lens in the scene, 20cm from the viewer, provides a challenging image to reproduce due to the highly view-dependent nature of the material, especially where parallax and focus effects are significant so close to the viewer. The glass ball brings an image into focus in a particular way, where rays that enter the ball further from the optical centre of the ball are more strongly refracted, this is shown in figure 5.31.

The effective focal length (EFL), where focus is defined as the distance between the plane through the centre of the lens and the beam waist (where the beam radius is at a minimum) of such a lens is given by:

$$EFL = \frac{nD}{4(n-1)} \quad (5.5)$$

Where  $n$  = refractive index and  $D$  = diameter of the ball lens. In the demonstration image, a refractive index of 1.85 is used, and with a diameter of 10cm, the  $EFL = \sim 5.5cm$ . The Back Focal Length (BFL) is the distance of the focal point from the lens surface itself - smaller than the EFL by  $D/2$ . As the centre of the ball is placed at 20cm in depth, a frontmost layer of 14.5cm in depth is allocated. The glass sphere, along with the metallic hanger, uses 4 layers of the 6 total assigned to the scene, as shown in figure 5.32.

Typically, such a low number of layers on any object close to the eye would result in image-breaking artefacts, as detailed in sections 4.4.1 and 5.3.3. The results here are a significant improvement, but could benefit further by increasing the number of layers for the glass object, which would see less of a reliance on the occlusion synthesis, which is still a

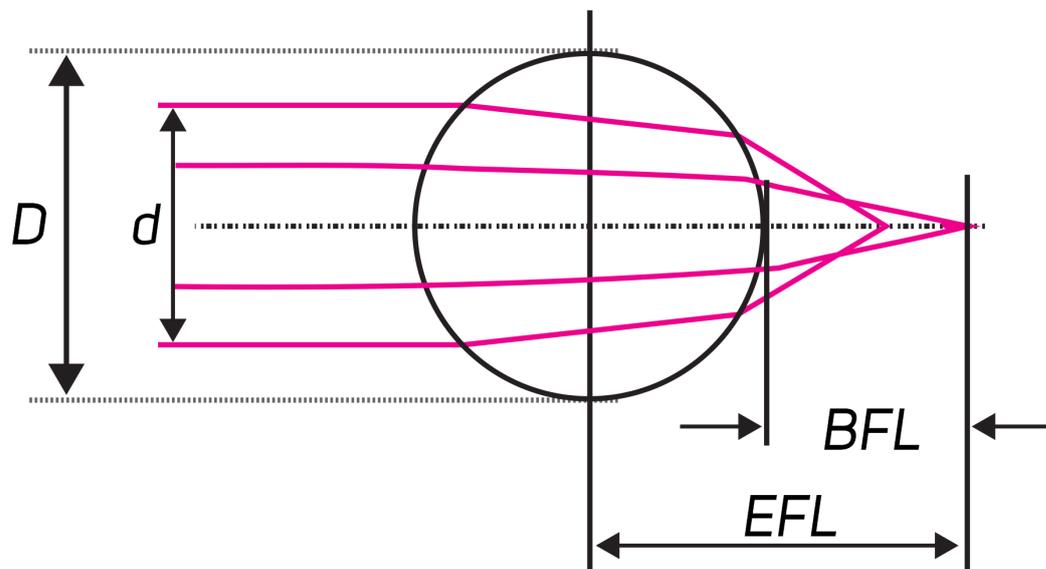


Fig. 5.31 An illustration of the focal length of a ball lens. The outermost rays are refracted more strongly than those which enter the ball closer to the optical axis - resulting in significant spherical aberrations.

simple image pyramid based in-fill - or by increasing the number of views. Both of these controls have been discussed in this chapter. The worst image errors are seen when focus pulls to the very closest layer when viewed from in-between views, shown in figure 5.33.

Though, it should be noted that objects so close to the eye start to become difficult to focus on for all viewers. Such a close layer distance, combined with the sparse number of layers and views, could be considered a sensible limit for viewable imagery. The scene used to demonstrate this work could benefit from more views for the refraction pass to reduce the perspective issues. A question also arises if perspective-doubling results in poor image perception when viewing a scene and only one eye falls on a view boundary - this appears to become negligible when viewed in stereo. This kind of image breakup is familiar in that lightfield type displays, at least where not providing a sufficiently large number of views, use this approximation of defocus in order to provide focal cues to the viewer. In a sufficiently high field of view display, these artefacts outside of the centre of view are likely to not even be perceived by the viewer.

Each full hologram of 81920 x 8192, with 6 layers and 9 occlusions per tile is estimated to be generated at interactive rates of  $\sim 1.2$  seconds per frame on a NVIDIA RTX2060 using extrapolated speed data and conservative estimates. In practice, the longest operation during the generation of these holograms was file IO to disk.

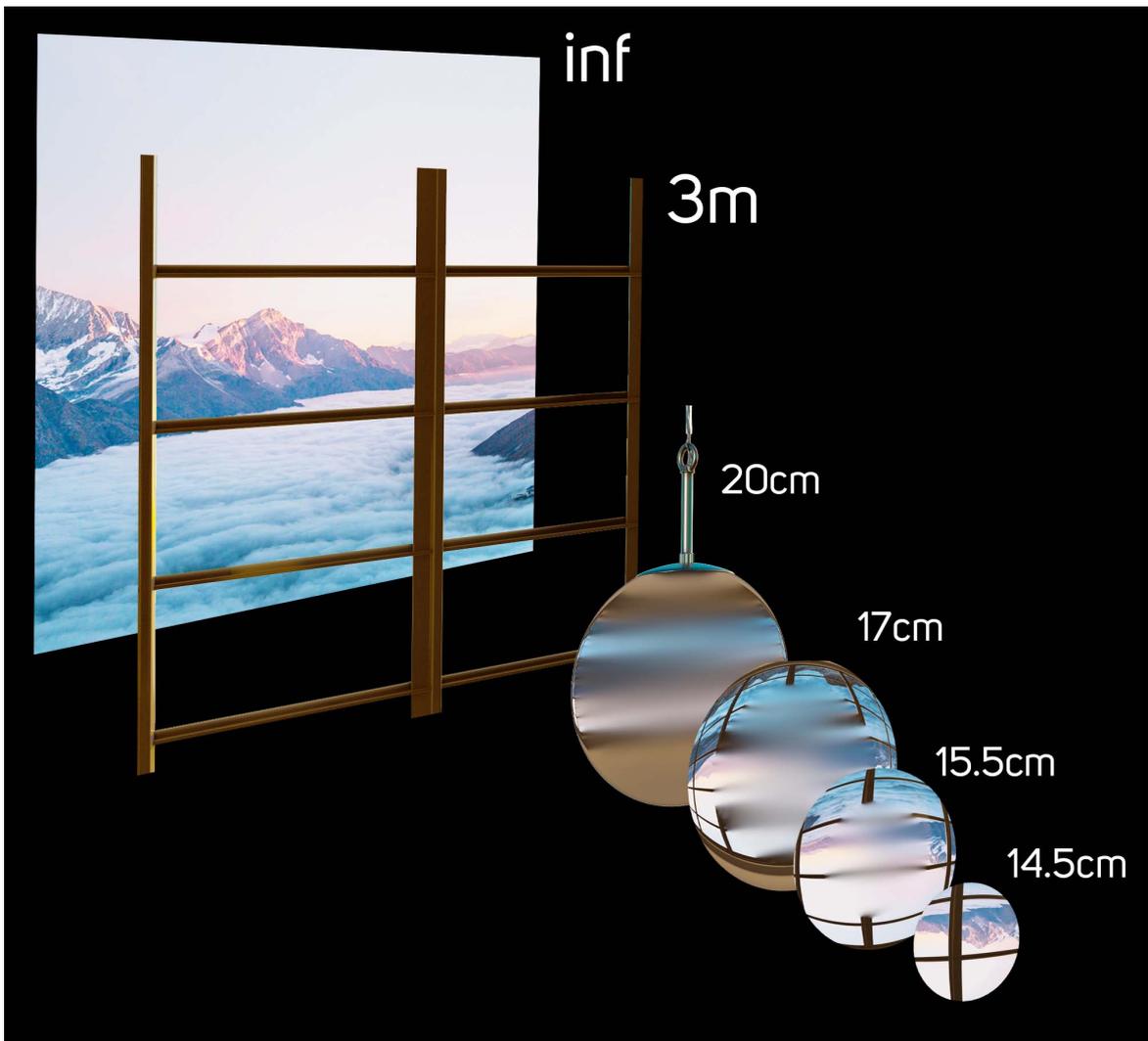


Fig. 5.32 An exploded view of the input layers of the demonstration scene, alongside their respective distances from the viewer.

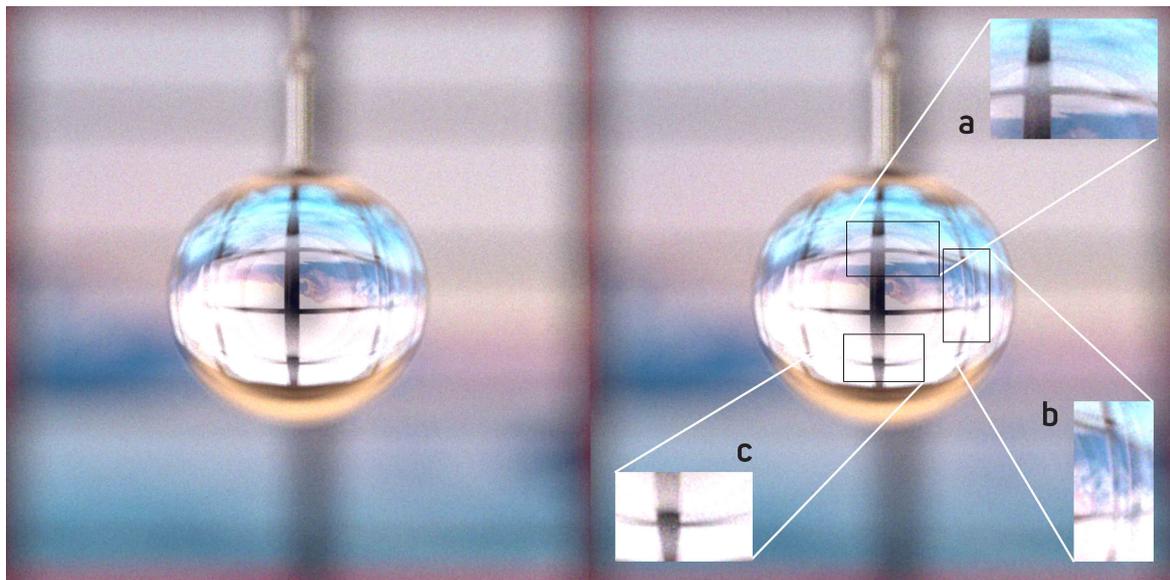


Fig. 5.33 Shows how lower numbers of layers on highly view-dependant segments of the image will result in the introduction of artefacts. The focus is pulled to 14.5cm from the viewer, and the eye position is between two views. Left: Un-annotated, Right: Annotated - (a) Shows perspective doubling where layer energies don't precisely match, and slight occlusion spilling with the bright edge shown at the window frame due to low accuracy infill. (b) Shows image doubling through an extreme refraction. (c) Insufficient occlusion synthesis quality, the source of which is seen in 5.32.

