Computational Projection Display

Towards Efficient High Brightness Projection in Cinema

by

Gerwin Damberg

Dipl.-Ing (FH), University of Applied Sciences, Karlsruhe, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Computer Science)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2017

Abstract

Cinema projectors need to compete with home theater displays in terms of image quality. High frame rate and high spatial resolution as well as stereoscopic 3D are common features today, but even the most advanced cinema projectors lack in-scene contrast and more importantly high peak luminance, both of which are essential perceptual attributes for images to look realistic. At the same time studies on High Dynamic Range (HDR) image statistics suggest that the average image intensity in a controlled ambient viewing environment such as cinema can be as low as 1% for cinematic HDR content and does not often exceed 18%, *middle gray* in photography. Traditional projection systems form images and colours by blocking the source light from a lamp, therefore attenuating on average between 99% and 82% of light before it reaches the screen. This inefficient use of light poses significant challenges for achieving higher peak brightness levels. We propose a new projector architecture built around commercially available components, in which light can be steered to form images. The gain in system efficiency significantly reduces the total cost of ownership of a projector (fewer components and lower operating cost) and at the same time increases peak luminance and improves black level beyond what is practically achievable with incumbent projector technologies. At the heart of this computational display technology is a new projector hardware design using phase-modulation in combination with new optimization algorithms for real-time phase retrieval. Based on this concept we propose and design a full featured projector prototype. To allow for display of legacy Standard Dynamic Range (SDR) as well as high brightness HDR content on light steering projectors we derive perceptually motivated, calibrated tone mapping and colour appearance models. We develop a calibrated optical forward model of the projector hardware and analyse the impact of content mapping parameters and algorithm choices on (light) power requirements.

Lay Summary

Movies look more appealing when they are brighter. In cinema, where the viewing environment is dark, small and bright highlights such as light sources and reflections or effects such as explosions are particularly important for directors to tell a story. Today's cinema projectors are limited in their ability to produce bright images. Why? Projection screens are large; projector light sources are expensive and hard to cool and projectors are inefficient in forming images: by blocking or wasting light not used in darker regions of an image.

This work analyses the problem and provides a solution: steering light that is not needed in dark areas of an image into bright areas improves the brightness of projectors dramatically and also allows for more visible detail in dark parts of a scene. We discuss and build several prototypes that show how this concept can be implemented with up to 20 times brighter highlights.

Preface

All publications that have resulted from or are related to the research presented in this work, along with the relative contributions of the collaborators are listed herein.

(1) **High Dynamic Range Projection Systems** - SID 2007 (Ref.: [19])
Authors: Gerwin Damberg, Helge Seetzen, Greg Ward, Wolfgang Heidrich, Lorne Whitehead
Contribution: As one of the early researchers at BrightSide Technologies Inc., a university start-up company, the author performed all of the research work discussed in the paper, including the image analysis, algorithms, and prototype work. The author wrote the first draft of the paper with select input from the co-authors. The high level concept of dual (amplitude) modulation in projectors was analogous to the Light Emitting Diode (LED) back light in the BrightSide HDR TV. Some of today's high-end cinema (Premium Large Format for Cinema (PLF)) projectors are based on the general architecture proposed in this paper.
The publication text is not included in this thesis, but results are summarized in Chapter 2.

(2) **Comparing Signal Detection Between Novel High-Luminance HDR and Standard Medical LCD Displays** - Journal of Display Technology 2008 (Ref.: [106])
Authors: M. Dylan Tisdall, Gerwin Damberg, Paul Wighton, Nhi Nguyen, Yan Tan, M. Stella Atkins, Hiroe Li, Helge Seetzen
Contribution: The research work described in the paper was jointly performed by Dr. Tisdall and the author whose work was focussed on the display hardware and image processing algorithms for the study, whereas Dr. Tisdall's work was focussed on preparing the Magnetic Resonance Imaging (MRI) image data and the study stimuli. The user study was executed jointly. The display prototype was built by the author. This user study in the field of medical imaging was suggested by Dr. Atkins and coordinated by the author.

The publication text is not included in this thesis, but results are summarized in Section 2.

(3) **A High Bit Depth Digital Imaging Pipeline for Vision Research** - APGV poster 2011 (Ref.: [58])
Authors: Timo Kunkel, Gerwin Damberg, Lewis Johnson
Contribution: The poster publication was initiated by Dr. Kunkel based on joint work with the author. The author's contribution are the development of Matlab scripts to incorporate the display experimentally into the PsychoPhysics Toolbox for vision research using the High Definition Serial Digital Interface (HD-SDI). Dr. Kunkel and the author prepared the poster. Mr. Johnson assisted in porting the Serial Digital Interface, SMPTE 292M (SDI)-card driver code.
This poster was presented by Dr. Kunkel at the Applied Perception in Graphics and Visualization (APGV) symposium in 2011. A modified version of the text is incorporated in Section 4 as it ties together the author's work on tone mapping and colour appearance (Section 3) and today's Society of Motion Picture & Television Engineers (SMPTE) HDR encoding and transmission standards. The research work on the universal mapping function referenced in Section 4.2 was initiated by the author and its final form was developed and published by Anders Ballestad and Andrey Kostin (of Dolby Canada) and further modified in the course of adoption in published video encoding standards by other Dolby employees.

(4) **Calibrated Image Appearance Reproduction** - Siggraph Asia 2012 (Ref.: [93])
Authors: Erik Reinhard, Tania Pouli, Timo Kunkel, Ben Long, Anders Ballestad, Gerwin Damberg
Contribution: As a Senior Research Engineer at Dolby Canada Research, the author's responsibilities included establishing new University relationships and guiding industry-relevant research work. This paper was a close collaboration between the author's research group at Dolby and the University of Bristol, UK. The author suggested and coordinated the work on combined colour appearance and tone mapping that is described in this paper and together with the first author, Dr. Reinhard formulated the need for a simple Colour Appearance Model (CAM) in the form of a tone mapping operator that functions over a large luminance range and omits the backward step of classical CAMs: the forward-only colour appearance model. Dr. Reinhard de-

rived the first concept of the model. Both Dr. Kunkel and the author implemented and evaluated early Matlab versions of the model. Dr. Kunkel focussed part of his PhD work on a variant of the model. Dr. Reinhard and Dr. Pouli later added the local adaptation parts of the model and wrote the first draft of the joint paper.
A version of this paper is included in Chapter 3 of this document.

(5) **State of the Art in Computational Fabrication and Display** - Eurographics 2013 (Ref.: [47])
Authors: Matthias Hullin, Ivo Ihrke, Wolfgang Heidrich, Tim Weyrich, Gerwin Damberg, Martin Fuchs
Contribution: In this high level State of the Art paper, the author provided a draft of the section on novel display devices.
The publication text is not included in this thesis, but parts are summarized in Section 2.

(6) **Efficient Freeform Lens Optimization for Computational Caustic Displays** - Optics Express 2015 (Ref.: [18])
Authors: Gerwin Damberg and Wolfgang Heidrich
Contributions: All of the research work and prototyping was performed by the author. The research topic as well as the approach was jointly suggested by the author and Dr. Heidrich. The physical lenses were designed by the author and 3D printed and polished by Dr. Heide who had access to a suitable 3D printer at the time.
A version of this publication is included in Sections 6.3 to 6.5 of this document.

(7) **High Brightness HDR Projection Using Dynamic Freeform Lensing** - Transactions on Graphics 2015 - presented at SIGGRAPH 2016 (Ref.: [17])
Authors: Gerwin Damberg, James Gregson, Wolfgang Heidrich
Contribution: The author performed all experimental work, developed the basic algorithm framework, implemented the image statistics work and wrote the first draft of the paper. Dr. Gregson later developed and implemented a real-time version of the algorithm which was included in the paper and is described in Section 6.6.3. Dr. Ballestad mapped the HDR images used in the HDR power survey. Dr. Ballestad and the author jointly initiated the image statistic research.
Parts of the paper are incorporated into Section 6.6 and into Section 5.3 of this document.

(8) **Temporal Considerations and Algorithm Architecture for Light Steering Projectors** - MTT Innovation Incorporated (MTT) Internal Reports 2016
Authors: James Gregson, Eric Kozak and Gerwin Damberg
Contributions: The work related to the RGB prototype was performed by the research team including the author (CTO) at MTT Innovation Inc., a UBC collaboration partner and demonstrated internally and at SIGGRAPH [15, 16]. The overall colourimetric calibration approach was provided by the author. Dr. James Gregson developed and implemented the specific calibration routines including capturing and digitizing the Point Spread Function (PSF) and non-steered components. Versions of the demo code were written by the author, Raveen Kumaran and James Gregson, who also provided the first draft of the algorithm write-up. The initial prototype hardware was conceptualized and prototyped by the author and significantly refined by the MTT research team, notably Raveen Kumaran, Johannes Minor, Erik Kozak and James Gregson. The temporal synchronization schemes were jointly developed and documented by the author and Eric Kozak.

Prototype Demonstrations

(9) **HDR Projector With Improved Contrast** - SIGGRAPH Emerging Technologies 2008 (invited)(Ref.: [20])
Authors: Gerwin Damberg, Peter Longhurst, Michael Kang

(10) **Light Steering Projector Monochromatic Proof of Concept** - SIGGRAPH Emerging Technologies 2014 (Ref.: [15])
Authors: Gerwin Damberg, Anders Ballestad, Erik Kozak, Johannes Minor, Raveen Kumaran

(11) **High Brightness HDR Projection Using Dynamic Phase Modulation** - SIGGRAPH Emerging Technologies 2015 (invited)(Ref.: [16])
Authors: Gerwin Damberg, Anders Ballestad, Erik Kozak, Johannes Minor, Raveen Kumaran, James Gregson, Wolfgang Heidrich

Patents and Patent Applications

(12) **Dynamic Freeform Lensing With Applications To High Dynamic Range Projection** - Provisional Patent Application US62007341
Inventors: Gerwin Damberg, Wolfgang Heidrich

(13) **Efficient, Dynamic, High Contrast Lensing with Applications To Imaging, Illumination and Projection** - PCT patent application CA2015050515
Inventors: Gerwin Damberg, James Gregson, Wolfgang Heidrich

(14) **Light Detection, Color Appearance Models, and Modifying Dynamic Range for Image Display** - US WO 2010/132237
Inventors: Gerwin Damberg, Erik Reinhard, Timo Kunkel, Anders Ballestad

(15) **Image Processing and Displaying Methods for Devices that Implement Color Appearance Models** - US WO 2010/083493
Inventors: Timo Kunkel, Erik Reinhard, Gerwin Damberg

Table of Contents

List of Tables

List of Figures

List of Acronyms

P

Q

R

S

T

U

Acknowledgements

I would like to thank the many people including past and present colleagues, friends and family that have in one way or another supported me throughout my studies at UBC and during my work in the start-up MTT Innovation Inc.

My research supervisor Dr. Wolfgang Heidrich has been a great inspiration throughout my graduate studies by providing just the right amount of guidance. He is one of the smartest people that I have gotten to know in this field and he made it possible for me in the first place to combine academic research with innovative and practical technology development at the start-up. I would further like to thank Roger Miller at the University Liaison Office for working with Wolfgang and me on the technology transfer of some of my research results at UBC to MTT.

I am thankful for many fruitful discussions with my lab mates James, Felix, Matthias, Evan, Lei, Brad, Shuochen and Qiang as well as with my academic friends and collaborators Karol, Erik, Timo, Greg and everyone else I had the pleasure of meeting over the years.

None of the momentum of maturing the light steering technology could have been maintained without the hard working and fun co-workers at MTT: my co-founders Anders and Eric as well as Johannes, Raveen, James, Sam, Eric, Andra, Ronan, Thomas and everyone else I got to meet and work with during this exciting time.

Finally I want to thank my parents for their continuous support and encouragement throughout the years.

To Astrid for her incredible love, patience and support during my work and studies; to Johannes, Philip and Mathea for encouraging me with playful enthusiasm to take enough breaks from the same.

Chapter 1

Introduction and Structure of the Thesis

1.1 Introduction

Display technologies have evolved increasingly fast over the past decades, always with the goal of providing the most realistic looking images, given cost and technology constraints. Advances in new display materials, faster and lower power semi-conductors, solid state illumination technologies as well as more powerful computational hardware have motivated the emergence of a new research discipline: computational displays. This field in computer science lies at the intersection of display optics, mathematical analysis and physical modelling, efficient computational processing, and, most importantly, visual perception. Computational displays aim to provide a visual experience beyond the capabilities of traditional systems by adding computational power to the display architecture.

While a variety of research concepts from the field have made their way into commercial display products, large screen projectors have predominantly been carved out from practical innovations in computational display. For realistic image appearance, arguably the most important visual property of a display system is the range and number of light levels and colours that can be displayed. Unfortunately, increasing this range significantly in projectors is prohibitively expensive, because peak luminance scales linearly with display power and light source cost, while brightness perception of luminance values is near logarithmic.

This thesis introduces our analysis and understanding of light requirements in the cinema production pipeline from capture and production on a reference display, encoding and distribution to display on a projector and finally light perception by the audience. We explore tone mapping and colour appearance in HDR image reproduction. We propose a new mathematical framework and a simple model to accurately map HDR images between a reference display and the projector and explore means to transmit video data

from source to display in a bit-efficient manner. Based on this perceptual understanding we propose and prototype a new, steerable light source and projector architecture based on phase modulation to efficiently achieve the required luminance and colour range for life-like images in cinema. Finally, we assemble and analyse HDR image statistics for theatrical high brightness HDR content. We work under the hypothesis that a much more optimized projector design can be derived based on understanding the light requirements for large screen cinema.

1.2 Contributions and Outline of the Dissertation

The scope of work presented herein aims at addressing the remaining bottlenecks in the current HDR pipeline in cinema including content creation (colour and tone mapping), content delivery and display (high brightness HDR projection) and finally the understanding of visual perception of light levels by the observer. Figure 1.1 provides a high-level overview of the scope of the Dissertation.

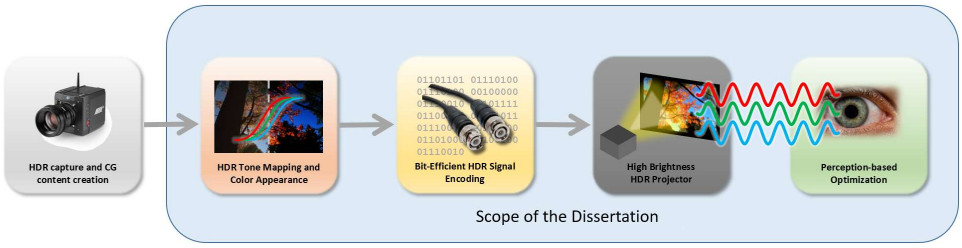


Figure 1.1: *Scope of the proposed work to address the major bottlenecks in the current cinematic pipeline to enable HDR. While HDR content capture/creation falls outside the scope of the work, we do perform statistical analysis on scene referred (camera captured) and display referred (colour corrected for presentation in a cinema) HDR content in order to better understand the requirements for an ideal HDR projector.*

The following overview outlines the remaining chapters of the thesis.

Chapter 2. Background and Related Work. We review basic concepts of high contrast projection together with the state of the art in computational displays as it relates to light steering, colour appearance models and tone reproduction operators.

Chapter 3. Visual Perception and Colour Appearance. This chapter introduces a new, combined colour appearance model and tone mapping operator that aids in mapping HDR content for arbitrary displays and viewing environments. The model, in a modified form was adapted in the SMPTE ST 2084:2014 standard that discusses the HDR Electro-Optical Transfer Function (EOTF) for reference displays.

Chapter 4. Digital Imaging Pipelines and Signal Quantization. In this chapter we discuss experiments related to utilizing video links with higher bandwidth to allow for artifact free psycho-physical vision experiments as well as encoding schemes that allow bit efficient quantization of video signals that represent the luminance range of modern HDR displays.

Chapter 5. Image Statistics and Power Requirements. In this chapter we discuss light steering projector architectures suitable for cinema taking into consideration cost, power levels and performance. We propose a new hybrid architecture consisting of a steered and a non-steered light path based on our analysis of HDR image data.

Chapter 6. Freeform Lensing and HDR Projector Proof of Concept. A new phase retrieval algorithm based on freeform lensing as well as our work on a new light steering HDR projector architecture concept is introduced.

Chapter 7. RGB Projector Prototype. We discuss the research and development that lead to a full featured HDR projector prototype, including suitable RGB laser light sources, temporal considerations of optically pulsed components in the light path and the optical forward model with algorithm implementations to control the system.

Chapter 8. Discussion and Conclusion. The contributions of this thesis are summarized and future directions of research are indicated.

Chapter 2

Background and Related Work

Our research draws from and builds on a number of different fields of related work, including display technologies and algorithms for freeform lens design, human visual perception, colour appearance and image tone reproduction operators. The following is a brief description of the state of the art in these related fields.

2.1 Spatial Light Modulators

Throughout this work we make use of Spatial Light Modulators (SLMs) for projection systems. In modern projectors these SLMs typically consist of pixelated micro displays in which each pixel can be electronically addressed to cause modulation of light. This section provides an overview of the key characteristics of some of the different types of common light modulators.

2.1.1 Spatial Resolution and Pixel Dimensions

Today, typical micro display spatial resolutions for consumer applications provide 1920 × 1080 pixels (High Definition (HD)) or 3840 × 2160 pixels (Ultra-High-Definition (UHD)) of spatial resolution [110]. In digital cinema, to accommodate a larger horizontal to vertical aspect ratio, the corresponding resolutions are 2048 × 1080 (Digital Cinema Initiative 2K standard (DCI2K)) and 4096 × 2160 (Digital Cinema Initiative 4K standard (DCI4K)) respectively [23]. Today, the dimension of a typical individual pixel can be between 1µm and 8µm in a projector [50]. There is a trend towards smaller micro displays, higher pixel count and smaller pixel pitch. Low power applications (e.g. mobile projectors and near-to-eye projectors) can utilize small micro displays. Higher power applications (e.g. conference room, large venue or cinema projectors) tend to require larger micro displays to manage heat dissipation and accommodate higher power light sources with higher divergence of the light within the optical system [108].

The micro displays that were used in this work have a spatial resolution of 1920 × 1080 and a pixel pitch of 8µm leading to active area display dimensions of 15.36 $mm\times$ 8.64 mm [42].

2.1.2 Types of Spatial Light Modulators

Conceptually, image formation in traditional projectors begins with a full white screen (Full Screen White (FSW)) created by a light source that uniformly illuminates a SLM (with the help of beam homogenization optics). Then, to create colours and tones of the image, light is attenuated on a per-pixel basis. A number of micro display technologies exist for this purpose. The three commercially most successful technologies are transmissive Liquid Crystal Display (LCD) or High Temperature Poly-Silicon (HTPS), reflective Digital Micromirror Device (DMD) and reflective Liquid Crystal on Silicon (LCoS). A comprehensive discussion of the nuances of each of the technologies is provided in [2] and [38].

Transmissive Liquid Crystal Displays. In the early days of digital projectors most systems were based on transmissive LCD technology which requires linearly polarized light at the input (e.g. with a polarization filter following the light source). The LCD can rotate the state of the polarization of light on a per-pixel basis (by applying an electric field across a cell filled with Liquid Crystal (LC) material). The per-pixel (polarization) modulated light then passes through another linear polarization filter (typically at 90° to the orientation of the first filter). Light at pixels with rotated polarization passes through the second polarization filter. Light at pixels with unchanged polarization does not pass the polarization filter (it gets absorbed). Light that has been partially rotated will partially pass the second filter and get partially absorbed by the second polarization filter. Switching the individual pixels at video rates requires active electronic components at each pixel, Thin Film Transistors (TFTs) as well as the required wiring to control them. These components within the clear or transmissive area of the display cause a limited pixel fill-factor for HTPS-based micro displays as well as a visible pixel structure on the projection screen (often referred to as the *screen door effect*).

Reflective Digital Micromirror Devices. In the late 1980's Texas Instruments Inc. (TI) developed the DMD technology, consisting of a Microelectromechanical System (MEMS)-based pixelated array of micro mirrors that can tilt into two distinct positions (typically across the diagonal of each

square pixel mirror) and either reflect incoming light towards a projection screen or into a *light dump* (heat sink) to absorb the light. The typical angle between the two states of each mirror is between $\pm 10°$ and $\pm 14°$. A large angle between the two mirror states enables a higher contrast in the resulting image due to higher separation of the incoming and outgoing illumination beams. Given the binary nature of the display device DMDs can only form binary image patterns. However, the resulting binary images can be updated at significantly higher speeds (between $20kHz$ and $80kHz$) compared to for example LCD (typically $30Hz$ to $240Hz$) and thus grey scale can be achieved with a binary/digital modulator. Benefits of DMDs include a relatively high pixel fill factor, the high degree of repeatability due to the digital drive scheme and well-defined mirror states, high reflectivity, and good power handling. Most cinema projectors are based on three DMD chips (one for each of the red, green and blue colour channels) and typical out-of-lens optical power is between $20,000$ and $60,000$ lumens. A draw back of DMD technology is the relatively high cost per device. In the consumer market projectors containing three DMDs are sparse. Lower cost single chip Digital Light Processor (DLP) projectors on the other hand use field sequential colour schemes which can cause image artifacts (or image deficiencies) related to colour break-up of the sequential colour image fields (also referred to as the *rainbow effect*) [2, 105].

Reflective Liquid Crystal on Silicon Displays. LCoS technology was developed to overcome some of the shortcomings of transmissive LCD projectors as well as the high cost of DLP technology predominantly in the rear-projection Television (TV) market. However with the rapid introduction of large, flat panel TVs based on plasma and LCD technology in the early 2000s there was less of an addressable market for the technology. A number of large companies in the semiconductor field (such as Intel, Thompson and Philips) as well as many small companies discontinued their investments into the development of LCoS technology. Other companies (such as SONY and JVC) pivoted and continued development of the technology for niche markets for example high-end home theatre projectors and cinema.
In LCoS technology a LC layer is sandwiched between a cover glass that acts as a global electrode and a silicon backplane with individually addressable electrodes. Similar to transmissive LCD technology a voltage across the LC cell rotates the crystals and with that changes the state of polarization of incoming light. Linearly polarized light initially passes through the LC layer and reflects off the backplane before passing through the LC layer a second

time to exit the display. Due to the light passing through the LC layer twice, the gap thickness of the LC cell can be designed to be small which in turn allows for a fast (polarization) switching speed and high frame rates. Instead of a transmissive linear polarization filter, LCoS-based devices employ a Polarizing Beam Splitter (PBS) positioned at 45° to the surface of the micro display. A PBS passes polarized light of one linear polarization (P-polarized light for *parallel*) and reflects (at 90°) light of the other linear polarization (S-polarized light for German *senkrecht* = perpendicular). LCoS devices can be produced at lower cost compared to DMDs. There are no moving parts and the backplane manufacturing process is comparable to the process used to produce semiconductors (e.g. Application-Specific Integrated Circuits (ASICs)). Other benefits include the high switching speed, a high pixel fill factor (93% and higher: all pixel switching circuitry is *hidden* behind the reflective electrodes) and high contrast optical systems that can be designed around the micro display [2].

Of the three technologies introduced here LCoS technology provides the highest native contrast (e.g. $50,000:1$) compared to DMD (e.g. $2,000:1$) and HTPS (e.g. $1,000:1$).

2.1.3 Amplitude and Phase-Only LCoS Modulators

The previous section discusses typical projection SLMs that attenuate light on a per-pixel basis to form images. In this thesis we introduce the use of *programmable lenses* to create a more efficient projection system (see Chapter 6). Similar to static lenses (for example a spherical glass lens) we would like to achieve a wavefront distortion effect to focus or spread light. The change of the state of polarization in a traditional LCoS display display always comes coupled with a retardation of the phase (in other words a distortion of the wavefront) of light. It is possible to design the LC material and the alignment layer of an LCoS device in such a way that the display can be operated in phase-only mode [31, 41, 61–63, 113]. Incoming light should be linearly polarized but the state of polarization will not change upon reflection off the micro display. In this configuration no PBS is required and hence light is not absorbed. Similar to the example of a static glass lens, this dynamically addressable phase retardation is accomplished by adjusting the effective refractive index along the light path in the LC material at each pixel. The change in refractive index is possible due to the use of a LC material with non-zero birefringence. In birefringent material the refractive index of the material depends on the polarization of light that

passed thought it. Birefringence can be quantified as the difference between refractive indices of a material and is defined as:

$$\Delta_{LC}n = n_e - n_o \tag{2.1}$$

where n_e is the extraordinary refractive index for incident light parallel to the preferred orientation direction of the LC material (the *director*) and n_o is the ordinary refractive index for light perpendicular to the director. A typical LC has a positive birefringence between 0.05 and 0.45 (see [21, 54, 114]). The amount of phase retardation of light as it passes through the LC material is given by:

$$\delta_{LC} = \frac{2\pi * \Delta n * d}{\lambda} \tag{2.2}$$

where d is the thickness of the LC material and λ is the wavelength of incident light. Given that an LCoS device is reflective and light passes through the LC material twice, the total phase retardation can be expressed as:

$$\delta_{LCoS} = \frac{2\pi * \Delta n * 2d}{\lambda} \tag{2.3}$$

For programmable lens and holographic applications it is often desirable to allow for a maximum phase retardation of exactly one wavelength (2π) for example for *phase-wrapping* of a phase function (see for example Figure 6.11). In this case the thickness of the LC device should be:

$$d = \frac{\lambda}{2 * \Delta n} \tag{2.4}$$

For broadband light device and material properties can be optimized around a center wavelength λ_c. Light with shorter or longer wavelength compared to λ_c will refract more or less which can cause wavelength dependant image magnification or demagnification resulting in chromatic aberration (see for example Figure 6.7).

Finally, in phase-only LCoS devices the pixel pitch plays an important role. A small pixel pitch is generally desirable as it allows for larger diffraction angels. However, given that the LC material thickness remains constant for a given desired phase retardation a small pixel pitch can cause inter-pixel

crosstalk (i.e. a mixed phase response) between individual pixels which can be challenging to predict.

We refer the interested reader to [117] and [98] for further details on phase-only LCoS devices and topics related to polarization engineering for projection applications.

2.2 Contrast Metrics

When comparing the contrast performance between different projectors it is important to differentiate between sequential contrast and in-scene contrast. The sequential contrast is a good indication of the native micro-display contrast, while the in-scene contrast is a much more meaningful contrast measure when judging the image quality of natural images as it takes into account the entire system's optics including for example scatter in the projection lens.

2.2.1 Sequential Projector Contrast

As the name suggests the sequential contrast is typically calculated from a set of sequential light measurements of the maximum (full white screen) and minimum (full black screen) amount of light that can be displayed by the projector. The sequential contrast serves as a good approximation of the native micro-display contrast, provided the projector light source is not globally dimmed/attenuated (see Section 2.2.1). While early video projectors based on liquid crystal micro displays were limited to on the order of 500:1 sequential contrast, today's projectors based on MEMS based devices (for example TI's DLP technology) or LCoS displays such as Sony's Silicon X-tal Reflective Display (SXRD) branded technology achieve on the order of 1,000:1 to 6,000:1 sequential contrast. JVC continues to push the envelope achieving on the order of 30,000:1 sequential contrast using their flavour of LCoS technology called Direct Drive Image Light Amplifier (D-ILA) [9]. The sequential contrast is an important data point in understanding a projector, but it is less relevant when predicting projector contrast performance for the reproduction of natural images.

Dynamic Irises and Global Light Source Dimming. The sequential contrast of a projector can be increased when the light source is dimmed or turned off entirely. High-pressure discharge lamps that are common in most projectors can neither be electrically dimmed by a large amount nor

can they be dimmed rapidly (e.g. at video frame rates). In order to globally modulate the lamp intensity, one or more dynamic irises are controlled in a content-dependant fashion. Similarly, with solid state light sources replacing lamps in more recent projector designs, an LED or laser array's global intensity can be dimmed electronically based on content. In overall bright scenes, the iris will be fully open (or the solid state light source driven at full intensity), whereas dark scenes will cause the light source to be off or dark for better (=darker) black level. The dynamic iris feature helps boost the sequential contrast metric for test patters, but provides little visual benefit for any but the darkest images - provided no image feature (including single white pixels) requires the peak luminance level. In this case, the iris will either remain fully open, which leads to an elevated black level, or the luminance of the bright image feature will be reduced by the amount that the light source is dimmed [45, 59]. In product marketing material the quoted sequential contrast numbers often exceed 1,000,000:1, in this context a rather meaningless measure for real image content.

2.2.2 In-scene Projector Contrast

For natural images, the in-scene contrast is a much more meaningful contrast measure. Here, a bright and a dark feature are measured simultaneously within the same image. There are several common test patterns, for example the American National Standards Institute (ANSI) 4x4 checker board contrast consisting of a grid of alternating white and black rectangles. The in-scene contrast, or simultaneous contrast, takes into account the entire optical path within the projector. The native micro display contrast, inter-pixel light scatter, scatter within the light path caused by reflections off of optical elements, undesired divergence of light within the projector as well as the optical quality of the projection lens all affect the in-scene contrast.

Aside from the native contrast of the micro display, the in-scene contrast of a projector is largely affected by the overall amount of light in an image (the mean luminance of an image). This is due to scattered light within the projector that raises the achievable black level more in brighter images compared to darker ones. In the case of for example a small white feature centered on a black background, a high in-scene contrast which is comparable to the sequential contrast can be measured, whereas a larger bright feature towards the edge of the same black image would cause a reduced in-scene contrast. In the example of the JVC projector (30,000:1 sequential contrast), the in-scene contrast for a dark image might be close to 30,000:1, whereas

the in-scene contrast for a bright scene (e.g. ANSI 4x4 checker board, with 50% mean luminance) would be closer to 500:1 due to scatter in the optics past the micro display. We refer to Section 5.3 for a first estimate of average image intensity for HDR cinema and note that for overall dark scenes a black level of 100,000× or less below the peak luminance can be appreciated by the observer.

To compare image quality of different HDR projectors the in-scene contrast of a test pattern with mean luminance comparable to target image content, or a selection of real images can be used. The capability of a projector to reproduce bright highlights is important for both, overall dark and overall bright scenes [13, 14, 118].

2.3 Dual Modulation Projection Displays

Over the last two decades, there have been several different proposals to implement dual modulation approaches in display applications to increase in-scene contrast. The availability of large flat panel TVs and high power LEDs led to an adoption of dual modulation techniques in consumer electronics [101]. Similar concepts to increase contrast in projectors include screens with spatially varying reflectivity (either statically [7] or dynamically [102]), and arrays of hundreds or even thousands of primitive projectors [102], proposed as means to increase on-screen luminance. Few of these concepts have made it past the research stage and small-scale prototypes. One exception are dual modulation projector designs using two traditional amplitude SLMs in sequence [8, 19, 20, 59]. These systems are typically intended for specialty applications requiring good black level and limited peak luminance. The low optical efficiency of amplitude SLMs results in both a low light intensity on screen and high power consumption, all at significantly increased system cost. Nevertheless dual amplitude attenuating projectors are being deployed not only in planetariums, training and simulation applications, but recently also in high end, *premium large format*, cinema.

To alleviate the problem of inefficient image formation, Hoskinson et al. [44, 46] introduced the notion of *light reallocation* using 2D arrays of tip-tilt mirrors in the light path of a small DLP projector, whereby the first modulator does not actually absorb much light, but moves it around within the image plane, so it can be reallocated from dark image regions to bright ones, essentially creating moving, bright spots of approximately constant size on the amplitude modulator. Hoskinson and co-authors used a continuously

tilting micro-mirror array to achieve this light reallocation. Unfortunately such mirror arrays are not easy to control accurately (achieving predictable tile-angles for a given drive signal) and are still only available as research prototypes at low spatial resolution (7×4 pixels in their work).

2.4 Holographic Displays

Holographic image formation models (e.g. [65]) have been adapted to create digital holograms [39] quite early in the history of phase SLMs. Holographic projection systems have been proposed in many flavours for research and speciality applications including projectors [10]. Some projection systems use diffraction patterns addressed on a phase SLM in combination with temporally and spatially coherent light for image generation. The challenges in holography for projectors lie in achieving sufficiently good image quality, the limited diffraction efficiency, often due to binary phase modulators [10], and the requirement for a Fourier lens, often resulting in an undesired, bright Direct Current (DC) spot within the active image area (zero-order diffraction, which is hard to eliminate completely).

2.5 Freeform Lenses

Recently there has been increased interest in freeform lens design, both for general lighting applications (e.g. [74]) and for goal-based caustics [6, 47]. In the latter application, we can distinguish between discrete optimization methods that work on a pixelated version of the problem (e.g. [83, 84, 116]), and those that optimize for continuous surfaces without obvious pixel structures (e.g. [33, 53, 87, 100, 115]). The current state of the art methods define an optimization problem on the gradients of the lens surface, which then have to be integrated up into a height field. This leads to a tension between satisfying a data term (the target caustic image) and maintaining the integrability of the gradient field.

Lens and Phase Function Equivalence. The effects of phase delays introduced by a smooth phase function can be related to an equivalent, physical refractive lens under the paraxial approximation, which can be derived using either geometric optics or from the Hyugens principle. The paraxial approximation is widely used in optics and holds when $\sin\theta \approx \theta$. For the projection system considered in this thesis, $|\theta| \leq 12°$, which corresponds to redirecting light from one side of the image to the other.

The error in the paraxial approximation is less than 1% for this case, which makes optimizing directly for the phase surface possible.

2.6 Tone Mapping Operators and Colour Appearance Models

Image reproduction has a history dating back to the start of photography, leading to significant advances in print reproduction [1], colour reproduction [48] and digital image reproduction [91]. The aim of image reproduction is to offer imagery for human consumption that looks natural, realistic as well as appealing and, above all, correct. It has long been understood that human visual perception plays a key role, and should therefore be taken into account in any image and video reproduction systems [48].

CAMs, for instance, focus on the perception of various attributes of colour, including hue, lightness and colourfulness [27]. Of particular importance for our work is lightness, which in colour science is loosely defined as the impression of how much light a patch of colour appears to emit, relative to a patch of colour perceived as *white*. In psychology, lightness is extensively studied as well, and it is found that its perception depends crucially on the presence of light sources, their intensities, distances as well as sizes ([37] and references therein). It is therefore a complex spatial phenomenon, not yet extensively explored in either colour appearance modeling or high dynamic range imaging.

The CIECAM02 colour appearance model is the most recent industry standard, adopted by the Commission International de l'Éclairage (CIE) [76]. CAMs can be used to adapt the tristimulus value of a patch of colour for observation under different viewing conditions. This is commonly achieved by operating the model in reverse, substituting parameters describing the new viewing environment. Although these models are accurate in their prediction of colour appearance, they are validated only under certain conditions, specifically for a limited range of illumination levels. One reason is that these models are derived from psychophysical data collected for a limited range of luminances [69]. A second reason is that by applying a CAM in both forward and reverse mode, significant discrepancies in lighting conditions are difficult to account for [92].

This can be remedied by deriving spatially varying colour appearance models, known as image appearance models [29, 56, 86]. These models are often closely related to CIECAM02, but include a spatially varying component to better deal with dynamic range mismatches.

In the area of dynamic range reduction, image appearance models are the exception in that their design naturally includes colour management. Tone reproduction operators are specifically designed to deal with large differences between the dynamic range of an image and that of the target display. Global operators compress the image by applying a single function to each pixel [24, 32, 109, 112]. The shape of this function is often sigmoidal [90, 94] or close to sigmoidal (see Chapter 7 of Reinhard et al. [95] for a discussion). Note that such response functions are known to describe photoreceptor output well [111].

Local operators add a spatially varying compressive function to take pixel neighbourhoods into consideration. This often takes the form of a Gaussian filter to approximate locally adaptative processes [12, 94], may use stacks of band-pass filters [67], or possibly employ edge-preserving smoothing operators that help minimize haloing artifacts [25]. They can also serve the purpose of local contrast management [30]. To our knowledge, only one tone reproduction operator is spatially variant by taking inspiration from lightness perception [55]. This method segments the image into separate regions, called frameworks, which each have a significantly different average luminance. This allows each framework to be treated semi-independently. Our model is also inspired by lightness perception, although we embed our technique into a colour management algorithm rather than a tone reproduction operator, allowing us to simultaneously manage lightness perception, dynamic range as well as colour appearance.

Chapter 3

Visual Perception and Colour Appearance

Managing the appearance of images across different display environments is a difficult problem, exacerbated by the proliferation of high dynamic range imaging technologies. Tone reproduction is often limited to luminance adjustment and is rarely calibrated against psychophysical data, while colour appearance modeling addresses colour reproduction in a calibrated manner, albeit over a limited luminance range. Only a few image appearance models bridge the gap, borrowing ideas from both areas. Our take on scene reproduction reduces computational complexity with respect to the state-of-the-art, and adds a spatially varying model of lightness perception. The predictive capabilities of the model are validated against all psychophysical data known to us, and visual comparisons show accurate and robust reproduction for challenging high dynamic range scenes.

3.1 Introduction

Traditional imaging pipelines are designed around the abilities of conventional capture and display devices, and therefore do not handle dynamic range beyond what can be represented with a single byte per pixel per colour channel. Light in the world around us cannot be well represented by such highly quantized values. This has led to the development of a collection of capture, processing and display technologies that are collectively termed HDR imaging [72, 95]. In particular capture and display hardware technologies are rapidly maturing [103, 107], opening up new opportunities in entertainment and broadcasting, but bringing new demands on encoding, storage and especially on display.

For instance, dynamic range reduction [77, 95] or expansion operators [5] address luminance mismatches between image and display, and may take display capabilities into account [71]. Many of these techniques offer sophisticated mechanisms and models to handle the extensive luminance range of

HDR, often inspired by aspects of the visual system. More often than not however, tone reproduction operators treat colour as a separate modality, and with much less precision. Usually, post-processing steps are applied to reduce the saturation of colours [70, 99]. Although this may lead to satisfying results in many cases, they are not sufficient for accurately modelling the appearance of colours under different conditions. For instance, a colour timer preparing a movie for the cinema will need to take into account the specific viewing conditions likely to occur (dark room, large display, varying viewing distances). On the other hand, when preparing the same movie for home viewing, to ensure that the viewer will have the same experience as the cinema goer, different viewing conditions are considered.

CAMs may be employed to predict the appearance of a given colour under different conditions (which are specified as inputs to the algorithm) [27]. As opposed to tonemapping, most CAMs are designed with a focus on colour and less so on dynamic range, making them less appropriate for dealing with HDR data as they offer little compression. Although CAMs offer high predictive power for single patches of colour, spatial relations in an image can greatly affect the appearance of colours, requiring spatially varying image appearance models instead [56].

Another mismatch between existing tonemapping and colour appearance solutions is the treatment of the viewing environment. In the former, display capabilities are rarely taken into account, albeit with one notable exception [71]. In CAMs on the other hand, the room illumination is taken into account in a symmetric manner to the original scene environment. Mixed adaptation to both room and display is not considered in any model that we are aware of.