## 5.4.2 Conclusion

Most approaches of forming a multi-view hologram of individual tiles, or hogels, require a dense sampling of views as when the eye is in between views, artefacts appear in the image related to occlusion and missing image information. However, when using the occlusion synthesis technique with depth information for aspects such as refraction, each tile can be viewed independently successfully, and tiled at significant sizes.

This chapter has presented a new “discrete-hybrid holographic rendering” method which achieves a far more efficient process of creating fully featured holograms than creating a holographic ray-tracer - by utilising multiple perspective images from a traditional renderer (not limited to offline renderers, provided the correct render passes are available) to create a view-sampled hologram for a fraction of the cost in compute.

The presented display method is suitable for both dividing up a high resolution SLM into subsections, for time-multiplexing multiple holograms, or for tiling multiple physical SLM devices in an optical setup designed for large-format display.



# Chapter 6

## Discussion & Further Work

This thesis has explored advancements in the creation of computer-generated holograms in the context of a modern media production workflow. This chapter provides specific summaries by project and a general discussion of the methods and results, alongside suggested paths for further developments.

### 6.1 Discussion

As presented, holography provides an exciting avenue for media creation, artistic expression, and furthering realism in display technology. Handling the computational requirements of holography is a significant step to overcome, and layer-based holography provides an immediate route to make the adoption of holographic display feasible - whilst delivering the visual quality and control that is expected from a display in existing production practice. The layer-based method undoubtedly has drawbacks, but as explored, there are numerous techniques that significantly improve the resulting display experience of this sparse scene representation which culminate in it being the most effective method of 3D display.

Offline productions have significantly more compute available than real-time applications, and in this context, could utilise a native holographic ray-tracing renderer if it were available. However, adoption of such a renderer would have significant knock-on effects on the production process after rendering - even if the holographic raytracer had met the featureset available in production today. Long render times make processes such as visual development difficult, so even if a ray-tracer is used for final-frame, it is likely that layers or similar quantisation schemes will be applicable for a significant time period. The actual final-frame process will be dependent on the requirements of the shot itself - many scenes may only require a certain number of layers to convey their intent.

In any case, a full production system, which may mix ray-traced and layered shots, will only provide significant benefit once larger format displays exist. For the time being, near-eye displays will likely dominate in use. For this purpose, layers are especially applicable, as lowering compute requirements for portable headset style devices is a priority - and all of the techniques explored here are suitable for adoption under this use-case. For layer-based holography from a production standpoint this thesis has provided feature-parity with existing processes, and is applicable for both real-time and offline productions. As hardware and software continues to develop, and compute availability increases, the developments presented here will allow layer-based holography to flourish.

In the following sections, the results of each section are discussed.

### **6.1.1 Production Workflow**

Having a clear, compatible production workflow for imagery and other scene data to be properly utilised in CGH is a key foundational requirement, and is presented by this work. This thesis has developed this area from only having SDR 8 bit 2.5D with un-managed colour data as input to a full framework for enabling high quality HDR floating point imagery.

The introduced techniques and standards allow for display which ultimately match the requirements of human vision and likely exceed the requirements of most media. This section, being the foundation of the display, was the most important to develop well to ensure an excellent viewing experience.

#### **Colour**

Colour is one of the most substantial parts to any display, and prior to this thesis, there has been no attention given to this in CGH - with most research appearing to incorrectly directly display BT.709 colourspace imagery in the much wider gamut provided by the monochromatic primaries.

A priority was matching what exists currently in the state-of-the-art in production practice, and this, besides supporting existing spaces such as BT.709, DCI-P3 and BT.2020, means ensuring a path exists for much wider gamuts such as ACES support. With the goal of ACES being to ensure compatibility of imagery made today with the wide-gamut displays of the future, it is a perfect match for CGH. Work continues to progress with ACES to ensure it is less visually characterised by the process, and benefits from this will be instantly realised in CGH when using tristimulus illumination.

Whilst colour perception is forgiving, utilising fully managed colourspaces from the beginning ensures that the media produced looks precisely how it was intended (whether

replicating a realistic scene, or providing an artistic intent) - and the techniques introduced here allow accurate WCG laser illumination. The methods presented also allow expansion to more than three primaries to increase display gamut area, and even are applicable to full spectral-display, which would deliver exceptional colour reproduction from image sources which can provide it.

Whichever direction this takes, having laid out the framework for colour managed CGH display in such a way that matches media production, it is then trivial to adapt the display system to follow.

### **Luminance**

One of the key selling points of CGH is its capability of presenting real-world luminances to the viewer. When combined with an algorithm and hardware that does not introduce a high noise floor, a vast dynamic range can be displayed to the viewer. When assuming a display of sufficient capability, high luminance display is only successfully achieved when the source imagery has a suitable latitude delivered in HDR, be it scene-linear or via an existing EOTF in a high bit-depth format, such as presented in the HQH standard. This thesis presents the correct techniques by which to do so and enable hologram generation from such.

Various limitations in conventional display technologies means that high luminances are difficult to reach - but the desired ranges can be met with CGH. Benefits of high brightness may include AR overlays on bright daylight objects/scenes or matching existing HDR specifications such as the EDR ranges specified for Dolby Vision, which would allow the same experience seen in the grading suite in the viewer's home.

Given the recommendations specified in this thesis, it is possible to create imagery which does not exhibit banding and is bright enough to cause discomfort, but the actual dynamic range of the display is fully dictated by the quality of the algorithm and hardware. The blessing, and curse, of holographic display is that all of the energy in the optical path gets placed into the viewable image - noise in the replay field receives this energy, caused by both errors in the algorithm and errors in values displayed on the SLM. The developments presented in this work will allow reductions in the noise floor in future to be instantly realised, and handle suitably determined maximum luminances.

### **Character**

Holographic display is ultimately a unique medium, and this thesis has introduced methods that allow control over aspects of the image presented by a hologram which have yet to have been explored.

Just as a lens and camera are chosen for a production for their visual qualities, similar choices can be made when displaying images via a hologram. As presented, a virtual aperture can be created that allows control over defocus just as a director of photography would make decisions about the lens, the bokeh and the speed. Creative decisions could extend to patterns in the hologram to provide a varying grain, and this extends to using varying algorithms to achieve various controls over noise.

## **Workflow**

Creating a workflow method that utilises the same formats and standards that existing productions and GPU accelerated architectures utilise is critical in order to encourage adoption, ease of use, and widest compatibility. This is delivered in this work via the use of the HQH standard, using EXR as a container for all of the data required to make high quality holograms. Primarily this enables holography to readily make use of the existing production-ready technology stack in studios, and allows images rendered via existing methods to be viewed on holographic displays today. Laying this foundation then enables multi-depth and multi-view representations to be delivered to the generation software, ensures that holograms can be generated in existing render infrastructure, and allows an entire holographic production to go through the same delivery process to consumers via streaming or physical media as a conventional 2D production.

None of this is limited to layer-based holography, and is applicable to every form of CGH, including ray-tracing or wave-based methods.

### **6.1.2 Multi-Sampled Depth**

Given that layer-based holography primarily relies on 2.5D imagery, it is inherently limiting to the viewing experience. Flaws include only pure diffuse material representation and gaps between layers visible in parallax. But just as 2D compositing in VFX for conventional media occurs in 3D space, the same techniques can be applied to provide great benefit to CGH. Once this 2.5D imagery is supplemented with additional information from the renderer, whether this is formed of additional AOVs, segmented depth renders via clipping planes or object IDs, synthesised RGBZ data, or complex volumetric render data from sources such as Deep, as presented by this thesis, significant gains can be made in improving the realism and viewing experience of the scene.

Adopting the techniques presented in this thesis enables layer-based holograms to exhibit semi-transparencies, reflections, refractions, fur, smoke/fog and anti-aliased edges with correct depth information for the first time. This occurs via the same pipeline used to provide

additional render data to a compositing application, and much of the data required by the holographic display may already be utilised by a compositing artist. This presents many visual effects which are only otherwise possible via the use of a holographic ray-tracer, whilst requiring a fraction of the compute.

The occlusion synthesis method introduced in this thesis allows larger viewing angles and more realistic defocus to be realised from the single render perspective that is provided to the holographic display - without the need for multiple perspectives to be rendered. These flaws are visible in the forms of dark edges and halos in every layer-based hologram method in the literature. Occlusion synthesis provides a convenient and efficient way of mitigating these issues whilst providing control over how much compute is required via the detail delivered in the synthesised, and the number of layers being occluded. These issues of course can be bypassed via rendering occluded areas in layers, but this is not always possible, either because the source material is only 2.5D, or because object representations (such as polygons, which represent objects as a hull rather than being “solid”) will still present visual errors when sliced as there is no information inside that object.

Other representations of objects which are not polygon based, such as volumes (such as the commonly used voxel-based VDB format) can be used to correctly provide representations of the internals of objects. These representations can be rendered and suitably sliced for layer-based holographic display via the techniques also presented in this thesis - the deep slicing technique presented allows multiple pixel values, which are recorded in Z for a single pixel, to be conveniently processed and displayed. This might include motion blur, hair, grass, fog, smoke or fire - though these can also be rendered as layers explicitly, and provide appropriate occlusion for other holographic layers also using methods presented here.

This section has presented techniques that reduce the visual artefacts of layer-based holography, and advances the types of materials, objects and representations that can be displayed in 3D space via CGH with the realism expected by the viewer - combined, they allow layer-based holography to be viable for use in practice.