3.2 Contributions

We therefore propose a novel, fully calibrated model for reproducing the appearance of images under a wide range of different scene and display/viewing conditions. As an example, the image in Figure 3.1 is processed for several different viewing and display conditions, leading to images that, when seen under the corresponding conditions, will have the same appearance. Algorithmically, the novelty of our approach is, that rather than matching the input to the visual system between scene and view environment, our algorithm crucially achieves relative computational simplicity by matching an intermediate state of visual processing, namely photoreceptor responses. We borrow from lightness perception to implement a spatially varying response

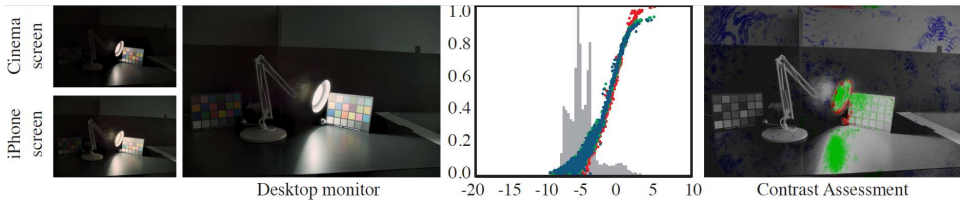


Figure 3.1: *Our appearance reproduction model faithfully reproduces colour, taking into account different viewing environments and display types as shown here for different combinations of viewing environment and display. The plot shows the image histogram (grey) as well as the input/output mapping (red, green and blue) for this particular image. The right panel shows that contrast is accurately reproduced for most pixels (shown grey), as determined by the dynamic range independent image quality metric [3].*

to the input, leading to a natural reproduction of scene appearance.

The algorithm is uniquely matched against all psychophysical corresponding colour and colour appearance data known to us. This allows us to prepare images as well as video for observation under known viewing conditions, as well as predict appearance correlates. We believe that this algorithm brings together best practices from research in colour appearance modeling, tone reproduction research as well as knowledge from human lightness perception, leading not only to calibrated, but also visually pleasing results. In summary, we offer the following contributions:

- We describe a calibrated and extensively evaluated global model that can match image appearance over a wide range of conditions and displays.
- We derive correlates for appearance characterization using a significantly simpler formulation than any existing models and achieve comparable predictive performance.
- We propose a novel formulation for modelling local aspects of lightness perception, which allows us to take display size, viewing distance and many other factors into account.

In Section 3.3 we describe our main model and in Section 3.4 we show how it is extended to include a spatially varying notion of lightness perception. The model is evaluated in the context of several applications in Section 3.5, while conclusions are drawn in Section 3.6.

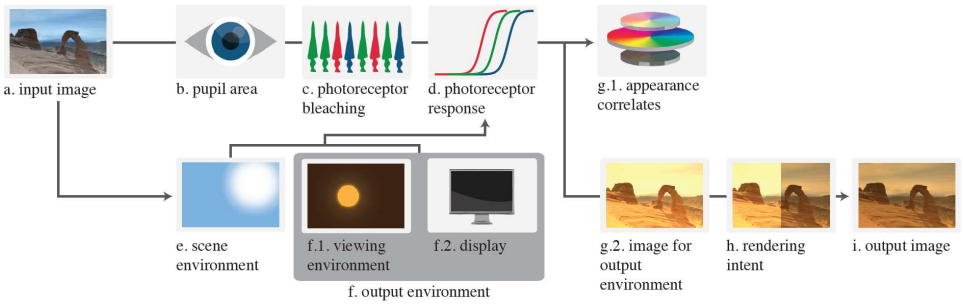


Figure 3.2: *An overview of the steps of our model.*

3.3 Our Model

We work under the hypothesis that it would be possible to match the output of photoreceptors, and possibly further stages of visual processing, to obtain a visual match. The aim is therefore to derive a model (a flow chart is shown in Figure 3.2) to process images such that when presented on a known display device under known illumination will elicit a perceptual response identical to the original scene environment in which the image was captured. This means that in addition to the input image (a), which we assume to be specified in calibrated photometric units[1] (cd/m^2), the model requires a set of parameters that characterizes the environment in which the image was taken (e), as well as a characterisation of the display device (f2) and the viewing environment (f1). These parameters are compatible with those used in other colour and image appearance models. The output of the model is then a new image (g2) which can be post-processed according to a desired rendering intent (h) and displayed on the specified display (i), generating the correct percept.

Additionally, it is possible to calculate a set of appearance correlates (g1) which describe how pixels in the input image would be perceived in the environment in which the image was captured. These are computed on the output of the global model, and this is discussed in Section 3.5.3. These correlates allow the model to serve as a standard colour appearance model.

We model the pathway that light takes from entering the eye until the computation of the neural response generated by the photoreceptors. This includes modulating the pupil size (Figure 3.2(b)), accounting for bleaching (c), as well as subsequent modeling of the neural response of the three

[1]The model produces plausible results even if the input image is not calibrated, although the resulting output image in that case is also not necessarily calibrated.

cone types (d). Although tone reproduction and colour appearance models often omit the effects of pupil size and bleaching, we have found that including these allows the model to accurately predict photoreceptor output with a straightforward model. The non-linear response is modulated by a spatially varying measure of scene adaptation (e), which is inspired by lightness perception and takes into account size, distance and strength of the light sources in the scene (Section 3.4). Before describing each stage, we begin by discussing the input parameters and their derivation.

3.3.1 Input Parameters

The first and most important input to the algorithm is an image $\mathbf{I}$ with N pixels $\mathbf{I}_n = (X_n, Y_n, Z_n)$ that ideally would be linear and specified in absolute values. We assume that the image is given in the CIE XYZ colour space, and that the Y channel is in cd/m^2. Like all colour appearance models, the scene in which the photograph was taken needs to be characterised. This is typically done by specifying an adapting luminance $L_{a,s}$ as well as the white point of the dominant luminant, specified as a CIE XYZ tristimulus value $\mathbf{W}_{a,s} = (X_{W,s}, Y_{W,x}, Z_{W,s})$ (with the Y channel normalized to 100)[2]. These parameters can be estimated from the image, and although it would be possible to derive one set of parameters for the entire image, we believe that for successful appearance reproduction it would be better to estimate these parameters in a spatially varying manner. This will allow us to take proximity of light sources, their size and intensity into account, thereby effectively approximating human lightness perception. This estimation technique is described in Section 3.4.

Following CIECAM02 we compute the degree of adaptation D_s, an interpolant which models to what extent the visual system is adapted [76]:

$$D_s = 1 - \frac{\exp((-L_{a,s} - 42)/92)}{3.6}. \tag{3.1}$$

In our calculations we also need a notion of the maximum values that occur in the image. However, directly measuring the maximum value will not lead to a robust algorithm. Instead, we compute the maximum scene luminance by taking the 90^{th} percentile, giving $L_{a,\max,s}$. This value was chosen such that small but extremely bright regions (e.g. sun) would not bias

[2]Note that we use 's' in subscripts to denote scene referred parameters, 'd' to refer to display parameters, and 'v' to refer to the viewing environment in which the display is located. If these identifiers are omitted, the variable applies equally to all scene, display and viewing conditions.

further computations. The associated maximum white point is $\mathbf{W}_{a,\max,s} = \mathbf{W}_{a,s}\ L_{a,\max,s}/L_{a,s}$, a scaled version of the log average white point.

In most colour appearance models, each of the input parameters has a counterpart that describes the display environment. However, the current and future range of display capabilities means that it may not be sufficient to only describe the viewing environment: humans may adapt to both the viewing environment as well as the display itself — an issue particularly important for HDR displays. We therefore specify two counterparts for each scene-related parameter; one describing the display and one describing the viewing environment. These parameters should be specified by the user, whereby we note that we typically use the display white point for $\mathbf{W}_{a,d} = (X_{W,d}, Y_{W,d}, Z_{W,d})$, and measure a white piece of paper using a chromameter to derive the viewing white point $\mathbf{W}_{a,v} = (X_{W,v}, Y_{W,v}, Z_{W,v})$. We then infer the following parameters: $L_{a,d} = Y_{W,d}$, $L_{a,v} = Y_{W,v}$, $L_{a,\max,d} = 5\,L_{a,d}$, $L_{a,\max,v} = 5\,L_{a,v}$ and finally, $Y_{b,d} = Y_{b,v} = 20$ and $D_d = D_v = 1$.

In the viewing environment the room illumination as well as the display may contribute to the state of adaptation of the viewer. Although partial adaptation is currently not a fully understood mechanism, we assume that the viewing and display illuminants contribute relative to their intensities, as well as the proportion of the field of view taken up by the display. As a result, we combine the room and display parameters into a unified set of parameters. This is achieved by measuring the visual angle α of the display relative to the entire visual field, and weighting $L_{a,d}$ and $L_{a,v}$ according to this fraction. This leads to a new $L_{a,v}$ defined as $\alpha\,L_{a,d} + (1-\alpha)\,L_{a,v}$. We weigh all other display and viewing parameters in the same manner, reducing the number of viewing/display related parameters by half.

Finally, we convert the image as well as all tristimulus values to LMS cones space by means of the Hunt-Pointer-Estévez transform [27]. In the following we describe the model of human vision that we employ using scene referred parameters. An exact counterpart can be assumed for the viewing environment. The output of these two models is then combined into a single algorithm that transforms the image such that appearance is preserved.

3.3.2 Pupil Size

The pupil size is normally interpreted as a function of the adapting luminance $L_{a,s}$ and thus could be seen as providing an optimal aperture for the given environment. However, it could also be interpreted as a protective device, constricting the pupil in bright environments to minimise the risk of damage to the retina. In that case, pupil area would be a function of the

maximum adapting luminance $L_{a,\max,s}$, used in the formulation for pupil area A_s [75]:

$$A_s = \pi \left(2.45 - 1.5 \tanh(0.4 \ln(L_{a,\max,s} + 1)\right)^2 mm^2. \tag{3.2}$$

The input image as well as all the white points and adapting luminances can be converted to trolands by multiplying by pupil area A_s, yielding retinal illuminance.

3.3.3 Bleaching

The opsins within the photoreceptors can change state if hit by too much light, stopping them temporarily from being able to transduce light. The probability that a photoreceptor is able to function is given by [43]:

$$\mathbf{p}(\mathbf{W}) = \frac{4.3}{4.3 + \ln(\mathbf{W})}. \tag{3.3}$$

The vector notation used here is to indicate that we compute the effect of bleaching for each of the three colour channels separately. We compute factors for $\mathbf{W}_{a,s}$, $\mathbf{W}_{a,\max,s}$ and $\mathbf{W}_{a,v}$. This means that the effective retinal illuminance for these two white points is:

$$\mathbf{W}^e_{a,s} = \mathbf{W}_{a,s} \odot \mathbf{p}(\mathbf{W}_{a,s})\ A_s \tag{3.4}$$

$$\mathbf{W}^e_{a,\max,s} = \mathbf{W}_{a,\max,s} \odot \mathbf{p}(\mathbf{W}_{a,\max,s})\ A_s. \tag{3.5}$$

While we could compute the effective retinal illuminance for the image as well, it would be computationally more efficient to compute a factor $\mathbf{f}_s = A_s\, \mathbf{p}(\mathbf{W}_{a,s})$ that is used to account for retinal illuminance in the final mapping function introduced in Section 3.3.5.

3.3.4 Photoreceptor Response

We choose a model of photoreceptor behavior that is similar to existing models as they have been applied to tone reproduction as well as colour appearance modelling, i.e. we use a model that is a variant of the well-known Michaelis-Menten equation[3]:

$$\mathbf{V_s} = \mathbf{V}_{\max,s}\ \frac{\mathbf{I}}{\mathbf{I} + \sigma_s\, \mathbf{V}_{\max,s}/\mathbf{f}_s} \tag{3.6}$$

[3]Often the Naka-Rushton equation is used instead, which has the same form, albeit with the $\mathbf{I}$ and σ_s terms exponentiated with $n \in [0, 1]$. In CIECAM02 n is 0.42, whereas Kim et al. [51] use $n = 0.73$. There is significant debate as to what value would be optimal, although measurements have shown that for the normal retina we have $n = 1$ ([34], page 17); hence our choice to omit the exponent.

where $\mathbf{V_s}$ is the neural response due to stimulus strength $\mathbf{I}$, the semi-saturation constant is given by σ_s and the maximum response is given by $\mathbf{V}_{\max,s}$. The factor $\mathbf{f}_s$ accounts for bleaching and pupil size, as discussed in Section 3.3.3. The main difference with the standard sigmoidal form is that the semi-saturation constant is multiplied by $\mathbf{V}_{\max,s}$. This allows us to match scene and viewing environments in a novel and interesting way, as shown in the following section. Moreover, this equation can be rewritten as

$$\mathbf{V_s} = \frac{\mathbf{I}}{\mathbf{I}\,\mathbf{V}_{\max,s}^{-1} + \sigma_s/\mathbf{f}_s}, \tag{3.7}$$

which shows that the output of this system is not normalized, but can vary according to the choice of $\mathbf{V}_{\max,s}$.

Most models of human vision choose $\mathbf{V}_{\max,s}$ to be constant. However, it has been shown that this value should vary according to overall illumination [43]. In particular, the maximum neural response reduces somewhat in bright environments. We surmise that its value is related to the maximum effective retinal illuminance, so that we can compute $\mathbf{V}_{\max,s}$ as:

$$\mathbf{V}_{\max,s} = k \left(\frac{\theta + \mathbf{W}^e_{a,\max,s}}{\theta} \right)^{-0.5}. \tag{3.8}$$

Although the literature does not specify values for k and θ, by optimising against a set of corresponding colour datasets [66] we have found that $k = 34$ and $\theta = 67$ yields consistently good results.

The semi-saturation constant σ_s models the neural mechanism of adaptation, and as such is a function of both the adapting luminance as well as the adapting white point. Following Kunkel and Reinhard [57], we find that interpolating between these values according to the degree of adaptation D_s produces an appropriate triplet of semi-saturation constants σ_s:

$$\sigma_s = D_s\ \mathbf{W}^e_{a,s} + (1 - D_s)\, A_s\, \mathbf{L}_{a,s}, \tag{3.9}$$

where $\mathbf{L}_{a,s} = L_{a,s}\,(1,1,1)$ is a scaled identity vector. The neural response for the scene environment $\mathbf{V}_s$ is now considered to be the quantity that gives rise to all further perceptual effects. It is therefore desirable to match this neural response across viewing conditions, which we discuss next.

3.3.5 Final Mapping Function

We assume that the neural response of the scene environment $\mathbf{V}_s$ needs to be recreated in a potentially different viewing environment. If we were to

display the image in a given room on a given display, this would elicit a neural response $\mathbf{V}_v$. We aim to modify the input image such that when observed in this viewing environment, the neural responses are matched.

The modification of the input image to the display image can be written as $\mathbf{L}_d = g(\mathbf{I})$. In general there is no guarantee that any mapping $g()$ leads to an appearance match. Alternatively, one could require that $\mathbf{V}_v(\mathbf{L}_d) = \mathbf{V}_s(\mathbf{I})$ — effecting a direct appearance match — for whichever functional form of $\mathbf{V}$ was chosen. Solving this equation for the tone reproduction operator $\mathbf{L}_d$ leads to a solution that encodes the forward and backward steps as seen in all colour appearance models and some tone reproduction operators. If the photoreceptor model was chosen as a sigmoidal compressive function, however, then it can be shown that the functional form of the tone reproduction operator follows a power law [92], which is fine for images of medium dynamic range but not sufficiently compressive for high dynamic range imagery.

On the other hand, sigmoidal operators that omit the backward step (e.g. [94]), have been shown to produce plausible images [64], despite being theoretically incorrect. The latter stems from the fact that sigmoidal compression produces perceived values which are then displayed as if they were luminances. This makes that the observer's visual system incorrectly perceives these values for a second time.

We can close this gap between theory and practice by noting that the precise form of our neural response computation allows us to match $\mathbf{V}_v$ and $\mathbf{V}_s$ in parts rather than as a single equation. We will show how this key idea leads to a desirable result in the following.

We begin by choosing a specific form for our mapping function $g()$, such that display values $\mathbf{L}_d$ are computed from the input $\mathbf{I}$ as follows:

$$\mathbf{L}_d = \mathbf{L}_{\max} \frac{\mathbf{I}}{\mathbf{I} + \tau \mathbf{L}_{\max}} = \frac{\mathbf{I}}{\mathbf{I}\mathbf{L}_{\max}^{-1} + \tau}. \tag{3.10}$$

Note that our computational model has exactly the same functional form as the assumed photoreceptor model of (3.6). Suppose we map the input image with this function and display the result, then the neural response of

an observer's photoreceptors would be:

$$\mathbf{V}_v = \mathbf{V}_{\max,v} \frac{\mathbf{L}_{\max}\frac{\mathbf{I}}{\mathbf{I}+\tau\,\mathbf{L}_{\max}}}{\mathbf{L}_{\max}\frac{\mathbf{I}}{\mathbf{I}+\tau\,\mathbf{L}_{\max}} + \frac{\sigma_v}{\mathbf{f}_v}\,\mathbf{V}_{\max,v}} \tag{3.11}$$

$$= \frac{\mathbf{V}_{\max,v}}{1+\frac{\sigma_v}{\mathbf{f}_v}\,\frac{\mathbf{V}_{\max,v}}{\mathbf{L}_{\max}}}\,\frac{\mathbf{I}}{\mathbf{I}+\tau\,\frac{\sigma_v}{\mathbf{f}_v}\,\frac{\mathbf{V}_{\max,v}}{\mathbf{L}_{\max}}} \tag{3.12}$$

$$= \mathbf{c}_1\,\frac{\mathbf{I}}{\mathbf{I}+\mathbf{c}_2}. \tag{3.13}$$

We note that (3.13) has once again the same form as (3.6). This is crucially important, as equating $\mathbf{V}_v$ to $\mathbf{V}_s$ can now be achieved by equating the constants $\mathbf{c}_1$ and $\mathbf{c}_2$ leading to two equations from which we can solve our two unknowns $\mathbf{L}_{\max}$ and τ:

$$\mathbf{V}_{\max,s}\frac{\mathbf{I}}{\mathbf{I}+\frac{\sigma_s}{\mathbf{f}_s}\,\mathbf{V}_{\max,s}} = \frac{\mathbf{V}_{\max,v}}{1+\frac{\sigma_v}{\mathbf{f}_v}\,\frac{\mathbf{V}_{\max,v}}{\mathbf{L}_{\max}}}\,\frac{\mathbf{I}}{\mathbf{I}+\tau\,\frac{\sigma_v}{\mathbf{f}_v}\,\frac{\mathbf{V}_{\max,v}}{\mathbf{L}_{\max}}}. \tag{3.14}$$

The two parts $\mathbf{c}_1$ and $\mathbf{c}_2$ correspond to the following two equations:

$$\mathbf{V}_{\max,s} = \frac{\mathbf{V}_{\max,v}}{1+\frac{\sigma_v}{\mathbf{f}_v}\,\frac{\mathbf{V}_{\max,v}}{\mathbf{L}_{\max}}} \tag{3.15}$$

$$\frac{\sigma_s}{\mathbf{f}_s}\,\mathbf{V}_{\max,s} = \tau\,\frac{\sigma_v}{\mathbf{f}_v}\,\frac{\mathbf{V}_{\max,v}}{\mathbf{L}_{\max}}. \tag{3.16}$$

Solving this system of equations for $\mathbf{L}_{\max}$ and τ gives:

$$\mathbf{L}_{\max} = \frac{\sigma_v}{\mathbf{f}_v}\frac{1}{1/\mathbf{V}_{\max,s} - 1/\mathbf{V}_{\max,v}} \tag{3.17}$$

$$\tau = \frac{\sigma_s}{\mathbf{f}_s}\,\frac{\mathbf{f}_v}{\sigma_v}. \tag{3.18}$$

In practice, we would implement the right-hand side of (3.10), and therefore need $\mathbf{L}_{\max}^{-1}$ which also has the mathematical advantage of approaching zero when the scene and viewing conditions are identical (and thereby $\mathbf{V}_{\max,s} = \mathbf{V}_{\max,v}$). Our method is formally not suitable to handle scene and display environments whereby the maximum scene luminance is smaller than the maximum luminance in the viewing environment. In that case we have $\mathbf{V}_{\max,s} > \mathbf{V}_{\max,v}$ (due to (3.8)) resulting in $\mathbf{L}_{\max} < 0$. Bearing in mind

that the range of values attained by $\mathbf{V}_{\max}$ is small, we therefore introduce a straightforward variation to $\mathbf{L}_{\max}^{-1}$ that will produce plausible images even when this condition is not met:

$$\mathbf{L}_{\max}^{-1} = \frac{\mathbf{f}_v}{\sigma_v} \left| \frac{1}{\mathbf{V}_{\max,s}} - \frac{1}{\mathbf{V}_{\max,v}} \right|. \tag{3.19}$$

An important observation is that our mapping function (3.10) reduces to the identity operator in case the scene and viewing environments are the same. This means that for images which do not require compression for the given viewing environment, the operator leaves pixels unaltered.

The algorithm calculates in cone response space three fully independent channels, which means that it is a strict von Kries model [27]. Although most colour appearance models compute chromatic adaptation in a separate sharpened colour space as a preprocess, we have folded chromatic adaptation into the computation of three different semi-saturation constants, a technique first shown to be viable by Kunkel and Reinhard [57].

As a result, the model can be directly evaluated against corresponding colour datasets. Such psychophysical datasets are obtained by showing observers a patch of colour under a specific illuminant. The patch is then shown under a different illuminant, and the observer is asked to adjust the tristimulus value of the patch until it appears identical. We evaluate the model against the CSAJ, Kuo and Luo, Lam and Rigg, Helson et al., Breneman, and the Braun and Fairchild datasets [66], as shown in Figure 3.3. For the three colour channels, the root mean square error is 2.0183, 1.6863 and 2.181 (overall 1.97), which is within 8% of CIECAM02 (but note that we achieve this performance without requiring a separate chromatic adaptation step).

3.3.6 The Hunt and Stevens effects

Two important colour appearance phenomena need to be taken into account, as they play a particularly important (yet mostly unexplored role) in high dynamic range imaging. The first is the Hunt effect, which predicts that the perception of saturation covaries with scene luminance [27]. This effect is implicitly handled in our model by defining independent compressive curves for each of the three colour channels.

The Stevens effect predicts that the perception of contrast co-varies with scene luminance [27]. We model this effect by re-introducing some of the contrast that was lost by applying (3.10). A reliable way to do this is by employing a bilateral filter-based unsharp mask [88]. Here, we use a bilateral

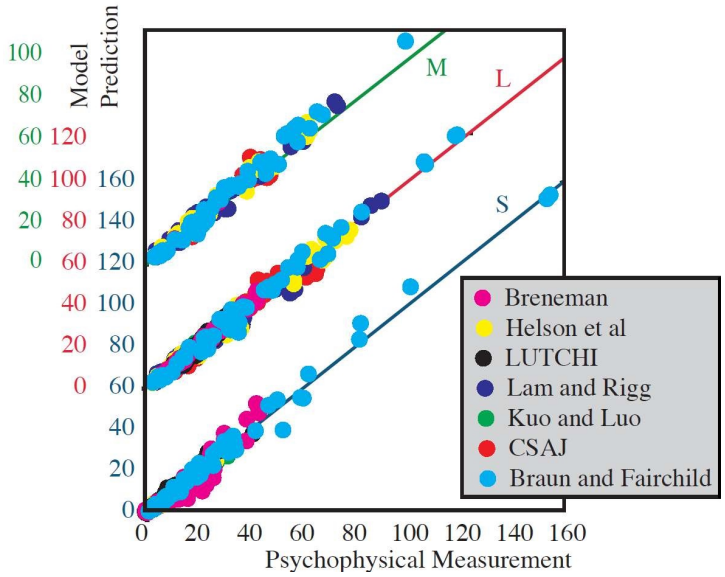


Figure 3.3: *The output of* (3.10) *for each of its channels plotted against seven different corresponding colour datasets.*

filter $B_{\rho_{\text{space}},\rho_{\text{L}}}$ with spatial constant $\rho_{\text{space}} = 0.1\, d_{\text{disp}}\ \max(1, L_{a,s}/L_{a,v})$, where d_{disp} is the display diagonal. The intensity constant is $\rho_{\text{L}} = 0.2\,\Delta L$, where ΔL is the scene's luminance range. The unsharp mask U is then computed as $U = Y_d - B_{\rho_{\text{space}},\rho_{\text{L}}}(Y_d)$, where Y_d is the luminance channel derived from $\mathbf{L}_d$, the output of (3.10), by conversion to Yxy space. The contrast enhanced image luminance is then given by $Y'_d = Y_d + 0.3\,U$. The computation is carried out in Yxy space for convenience, but could be carried out directly in LMS space.

3.3.7 Post-Processing

The image obtained by executing (3.10) will create a visual match in the chosen viewing environment, under the condition that the display range is sufficient. In essence the image is prepared for a particular viewing condition, as opposed to a particular display device. As a result, pixels may be mapped to values outside the display's range. This would occur when the display is dim relative to the viewing environment (e.g. when taking a laptop outside into the sun).

In such cases it is inevitable that some information is lost. To decide how to display an image we introduce two rendering intents that post-process

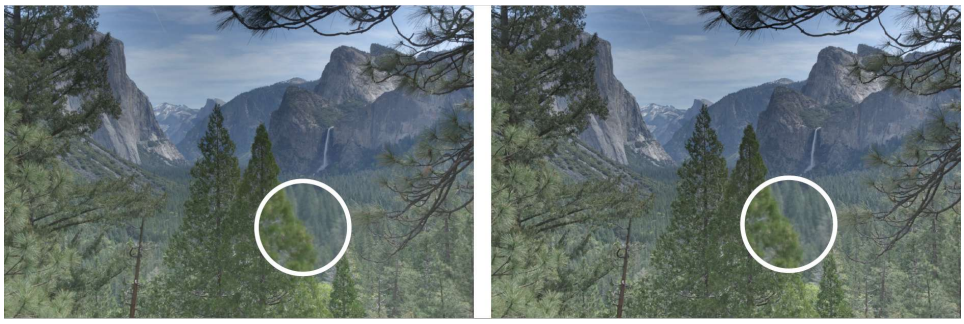


Figure 3.4: *An image without (left) and with (right) inclusion of our model of the Stevens effect, showing that the perception of contrast is maintained better in the right image.*

the image according to the wishes of the observer. The most precise way to display pixels is to simply clamp all pixels that are above the maximum display value. We call this rendering intent *physical*. It maps as many pixels as accurately as possible. Alternatively we can normalise the image to the display range, a rendering intent we call *linear*. This produces plausible images despite mapping pixels to values that are potentially either too high or too low.

Note that such issues would not occur if the observer chose the illumination in the viewing environment to be appropriately matched to the display capabilities. As such, these rendering intents are created to allow the display of images under less than ideal circumstances.

Finally, as per normal, the image should be gamma corrected prior to display to account for display non-linearities.

3.4 Local Lightness Perception

The model described so far affects all pixels in the image in the same way as it produces a single tone curve per channel. Although this is sufficient for many images, in more extreme scenarios, such as the image shown in Figure 3.1, the resulting tone curve is appropriate for neither the bright or the darker parts of the image. A better solution would be to adapt the curve to take into account local trends in the image.

In the human retina, lateral interconnections allow such adaptation to occur. Examples are horizontal cells that modify photoreceptor output [60] as well as various amacrine cell types. These cells form neural substrates

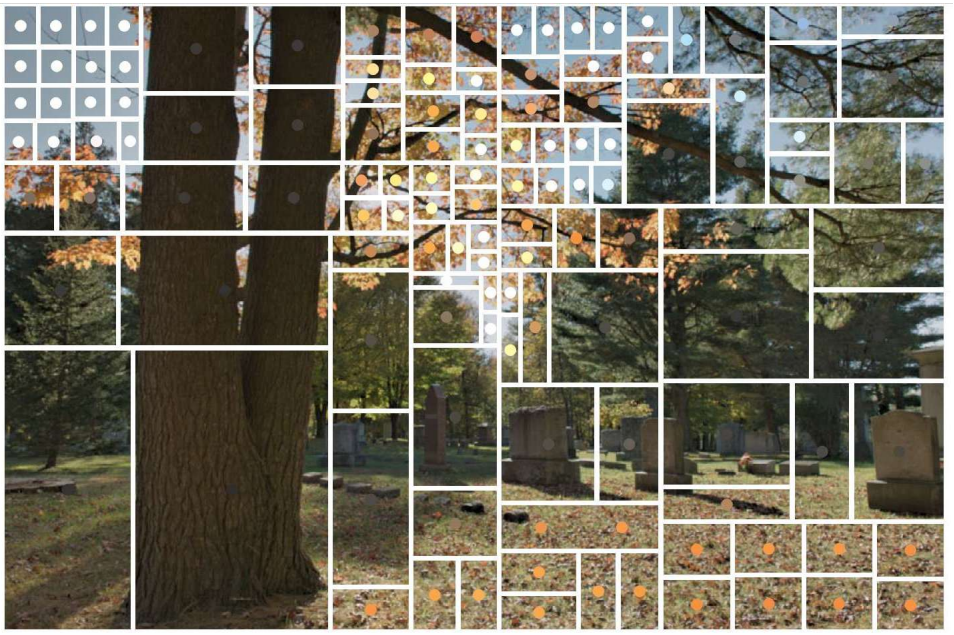


Figure 3.5: *A visualization of the median cut algorithm [22], creating regions which can be represented with a single point acting as its representative light source.*

which compute some form of spatial adaptation. Many other lateral, forward, and backward interconnections occur further along the visual pathway. These together give rise to various perceptual phenomena, one of which is the notion of lightness.

Studies of lightness perception have shown that the size of a region in a scene, in addition to distance and strength, plays a crucial role in how that part of the scene is perceived. Gilchrist et al. [36] formalised these effects in the following rule, known as the Area Rule: *"In a simple display, when the darker of the two regions has the greater relative area, as the darker region grows in area, its lightness value goes up in direct proportion. At the same time the lighter region first appears white, then a fluorescent white and finally, self-luminous."*

With this in mind, our model of lightness perception considers luminance, distance as well as region size. The first issue then is naturally the detection of appropriate regions in the input image. Although some form of segmentation could be used [55], this would be a costly operation and would require an additional cleanup step.

For our solution, we draw inspiration from importance sampling techniques and specifically the median cut algorithm [22], which recursively subdivides the image into R regions r of equal luminous energy for a given number of iterations. A light source is then placed in the centre of each region (or in a different position, such as a centroid determined according to the energy distribution within that region) and its colour is computed as the region's average pixel value (see Figure 3.5).

This method is both robust and efficient, subdividing the image into regions of approximately equal energy. These regions can naturally serve as neighborhoods over which the scene parameters $L_{a,s}$, $\mathbf{W}_{a,s}$ and D_s are estimated locally. In addition, the sizes of the different regions are implicitly computed in the subdivision process. For instance, Figure 3.5 shows that approximately half the regions are allocated to light parts of the scene, but these are much smaller than their darker counterparts. This encodes the relation between light and dark parts of the image and allows us to approximate the area rule. We have found that 7 levels of recursion (128 regions) creates regions that represent their local areas well. Subdividing into more regions only creates more computations without further benefit, while using fewer regions gracefully degrades the result, to obtain a global operator in the limit of using the image as a single region.

We restrict the minimum size of a region to 10 pixels to ensure that small, bright regions such as highlights are not over-represented in the estimation of adaptation levels. We position the virtual light source representing each region at its center. For each region $r \in R$, the pixels within that region are used to compute local adaptation levels $L_{a,s,r}$ (computed as the region's geometric mean), degree of adaptation $D_{s,r}$ (Equation (3.1)) and effective retinal illuminance $\mathbf{W}^e_{a,s,r}$ (Equation (3.4)). Note that the pupil size used for the computation of $\mathbf{W}^e_{a,s,r}$ is computed based on global scene parameters as the viewer will be observing the scene as a whole (Equation (3.2)). This is also the case for $\mathbf{V}_{\max,s}$ (and consequently $\mathbf{L}^{-1}_{\max}$) as it represents a global scaling to the maximum response of the photoreceptors (Equations (3.8) and (3.19)).

The semi-saturation value that drives the compressive part of our model is then computed for each pixel based on its distance to each of the virtual light sources. This computation also takes into account the size of the display and the viewing distance, effectively including visual angle as a factor for determining the contribution of each light source for a particular pixel.

Thus, for any image pixel $\mathbf{I}_n$, a weight $w_{r,n} \in \mathbf{w}_n$ is computed per region r, based on the pixel distance $d_{r,n}$ between the pixel and the center of each

region r:

$$w_{n,r} = \left(1 - \cos\left(\frac{\pi}{2} - \frac{d_{n,r}}{d_{\text{diag}}}\,\omega\right)\right)^4. \tag{3.20}$$

Here, d_{diag} is the display diagonal in pixels, which is used to normalize the pixel distances for each region, and $\omega = \theta_D/120$ is the ratio between the visual angle covered by the display itself and the full (binocular) visual field of the viewer, which we set to 120° [104]. The visual angle θ_D of the display is computed based on its physical diagonal size d_{size} and the expected viewing distance d_{view} as $\theta_d = 2\tan^{-1}(0.5\,d_{\text{size}}/d_{\text{view}})$. The weights $\mathbf{w}_n$ are finally normalized so that they sum to 1 for each pixel. Pixels are then assigned their own values for adaptation luminance, adaptation white point and degree of adaptation. For instance, the adapting luminance is:

$$L_{a,s,n} = \sum_{r=1}^{R} w_{n,r}\; L_{a,s,r}. \tag{3.21}$$

The per pixel degree of adaptation $D_{s,n}$ and white points $\mathbf{W}^e_{a,s,n}$ are computed analogously. A semi-saturation triplet $\sigma_{s,n}$ is then computed for each pixel $\mathbf{I}_n$ using a spatially varying version of (3.9):

$$\sigma_{s,n} = D_{s,n}\;\mathbf{W}^e_{a,s,n} + (1 - D_{s,n})\,A_s\,\mathbf{L}_{a,s,n}, \tag{3.22}$$

where $\mathbf{L}_{a,s,n} = L_{a,s,n}\,(1,1,1)$. With this per pixel parameter the image is then processed using the model described in Section 3.3.5.

Our approach allows us to model spatial aspects of lightness perception and preserves sharp edges in the image without introducing artefacts. Spatial modulations are overall very low frequency and their intensity depends on the specific viewing conditions. If the image is intended to be viewed on a laptop monitor in typical office conditions, it will only fill a relatively small part of the viewer's field of view and other light sources are likely to play a role in their overall adaptation, meaning that all parts of the image will be perceived in a similar way. On the other hand, if the target viewing condition is a cinema, the surrounding illumination is likely to be extremely low and the screen very large. Consequently, the displayed image will largely determine the adaptation of the viewer. If the original scene was much brighter than the display or projector can handle, effects due to lightness perception that would be apparent if the viewer was in the original environment, need to be simulated in the compressed output to match the scene appearance.

Figure 3.6: *This scene was rendered with our algorithm, using a global set of parameters (left) and per pixel parameters (right), approximating lightness perception.*

3.5 Applications

In this section we describe three usage scenarios for which our model is applicable. The first is what we call scene reproduction, whereby (un-)calibrated images are prepared for viewing in a specific viewing environment. The second is video reproduction, demonstrating the use of our technique on high dynamic range video, and finally we present our take on appearance prediction, which requires the calculation of appearance correlates. In these sections we compare our results to the state-of-the-art in tone reproduction, colour appearance and image appearance models.

3.5.1 Scene Reproduction

The main aim of our work is to reproduce scene appearance across a variety of viewing conditions. With this we mean to reproduce the overall look and feel of a scene. This is not the same as reproducing a photograph of a scene, which would produce an altogether different look. An example demonstrating this subtle but important difference is shown in Figure 3.6. The left image was created with our algorithm, although we have set $L_{a,s}$ and $\mathbf{W}_{a,s}$ to the image's global geometric mean, rather than computing these values per pixel, as our standard algorithm would do (right image). We note that the left image looks more like a photograph of a scene than like the scene itself. In our opinion the right image is a perceptually more correct representation of the scene. In particular, note the bark on the tree in the right image. The part of the tree trunk shown against the sky is both darker and less rich in detail than the bottom part of the tree trunk. This is

indeed how such a scene appears as a result of human lightness perception, although this is rarely captured in photographs.

It is important to note that we intend to match appearances, rather than maximize image attributes. This means that our aim is not to produce the most punchy and contrast-rich image. We are, however, aiming to produce well-balanced images with the correct distribution of intensities, contrasts and colours. Consider Figure 3.7, where we compare our result with a set of state-of-the-art colour appearance models and tone reproduction operators. This well-known image is difficult to reproduce accurately due to its large dynamic range, as well as its strong colour cast. Notice that in our scene reproduction (top left) the colours are reproduced accurately. The marble of the stairs can be clearly differentiated from the rest of the floor, the gold on the walls appears as gold, and the wood on the ceiling looks like wood.

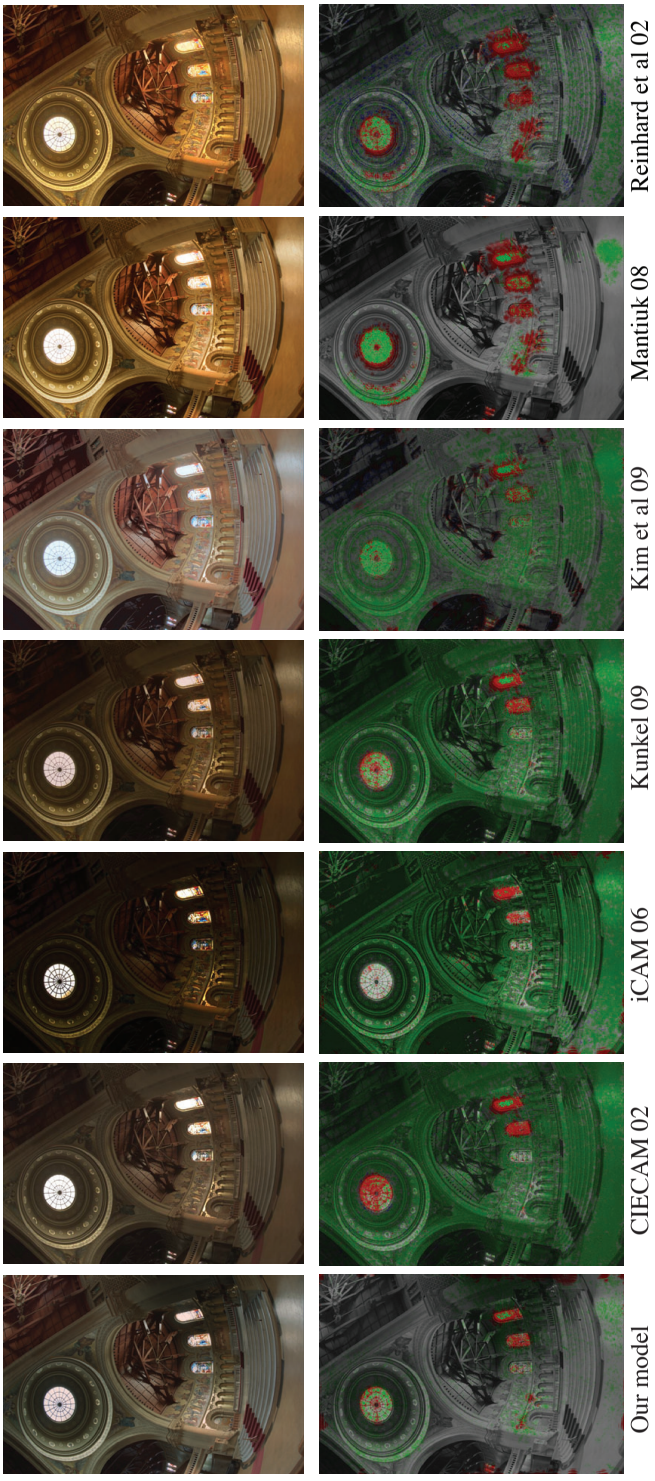


Figure 3.7: *A visual comparison of our results with a variety of colour appearance, image appearance and tone reproduction operators. The mapped image is shown in the top row, and the bottom row shows the difference between the input high dynamic range image and the output image, according to the dynamic range independent quality assessment metric [3]. This metric shows where contrasts have changed above visible threshold (see text).*

The dynamic range independent quality assessment metric [3] shows that comparing input with our output yields minimal distortions (bottom left in Figure 3.7. Our result was created using default parameters that describe an average laptop in an office environment[4]. The encoding of colour in the quality metric images is as follows: green pixels show loss of contrast, red indicates amplification of contrast, and blue indicates a contrast reversal. Grey shows image areas without artifacts.