### 6.1.3 Multi-View Rendering

Utilising the techniques presented in the previous sections allows a seamless multi-view experience to be presented. This multi-view section expands the introduced workflow to support generation of holograms that are composed of multiple render perspectives in such a way to allow a large viewing region whilst uniquely providing low-compute capabilities and sufficient controls over quality vs generation time. Here, an efficient “discrete-hybrid” multi-view method is presented that supports existing display hardware proposals in the

literature, whilst still making use of the existing production workflows as per the rest of this thesis.

The largest issue with existing methods of multi-view is primarily that they require a significant number of rendered views, akin to lightfield rendering, in order to have acceptable image quality. Given a large scale holographic display, there is a sliding scale of algorithmic support which enables “multi-view”, this ranges from a single render view input with a large viewing area which has an unacceptable amount of error, to a ray-traced hologram, where each pixel is effectively a view.

Sparse sampling of input views would normally result in significant artefacts, either due to occlusion errors or due to significant jumps in perspective - but by utilising the methods introduced, it is possible to create holograms that form a balance in the number of input views required with the compute required for a given viewing area. Incorporating a combination of additional render layers and occlusion synthesis alleviates occlusion shadow duplications in the holographic image, and utilising the depth information of reflections and refractions reduces the errors visible in view-dependent materials when the viewer’s pupil is traversing the viewing area.

In this method, by combining the techniques of occlusion synthesis and multi-sampled depths, artefact free imagery is formed in such a way that presents seamless viewing - a surprising outcome. This can be further improved in visual quality by the utilisation of an increased number of layers during hologram generation, which reduces parallax errors. This technique allows interactive rates to be achieved on existing consumer hardware at a preview number of layers, which can then be significantly increased for higher quality final-frame hologram generation.

## 6.2 Further Work

Having touched on a not insignificant selection of topics in this work, there were a lot of ideas along the way that were not able to be explored, or a little tangential to the research goals.

### 6.2.1 Layers and Other Scene Quantisation Schemes

This EngD has been concerned with methods specific to layer-based holography, primarily because it is the method by which interactive rates of hologram computation is most easily achieved on current widely available GPU hardware. But even when not concerned about the computation required for the generation of holograms, making decisions about how scenes

should be allocated to layers is not always easy. Splitting objects, especially continuous surfaces which have a significant range of depths (such as the surface of a table top), into a smaller number of layers always seems to result in undesirable effects. This occurs even when utilising techniques such as occlusion synthesis to fill gaps between layers when automating layer allocation which is not sufficiently scene-aware, such as via some function of depth alone. It is not always possible to alleviate this manually - so it would be valuable for further work to consider scene-aware layer allocation.

Layers are imperfect by design, and the viewing experience is especially affected where the viewer is not central to the optical axis of the hologram, as explored in section 4.4.2 - which is often the case when viewing a hologram on a benchtop projector. Such a projector in the lab displays a hologram of a single render perspective to emulate the application of an AR or VR headset, and designed for an eye which is roughly fixed in position. Where “roughly” refers to a small eyebox, with some room for the eye to move around, either in viewing direction or translation (such as of the viewer’s head in relation to the viewing device). Occlusion synthesis assists these scenarios, however, only if multiple layers are used.

Research has indicated that a single layer, with suitable eye tracking in order to move this single holographic layer through depth [138], and defocus simulated in the image plane, is sufficient to produce a flawless visual experience. This is commonly referred to as gaze-contingent varifocal display, and typically is used with conventional displays and focus-tunable lenses [159]. While simulated blur techniques have been shown to not provide the correct cues to drive accommodation as effectively as real defocus [140], certain methods such as "ChromaBlur" can even begin to drive accommodation on 2D displays [35] - and whilst there is this comparison between perceptual blur simulations [140] and separate comparison of the effect of different CGH techniques on accommodation [115], there has not been a study that compares the two categories together.

Generally, having an eye tracking system would also allow for foveated display to occur (such as in “Metameric Varifocal Holograms”)[212], which could have significant benefits for power consumption on portable devices - ensuring that unnecessary rendering does not occur. Further developments could be contributed in this area by integrating such a display to the workflow presented in this thesis to allow for fast scene previewing at high image quality, as well as develop methods that allow a position received from an eye tracking system to correspond to a given depth or view in complex rendering cases such as via reflections or refractions.

Other methods of representing scenes will require attention as holographic displays get larger and allow non-fixed position viewing. Work could include decomposing scenes into

level of detail models in order to be represented via low-polygon holographic meshes or a combination of planar layers and tilted planes. This mesh could come directly from the DCC software or be generated via techniques similar to relighting geometry in compositing via the use of normal maps, and simply have the standard render projected onto the holographic planes. Adding such a method to the range of holographic display techniques would mean that certain objects could have more complex holograms generated where required. This could then be holographically composited with other elements generated via other methods, in the same fashion as points and layers have previously been combined[71], or by itself provide an additional quality level on the scale from single layer to full raytracing.

Eventually, the compute budget on devices will allow full holographic ray-tracing, either via dedicated acceleration silicon[116] or via the increase in compute available on graphics processors. But it is likely that in practice, not every display will need such levels of accuracy at all times - and introducing methods that can suitably indicate which elements of a scene require a given hologram generation technique would allow for efficient use of power budgets whilst providing the required visual results.

### 6.2.2 High Quality Display

CGH can be pushed far beyond the existing conventions for display by delivering more image quality than is physically viewable, and exploring these boundaries is an exciting direction for the research. Even as issues with noise and resolution continue to be reduced by the development of improved SLM hardware, algorithms and laser speckle reduction techniques - in terms of media production there will still be many issues to tackle.

Given that a holographic display is illuminated by monochromatic sources, spectral display becomes a real possibility, and allows for more than 3 primaries to be used to form the holographic image. Existing research shows that methods of optimising the number of primaries for the spectral locus can assist in choosing laser wavelengths[192], but there are no recommendations on performing this kind of display in holography, nor how such a workflow should occur even in conventional media. Research in rendering and storing spectral imagery[102, 215, 62] could be adapted into a holographic workflow to enable extremely wide-gamut display systems. Delivering an increased range of colour sub-frames would also reduce the appearance of noise in the system. Recent work in wavelength translation for holograms[129] could also enable these sub-frames to vary in wavelength (if the additional primaries do not stray too far from the correct colour transform) without requiring additional rendering or storage.

At a minimum, a system optimised for existing wide colour standards, defined with 3 monochromatic primaries, could provide accurate mastering displays and enable further

research into topics such as observer metamerism. Improvements or calibration profiles could be developed for individual viewers, and holography could ensure such a calibration is visible by only that viewer.

Looking further into the future, defining how a holographic ray-tracing system should be built to successfully integrate with existing pipelines is a challenge, especially in such a way that would allow parity with existing renderers or provide a means by which to provide both 2D and holographic outputs. Recent developments such as USD[208] and MaterialX[141] could provide the foundation for such a system to exist, and more work is required to explore how this could function successfully.

Another area of the research which will need to advance, not just for the benefit of CGH, is the efficient render, encode, storage and transmission of multi-view imagery for the selection of emerging three-dimensional display technologies. Techniques such as those explored in “Compositing light field video” [56], where a lightfield is encoded into a Deep EXR file could directly be used by the workflow presented in this thesis for display.

More work is also required in delivering full-complex modulation of light in hardware to achieve improved image quality, where SLMs are either currently phase or amplitude only.

### 6.2.3 Occlusion

Occlusion proves to be a costly operation, as each layer that needs to be occluded adds another Fourier transform to every other layer per frame, more research is certainly needed to reduce this burden, which adds significantly to compute times for scenes with more layers. There are various avenues which could tackle this - utilising eye tracking to continuously present a correctly geometrically occluded view to the pupil, or methods of combining layers to reduce the total count.

In the context of applications such as AR displays, a key benefit of holography is that it allows objects to be placed seamlessly alongside real world objects - but of course this can only add light to the visible scene, and not yet occlude real-world objects. This will be an important factor to make AR more usable, especially when considering that such a display may be used to consume media including footage of people - this is extremely problematic as the display cannot replicate a full range of skin tones as an additive only display system. Occlusion using an additional LCD in the optical system would allow the Over compositing operation to occur on real-world views - not only occluding real-world objects, but also allowing virtual objects to be properly represented in Augmented Reality [98]. Work is required in order to determine how such a display system should be built, how the masking data should be delivered to the occluding hardware, either an LCD or otherwise, and how such a physical occluding layer interacts with the depth perception of imagery displayed.

### 6.2.4 Quality of Life

In terms of improving every day operation of holographic displays, it will be necessary to introduce, and automate, best practices in calibration. Whilst this was touched upon in section 3.8, this process so far has been manual - research in performing this automatically, either via internal optics or via an additional sensing device in front of the display will solve a large pain point. This includes calibration of holographic zernike polynomials, such as aberration corrections to reduce the effects of optical elements in the projection system [106], luminance and colour calibration to match given standards, especially as the lasers age or vary in operating temperatures, and calibration of depth values to ensure that objects appear at the intended distances from the viewer. Given that holographic displays are not currently in production in consumer products, and their total adoption will likely not be immediate, it will also be necessary to have a suitable simulation system, allowing artists working on holographic shots to emulate the viewing experience in 2D. Some work has already been contributed to this area of research in the form of the JPEG “Pleno” group, with “Numerical Reconstruction Software for Holography” being presented [105], but such a system will need to be able to simulate various types of optical setups to be able to match the plethora of real hardware that will become available.

In a wider view of improving the experience of generating and viewing holograms, speed again is a factor, but by making use of advances of specialised hardware, such as tensor cores to increase FFT speed[131], ray-tracing hardware in GPUs, or FPGA [220] ASIC devices, or even edge compute [213], the ease at which imagery can be created and iterated upon will be increased. The work presented in this thesis provides a significant strengthening to the foundation for delivering this further research, and it will be exciting to both contribute personally and see how others assist in developing this exciting display technology.

# Chapter 7

## Conclusions & Final Note

### 7.1 Conclusion

CGH is not new, but the prospect of it becoming a feasible display technology with widespread adoption is. This work has laid the foundation for digital holographic display to utilise the same imagery produced for current media, by enabling compatibility with modern production practice to enable existing standards to be realised by CGH and then surpassed.

The primary motivation for this work was to make use of the advancements in computer graphics techniques and apply them to CGH. While this was originally intended to be limited to advancing support for concepts such as semi-transparencies, a much broader range of benefits have been achieved by working through the various bottlenecks in existing CGH workflow techniques. There was a significant gap in the application of conventional media production knowledge, which has now been filled, alongside solving the issues listed in the Research Problems Overview in section 1.1.

In summary, the novel contributions to the development of CGH by this work include:

1. The HQH standard format, enabling HDR, WCG and suitable layer, view and AOV storage for CGH generation, alongside a theoretical framework presented for its use. This is the first proposed input format standard for CGH, and first complete analysis of application of existing media standards to CGH.
2. The concept of Artistic optical controls for CGH, where previously "in-software" optical corrections for CGH had been purely for reversing optical errors in the projector.
3. The Holographic Compositing technique, utilising diffractive occlusion algorithms to allow multi-sampled depths - including occlusion synthesis, reflective/refractive

display methods and deep render format display. Enabling semi-transparencies such as anti-aliased edges and volumetric representations for the first time. This makes layer-based CGH viable by removing visible artefacts, and advances display qualities by enabling scene representations (such as volumes) not previously demonstrated in any CGH method.

4. Discrete-Hybrid Holography, enabling surprisingly low compute requirements for continuously perceived large area displays - alongside the tools to render and handle the required scene data. This contribution presents a culmination of the novel work in this thesis and enables an effective method of large-format display for future generation CGH devices.

The most unique factor to this work is that it is the only research in CGH that has tackled issues from the production perspective. Conventional display workflows are well established, and continue to advance and iterate - but given the eventual availability of CGH consumer devices, productions will have to target both 2D and 3D displays. The burden of this should be as minimal as possible on the VFX artist - and the techniques presented here help provide a strong foundation for a new exciting area of research in computer graphics which improve image quality and reduce complexity in CGH production.

The work presented here is already being used internally at VividQ for demonstration display systems and customer-driven projects, and is being integrated into the Core product GPU accelerated package for hologram generation. Some techniques, most notably occlusion synthesis and the other holographic compositing methods, have already been adapted for real-time use in game engines for AR and VR devices.

Methods introduced here allow real world luminances and wide colour gamuts to be realised in CGH for the first time, alongside suitable delivery and workflow specifications. Limitations with layer-based holography have been identified and solutions then proposed and demonstrated. The primary contributions include the concept of holographic compositing; which allows complex scenes to be better represented in CGH with transparencies, reflections, refractions, volumes such as fog, fire and smoke, motion blur, and fine detailed objects such as hair and fur. Visual artefacts are resolved via the use of occlusion synthesis for both single and multi-view display cases, and a “discrete-hybrid” multi-view algorithm is proposed to allow larger holographic displays to make use of layer-based holography efficiently.

Despite layer-based holograms being intrinsically designed to provide an approximation of the 3D scene to the viewer, this thesis has introduced a number of valuable techniques that allows the discrete sampling of layer-based holography to be perceived as continuous, and make the most of the realism produced by existing production renderers. Scenes can

now be more accurately represented than previously possible, and errors introduced by the approximation of such scenes have been sufficiently resolved.

## **7.2 EngD Experience Final Note**

The progress made in CGH, computer graphics and processing power has been immense throughout the duration of this EngD. When I first joined VividQ and was presented the state-of-the-art holographic display, I was honestly a little disappointed by the noise, the limited field of view, and the compute requirements - not to mention the lack of useful integration and ease of use. On one hand, this of course presented a great opportunity for this EngD research project, but was also not an insignificant amount of work. Fortunately, work has progressed so much that today there are now already CGH displays in development that are able to match the display quality of the 2D panels we are used to using daily, and perform at realtime frame-rates on consumer hardware.

The amount of interest in CGH has increased dramatically, and in the four years of this work gone from very few knowing what CGH was, to CGH related works being centre-front of conferences. It's great to see several research groups helping push the boundaries of this exciting display technology, and great to see this work being realised in practice by companies aiming to use CGH at first for AR/VR applications - with the first devices becoming available in the next few years.

The breadth of knowledge required for making good holograms has been a challenge to learn, and I'm still learning - hence the additional year I needed to finish writing up this thesis! Despite the many long days, it has been inspiring working on this new format of display, and a great pleasure to contribute to this exciting field.



# References

- [1] *ACES Case Study – Screen Gems (tech level: low)*, Alex Forsythe (Academy ACES) [n.d.], [https://www.youtube.com/watch?v=jnPQ\\_HCkw4M](https://www.youtube.com/watch?v=jnPQ_HCkw4M). Accessed: 11-09-2022.
- [2] *ACES Virtual Working Group – Output Transforms* [n.d.], <https://community.acescentral.com/c/aces-development-acesnext/vwg-output-transforms/>. Accessed: 11-09-2022.
- [3] Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J. and Ogden, J. M. [1984], ‘Pyramid methods in image processing’, *RCA engineer* **29**(6), 33–41.
- [4] Akşit, K. and Itoh, Y. [2023], Holobeam: Paper-thin near-eye displays, in ‘2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)’, IEEE, pp. 581–591.
- [5] Banks, M. S., Sprague, W. W., Schmoll, J., Parnell, J. A. and Love, G. D. [2015], ‘Why do animal eyes have pupils of different shapes?’, *Science advances* **1**(7), e1500391.
- [6] Barnes, C., Shechtman, E., Finkelstein, A. and Goldman, D. B. [2009], ‘Patchmatch: A randomized correspondence algorithm for structural image editing’, *ACM Trans. Graph.* **28**(3), 24.
- [7] Barsky, B. A. [2004], Vision-realistic rendering: simulation of the scanned foveal image from wavefront data of human subjects, in ‘Proceedings of the 1st Symposium on Applied perception in graphics and visualization’, pp. 73–81.
- [8] Barten, P. G. [1999], *Contrast sensitivity of the human eye and its effects on image quality*, SPIE press.
- [9] Barten, P. G. [2003], Formula for the contrast sensitivity of the human eye, in ‘Image Quality and System Performance’, Vol. 5294, SPIE, pp. 231–238.
- [10] *Baselight, FilmLight Ltd.* [n.d.], <https://www.filmlight.ltd.uk/>. Accessed: 11-09-2022.
- [11] Bayraktar, M. and Özcan, M. [2010], ‘Method to calculate the far field of three-dimensional objects for computer-generated holography’, *Applied optics* **49**(24), 4647–4654.
- [12] Benton, S. A. [1983], Survey of holographic stereograms, in ‘Processing and Display of Three-Dimensional Data’, Vol. 367, SPIE, pp. 15–19.
- [13] Benton, S. A. and Bove Jr, V. M. [2007], *Holographic Imaging*, John Wiley & Sons, Inc.

- [14] Bérard, P., Bradley, D., Nitti, M., Beeler, T. and Gross, M. H. [2014], ‘High-quality capture of eyes.’, *ACM Trans. Graph.* **33**(6), 223–1.
- [15] Birnbaum, T., Blinder, D., Muhamad, R. K., Schretter, C., Symeonidou, A. and Schelkens, P. [2020], ‘Object-based digital hologram segmentation and motion compensation’, *Optics Express* **28**(8), 11861–11882.
- [16] Blinn, J. F. [1977], Models of light reflection for computer synthesized pictures, in ‘Proceedings of the 4th Annual Conference on Computer Graphics and Interactive Techniques’, SIGGRAPH ’77, Association for Computing Machinery, New York, NY, USA, p. 192–198.  
**URL:** <https://doi.org/10.1145/563858.563893>
- [17] Borer, T. and Cotton, A. [2015], ‘Arib std-b67-essential parameter values for the extended image dynamic range television system for programme production’, *ARIB Standard. Association of Radio Industries and Businesses* .
- [18] *Breaking Bad*, Copyright AMC Network Entertainment. [2008-].
- [19] Buckley, E. [2007], Computer-generated holograms for real-time image display and sensor applications, PhD thesis, Cambridge University.  
**URL:** <https://doi.org/10.17863/CAM.31430>
- [20] Burgess, J. [2020], ‘RTX on—The NVIDIA Turing GPU’, *IEEE Micro* **40**(2), 36–44.
- [21] Burley, B. and Studios, W. D. A. [2012], Physically-based shading at disney, in ‘Acm Siggraph’, Vol. 2012, vol. 2012, pp. 1–7.
- [22] Cable, A. [2007], Real-time high-quality two and three-dimensional holographic video projection using the one-step phase retrieval (OSPR) approach, PhD thesis, Cambridge University.  
**URL:** <https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.597206>
- [23] Catmull, E. E. [1974], A Subdivision Algorithm for Computer Display of Curved Surfaces., PhD thesis. AAI7504786.
- [24] Chakravarthula, P., Tseng, E., Fuchs, H. and Heide, F. [2022], ‘Hogel-free holography’, *ACM Transactions on Graphics* **41**(5), 1–16.
- [25] Chang, J., Kauvar, I., Hu, X. and Wetzstein, G. [2016], Variable aperture light field photography: overcoming the diffraction-limited spatio-angular resolution tradeoff, in ‘Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition’, pp. 3737–3745.
- [26] *Chaos VRAY, RGB Colour Back to Beauty Compositing* [n.d.], [https://docs.chaos.com/display/VMAYA/RGB\\_Color](https://docs.chaos.com/display/VMAYA/RGB_Color). Accessed: 17-02-2023.
- [27] Chaves, J. [2017], *Introduction to nonimaging optics*, CRC press.
- [28] Chen, J.-S. and Chu, D. P. [2015], ‘Improved layer-based method for rapid hologram generation and real-time interactive holographic display applications’, *Opt. Express* **23**(14), 18143–18155.  
**URL:** <http://opg.optica.org/oe/abstract.cfm?URI=oe-23-14-18143>