The same figure shows renderings and quality assessments for a set of state-of-the-art colour appearance models and tone reproduction operators. These include CIECAM02 [76], Kunkel's [57] and Kim's [51] colour appearance models, the iCAM06 image appearance model [56], and Mantiuk's display adaptive operator [71] as well as the photographic operator [94]. Where possible, we have used default parameters to create these images. For CIECAM02 this was not possible, requiring hand tuning to find the parameters that best represent the scene.

Each of these algorithms create images with very different appearance. Characteristic for tone reproduction operators, the display adaptive and photographic operators do not reproduce colours appropriately, making the scene overly yellow. As expected, the colour and image appearance models do better in that respect, but tend to make the image too flat – the quality metric shows significant loss of contrast for these results. They also make the image either a little light or dark. It may be that default parameters are not optimal for this image, although it is more likely, at least for CIECAM02 and Kunkel's method, that the inverse step largely undoes the compression effected by the forward step. This is the reason that we have taken a different approach, having a single compressive step that takes both scene and viewing parameters into account.

A further comparison with Mantiuk et al.[71], iCAM06 [56] and Kim et al.[51] is given in Figure 3.8 for a set of four calibrated images from the HDR survey [26]. All algorithms in this figure take the display environment into account, which was set to an average laptop display, assumed to have a D65 white point. Thus, this figure is best viewed on a D65 display, ideally under a D65 illuminant, zoomed in such that each panel takes the full size of the display. The viewing distance should then be about 30 cm. Note that our method produces even results without over-saturating the images (a problem common in tone reproduction [70]). Our algorithm produces

[4]Display: $L_{a,d} = 60\ cd/m^2$, $L_{a,\max,d} = 191$, $\mathbf{W}_{a,d} = (55, 60, 65)$, $\mathbf{W}_{a,\max,d} = (172, 191, 190)$; Viewing environment: $L_{a,v} = 800\ cd/m^2$, $L_{a,\max,v} = 7010$, $\mathbf{W}_{a,v} = (760, 800, 871)$, $\mathbf{W}_{a,\max,v} = (6662, 7010, 7632)$.

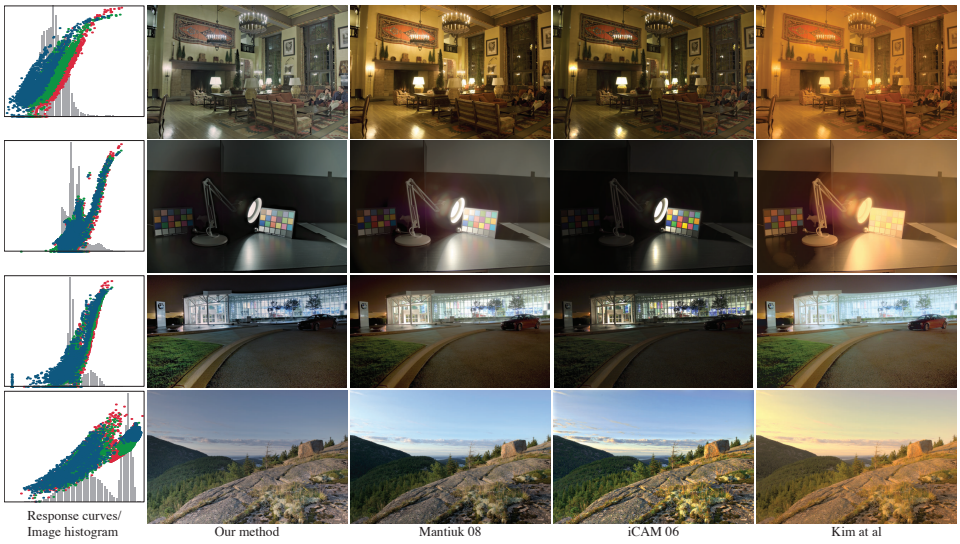


Figure 3.8: *A visual comparison of our output with several models that take the display environment into account. The left column shows the response curves for our method (log-linear plots), showing how they adapt to the image histogram (shown in grey). As each channel is processed independently, strong colour casts are automatically removed, as shown in the top image.*

sufficient contrast to show both colour checkers in the second image from the top with good fidelity. The early sunrise scene shows a dark sky, as expected of a scene captured at 6:19AM. The night scene produces a well balanced result. Here, the car appears relatively dark as it is depicted against a light store front, simulating the effects of lightness perception.

Finally, Figures 3.1 and 3.9 show the effect of preparing an image for different display and viewing environments. Figure 3.9 shows an image prepared for an average laptop (parameters as before) as well as an iPhone screen[5]. Although the differences are subtle, they do allow the snow to be reproduced accurately in each case.

3.5.2 Video Reproduction

The proposed model is suited for managing the appearance of video content in addition to images. We have experimented with a range of high dynamic

[5]Display: $L_{a,d} = 100\ cd/m^2$, $L_{a,\max,d} = 250$, $\mathbf{W}_{a,d} = (91, 100, 110)$, $\mathbf{W}_{a,\max,d} = (227, 250, 276)$; Viewing environment: $L_{a,v} = 2000\ cd/m^2$, $L_{a,\max,v} = 5440$, $\mathbf{W}_{a,v} = (1907, 2000, 1907)$, $\mathbf{W}_{a,\max,v} = (5135, 5440, 5472)$.

Figure 3.9: *An image prepared for display on a laptop (left) and an iphone screen (right), taking into account differences in size, display properties, as well as viewing environments.*

Figure 3.10: *A series of frames from an HDR video processed with the proposed model. The adapting white point is progressively changed in the video to demonstrate how the image is modified to simulate adaptation to the varying white point.*

range video segments [107]. A series of frames from a video processed with our model is shown in Figure 3.10. For this video, we simulated a temporally varying illumination environment by using a strongly coloured border (ranging from CIE D75 to CIE A) to which the observer will adapt. Using the colour of this border as input to the algorithm, we prepared the video for observation with this border on a display with 15 inch diagonal screen area and for a viewing distance of 15 inches. Note that this allows the overall appearance of the video to co-vary with the border, thus conveying the correct appearance.

Although the method cannot offer theoretical guarantees of temporal stability, we note that, in practice, the algorithm did not introduce flicker in the videos we have tested. If temporal stability needs to be guaranteed, it should be possible to apply a leaky integrator to $\sigma_{s,n}$ [52].

3.5.3 Appearance Prediction

The model can be used to derive appearance correlates, including lightness, hue and colourfulness, that describe colour under specific viewing conditions [27]. In colour science, such predictions are highly valuable as they allow inferences to be made as to how humans perceive colour in context. In our model, computing appearance correlates requires specifying the scene environment and resetting the viewing and display environments. The latter is done by setting the following parameters $L_{a,v} = 50$, $Y_{b,v} = 20$, $L_{a,\max,v} = L_{a,v}\ 100/Y_{b,v}$, $\mathbf{W}_{a,v} = (50, 50, 50)$ and $\mathbf{W}_{a,\max,v} = \mathbf{W}_{a,v}\ L_{a,\max,v}/L_{a,v}$. The scene environment should be measured, for instance with a photometer, leading to a specification of the scene parameters $L_{a,s}$ and $\mathbf{W}_{a,s}$. Colour appearance models normally also specify a relative background $Y_{b,s}$, which we set to 20. This allows us to compute the maximum adapting luminance as $L_{a,\max,s} = L_{a,s}\ 100/Y_{b,s}$.

Under these conditions, the correlate of lightness J is computed as the ratio of the achromatic signal and achromatic white:

$$J = 121 \left(\frac{L_{a,s}}{L_{a,s} + 168} \frac{\mathbf{L}_d \cdot \mathbf{w}_j}{\mathbf{L}_{d,W} \cdot \mathbf{w}_j} \right)^{0.29\, z}, \tag{3.23}$$

where $\mathbf{L}_{d,W}$ is the result of passing the scene white point $\mathbf{W}_{a,s}$ through (3.10), $z = 1.48 + (Y_{b,s}\, L_{a,s}/100)^{0.5}$ and $\mathbf{w}_j$ is a vector of weights given by (7.53, 6.93, 1.43). Following other appearance models [51, 57, 76], the correlate of hue H is computed on the basis of opponent signals a and b:

$$H = \frac{180}{\pi} \tan^{-1} \left(\frac{\mathbf{L}_d \cdot \mathbf{b_h}}{\mathbf{L}_d \cdot \mathbf{a_h}} \right), \tag{3.24}$$

where $\mathbf{a_h} = (12.59, -14.15, 1.68)$ and $\mathbf{b_h} = (3.38, -1.37, -2.62)$. Finally, the correlate of colourfulness M can be computed with a method inspired by Kunkel and Reinhard [57]:

$$M = 93 \left(\frac{(\mathbf{L}_d \cdot \mathbf{a}_m)^2 + (\mathbf{L}_d \cdot \mathbf{b}_m)^2}{10^{-3} + \max(0, \mathbf{L}_d \cdot \mathbf{d}_m)} \right)^{0.73} \sqrt{J}, \tag{3.25}$$

where $\mathbf{a}_m = (2.10, -2.84, 0.82)$, $\mathbf{b}_m = (2.63, -2.26, -0.40)$ and $\mathbf{d}_m = (24.46, 0.93, 11.01)$.

This relatively straightforward model can be directly compared against psychophysical colour appearance datasets. There are two such datasets that are relevant to colour appearance modeling and high dynamic range imaging, which are the LUTCHI dataset [69] as well as the Kim et al. dataset [51].

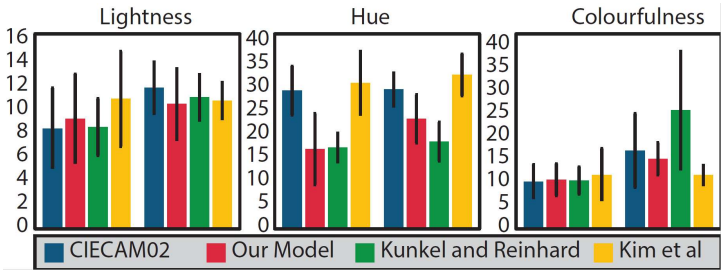


Figure 3.11: *The performance as a predictive appearance model is confirmed in this comparison against three other state-of-the-art colour appearance models, for three appearance correlates and two psychophysical datasets (RMS error; left four bars: LUTCHI, right four bars: Kim et al.).*

These sets are the result of psychophysical experimentation in which human observers assess patches of colour as presented under specific viewing conditions. These conditions are described by parameters that are compatible with the input parameters to our algorithm.

The three relevant appearance correlates measured in these datasets are lightness, hue and colourfulness. Averaging over all different observation conditions in each of the datasets, we obtain the results as shown in Figure 3.11. We compare our results for the two datasets against three other colour appearance models, namely CIECAM02 [76], Kim et al [51] and Kunkel and Reinhard [57]. For each of the three colour appearance correlates we find that our results are comparable to or better than the other models which we deem the current state-of-the-art. It can thus serve as a predictive colour appearance model, even over the extended dynamic range of Kim's dataset. This is achieved while obviating the need for an inverse step, and therefore comes with the benefit of relative computational simplicity.

3.5.4 Limitations

Our method is generally robust and achieves calibrated results when tested against colour appearance datasets as well as corresponding colour datasets. For images we use an estimation technique to derive input parameters. Part of this estimation involves finding the 90^{th} percentile of the image to robustly infer these parameters. While this gets the image right in most cases, of the more than 100 calibrated images we have tested our algorithm against [26], we have found that in less than 3% the image is reproduced somewhat too

dark. For those images, choosing a higher percentile would fix this problem, although estimating this parameter automatically for now remains an open problem.

3.6 Conclusions

Drawing inspiration from colour appearance modeling, tone reproduction research as well as studies of lightness perception, we present an appearance reproduction algorithm which prepares images and video for display under specific, measured viewing conditions. Importantly, we take the room illumination as well as the characteristics of the display device into account, leading to precise appearance reproduction. The model requires input parameters describing the viewing environment and the display, but rather than leaving its settings as guesswork, the correct inputs can be obtained through a few direct measurements of the environment and display.

For calibrated scenes it would be possible to measure the scene parameters and supply these to the algorithm. However, we have found that a straightforward estimation technique, embedded into a model of lightness perception and implemented by means of the median cut algorithm, is able to robustly set all scene-referred parameters. The model also returns plausible imagery in case the input image is not calibrated.

To ensure correct appearance reproduction, the algorithm was successfully validated against seven corresponding colour datasets, the LUTCHI colour appearance dataset as well as Kim et al's high dynamic range colour appearance dataset. Combining a spatial model of lightness perception with validated colour management has led to an algorithm that allows images and videos to be appreciated as scenes, rather than photographs.

Chapter 4

Digital Imaging Pipelines and Signal Quantization

This chapter first discusses our experimental work to enable psycho-physical vision experiments with a bus between Personal Computer (PC) and display that is sufficiently wide to avoid visible banding artifacts by using professional display hardware. We then introduce our early work on a universal tone mapping function and how it leads towards a perceptually meaningful luminance encoding scheme for video signals in today's HDR standards.

4.1 Experiments With A High Bit Depth Digital Imaging Pipeline for Vision Research

In order to achieve accurate results in user studies in the fields of Psychophysics, Experimental Psychology, Ophthalmology and clinical studies there are high demands towards an imaging pipeline presenting these stimuli in an experiment (as illustrated in Figure 4.1). For example, display stability and repeatability, both short term and long term, are crucial when conducting research leading to robust results. Further important factors are the perceptual limits of a graphics pipeline. Here, two important elements are the achievable dynamic range and the colour gamut, which would ideally approximate or exceed the capabilities of the Human Visual System (HVS). In an optimal solution, those stimulus dimensions would be displayed with continuous intensity levels between their respective extrema (e.g. from dark to light) when presenting them to participants.

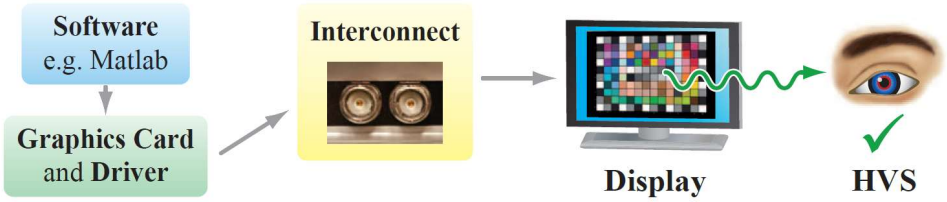


Figure 4.1: *Components within the imaging pipeline for vision science.*

When using conventional display devices, the dynamic range as well as the colour gamut is usually limited and much narrower than the capabilities of the HVS. Also, signals representing those dimensions in the digital domain need to be quantized which is referred to as bit depth. Unfortunately, common graphics hardware is only capable of outputting a maximum of 8 bits per channel which results in 256 interval steps from the devices black to white level. This is often not sufficient, especially when using wide dynamic ranges and colour gamuts leading e.g. to false contour artifacts (banding). Therefore there is demand to use 10 or even 12 bit systems allowing for 1024 or 4096 steps, respectively. To achieve higher bit depths (approximating continuous levels), researchers have been using analog Cathode Ray Tube (CRT) displays that are connected via Video Graphics Array (VGA) to specialized research graphics cards. However, it is getting more difficult to come by high-end analog monitors. Therefore, a digital successor is desirable. Such a successor would offer a fully digital pipeline from software to display, which can be calibrated throughout.

There are three major elements in an imaging pipeline targeted towards vision science: the software to control the study, the display interface and the actual display device.

The Matlab technical computing software package as well as the Psych-Toolbox are popular environments to create experiments. Currently, Matlab and PsychToolbox only support 8 bit output via Digital Visual Interface (DVI), High-Definition Multimedia Interface (HDMI) or DisplayPort (DP) on a standard graphics card. The DP specifications include the option to output 10 bits, but so far this is not reliable at the Operating System (OS) level. Although there exist software modules that allow greater than 8 bit output on specialized hardware, they are usually not using standardized hardware, increasing the complexity of experimental designs. The SDI is particularly suitable for image bit depths greater than 8 bit (see interconnect

in Figure 4.1). It is a robust and thoroughly tested professional broadcast interface, which currently supports up to 12 bits per channel with resolutions of up to 2K or the support for 3D (SMPTE 424M). One limitation of the standard is the maximum refresh rate of $60Hz$, which makes this interface less favourable for perceptual experiments requiring high frame rates. SDI interface cards for desktop and laptop computers have become reasonably affordable in recent years. Manufacturers of such cards are Nvidia, Matrox and Black Magic Design to name a few.

A further crucial component is the display. Aside from CRT displays, the vision science community has been using various high end display systems such as SpectraView Reference 2690 and HP Dream- Colour monitors. Also, the BrightSide DR37-P HDR display has been used in many studies. If precision and accuracy are crucial to an experimental design, suitable display devices can be sourced in the professional film and broadcast industry. Offering a wide colour gamut with dark black levels as well as high colour accuracy and temporal stability required for colour grading of motion pictures and TV series make these displays useful for vision research. Many of these professional displays support a high bit depth video input: HD-SDI.