- [29] Chen, J.-S., Smithwick, Q. and Chu, D. [2013a], Implementation of shading effect for reconstruction of smooth layer-based 3D holographic images, in A. J. Woods, N. S. Holliman and G. E. Favalora, eds, 'Stereoscopic Displays and Applications XXIV', Vol. 8648, International Society for Optics and Photonics, SPIE, pp. 218–226.  
**URL:** <https://doi.org/10.1117/12.2004704>
- [30] Chen, J.-S., Smithwick, Q. and Chu, D. [2013b], Implementation of shading effect for reconstruction of smooth layer-based 3D holographic images, in A. J. Woods, N. S. Holliman and G. E. Favalora, eds, 'Stereoscopic Displays and Applications XXIV', Vol. 8648, International Society for Optics and Photonics, SPIE, p. 86480R.  
**URL:** <https://doi.org/10.1117/12.2004704>
- [31] Chen, J.-S., Smithwick, Q. and Chu, D. [2013c], Implementation of shading effect for reconstruction of smooth layer-based 3d holographic images, in 'Stereoscopic Displays and Applications XXIV', Vol. 8648, SPIE, pp. 218–226.
- [32] Chen, R. H.-y. [2010], Computer Generated Holograms for 3-D Holographic Display, Phd, Cambridge University.  
**URL:** <https://doi.org/10.17863/CAM.14125>
- [33] Chen, R. H.-Y. and Wilkinson, T. D. [2009], 'Computer generated hologram with geometric occlusion using gpu-accelerated depth buffer rasterization for three-dimensional display', *Appl. Opt.* **48**(21), 4246–4255.  
**URL:** <http://opg.optica.org/ao/abstract.cfm?URI=ao-48-21-4246>
- [34] Choi, S., Gopakumar, M., Peng, Y., Kim, J., O'Toole, M. and Wetzstein, G. [2022], Time-multiplexed neural holography: a flexible framework for holographic near-eye displays with fast heavily-quantized spatial light modulators, in 'ACM SIGGRAPH 2022 Conference Proceedings', pp. 1–9.
- [35] Cholewiak, S. A., Love, G. D., Srinivasan, P. P., Ng, R. and Banks, M. S. [2017], 'Chromablur: Rendering chromatic eye aberration improves accommodation and realism', *ACM Transactions on Graphics (TOG)* **36**(6), 1–12.
- [36] Ciuffreda, K. J., Wang, B. and Vasudevan, B. [2007], 'Conceptual model of human blur perception', *Vision Research* **47**(9), 1245–1252.  
**URL:** <https://www.sciencedirect.com/science/article/pii/S0042698906005621>
- [37] *Colour*, Thomas Mansencal [n.d.], <https://www.colour-science.org/>. Accessed: 11-09-2022.
- [38] Cook, R. L., Carpenter, L. and Catmull, E. [1987], 'The reyes image rendering architecture', *ACM SIGGRAPH Computer Graphics* **21**(4), 95–102.
- [39] Cook, R. L., Porter, T. and Carpenter, L. [1984], Distributed ray tracing, in 'Proceedings of the 11th annual conference on Computer graphics and interactive techniques', pp. 137–145.
- [40] Cooper, S., Achard, R., Fry, A., Molholm, R., Borg, L., Fostier, L., Langlands, A., Mansencal, T., Agland, S., Karis, B. et al. [2017], 'Retrospective and enhancements'.

- [41] Cuervo, E., Chintalapudi, K. and Kotaru, M. [2018], Creating the perfect illusion: What will it take to create life-like virtual reality headsets?, in 'Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications', pp. 7–12.
- [42] Cutting, J. E. [1995], Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth Wayfinding View project Narrative theory and the dynamics of popular movies View project, Technical report.  
**URL:** <https://www.researchgate.net/publication/236964257>
- [43] Dainty, J. and Welford, W. [1971], 'Reduction of speckle in image plane hologram reconstruction by moving pupils', *Optics communications* **3**(5), 289–294.
- [44] *Davinci Resolve, Blackmagic Design* [n.d.], <https://www.blackmagicdesign.com/uk/products/davinciresolve>. Accessed: 11-09-2022.
- [45] de la Perriere, V. B., Drazic, V., Doyen, D. and Schubert, A. [2020], Synthesis of computer generated holograms using layer-based method and perspective projection images, in '2020 International Conference on 3D Immersion (IC3D)', IEEE, pp. 1–5.
- [46] *Deadline, AWS Thinkbox* [n.d.], <https://www.awsthinkbox.com/deadline>. Accessed: 11-09-2022.
- [47] Demolder, A. [2022], 'Toward the standardization of high-quality computer-generated holography media production workflow', *SMPTE Motion Imaging Journal* **131**(1), 48–58.
- [48] Demolder, A. [UK Patent Application GB2214017.2, Sept. 2022 (withdrawn). Application GB2314566.7, Sept. 2023], 'Method, apparatus, computer program and computer-readable medium for determining a computer-generated hologram'.
- [49] Demolder, A., Kaczorowski, A., Newman, A. and Nait-Charif, H. [2020], 'Artistic control of defocus in computer-generated holography', *ACM CVMP* **20**.
- [50] Demolder, A., Newman, A., Durrant, T., Nait-Charif, H., Adzhiev, V. and Kaczorowski, A. [2021], Enabling reflective & refractive depth representation in computer-generated holography, in 'ACM SIGGRAPH 2021 Posters', pp. 1–2.
- [51] Disney, W. E. [1940], 'Art of animation'. US Patent 2,201,689.
- [52] *District 9, Copyright TriStar Pictures Inc.* [2009].
- [53] *Dolby Vision Whitepaper – An Introduction to Dolby Vision, Dolby Laboratories* [n.d.], [https://professional.dolby.com/siteassets/pdfs/dolby-vision-whitepaper\\_an-introduction-to-dolby-vision\\_0916.pdf](https://professional.dolby.com/siteassets/pdfs/dolby-vision-whitepaper_an-introduction-to-dolby-vision_0916.pdf). Accessed: 11-09-2022.
- [54] Dorsch, R. G., Lohmann, A. W. and Sinzinger, S. [1994], 'Fresnel ping-pong algorithm for two-plane computer-generated hologram display', *Applied optics* **33**(5), 869–875.
- [55] Duane, A. [1922], 'Studies in monocular and binocular accommodation, with their clinical application', *Transactions of the American Ophthalmological Society* **20**, 132.

- [56] DuVall, M., Flynn, J., Broxton, M. and Debevec, P. [2019], Compositing light field video using multiplane images, in ‘ACM SIGGRAPH 2019 Posters’, pp. 1–2.
- [57] Ebner, C., Mori, S., Mohr, P., Peng, Y., Schmalstieg, D., Wetzstein, G. and Kalkofen, D. [2022], ‘Video see-through mixed reality with focus cues’, *IEEE Transactions on Visualization and Computer Graphics* **28**(5), 2256–2266.
- [58] Egstad, J., Davis, M. and Lacewell, D. [2015], Improved deep image compositing using subpixel masks, in ‘Proceedings of the 2015 Symposium on Digital Production’, pp. 21–27.
- [59] Elkoura, G., Grassia, S., Boonyatera, S., Mohr, A., Jeremias-Vila, P. and Kuruc, M. [2019], A deep dive into universal scene description and hydra, in ‘ACM SIGGRAPH 2019 Courses’, pp. 1–48.
- [60] Fascione, L., Hanika, J., Leone, M., Droske, M., Schwarzhaupt, J., Davidovič, T., Weidlich, A. and Meng, J. [2018], ‘Manuka: A batch-shading architecture for spectral path tracing in movie production’, *ACM Transactions on Graphics (TOG)* **37**(3), 1–18.
- [61] Fascione, L., Hanika, J., Pieké, R., Villemin, R., Hery, C., Gamito, M., Emrose, L. and Mazzone, A. [2018], Path tracing in production, in ‘ACM SIGGRAPH 2018 Courses’, SIGGRAPH ’18, Association for Computing Machinery, New York, NY, USA.  
**URL:** <https://doi.org/10.1145/3214834.3214864>
- [62] Fichet, A., Pacanowski, R. and Wilkie, A. [2021], ‘An openexr layout for spectral images’, *Journal of Computer Graphics Techniques* .
- [63] *Fireside Chat with VividQ, MultiDepth Sampling for CGH, Holographic Compositing* [n.d.], <https://spie.org/conferences-and-exhibitions/spie-online/spie-online-on-demand/ar-vr-mr-xr>. Accessed: 26-02-2023.
- [64] Flynn, J., Broxton, M., Debevec, P., DuVall, M., Fyffe, G., Overbeck, R., Snavely, N. and Tucker, R. [2019], Deepview: View synthesis with learned gradient descent, in ‘Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition’, pp. 2367–2376.
- [65] *Foundry Nuke - User Guide, Compositing with Nuke, Merging Image, Merge Operations* [n.d.], [https://learn.foundry.com/nuke/content/comp\\_environment/merging/merge\\_operations.html](https://learn.foundry.com/nuke/content/comp_environment/merging/merge_operations.html). Accessed: 17-02-2023.
- [66] Freeman, J. and Simoncelli, E. P. [2011], ‘Metamers of the ventral stream’, *Nature neuroscience* **14**(9), 1195–1201.
- [67] Friedman, J. and Jones, A. C. [2015], Fully automatic id mattes with support for motion blur and transparency, in ‘ACM SIGGRAPH 2015 Posters’, pp. 1–1.
- [68] Gabor, D. [1948], ‘A new microscopic principle’, *nature* **161**, 777–778.
- [69] Gerchberg, R. W. and Saxton, W. O. [1972], ‘A practical algorithm for the determination of phase from image by 103.240. 126.9 on 10/21/18. Re-use and distribution is strictly not permitted, except for Open Access articles. and diffraction plane pictures,”’, *Optik* **35**, 227–246.

- [70] Gerveshi, A. and Looper, S. [2019], Distributed multi-context interactive rendering, in 'Proceedings of the 2019 Digital Production Symposium', DigiPro '19, Association for Computing Machinery, New York, NY, USA.  
**URL:** <https://doi.org/10.1145/3329715.3338878>
- [71] Gilles, A. [2016], Fast hologram synthesis methods for realistic 3D visualization, PhD thesis, Rennes, INSA.
- [72] Golan, L. and Shoham, S. [2009], 'Speckle elimination using shift-averaging in high-rate holographic projection', *Opt. Express* **17**(3), 1330–1339.  
**URL:** <http://www.opticsexpress.org/abstract.cfm?URI=oe-17-3-1330>
- [73] Goodman, J. W. [2005], *Introduction to Fourier optics*, Roberts and Company Publishers.
- [74] Hahn, J., Kim, H., Lim, Y., Park, G. and Lee, B. [2008], 'Wide viewing angle dynamic holographic stereogram with a curved array of spatial light modulators', *Optics express* **16**(16), 12372–12386.
- [75] Hainich, R. R. and Bimber, O. [2016], *Displays: fundamentals & applications*, AK Peters/CRC Press.
- [76] Halle, M. W. [1991], The generalized holographic stereogram, PhD thesis, Massachusetts Institute of Technology.
- [77] Halle, M. W. [1994], Holographic stereograms as discrete imaging systems, in 'Practical Holography VIII', Vol. 2176, SPIE, pp. 73–84.
- [78] Halle, M. W. [1997], Multiple viewpoint rendering for three-dimensional displays, PhD thesis, Massachusetts Institute of Technology.
- [79] Hannan, W. J., Flory, R. E., Lurie, M. and Ryan, R. J. [1973], 'Holotape®: A low-cost prerecorded television system using holographic storage', *Journal of the SMPTE* **82**(11), 905–915.
- [80] Hardy, A. C. [1920], 'A study of the persistence of vision', *Proceedings of the National Academy of Sciences of the United States of America* **6**(4), 221.
- [81] Hariharan, P. [1996], *Optical Holography: Principles, Techniques and Applications*, Cambridge Studies in Modern Optics, 2 edn, Cambridge University Press.
- [82] Hayashi, N., Sakamoto, Y. and Honda, Y. [2011], Improvement of camera arrangement in computer-generated holograms synthesized from multi-view images, in 'Practical Holography XXV: Materials and Applications', Vol. 7957, SPIE, pp. 284–292.
- [83] Heidrich, W., Slusallek, P. and Seidel, H.-P. [1997], An image-based model for realistic lens systems in interactive computer graphics, in 'Graphics Interface', Vol. 97, Citeseer, pp. 68–75.
- [84] Herman, I. P. [2016], *Physics of the human body*, Springer.