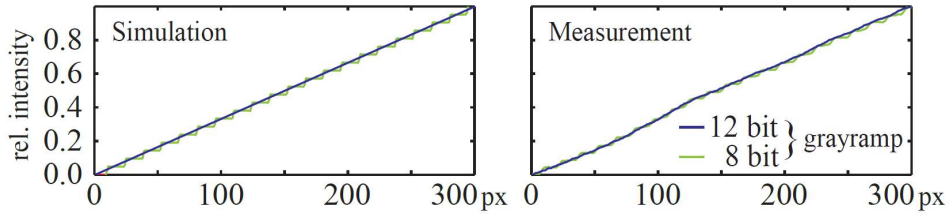


Figure 4.2: *Simulated and measured intensity using varying bit-depths.*

We have prototyped a software module to directly send 16 bit integer Additive Colour Red Green Blue (RGB) matrices via HD-SDI to a display. As HD-SDI currently only supports 12 bits, the least significant bits are truncated. The RGB matrices in this approach represent drive values for the display. Thus, it is display device independent and allows for display calibration and characterizations tailored to a given monitor e.g. by using gamma functions or lookup tables. A comparison of gray ramps with varying bit-depths is given in Figure 4.2. With such a set-up, bit-depth, colour gamut and dynamic range limitations can be alleviated forming a powerful imaging pipeline for psychophysics. At a high level the work in

this section merely constitutes the exercise of sending a discretized video signal over a more capable bus (HD-SDI instead of DVI) to a display. Optimizing the way a signal is discretized for transmission over a bus with finite bandwidth is a different and possibly harder to solve problem. We reference gamma encoding above, which has been the default encoding for consumer and professional display devices for many years, due to the nature of the CRT phosphor response and the HVSs perception of luminance values in the rage of a CRT display (approximately 0.5 to 100 cd/m^2). A breadth of literature discusses the topic (for example see [89]). Gamma encoding can be interpreted as a simple exponential function applied to the normalized linear input signal, often using a power of $\gamma = 2.2$ for consumer or $\gamma = 2.6$ for cinema applications. This encoding scheme historically worked well for two reasons: first the HVSs response to luminance and the inverse response of the CRT material properties matched well (by accident) and secondly the luminance range and contrast capabilities of most displays were comparable and thus a relative interpretation (i.e. 0 and 1 mapping to the minimum and the maximum display luminance) with a shared encoding scheme was feasible for a large variety of display devices. Today, there is a larger variety of display devices (e.g. Plasma, LCD, Organic Light Emitting Diodes (OLEDs) or direct view LEDs). Many displays natively do not match the gamma approximation very well any more (for instance direct view LED walls exhibit a proportional light output relative to drive current). Furthermore the variation in luminance capabilities between different display types has increased, and hence a single (relative) drive scheme that maps 0 to the minimum and 1 to the maximum display luminance is not ideal. Finally, the overall peak luminance of most displays today is significantly higher (often 1000 cd/m^2 and more) than the previous CRT target peak luminance of 100 cd/m^2, and thus the gamma exponent is not a good approximation of the HVSs response within this new luminance range. Two aspects of our work in Chapter 3 are relevant for video encoding for new displays: first a video signal should represent absolute luminance and colour information rather than map to the relative minimum and maximum luminance and chromaticity coordinates of the display and secondly, an exponential function such as the gamma curve is no longer a good perceptual match when modelling the HVSs response to the (larger) luminance range of modern displays. A better approach are curves inspired by the photoreceptor response that are briefly introduced in the Section 4.2.

4.2 Perceptually Optimal Luminance Quantization

The work in Chapter 3, especially the model developed to predict lightness perception inspired the development of a practical and perceptually accurate tone mapping operator for use in the studio post-production environment and for mapping of HDR content to consumer displays. The final, universal mapping function is modelled based on a modified version of the Naka-Rushton cone response equation [78] and takes the following form:

$$L_{out} = \frac{c_1 + c_2 * L^n}{1 + c_3 * L^n}, \tag{4.1}$$

where L is the input (scene) absolute luminance, L_{out} is the output target display luminance, and the constants c_1, c_2, and c_3 are calculated based on absolute luminance levels for:

- L_{min}: minimum scene luminance
- L_{as}: scene adaptation level
- L_{max}: maximum scene luminance
- Ld_{min}: minimum target (display) luminance
- L_{at}: target adaptation level
- Ld_{max}: maximum target luminance
- n: contrast or slope of the resulting mapping function, typically 1

A perceptually optimal video signal is encoded in a non-linear fashion, i.e. the relationship between code words (store and transmit pixels) and the light levels that those code word represent is non-linear. For traditional display systems that operate within a limited luminance and contrast range, an exponential (or *gamma*) encoding or a logarithmic encoding was suitable. For HDR content, in which absolute luminance information across a large brightness range is required, neither gamma nor logarithmic encoding are ideal, when the available number of bits are limited. A curve that approximately describes Just Noticeable Differences (JNDs) of neighbouring luminance levels across the entire range would be ideal. A simple approximation of such a curve is central to the colour appearance model in Chapter 3 and was important in Ballestad and Kostin's follow-up work

towards Equation 4.1 [4], that eventually lead to a perceptually optimal discretization function also known as Perceptual Quantization (PQ) which was published in the SMPTE ST 2084:2014 standard (High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays), where it is described as:

$$N = \left(\frac{c_1 + c_2 * L^{m1}}{1 + c_3 * L^{m1}} \right)^{m_2}, \tag{4.2}$$

where:

- $m1 = 2610/4096 * 1/4$
- $m2 = 2523/4096 * 128$
- $c1 = 3424/4096$
- $c2 = 2413/4096 * 32$
- $c3 = 2392/4096 * 32$,

and where L is the absolute input luminance between 0 and $10,000 cd/m^2$, N is the signal between 0 and 1 ready for quantization to, for example, the 12 bit per colour channel range. The interested reader is referred to [73] and [79] for the reasoning that led to the standard, as well as to [35] and [68] for related proposals on efficient encoding of colour difference signals.

The general form of Equation 4.1 is used in Chapter 5 to adjust image luminance levels, in particular the average luminance level of video content mastered for HDR TVs to our light steering HDR projector in the cinema environment.

Chapter 5

Image Statistics and Power Requirements

The research in this chapter outlines our approach to determine an ideal projector configuration based on an optical model of the new HDR projector architecture and image statistics of target HDR content. We operate under the hypothesis that the traditionally conflicting goals of reproducing high brightness, high contrast HDR content and minimizing power and cost (light source) requirements for such a system can be balanced by understanding content characteristics. The work makes use of an optical simulation and cost model of several projector architectures: the Light Budget Estimator (LBE) in combination with statistical analysis of a variety of representative theatrical HDR content.

5.1 Introduction

Lamp power in a traditional (light attenuating) projection systems linearly scales with the desired peak luminance on screen, whereas the HVS's brightness perception of these luminance values is non-linear (near-logarithmic), see Chapter 3. The projector architecture in Chapter 6 was developed to enable image features that can be reproduced at significantly higher than FSW luminance levels. Figure 5.1 compares on a logarithmic scale today's standardized luminance range in cinema (see Digital Cinema Initiative (DCI) specifications: peak luminance white is defined at $14fL = 48cd/m^2$) with an example of the luminance range that is achievable in a projector architecture using light steering.

5.1.1 Light Steering Efficiency

In an ideal light steering projector, *all* the available source light is used to form the image. In such an architecture the power required to form an image is equal to the average (mean) luminance, or light power of the target image. Two terms that are commonly used in industry and relate to this mean

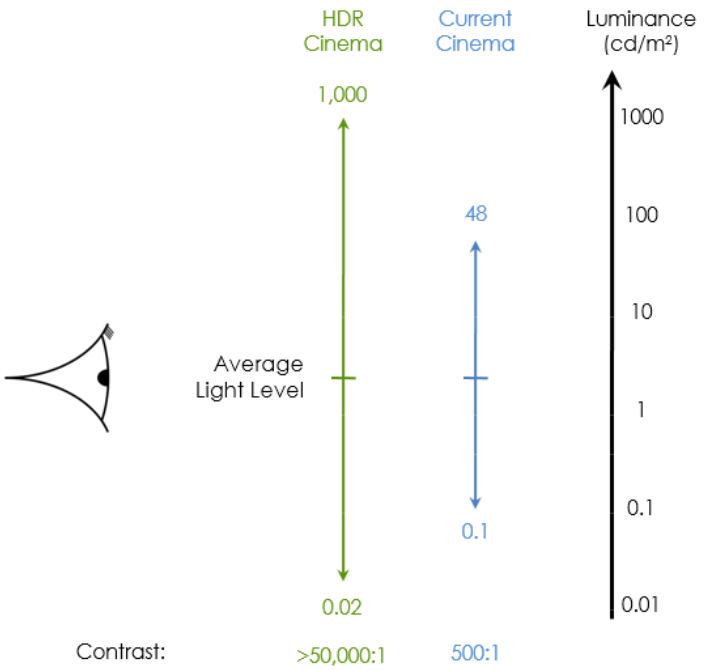


Figure 5.1: *Standardized cinematic SDR luminance range (yellow), compared to an example of the HDR luminance range that could be achieved with a light steering projector architecture (green) on a logarithmic scale to visualize approximate perceptual brightness impact.*

luminance are the Average Picture Level (APL) and the Average Luminance Level (ALL). Since the APL often refers to the average of all code words in a discretized signal or a percentage relative to to the maximum signal value we prefer to use the term ALL in this work. The ALL is the mean luminance of an image. In a practical implementation, such as the prototype in Chapter 6 the ability to steer light into very small bright features is bound for example by the system PSF. Figure 5.2 shows a selection of test patterns with identical minimum and maximum luminance in the source signal, but different mean luminance (power), reproduced on a light steering projector and on a traditional, amplitude attenuating projector. The solid curve shows the theoretical possible maximum peak luminance in an ideal light steering projector, which is:

$$L_{max} = \frac{1}{L_{mean}}. \tag{5.1}$$

The blue curve shows actual measured data on an early light steering prototype. It can be seen that the steering efficiency is not 100%. The non-steered light is in part scattered to areas outside the projection screen. Some of the non-steered light is present on the screen and elevates the black level.

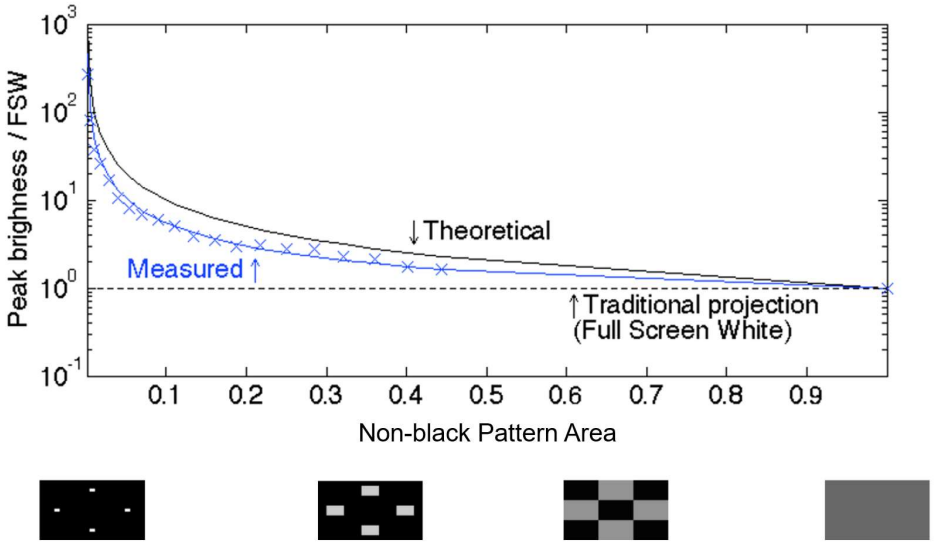


Figure 5.2: *Theoretical and measured steering efficiency for test patterns with varying mean luminance compared to a traditional projection system that uses amplitude modulation to form images.*

5.1.2 Component Efficiency

Introducing additional components into the light path of an existing projection system almost always reduces the total light throughput. In the case of the new proposed projector architecture we have to take into account the reflectivity of the phase modulator ($65 - 80\%$), losses due to higher order diffraction effects off the pixel grid as they appear on any regularly structured surface in the light path ($40 - 50\%$ losses), as well as additionally required optical elements such as relay lenses, diffuser, broadband mirrors to fold the optical path and dichroic mirrors to combine or split colour channels. In an early prototype the total light efficiency of the system was measured at around 5% from source light to screen. This inefficiency is balanced in part with the large gains in peak luminance due to light steering, but is overall

costly, especially for images that have a high ALL. We estimate that in a optimized system (custom coatings to light source wavelengths) the overall efficiency can be as high as $15 - 25\%$, however this is still lower compared to a traditional projector architecture in which, for a full screen white test pattern, around $35 - 45\%$ of source light can reach the screen.

5.1.3 Narrowband Light sources

In the new HDR projector architecture the light steering efficiency, and with it the maximum peak luminance, is highest for well collimated, narrowband light. Laser diodes are well suited to the application. There are however several pitfalls when using laser diodes. These include:

- Cost: other light sources, such as LEDs, lamps and laser+phosphor systems are as of today still more cost-effective in terms of output lumens per dollar.
- Interobserver metamerism variations: light from narrowband light sources has a higher chance of being perceived differently by different observers compared to broadband light. This effect is less pronounced for light that is not narrowband.
- Screen speckle: Small scale interference patterns from laser light can cause small spatial intensity variations on the projection screen that are disturbing. This effect is not visible for broadband light.

5.2 Full Light Steering and Hybrid Light Steering Architectures

Figure 5.3 shows at a high level the proposed system architecture introduced in Chapter 6.

Some of these negative effects of using a laser light source can be countered by either breaking up some of the laser properties (for example by introducing a larger angular diversity) or by mixing light from laser and non-laser based light sources. More interestingly, many overall bright scenes, do not require a low black level, due to adaptation of the human visual system to the brighter parts of the image.

We introduce an efficient hybrid projector architecture, in which light steering narrow-band light is combined with broadband non-steered light (uniformly illuminating an amplitude modulator) into one system to reduce

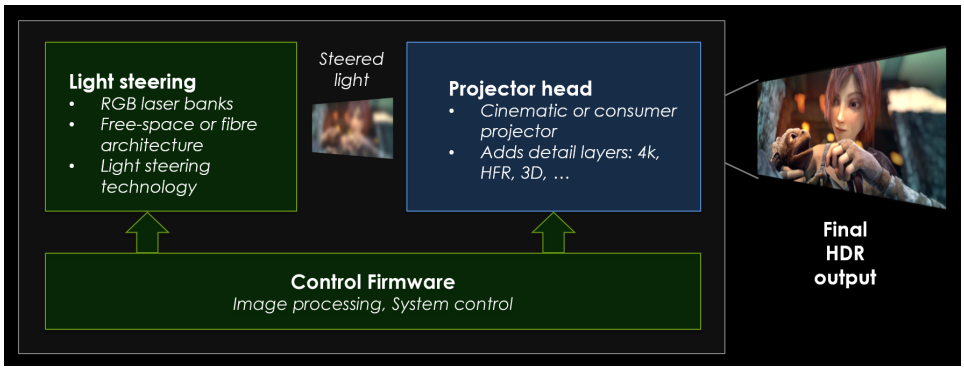


Figure 5.3: *Full light steering architecture with a total possible system efficiency of between* $15 - 25\%$.

system cost further and mitigate certain image artifacts. Figure 5.4 depicts at a high level an example of this hybrid architecture.

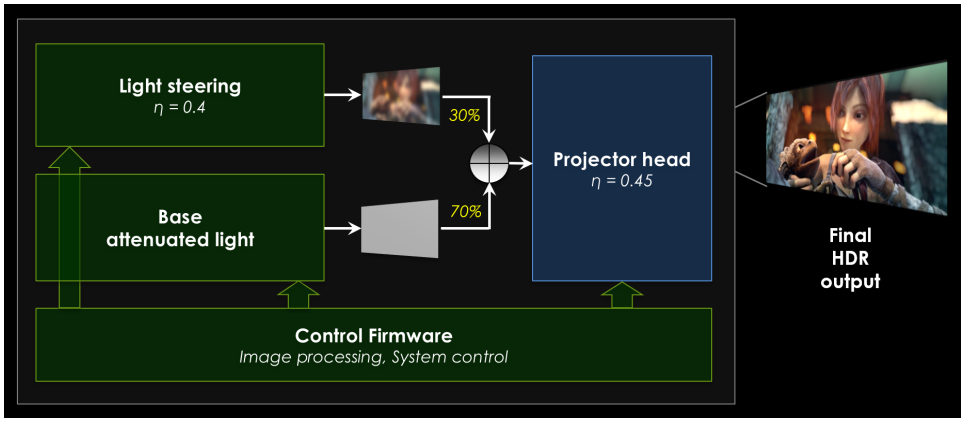


Figure 5.4: *Hybrid architecture to increase overall system light efficiency and reduce cost, while achieving comparable image quality to a full light steering projector. In this example the light efficiency of the steered light path is* 13.5% *and the light efficiency of the non-steered light path is* 31.5%.

Hardware parameters can include the split of source light between the steering and non-steering parts of the system, as well as types of light source and associated cost and spectral properties.

Software parameters can include aspects of the algorithm used to drive the system and allocate light between the steering and non-steering stages. Figures 5.5 shows an example of splitting a linear input signal between a

classical projector light path and a light steering path.

Finally a perceptually meaningful image quality metric needs to be selected and implemented to determine to what extent the desired target image colour and intensity has been faithfully reproduced.

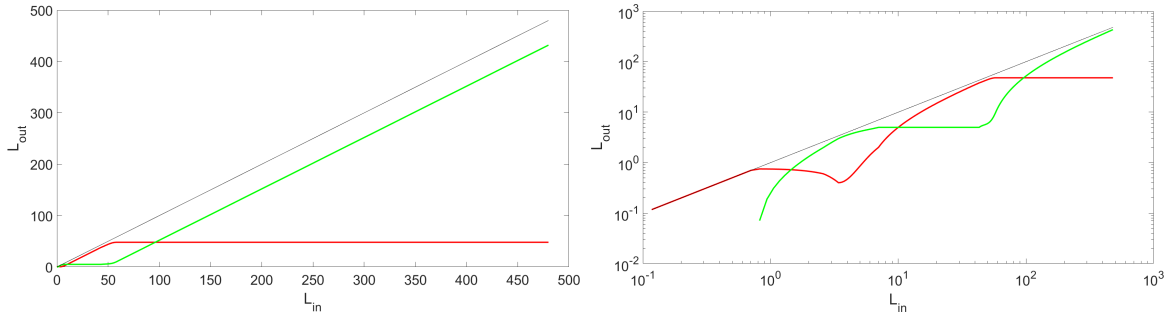


Figure 5.5: *Linear (left) and log (right) plot showing a possible split of an input signal (black) between a light steering projector (green) and a non-light steering projector (red).*

In this section we discuss image statistics that were collected based on non-theatrical HDR image data (mapped to an approximate luminance range) as well as a (very) simplified system model of a full light steering projector.

5.3 Average Luminance of HDR Images in Cinema

Little high-brightness HDR video content is publicly available that has been colour graded for a theatrical viewing environment. Partially this is of course due to the current lack of sufficiently capable large screen projection systems. In this section we attempt to estimate the relative power required to reproduce HDR luminance levels up to 10 times above current peak luminance in cinema using colour graded HDR still images. An analysis of 104 HDR images has been performed, and power requirements for a light steering projector as in the proposed architecture has been estimated. In this theoretical exercise it was found that a light steering projector with less power than a traditional cinema projector can directly reproduce all images up to $48cd/m^2$ and almost all of the surveyed HDR images up to $480cd/m^2$ without the need for additional tonal compression. Table 5.1 summarizes the results.

5.3.1 Methodology

Mark Fairchilds [28] set of 104 scene-referred HDR images was analyzed (see Figure 5.6 for examples). The images differ in dynamic range from less than $1,000 : 1$ to over $10^9 : 1$. Most images are outdoor scenes. While the image data represents measured, scene-referred HDR (actual scene luminance levels) and is not intended for viewing on a cinema projector, an initial guess for a cinema-suitable rendering can be established by shifting the image intensity, so that the ALL approximately matches the estimated viewer adaptation level in cinema. A simple linear scaling operator, $S_{adaptation}$, was determined manually for each image. Images were hand-tuned in a dark viewing environment, on a calibrated $27''$ reference monitor (Dell U2713HMt, calibration confirmed using a Photo Research Inc. PR-650 spectro-radiometer) which was set to a peak white luminance of $48cd/m^2$ (D65 white point). While adjusting the intensity the images were viewed from a distance of approximately 3-5 screen-heights.

Figure 5.6: *Examples from the image set that was used in this study.*

Once an adequate brightness scaling factor had been determined, luminance levels above 10 times that of FSW, $480cd/m^2$, were clipped. Next, the steering efficiency of the proposed projector architecture was accounted for via a system PSF approximation (in this case a somewhat conservative, large Gaussian kernel spanning effectively 81 pixels of 1920 horizontal image pixels). The mean intensity across all pixels of the resulting luminance profile serves as an approximate metric for power requirements of a light steering projector.