- [85] Hexley, A. C., Yöntem, A. Ö., Spitschan, M., Smithson, H. E. and Mantiuk, R. [2020], ‘Demonstrating a multi-primary high dynamic range display system for vision experiments’, *JOSA A* **37**(4), A271–A284.
- [86] *High Dynamic Range Color Grading and Display in Frostbite*, Alex Fry (EA) [n.d.], <https://www.ea.com/frostbite/news/high-dynamic-range-color-grading-and-display-in-frostbite>. Accessed: 11-09-2022.
- [87] Hillman, P. [2013], ‘The theory of openxr deep samples’, *Retrieved April 29*, 2018.
- [88] Hillman, P. [2018], A scheme for storing object id manifests in openxr images, in ‘Proceedings of the 8th Annual Digital Production Symposium’, pp. 1–8.
- [89] Hsueh, C. K. and Sawchuk, A. A. [1978], ‘Computer-generated double-phase holograms’, *Appl. Opt.* **17**(24), 3874–3883.  
**URL:** <https://opg.optica.org/ao/abstract.cfm?URI=ao-17-24-3874>
- [90] Hullin, M. B., Eisemann, E., Seidel, H.-P. and Lee, S. [2011], ‘Physically-based real-time lens flare rendering’, *ACM Trans. Graph. (Proc. SIGGRAPH 2011)* **30**(4), 108:1–108:9.
- [91] Ichikawa, T., Yamaguchi, K. and Sakamoto, Y. [2013a], Realistic expression for full-parallax computer-generated holograms with the ray-tracing method, in ‘Applied Optics’, Vol. 52, Hokkaido University.  
**URL:** <https://doi.org/10.1364/AO.52.00A201>
- [92] Ichikawa, T., Yamaguchi, K. and Sakamoto, Y. [2013b], ‘Realistic expression for full-parallax computer-generated holograms with the ray-tracing method’, *Applied Optics* **52**(1), A201–A209.
- [93] Ichioka, Y., Izumi, M. and Suzuki, T. [1971], ‘Scanning halftone plotter and computer-generated continuous-tone hologram’, *Appl. Opt.* **10**(2), 403–411.  
**URL:** <http://opg.optica.org/ao/abstract.cfm?URI=ao-10-2-403>
- [94] Igarashi, S., Nakamura, T. and Yamaguchi, M. [2016], ‘Fast method of calculating a photorealistic hologram based on orthographic ray–wavefront conversion’, *Optics Letters* **41**(7), 1396–1399.
- [95] Im, D., Cho, J., Hahn, J., Lee, B. and Kim, H. [2015], ‘Accelerated synthesis algorithm of polygon computer-generated holograms’, *Opt. Express* **23**(3), 2863–2871.  
**URL:** <http://opg.optica.org/oe/abstract.cfm?URI=oe-23-3-2863>
- [96] *Industrialised Cinematography*, Craig Welsh [n.d.], <https://www.expandedcinematography.com/industrialised-cinematography.html>. Accessed: 11-09-2022.
- [97] *Information Display Measurements Standard, International Committee for Display Metrology* [n.d.], <https://www.sid.org/Standards/ICDM#8271483-idms-download>. Accessed: 11-09-2022.

- [98] Itoh, Y., Hamasaki, T. and Sugimoto, M. [2017], ‘Occlusion leak compensation for optical see-through displays using a single-layer transmissive spatial light modulator’, *IEEE transactions on visualization and computer graphics* **23**(11), 2463–2473.
- [99] ITU [1993], ITU-R BT.709 Parameter values for the HDTV standards for production and international programme exchange, Standard, International Telecommunications Union, Geneva, CH.
- [100] ITU [2012], ITU-R BT.2020 Parameter values for ultra-high definition television systems for production and international programme exchange, Standard, International Telecommunications Union, Geneva, CH.
- [101] ITU [2016], ITU-R BT.2100 Parameter values for ultra-high definition television systems for production and international programme exchange, Standard, International Telecommunications Union, Geneva, CH.
- [102] Jakob, W. [2010], ‘Mitsuba renderer’. <http://www.mitsuba-renderer.org>.
- [103] *Jason and the Argonauts (1963)* [n.d.], <https://www.imdb.com/title/tt0057197/>. Accessed: 17-02-2023.
- [104] Jendral, A., Bräuer, R. and Bryngdahl, O. [1995], ‘Computer-generated image holograms for 3d-display’, *Optical Review* **2**, 187–188.
- [105] *JPEG Pleno, Joint Photographic Experts Group (JPEG)* [n.d.], <https://jpeg.org/jpegpleno/>. Accessed: 11-09-2022.
- [106] Kaczorowski, A. [2018], Adaptive aberration correction for holographic projectors, PhD thesis, University of Cambridge.
- [107] Kainz, F. [2013], ‘Interpreting openexr deep pixels’, *Retrieved April 29*, 2018.
- [108] Kainz, F., Bogart, R. and Hess, D. [2004], ‘The openexr image file format. gpu gems: Programming techniques, tips and tricks for real-time graphics, r. fernando, ed’.
- [109] Kang, H., Stoykova, E. and Yoshikawa, H. [2016], ‘Fast phase-added stereogram algorithm for generation of photorealistic 3d content’, *Applied optics* **55**(3), A135–A143.
- [110] Kang, H., Yamaguchi, T. and Yoshikawa, H. [2008], ‘Accurate phase-added stereogram to improve the coherent stereogram’, *Applied optics* **47**(19), D44–D54.
- [111] Katayama, A., Tanaka, K., Oshino, T. and Tamura, H. [1995], dependent stereoscopic display using interpolation of multiviewpoint images, in ‘Stereoscopic displays and virtual reality systems II’, Vol. 2409, SPIE, pp. 11–20.
- [112] Kavaklı, K., Itoh, Y., Urey, H. and Akşit, K. [2022], ‘Realistic defocus blur for multiplane computer-generated holography’, *arXiv preprint arXiv:2205.07030*.
- [113] Kavaklı, K., Urey, H. and Akşit, K. [2022], ‘Learned holographic light transport’, *Applied Optics* **61**(5), B50–B55.

- [114] Kavaklı, K., Shi, L., Urey, H., Matusik, W. and Akşit, K. [2023], ‘Holohdr: Multi-color holograms improve dynamic range’, *arxiv* .
- [115] Kim, D., Nam, S.-W., Lee, B., Seo, J.-M. and Lee, B. [2022], ‘Accommodative holography: improving accommodation response for perceptually realistic holographic displays’, *ACM Transactions on Graphics (TOG)* **41**(4), 1–15.
- [116] Kim, H., Kim, Y., Ji, H., Park, H., An, J., Song, H., Kim, Y. T., Lee, H.-S. and Kim, K. [2019], ‘A single-chip fpga holographic video processor’, *IEEE Transactions on Industrial Electronics* **66**(3), 2066–2073.
- [117] Kim, H. R., Angelaki, D. E. and DeAngelis, G. C. [2016], ‘The neural basis of depth perception from motion parallax’, *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**(1697), 20150256.
- [118] Kim, J., Gopakumar, M., Choi, S., Peng, Y., Lopes, W. and Wetzstein, G. [2022], Holographic glasses for virtual reality, in ‘Proceedings of the ACM SIGGRAPH’, p. 1–8.
- [119] *King Kong (1933)* [n.d.], <https://www.imdb.com/title/tt0024216/>. Accessed: 17-02-2023.
- [120] Kinoshita, M. and Sakamoto, Y. [2009], Computer-generated holograms at an arbitrary viewpoint synthesized from multi-view images, in ‘Practical Holography XXIII: Materials and Applications’, Vol. 7233, SPIE, pp. 237–244.
- [121] Kopf, J., Matzen, K., Alsisan, S., Quigley, O., Ge, F., Chong, Y., Patterson, J., Frahm, J.-M., Wu, S., Yu, M. et al. [2020], ‘One shot 3d photography’, *ACM Transactions on Graphics (TOG)* **39**(4), 76–1.
- [122] Kozacki, T., Finke, G., Garbat, P., Zaperty, W. and Kujawińska, M. [2012], ‘Wide angle holographic display system with spatiotemporal multiplexing’, *Optics express* **20**(25), 27473–27481.
- [123] Kunkel, T., Daly, S., Miller, S. and Froehlich, J. [2016], Perceptual design for high dynamic range systems, in ‘High Dynamic Range Video’, Elsevier, pp. 391–430.
- [124] Lakshminarayanan, V. and Fleck, A. [2011], ‘Zernike polynomials: a guide’, *Journal of Modern Optics* **58**(7), 545–561.
- [125] Lanman, D. and Luebke, D. [2013], ‘Near-eye light field displays’, *ACM Transactions on Graphics (TOG)* **32**(6), 1–10.
- [126] *Laser technical FAQ*, Christie Digital Systems USA, Inc. [n.d.], <https://www.christiedigital.com/globalassets/resources/public/christie-reallaser-faq.pdf>. Accessed: 11-09-2022.
- [127] Lazarev, G., Hermerschmidt, A., Krüger, S. and Osten, S. [2012], ‘LCOS spatial light modulators: trends and applications’, *Optical Imaging and Metrology: Advanced Technologies* pp. 1–29.