Computational steps:

- Compute scaled luminance: $Y_s = Y_{hdr} \times S_{adaptation}$
- Clip Y_s to $10 \times 48cd/m^2 = 480cd/m^2$: $Y_{sc} = min(480, Y_s)$
- Account for steering efficiency: $Y_{scm} = Y_{sc} * g$
- Determine arithmetic mean luminance: $\bar{Y}_{scm} = mean(Y_{scm})$

- Scale to reference ($48cd/m^2$): $P_{rel} = \frac{\bar{Y}_{scm}}{48cd/m^2}$

5.3.2 Results

Figure 5.7 shows the estimated power required to reproduce each HDR image on a light steering projector with peak luminance identical to that of a traditional cinema projector, $48cd/m^2$, and of a light steering projector with a peak luminance one order of magnitude greater than cinema reference systems: $480cd/m^2$. All images can be reproduced on the $48cd/m^2$ light steering architecture using only a fraction of the power (13%) of a traditional projector. More importantly, almost all images can be reproduced up to $480cd/m^2$ (10x higher peak luminance) using the same or less power compared to a traditional projector.

We note that the ALL of our data set when using the scale and clip operations described above with no further artistic colour corrections appears higher than what might be expected from cinema-ready high brightness HDR content. We point the interested reader to [118] for a recent introduction to HDR content production in which significantly lower ALLs (approximately 3% and less) have been reported. This in turn would suggest that the power requirements for a light steering projector architecture could be even lower (or peak luminance and contrast higher) than proposed in our work. For our comparisons on the HDR prototype projector in Section 6.6 we select test images within a range of relatively conservative (=high) ALLs from 7% - 45% (see Table 6.2, second column).

Table 5.1: *Power required to reproduce the images from the HDR data set on three different projectors (relative to a standard cinema projector in the first row).*

L_{peak}	Steering?	P_{rel} (min)	P_{rel} (median)	P_{rel} 90%tile
$48cd/m^2$	no	1	1	1
$48cd/m^2$	yes	0.0107	0.1079	0.2595
$480cd/m^2$	yes	0.0107	0.1832	0.8554

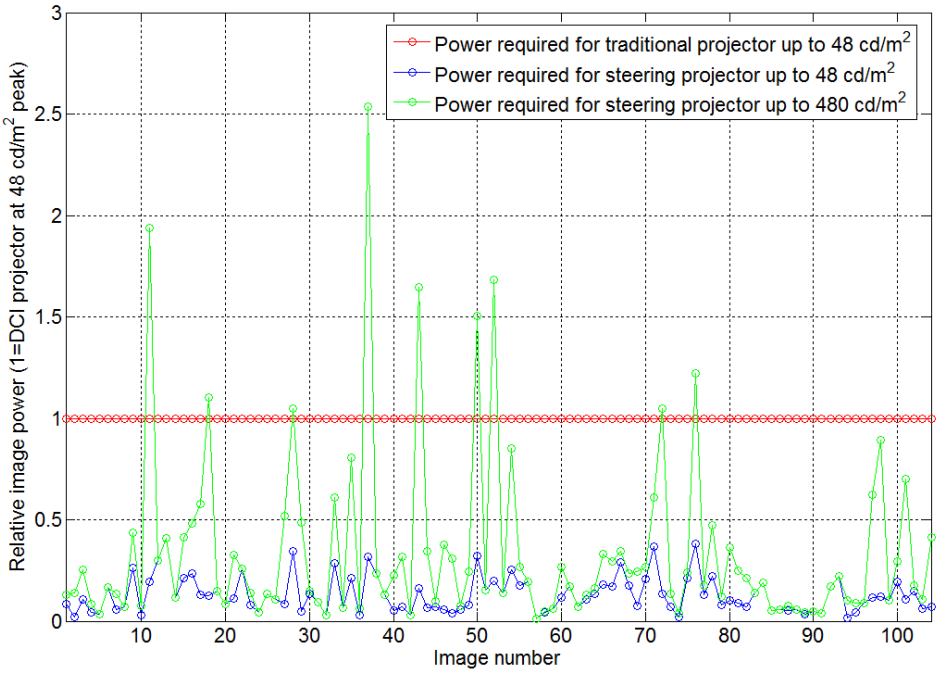


Figure 5.7: *Relative power required to reproduce each of 104 HDR images on three hypothetical projectors: two light steering projectors with a peak luminance of* $48cd/m^2$ *(blue) and* $480cd/m^2$ *(green) relative to a traditional, light blocking cinema projector with peak luminance of* $48cd/m^2$ *(red). The average power required to achieve identical peak luminance (*$48cd/m^2$*) is on the order of 13% of a traditional projector. More importantly, all but the very brightest images (approximately 9% of all images under test), can be re-produced up to a peak luminance of* $480cd/m^2$ *while using less or the same amount of power.*

Chapter 6

Freeform Lensing Phase Retrieval and Light Steering HDR Projector Proof of Concept

6.1 Introduction

Ideally, HDR projectors *should* produce both darker black and (much) brighter highlights while at the same time maintaining an appropriate-for-the-viewing-environment ALL. However, todays HDR projectors predominantly focus on improving black level (for example recently demonstrated laser projection systems by Kodak, IMAX, Zeiss, Dolby, Barco, Christie and others). Improved contrast *and* peak luminance are vital for higher perceived image quality (brightness, colourfulness) [96]. Brightness perception of luminance levels is near-logarithmic in the photopic range. Doubling the luminance of an image feature on a projection screen (e.g. by increasing the lamp power of a traditional projector by 2×) does not result in a significant improvement in perceived brightness.

Results in [93, 97, 118] suggest that 10×, 20× or even 100× increases in peak luminance would be desirable, even if most images only require a very small percentage of pixels to be this bright (also see Section 5.3).

Such drastic improvements cannot be achieved with conventional projector designs, which use amplitude SLMs to generate images by pixel-selectively blocking light. For a typical scene, this process destroys between 82% and 99% of the light that *could* reach the screen, with the energy being dissipated as heat. This causes a number of engineering challenges, including excessive power consumption, thermal engineering, and cost, which ultimately limit the peak luminance in current projector designs.

We explore the use of *dynamic* freeform lenses in the context of light efficient, high (local) peak luminance, and high contrast (high dynamic

range, HDR) projection systems. Freeform lenses, i.e. aspherical, asymmetric lenses have recently received a lot of attention in optics as well as computer graphics. In the latter community, freeform lenses have mostly been considered under the auspices of *goal-based caustics*, i.e. the design of lenses that generate a specific caustic image under pre-defined illumination conditions [33, 84, 100, 115].

We implement dynamic freeform lensing on a *phase-only* SLM, which is combined with a conventional light blocking device such as a reflective LCD in a new type of cascaded modulation approach. The phase modulator in our approach creates a smooth, but still quite detailed "caustic" image on the amplitude modulator. Since the caustic image merely redistributes, or "reallocates", light [46], this approach produces both a higher dynamic range as well as an improved (local) peak luminance as compared to conventional projectors.

6.2 Contributions

We offer the following contributions:

- A new approach to generate freeform lenses or goal driven caustics using common approximations in optics to directly optimize the phase modulation pattern or lens shape of a freeform lens ([18]).
- A new dual-modulation projector design that combines one phase and one amplitude modulator for image generation and enables high brightness, high contrast images.

To our knowledge this is both the first time that practical high-resolution dynamic light redistribution has been shown using *commercially available* hardware, as well as the first time that phase-only SLMs have been used for dynamic freeform lensing.

6.3 Phase Modulation Image Formation

To derive the image formation model for a phase modulation display, we consider the geometric configuration shown in Figure 6.1: a lens plane and an image plane (e.g. a screen) are placed parallel to each other at focal distance f. Collimated light is incident at the lens plane from the normal direction, but a phase modulator (or lens) in the lens plane distorts the phase of the light, resulting in a curved phase function $p(\mathbf{x})$, corresponding to a local deflection of the light rays.

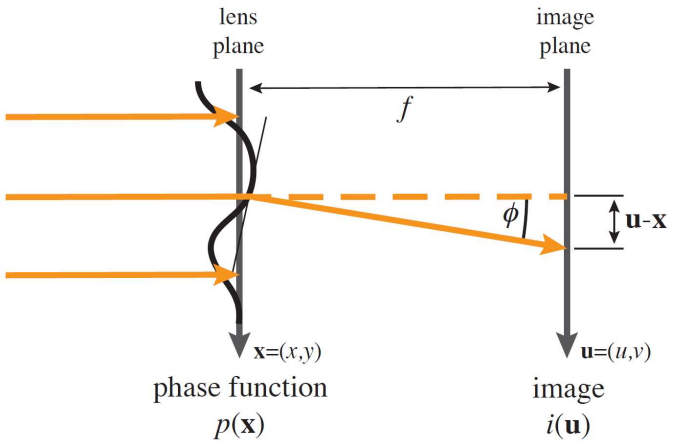


Figure 6.1: *Geometry for the image formation model. Phase modulation takes place in the lens plane, which is placed at a focal distance of f from the image plane. This results in a curvature of the wavefront, represented by a phase function $p(\mathbf{x})$.*

Light deflection. With the *paraxial approximation* $\sin\phi \approx \phi$, which is valid for small deflection angles, we obtain in 2D that

$$u - x = f \cdot \sin\phi \approx f \cdot \frac{\partial p(x,y)}{\partial x}. \tag{6.1}$$

In 3D this leads to the following equation for the mapping between $\mathbf{x}$ on the lens plane and $\mathbf{u}$ on the image plane:

$$\mathbf{u}(\mathbf{x}) = \mathbf{x} + f \cdot \nabla p(\mathbf{x}). \tag{6.2}$$

Intensity modulation. With the above mapping, we now need to derive the intensity change associated with this distortion. Let $d\mathbf{x}$ be a differential area on the lens plane, and let $d\mathbf{u} = m(\mathbf{x}) \cdot d\mathbf{x}$ be the differential area of the corresponding region on the image plane, where $m(.)$ is a spatially varying magnification factor. The intensity on the image plane is then given as

$$i(\mathbf{u}(\mathbf{x})) = \frac{d\mathbf{x}}{d\mathbf{u}} i_0 = \frac{1}{m(\mathbf{x})} i_0, \tag{6.3}$$

where i_0 is the intensity of the collimated light incident at the lens plane. In the following we set $i_0 = 1$ for simplicity of notation.

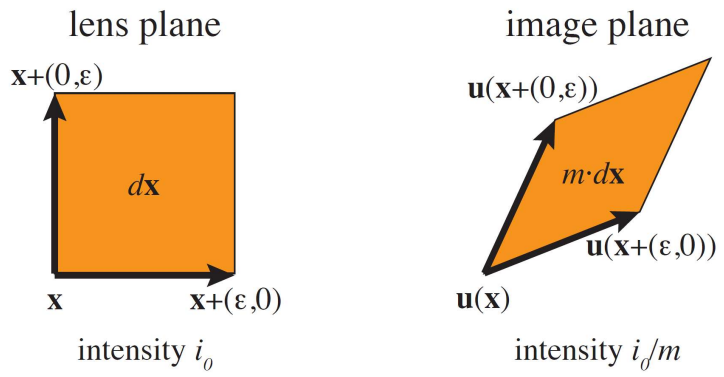


Figure 6.2: *Intensity change due to the distortion of a differential area* $d\mathbf{x}$.

The magnification factor $m(.)$ can be expressed in terms of the derivatives of the mapping between the lens and image planes (also compare Figure 6.2):

$$\begin{aligned} m(\mathbf{x}) &= \left(\frac{\partial}{\partial x}\mathbf{u}(\mathbf{x})\right) \times \left(\frac{\partial}{\partial y}\mathbf{u}(\mathbf{x})\right) \\ &\approx 1 + f\frac{\partial^2}{\partial x^2}p(\mathbf{x}) + f\frac{\partial^2}{\partial y^2}p(\mathbf{x}) \\ &= 1 + f \cdot \nabla^2 p(\mathbf{x}). \end{aligned} \tag{6.4}$$

This yields the following expression for the intensity distribution on the image plane:

$$i(\mathbf{x} + f \cdot \nabla p(\mathbf{x})) = \frac{1}{1 + f \cdot \nabla^2 p(\mathbf{x})}. \tag{6.5}$$

In other words, the magnification m, and therefore the intensity $i(\mathbf{u})$ on the image plane can be directly computed from the Laplacian of the scalar phase function *on the lens plane*.

6.4 Optimization Problem

While it is possible to directly turn the image formation mode from Equation 6.5 into an optimization problem, we found that we can achieve better convergence by first linearizing the equation with a first-order Taylor approximation, which yields

$$i(\mathbf{x} + f \cdot \nabla p(\mathbf{x})) \approx 1 - f \cdot \nabla^2 p(\mathbf{x}), \tag{6.6}$$

where the left hand side can be interpreted as a *warped image* $i_p(\mathbf{x}) = i(\mathbf{x} + f \cdot \nabla p(\mathbf{x}))$ where the target intensity $i(\mathbf{u})$ in the image plane has been warped backwards onto the lens plane using the distortion $\mathbf{u}(\mathbf{x})$ produced by a given phase function $p(\mathbf{x})$.

With this parameterization, the continuous least-square optimization problem for determining the desired phase function becomes

$$\hat{p}(\mathbf{x}) = \text{argmin}_{p(\mathbf{x})} \int_{\mathbf{x}} \left(i_p(\mathbf{x}) - 1 + f \cdot \nabla^2 p(\mathbf{x}) \right)^2 \, d\mathbf{x}. \tag{6.7}$$

This problem can be solved by iterating between updates to the phase function and updates to the warped image, as shown in Algorithm 1. The algorithm is initialized with the target image intensity. From this, the first phase pattern is computed, which in turn is used to warp the original target image intensity to provide a distorted intensity image for use in the next iteration.

Algorithm 1 Freeform lens optimization

```
// Initialization
i_p^0(x) = i(u)
while not converged do
  // phase update
  p^(k)(x)=argmin_p(x) ∫_x ( i_p^(k-1)(x) − 1 + f · ∇²p(x) )² dx
  // image warp
  i_p^(k)(x)=i(x + f · ∇p^k(x))
end while
```

After discretization of $i(.)$ and $p(.)$ into pixels, the phase update corresponds to solving a linear least squares problem with a discrete Laplace operator as the system matrix. We can solve this positive semi-definite system using a number of different algorithms, including Conjugate Gradient, BICGSTAB and Quasi Minimal Residual (QMR). The image warp corresponds to a texture mapping operation and can be implemented on a GPU. We implement a non-optimized prototype of the algorithm in the Matlab programming environment using QMR as the least squares solver. Table 6.1 shows run times for Algorithm 1 and a selection of artificial and natural test images at different resolution. It was executed on a single core of a mobile Intel Core i7 clocked at 1.9 GHz with 8 GByte of memory.

We note that due to the continuous nature of the resulting lens surfaces, computation of the phase with resolutions as low as 128×64 are sufficient

Table 6.1: *Run times of Algorithm 1 using five iterations for a set of different test images and image resolutions.*

Image	Resolution	Runtime
Logo	128×64	2.62 s
Lena	128×64	2.14 s
Wave	128×64	1.81 s
Logo	256×128	4.03 s
Lena	256×128	4.75 s
Wave	256×128	3.23 s
Logo	512×256	9.37 s
Lena	512×256	10.22 s
Wave	512×256	5.27 s

for applications such as structured illumination in a projector. We also note that the algorithm could, with slight modifications, be rewritten as a convolution in the Fourier domain which would result in orders of magnitude shorter computation time for single threaded CPU implementations and even further speed-ups on parallel hardware such as GPUs. With these improvements, computations at, for example, 1920×1080 resolution will be possible at video frame rates. In addition both, the resulting contrast of the caustic image as well as the sharpness (effective resolution), benefit from higher working resolution.

The progression of this algorithm is depicted in Fig. 6.3. We show the undistorted target image, from which we optimize an initial phase function. Using this phase function, we update the target image in the lens plane by backward warping the image-plane target. This process increasingly distorts the target image for the modulator plane as the phase function converges. The backward warping step implies a non-convex objective function, but we empirically find that we achieve convergence in only a small number of iterations (5-10).

6.5 Simulation Results

We evaluate the performance of our algorithm by utilizing different simulation techniques: a common computer graphics ray tracer and a custom wavefront model based on the Huygens-Fresnel principle to simulate diffraction effects at a spectral resolution of 5nm.

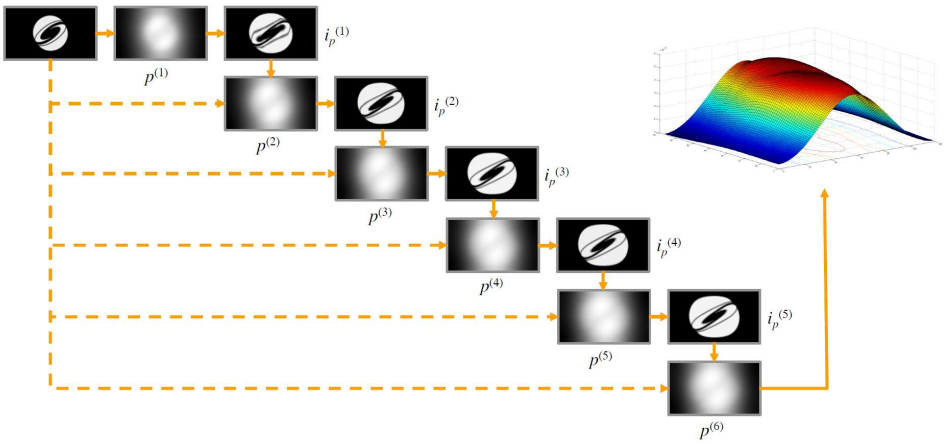


Figure 6.3: *Algorithm progression for six iterations: target i gets progressively distorted by backwards warping onto lens plane $i_p^{(k)}$ as phase function $p^{(k)}$ converges towards a solution. The 3D graphic depicts the final lens height field.*

6.5.1 Ray Tracer Simulation

For the ray tracer simulation we use the LuxRender framework, an unbiased, physically-based rendering engine for the Blender tool. The setup of the simulation is quite straightforward: the freeform lens is imported as a mesh, and material properties are set to mimic a physical lens manufactured out of acrylic.

A distant spot light provides approximately collimated illumination, a white surface with Lambertian reflectance properties serves as screen. The linear, high dynamic range data output from the simulation is tone mapped for display. The results (see Fig. 6.4) visually match the target well.

6.5.2 Physical Optics Simulation

To analyze possible diffraction effects that cannot be modeled in a ray tracer based on geometric optics principles, we perform a wave optics simulation based on the Huygens–Fresnel principle. We compute a freeform lens surface for a binary test image (see Fig. 6.5) and illuminate it in simulation with light from a common 3-LED (RGB) white light source (see Fig. 6.6, dotted line) in 5nm steps. We integrate over spectrum using the luminous efficiency of the LED and the spectral sensitivity curves of the CIE colour matching

Figure 6.4: *LuxRender simulation results of a caustic image caused by an acrylic freeform lens. The inset shows the absolute intensity difference between simulated and original image, where the original image is encoded in the interval [0-1], and 0 in the difference map (green) means no difference. There are three possible sources of error: reflections off the edges of the physically thick lens (vertical and horizontal lines), misalignment and scaling of the output relative to the original (manual alignment) and the nature of the light source (not perfectly collimated).*

functions (see Fig. 6.6, solid line), as well as a 3x3 transformation matrix and a 2.2 gamma to map tristimulus values to display/print RGB primaries for each LED die and for the combined white light source (see Fig.6.7). As expected, the wavefront simulation reveals chromatic aberrations within the pattern and diffraction off the edge of the modulator, which can be (partially) mitigated, for example, by computing separate lens surfaces for each of R,G and B.