- [128] Lee, B., Kim, D., Lee, S., Chen, C. and Lee, B. [2022], ‘High-contrast, speckle-free, true 3d holography via binary cgh optimization’, *Scientific reports* **12**(1), 1–12.
- [129] Lee, J. and Lee, B. [2022], ‘Deep neural network and its training strategy for converting computer-generated hologram between different display systems’, *OPTICA: Digital Holography and Three-Dimensional Imaging* p. W2A–11.
- [130] Levoy, M. and Hanrahan, P. [1996], Light field rendering, in ‘Proceedings of the 23rd annual conference on Computer graphics and interactive techniques’, pp. 31–42.
- [131] Li, B., Cheng, S. and Lin, J. [2021], tcfft: A fast half-precision fft library for nvidia tensor cores, in ‘2021 IEEE International Conference on Cluster Computing (CLUSTER)’, pp. 1–11.
- [132] Li, J., Smithwick, Q. and Chu, D. [2022], ‘Holobricks: modular coarse integral holographic displays’, *Light: Science & Applications* **11**(1), 1–15.
- [133] Lippmann, G. [1908], ‘La photographie integrale’, *Comptes-Rendus* **146**, 446–451.
- [134] Lohmann, A. W. and Paris, D. P. [1967], ‘Binary fraunhofer holograms, generated by computer’, *Appl. Opt.* **6**(10), 1739–1748.  
**URL:** <http://opg.optica.org/ao/abstract.cfm?URI=ao-6-10-1739>
- [135] Lokovic, T. and Veach, E. [2000], Deep shadow maps, in ‘Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques’, SIGGRAPH ’00, ACM Press/Addison-Wesley Publishing Co., USA, p. 385–392.  
**URL:** <https://doi.org/10.1145/344779.344958>
- [136] Lucente, M. E. [1993], ‘Interactive computation of holograms using a look-up table’, *Journal of Electronic Imaging* **2**(1), 28–34.
- [137] Lucente, M. and Galyean, T. A. [1995], Rendering interactive holographic images, in ‘Proceedings of the 22nd annual conference on Computer graphics and interactive techniques’, pp. 387–394.
- [138] Maimone, A., Georgiou, A. and Kollin, J. S. [2017], ‘Holographic near-eye displays for virtual and augmented reality’, *ACM Trans. Graph.* **36**(4).  
**URL:** <https://doi.org/10.1145/3072959.3073624>
- [139] Mankowska, N. D., Marcinkowska, A. B., Waskow, M., Sharma, R. I., Kot, J. and Winklewski, P. J. [2021], ‘Critical flicker fusion frequency: A narrative review’, *Medicina* **57**(10), 1096.
- [140] March, J., Krishnan, A., Watt, S., Wernikowski, M., Gao, H., Yöntem, A. Ö. and Mantiuk, R. [2022], Impact of correct and simulated focus cues on perceived realism, in ‘SIGGRAPH Asia 2022 Conference Papers’, pp. 1–9.
- [141] *MaterialX, Industrial Light & Magic* [n.d.], <https://materialx.org/>. Accessed: 11-09-2022.
- [142] Matsushima, K., Schimmel, H. and Wyrowski, F. [2003], ‘Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves’, *JOSA A* **20**(9), 1755–1762.

- [143] Max, N. L. and Lerner, D. M. [1985], A two-and-a-half-d motion-blur algorithm, in 'Proceedings of the 12th annual conference on Computer graphics and interactive techniques', pp. 85–93.
- [144] Meier, R. W. [1965], 'Depth of focus and depth of field in holography', *JOSA* **55**(12), 1693–1694.
- [145] Mellish, B. [2002], 'Diagram of the holographic recording process'.  
**URL:** <https://commons.wikimedia.org/wiki/File:Holography-record.png>
- [146] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R. and Ng, R. [2021], 'Nerf: Representing scenes as neural radiance fields for view synthesis', *Communications of the ACM* **65**(1), 99–106.
- [147] Muhamad, R. K., Blinder, D., Symeonidou, A., Birnbaum, T., Watanabe, O., Schretter, C. and Schelkens, P. [2019], 'Exact global motion compensation for holographic video compression', *Applied Optics* **58**(34), G204–G217.
- [148] Netflix [2022], 'Netflix open content library'.  
**URL:** <https://opencontent.netflix.com/>
- [149] *New CIE XYZ functions transformed from the CIE (2006) LMS functions*, CVRL [n.d.], <http://cvrl.ioo.ucl.ac.uk/ciexyzpr.htm>. Accessed: 11-09-2022.
- [150] Newell, M. E., Newell, R. and Sancha, T. L. [1972], A solution to the hidden surface problem, in 'Proceedings of the ACM annual conference-Volume 1', pp. 443–450.
- [151] Newman, A. J., Widjanarko, T., Spiess, A. O., Milne, D. F., Wilkinson, T. D. and Kaczorowski, A. [2021], 'An intrinsically eye safe approach to high apparent brightness augmented reality displays using computer-generated holography', *Journal of the Society for Information Display* **29**(11), 840–851.
- [152] Nishi, H. and Matsushima, K. [2017], 'Rendering of specular curved objects in polygon-based computer holography', *Applied optics* **56**(13), F37–F44.
- [153] Nishi, H., Matsushima, K. and Nakahara, S. [2011], 'Rendering of specular surfaces in polygon-based computer-generated holograms', *Applied optics* **50**(34), H245–H252.
- [154] Noll, A. M. [1971], 'Scanned-display computer graphics', *Commun. ACM* **14**(3), 143–150.  
**URL:** <https://doi.org/10.1145/362566.362567>
- [155] Okun, J. A., Susan Zwerman, V. et al. [2020], *The VES handbook of visual effects: industry standard VFX practices and procedures*, Routledge.
- [156] *OpenCue*, ASWF [n.d.], <https://www.opencue.io/>. Accessed: 11-09-2022.
- [157] *OpenEXR*, *Industrial Light & Magic* [n.d.], <https://www.openexr.com/>. Accessed: 11-09-2022.
- [158] *OpenImageIO* [n.d.], <https://openimageio.org/>. Accessed: 17-02-2023.

- [159] Padmanaban, N., Konrad, R., Stramer, T., Cooper, E. A. and Wetzstein, G. [2017], ‘Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays’, *Proceedings of the National Academy of Sciences* **114**(9), 2183–2188.
- [160] Padmanaban, N., Peng, Y. and Wetzstein, G. [2019], ‘Holographic near-eye displays based on overlap-add stereograms’, *ACM Transactions on Graphics (TOG)* **38**(6), 1–13.
- [161] Pan, Y., Wang, Y., Liu, J., Li, X. and Jia, J. [2013], ‘Fast polygon-based method for calculating computer-generated holograms in three-dimensional display’, *Appl. Opt.* **52**(1), A290–A299.  
**URL:** <http://opg.optica.org/ao/abstract.cfm?URI=ao-52-1-A290>
- [162] Pan, Y., Xu, X. and Liang, X. [2013], ‘Fast distributed large-pixel-count hologram computation using a GPU cluster’, *Applied optics* **52**(26), 6562–6571.
- [163] Pedrotti, F. L., Pedrotti, L. M. and Pedrotti, L. S. [2017], *Introduction to optics*, Cambridge University Press.
- [164] Pekkarinen, E. and Balzer, M. [2019], Physically based lens flare rendering in " the lego movie 2", in ‘Proceedings of the 2019 Digital Production Symposium’, pp. 1–3.
- [165] Peng, Y., Choi, S., Padmanaban, N., Kim, J. and Wetzstein, G. [2020], Neural Holography, in ‘Special Interest Group on Computer Graphics and Interactive Techniques Conference Emerging Technologies’, ACM, New York, NY, USA, pp. 1–2.  
**URL:** <https://dl.acm.org/doi/10.1145/3388534.3407295>
- [166] Peng, Y., Choi, S., Padmanaban, N. and Wetzstein, G. [2020], ‘Neural holography with camera-in-the-loop training’, *ACM Trans. Graph.* **39**(6).
- [167] *Pixar Renderman Documentation - Deep Compositing* [n.d.], [https://renderman.pixar.com/resources/RenderMan\\_20/deepCompositing.html](https://renderman.pixar.com/resources/RenderMan_20/deepCompositing.html). Accessed: 17-02-2023.
- [168] Pointer, M. R. [1980], ‘The gamut of real surface colours’, *Color Research & Application* **5**(3), 145–155.
- [169] Porter, T. and Duff, T. [1984], Compositing digital images, in ‘Proceedings of the 11th annual conference on Computer graphics and interactive techniques’, pp. 253–259.
- [170] Potmesil, M. and Chakravarty, I. [1981], ‘A lens and aperture camera model for synthetic image generation’, *ACM SIGGRAPH Computer Graphics* **15**(3), 297–305.
- [171] Potmesil, M. and Chakravarty, I. [1983], ‘Modeling motion blur in computer-generated images’, *ACM SIGGRAPH Computer Graphics* **17**(3), 389–399.
- [172] Poynton, C. [2018], Colour appearance issues in digital video, HD/UHD, and D-cinema, PhD thesis, : Individualized Interdisciplinary Studies.
- [173] Ramberg, E. [1972], ‘Holographic information storage’, *RCA Rev.* **33**(1), 5–53.
- [174] Reeves, P. [1918], ‘Rate of pupillary dilation and contraction.’, *Psychological Review* **25**(4), 330.

- [175] Ritschel, T., Ihrke, M., Frisvad, J. R., Coppens, J., Myszkowski, K. and Seidel, H.-P. [2009], Temporal glare: Real-time dynamic simulation of the scattering in the human eye, *in* ‘Computer Graphics Forum’, Vol. 28, Wiley Online Library, pp. 183–192.
- [176] Roth, S., Ben-David, I., Ben-Chorin, M., Eliav, D. and Ben-David, O. [2003], 10.2: wide gamut, high brightness multiple primaries single panel projection displays, *in* ‘SID Symposium Digest of Technical Papers’, Vol. 34, Wiley Online Library, pp. 118–121.
- [177] Sahin, E., Stoykova, E., Mäkinen, J. and Gotchev, A. [2020], ‘Computer-generated holograms for 3d imaging: A survey’, *ACM Comput. Surv.* **53**(2).
- [178] Sato, H., Kakue, T., Ichihashi, Y., Endo, Y., Wakunami, K., Oi, R., Yamamoto, K., Nakayama, H., Shimobaba, T. and Ito, T. [2018], ‘Real-time colour hologram generation based on ray-sampling plane with multi-GPU acceleration’, *Scientific reports* **8**(1), 1–10.
- [179] Schelkens, P., Blinder, D., Birnbaum, T., Symeonidou, A., Ahar, A. and Schretter, C. [2018], Source coding of holographic data: challenges, algorithms and standardization efforts, *in* ‘Speckle 2018: VII International Conference on Speckle Metrology’, Vol. 10834, SPIE, pp. 257–262.
- [180] Selan, J. [2012], Cinematic color: from your monitor to the big screen, *in* ‘ACM SIGGRAPH 2012 Courses’, pp. 1–54.
- [181] Shi, L., Huang, F.-C., Lopes, W., Matusik, W. and Luebke, D. [2017a], ‘Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3d computer graphics’, *ACM Trans. Graph.* **36**(6).  
**URL:** <https://doi.org/10.1145/3130800.3130832>
- [182] Shi, L., Huang, F.-C., Lopes, W., Matusik, W. and Luebke, D. [2017b], ‘Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3d computer graphics’, *ACM Transactions on Graphics (TOG)* **36**(6), 1–17.
- [183] Shi, L., Li, B., Kim, C., Kellnhofer, P. and Matusik, W. [2021], ‘Towards real-time photorealistic 3d holography with deep neural networks’, *Nature* **591**(7849).
- [184] Shi, L., Li, B. and Matusik, W. [2022], ‘End-to-end learning of 3d phase-only holograms for holographic display’, *Light: Science & Applications* **11**(1), 1–18.
- [185] Shimobaba, T. and Ito, T. [2015], ‘Random phase-free computer-generated hologram’, *Opt. Express* **23**(7), 9549–9554.  
**URL:** <http://opg.optica.org/oe/abstract.cfm?URI=oe-23-7-9549>
- [186] Shimobaba, T. and Ito, T. [2019], *Computer Holography: Acceleration Algorithms and Hardware Implementations*, CRC press.
- [187] Shimobaba, T., Makowski, M., Nagahama, Y., Endo, Y., Hirayama, R., Hiyama, D., Hasegawa, S., Sano, M., Kakue, T., Oikawa, M., Sugie, T., Takada, N. and Ito, T. [2016], ‘Color computer-generated hologram generation using the random phase-free method and color space conversion’, *Appl. Opt.* **55**(15), 4159–4165.  
**URL:** <http://opg.optica.org/ao/abstract.cfm?URI=ao-55-15-4159>

- [188] Shirley, P. [2016], ‘Ray tracing in one weekend’, *Amazon Digital Services LLC* **1**.
- [189] Slinger, C., Cameron, C. and Stanley, M. [2005], ‘Computer-generated holography as a generic display technology’, *Computer* **38**(8), 46–53.
- [190] Smithwick, Q., Chen, J.-S. and Chu, D. [2013], 26.1: A coarse integral holographic display, in ‘SID Symposium Digest of Technical Papers’, Vol. 44, Wiley Online Library, pp. 310–313.
- [191] SMPTE, *ST 2065-4, “ACES Image Container File Layout” Informative Notes: TB-2014-006*. [n.d.], [j.mp/TB-2014-006](http://j.mp/TB-2014-006). Accessed: 11-09-2022.
- [192] Song, H., Li, H. and Liu, X. [2018], ‘Studies on different primaries for a nearly-ultimate gamut in a laser display’, *Opt. Express* **26**(18), 23436–23448.  
**URL:** <http://opg.optica.org/oe/abstract.cfm?URI=oe-26-18-23436>
- [193] *ST 2065-1:2012 - SMPTE Standard - Academy Color Encoding Specification (ACES)* [2012], *ST 2065-1:2012* pp. 1–23.
- [194] *ST 2065-4:2013 - SMPTE Standard - ACES Image Container File Layout* [2013], *ST 2065-4:2013* pp. 1–36.
- [195] *ST 2067-2:2020 - SMPTE Standard - Interoperable Master Format — Core Constraints* [2020], *ST 2067-2:2020* pp. 1–43.
- [196] *ST 2067-50:2018 - SMPTE Standard - Interoperable Master Format — Application #5 ACES* [2018], *ST 2067-50:2018* pp. 1–28.
- [197] *ST 2084:2014 - SMPTE Standard - High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays* [2014], *ST 2084:2014* pp. 1–14.
- [198] *ST 429-2:2020 - SMPTE Standard - D-Cinema Packaging — DCP Operational Constraints* [2020], *ST 429-2:2020* pp. 1–31.
- [199] Stockman, A. [2019], ‘Cone fundamentals and cie standards’, *Current Opinion in Behavioral Sciences* **30**, 87–93.
- [200] Stockman, A. and Sharpe, L. T. [2000], ‘The spectral sensitivities of the middle-and long-wavelength-sensitive cones derived from measurements in observers of known genotype’, *Vision research* **40**(13), 1711–1737.
- [201] Strasburger, H. [2020], ‘Seven myths on crowding and peripheral vision’, *i-Perception* **11**(3), 2041669520913052.
- [202] *The Lego Batman Movie*, Copyright Warner Bros. Pictures. [2017].
- [203] *The Lego Movie*, Copyright Warner Bros. Pictures. [2014].
- [204] *Transformers: Revenge of the Fallen*, Copyright Paramount Pictures. [2009].
- [205] *Tricks of Our Trade (1957) from the Series The Magical World of Disney*, aired 1957 [n.d.], <https://www.imdb.com/title/tt0833037/>. Accessed: 17-02-2023.

- [206] Tsang, P., Poon, T.-C. and Wu, Y. [2018], ‘Review of fast methods for point-based computer-generated holography’, *Photonics Research* **6**(9), 837–846.
- [207] Universale, C. O. [1984], ‘Visual acuity measurement standard’, *Visual Functions Committee* .
- [208] *Universal Scene Description, Pixar* [n.d.], <https://graphics.pixar.com/usd/release/index.html>. Accessed: 11-09-2022.
- [209] Utsugi, T. and Yamaguchi, M. [2013], ‘Reduction of the recorded speckle noise in holographic 3d printer’, *Opt. Express* **21**(1), 662–674.  
**URL:** <http://opg.optica.org/oe/abstract.cfm?URI=oe-21-1-662>
- [210] *VFX Reference Platform, VES Technology Committee* [n.d.], <https://vfxplatform.com/>. Accessed: 11-09-2022.
- [211] Walton, D. R., Dos Anjos, R. K., Friston, S., Swapp, D., Akşit, K., Steed, A. and Ritschel, T. [2021], ‘Beyond blur: Real-time ventral metamers for foveated rendering’, *ACM Transactions on Graphics* **40**(4), 1–14.
- [212] Walton, D. R., Kavakli, K., Anjos, R. K. D., Swapp, D., Weyrich, T., Urey, H., Steed, A., Ritschel, T. and Aksit, K. [2022], Metameric varifocal holograms, in ‘2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)’, IEEE.  
**URL:** <https://doi.org/10.1109/vr51125.2022.00096>
- [213] Wang, Y., Chakravarthula, P., Sun, Q. and Chen, B. [2022], ‘Joint neural phase retrieval and compression for energy- and computation-efficient holography on the edge’, *ACM Transactions on Graphics* **41**(4).
- [214] Warnock, J. E. [1969], *A hidden surface algorithm for computer generated halftone pictures*, The University of Utah.
- [215] Weidlich, A., Forsythe, A., Dyer, S., Mansencal, T., Hanika, J., Wilkie, A., Emrose, L. and Langlands, A. [2021], Spectral imaging in production: Course notes siggraph 2021, in ‘ACM SIGGRAPH 2021 Courses’, SIGGRAPH ’21, Association for Computing Machinery, New York, NY, USA.  
**URL:** <https://doi.org/10.1145/3450508.3464582>
- [216] Whitted, T. [1979], An improved illumination model for shaded display, in ‘Proceedings of the 6th annual conference on Computer graphics and interactive techniques’, p. 14.
- [217] *Who Framed Roger Rabbit (1988)* [n.d.], <https://www.imdb.com/title/tt0096438/>. Accessed: 17-02-2023.
- [218] Wilkinson, P. T. [n.d.], ‘Holographic projection displays: Beyond star wars’.  
**URL:** <https://www.youtube.com/watch?v=-nNjc5BLV90>
- [219] Yamaguchi, M., Hoshino, H., Honda, T. and Ohyama, N. [1993], Phase-added stereogram: calculation of hologram using computer graphics technique, in ‘Practical Holography VII: Imaging and Materials’, Vol. 1914, SPIE, pp. 25–31.

- [220] Yamamoto, Y., Shimobaba, T. and Ito, T. [2022], ‘Horn-9: Special-purpose computer for electroholography with the hilbert transform’, *Optics Express* **30**(21), 38115–38127.
- [221] Yang, D., Seo, W., Yu, H., Kim, S. I., Shin, B., Lee, C.-K., Moon, S., An, J., Hong, J.-Y., Sung, G. et al. [2022], ‘Diffraction-engineered holography: Beyond the depth representation limit of holographic displays’, *Nature Communications* **13**(1), 6012.
- [222] Yang, F., Kadis, A., Mouthaan, R., Wetherfield, B., Kaczorowski, A. and Wilkinson, T. D. [2022], ‘Perceptually motivated loss functions for computer generated holographic displays’, *Scientific reports* **12**(1), 1–12.
- [223] Yang, L., Liu, S. and Salvi, M. [2020], A survey of temporal antialiasing techniques, in ‘Computer graphics forum’, Vol. 39, Wiley Online Library, pp. 607–621.
- [224] Yaraş, F., Kang, H. and Onural, L. [2011], ‘Circular holographic video display system’, *Optics Express* **19**(10), 9147–9156.
- [225] Yoshida, A., Ihrke, M., Mantiuk, R. and Seidel, H.-P. [2008], Brightness of the glare illusion, in ‘Proceedings of the 5th symposium on Applied perception in graphics and visualization’, pp. 83–90.
- [226] Young, T. [1804], ‘I. The Bakerian Lecture. Experiments and calculations relative to physical optics’, *Philosophical transactions of the Royal Society of London* (94), 1–16.
- [227] Zhang, H., Cao, L. and Jin, G. [2017], ‘Computer-generated hologram with occlusion effect using layer-based processing’, *Applied optics* **56**(13), F138–F143.
- [228] Zhang, H., Tan, Q. and Jin, G. [2013], ‘Full parallax three-dimensional computer generated hologram with occlusion effect using ray casting technique’, *Journal of Physics: Conference Series* **415**, 012048.  
**URL:** <https://doi.org/10.1088/1742-6596/415/1/012048>
- [229] Ziegler, R., Croci, S. and Gross, M. [2008], Lighting and occlusion in a wave-based framework, in ‘Computer Graphics Forum’, Vol. 27, Wiley Online Library, pp. 211–220.