The phase function $p(\mathbf{x})$ can be used directly to drive a digital phase modulation display. However, if instead we would like to create a refractive lens surface out of a transparent material, then this phase function needs to be converted to a geometric model for the lens shape.

Similar to other research on goal-based caustics [84], we assume a lens shape that is flat on one side, and a freeform height field $h(\mathbf{x})$ on the other side (see Figure 6.9). In the (x, z) plane, the deflection angle ϕ is related to the incident (θ_i) and the exitant (θ_o) angles at the height field (also see

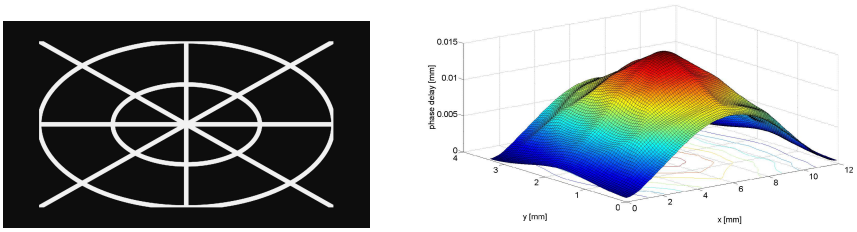


Figure 6.5: *Binary test pattern (left) and resulting lens height field (right) used in the wave optics simulation.*

Figure 6.1) as follows:

$$\frac{\partial p(\mathbf{x})}{\partial x} \approx \phi = \theta_o - \theta_i. \tag{6.8}$$

The analogous relationship holds in the (y, z) plane.

In addition, the lens material has a refractive index of n. Using Snell's law, and again the paraxial approximation, we obtain

$$\frac{1}{n} = \frac{\sin\theta_i}{\sin\theta_o} \approx \frac{\theta_i}{\theta_o}. \tag{6.9}$$

Using Equations 6.8 and 6.9, as well as $\theta_i \approx \partial h(\mathbf{x})/\partial x$, we can derive the lens shape as

$$h(\mathbf{x}) = h_0 + \frac{1}{n-1} p(\mathbf{x}), \tag{6.10}$$

where h_0 is a base thickness for the lens.

We note that the height is a linear function of the phase, and the refractive index n shows up only as a scalar multiplier to the phase function $p(.)$.

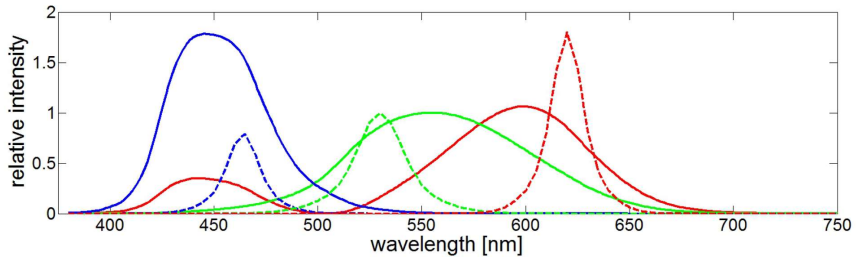
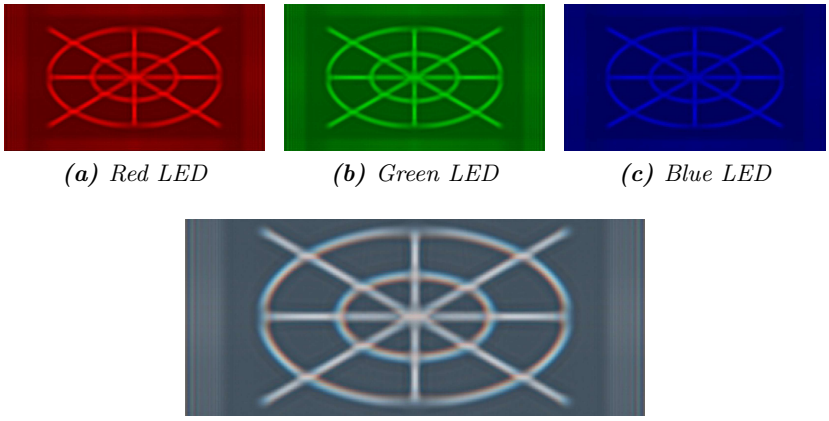


Figure 6.6: *Spectra of standard white 3-LED (RGB) [82] (dotted graph) and the CIE standard observer colour matching functions (solid graph) used in the wave optics simulation.*



(a) Red LED *(b) Green LED* *(c) Blue LED*

(d) Combined (RGB) White LED

Figure 6.7: *Wave optics simulation for a test lens using standard white 3-LED (RGB) spectra. The simulation was performed at 5nm intervals and mapped to a RGB colour space for print.*

Since p itself is approximately linear in the focus distance f (Equation 6.5), we can see that uniform scaling of the height field and uniform changes of the refractive index simply manifest themselves as a refocusing of the lens. This also shows that it is equivalently possible to adjust the optimization procedure to directly optimize for $h(.)$ instead of $p(.)$. We chose to optimize for phase because we desire the use of spatial phase modulators for applications in video projectors.

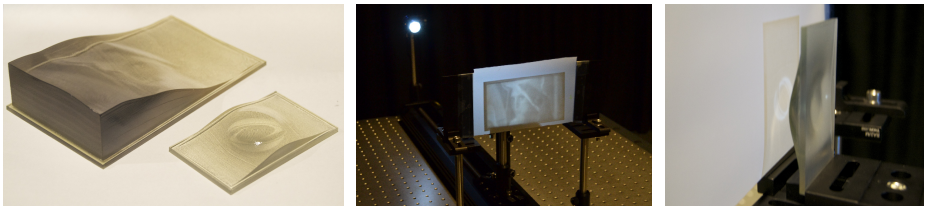


Figure 6.8: *Results using 3D-printed refractive lenses. The left image shows the lenses themselves, while the center and right images show the caustics generated by them: the Lena image and a Siggraph logo. Due to resolution limits on the 3D printer, the lens dimensions have been optimized for large feature scales, which results in a short focal length.*

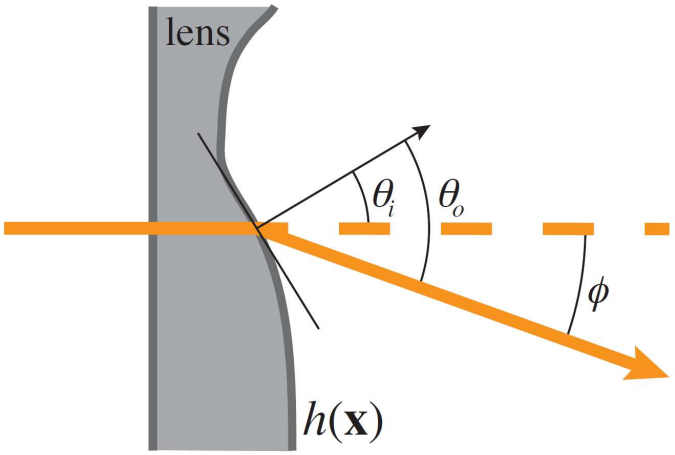


Figure 6.9: *Geometry for refraction in a freeform lens defined by a height field* $h(\mathbf{x})$.

6.5.3 Refractive 3D Printed Lens Results

Figure 6.8 shows results for goal-based caustics using refractive freeform lenses generated with our method. The lenses (shown on the left of Figure 6.8) were 3D printed on an Objet Connex 260 rapid prototyping machine using VeroClear material. Afterwards, the lenses were thoroughly cleaned and the flat side was manually polished using fine grained sand paper and polishing paste. This type of 3D printer has a layer thickness of $42\mu m$, which limits the feature size that we can create.

As discussed above, the model can be re-scaled to achieve different focal distances. To accommodate the resolution limits of the fabrication method, we chose very short focal distances of (about 1" for the Siggraph logo and 5" for the Lena image). Although these scales test the very limits of the paraxial approximation used in the derivation of our image formation model, the image quality is still quite good. With better fabrication methods such as injection molding, high precision milling [87], or even detailed manual polishing of a 3D printed surface, one could both improve the image quality and reduce the feature size, so that far field projection becomes feasible.

6.5.4 Static Phase Plates

We evaluate a selection of phase patterns using static phase plates with dimensions comparable to that of a phase-only SLM (approximately 12mm

x 7mm). Figure 6.10 shows two phase plates that were manufactured on a fused silica wafer using a lithography based process as well as the resulting light fields when illuminated with a collimated laser beam. The spatial resolution of the phase pattern is high at 12,288 x 6912 pixels per lens (1.00 µm pixel pitch). The phase resolution was limited to 8 phase levels (4 masks in the lithography process resulted in 8 phase levels between 0 and 2 pi). The static phase plates were manufactured to evaluate our freeform lenses for static beam shaping applications and to test the method using high power lasers (in this case a fiber coupled laser) in which spatial coherence properties are not preserved. The laser used for experiments (see results in Figure 6.10) is a 638nm laser with up to 60W optical power. A collimation lens doublet is used to expand the beam and to provide an approximately collimated (but slightly diverging) beam of light to illuminate the phase plates. Spatial coherence is partially broken up as multiple laser sources within the module are combined and coupled into a 400 µm fiber which integrates light travelling along multiple light paths. The ability to focus light from this source is limited (this applies not only to our phase plates, but also to ordinary glass lenses), and thus the sharpness of the resulting image is affected. For the proposed application as a structured light source in a projection system, a small amount of blur in the image is acceptable, if not desirable. The ray-tracing simulations in Section 6.5.1 make use of a perfectly collimated light source and hence blur is not modeled, however the overall geometry of the real light distribution is visually undistorted and the contrast is comparable to the simulations. For the projector prototype described in Section 6.6 we use a free space laser and achieve a similar amount of blur with a diffuser.

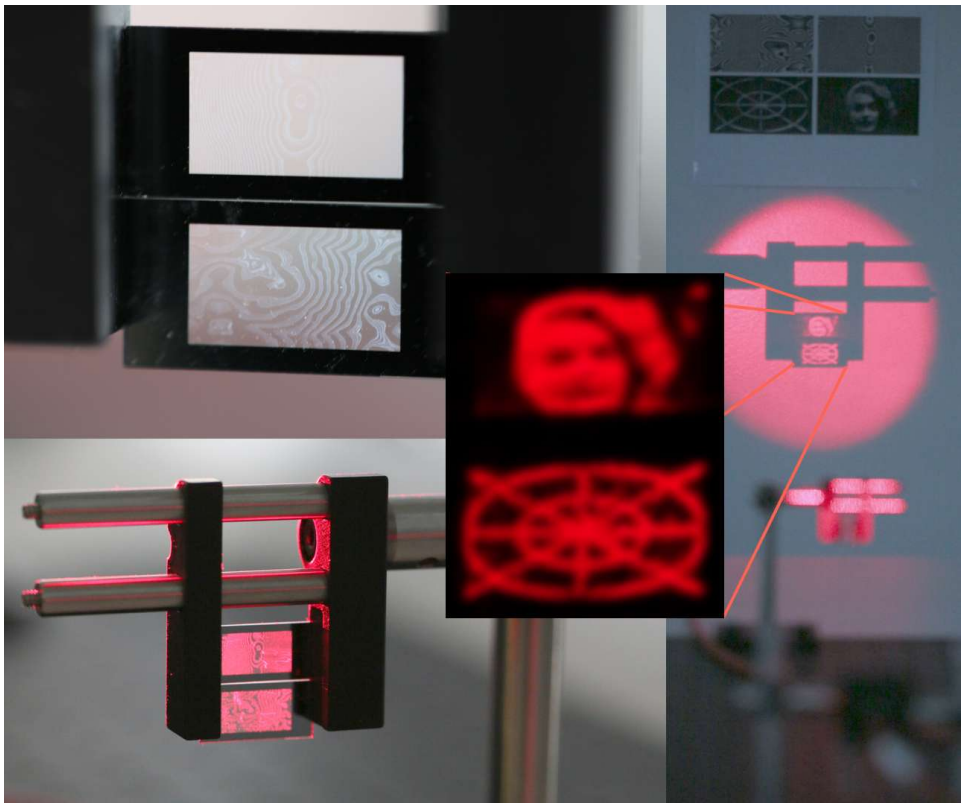


Figure 6.10: *Static phase plates manufactured on a fused silica wafer for two test patterns (Marilyn and Align). Left: phase plates without and with laser illumination. The top phase plate reproduces Marilyn, the lower phase plate focuses the Align pattern. Right: The fiber-coupled and beam expanded laser light source as well as two phase plates mounted in the light path (out of focus in the photo) are shown in the foreground. The projected structured illumination pattern is focused on the screen. For reference, above the red projection is a printed copy of the phase pattern etched into the wafer as well as a print-out of a wave-front simulation of the expected intensity distribution on screen. The light pattern visually matches the simulation well.*

6.6 Dynamic Lensing in Projection Systems

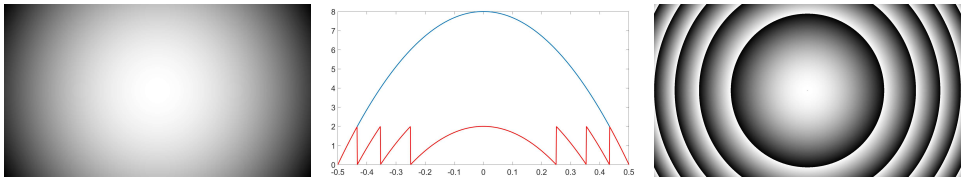
In order to apply the freeform lens concept in projection displays, we require a spatial light modulator that can manipulate the shape of the wavefront of

reflected or transmitted light. We first provide a brief overview of the different technologies available for this purpose, and then describe experiments and prototypes using this technology.

There are several commercially available technologies that can rapidly manipulate the phase of an incident wavefront on a per-pixel basis. These devices include MEMS displays such as analog mirror arrays [44] or deformable mirrors for use in wavefront sensing and correction applications. The benefit of MEMS-based devices is the very fast temporal response. Disadvantages include cost and availability as well as the relatively low spatial resolution: devices with 4096 actuators currently mark the upper end of the range in this domain leaving a resolution gap of 3 orders of magnitude to common projection displays.

An alternative to MEMS based mirrors is offered by liquid crystal displays either in form of a transmissive LCD or in a reflective configuration: LCoS. Liquid crystal displays can retard the phase of light and offer high spatial resolution. Reflective LCoS devices can update at higher switching speeds compared to transmissive LCD due to the reduced cell gap and provide a high pixel fill factor. Omitting the input/output polarizing beam splitter and careful management of the polarization of incoming light allows for the operation in phase-only mode in which phase is retarded based on the rotation of the liquid crystals in each pixel. Although standard LCoS modules can in principle be used as phase modulators, dedicated SLMs are available that can be calibrated to shift phase by one wavelength or more which allows for the implementation of *steeper* lenses that steer light more aggressively. The pixel values of the LCoS module then correspond directly to the wavelength modulated phase function, i.e. $\mathrm{mod}(p(x), \lambda)$. For more on this topic we refer the reader to [98].

Our choice of phase SLM is a reflective LCoS chip distributed by [42]. It provides a spatial resolution of $1920 * 1080$ discrete pixels at a pixel pitch of $6.4\mu m$, and can be updated at up to $60Hz$. Access to a look-up-table allows for calibration of the modulator for different working wavelengths. The fill factor and reflectivity of the display are high compared to other technologies at 93% and 75% respectively. The phase retardation is calibrated to between 0 and 2π, equivalent to one wavelength of light. This is sufficient to generate freeform lenses with a long focal distance. For shorter focal distances, we require more strongly curved wavefronts, which creates larger values for $p(.)$. We can address this issue by *phase wrapping*, i.e. using only the fractional part of $p(.)$ as drive signal for the phase SLM. This results in a pattern similar to a Fresnel lens (see Figure 6.11 and also red box in Figure 6.12 as well as the phase pattern in Figure 6.13).



(a) 0 *to* 8π *spherical lens* *(b)* *cross section* *(c)* *phase-wrapped at* 2π

Figure 6.11: *Phase wrapping example: (a) phase function of a spherical lens with a height of* 8π*, (b) a plot of the cross section of the original and the phase wrapped lens, and (c) the same lens wrapped at intervals of* 2π*.*

6.6.1 Monochromatic Prototype

We initially analyze image results from the first (phase) modulation stage in isolation (see Figure 6.12) and later relay the resulting light profile into the second stage for amplitude attenuation.

The use of single frequency lasers causes small-scale artifacts including screen speckle and diffraction *fringes* due to interference (Figure 6.13, center photo). As previously mentioned these artifacts can be reduced below the noticeable visible threshold by using for example a set of lasers with different center wavelengths or broadband light source such as LED and lamps. When constraint to using a narrowband light source such as in our test setup, a similar image *smoothing* effect can be achieved by spatially or temporally averaging the image using for example a diffuser or commercially available continuous deformable mirrors that introduces slight angular diversity in a pseudo-random fashion at high speeds. For ease of implementation we choose to use a thin film diffuser placed in an intermediate image plane following the phase SLM. A photo of the *cleaned-up* intensity profiles can be seen in Figure 6.13, right.

We demonstrate a first prototype of a high brightness, high dynamic range projection system, in which we form a first image based on our dynamic lensing method and provide additional sharpness and contrast using a traditional LCoS-based amplitude modulating display.

At a high level, the light path of a traditional projection system consist of a high intensity light source and some form of beam homogenization, for example beam expansion, collimation and homogenization, colour separation and recombining optics. At the heart of a typical projector, a small SLM attenuates the amplitude of light per pixel. The resulting image is then magnified and imaged onto the projection screen.

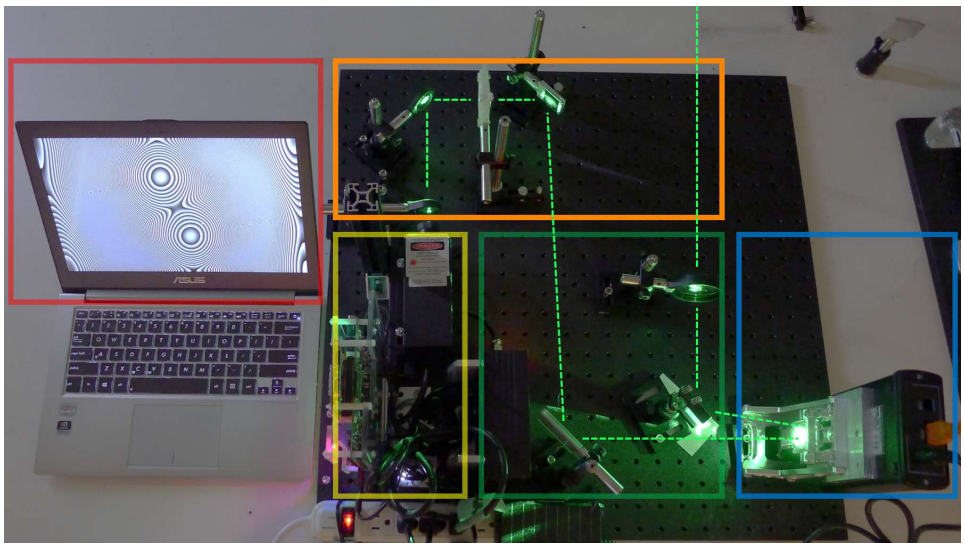


Figure 6.12: *Single modulation test setup for lasers consisting of a light source (yellow box, 532nm DPSS laser and laser controller), beam expansion and collimation optics (orange box), the reflective phase SLM (blue), various folding mirrors and a simple projection lens to relay the image from and intermediate image plane onto the projection screen (green). The phase pattern shown on the computer screen correlates linearly to the desired phase retardation in the optical path to form the image. It has been phase-wrapped at multiples of one wavelength and can be addressed directly onto the micro display SLM.*

We largely keep this architecture intact, but replace the uniform illumination module with both, the collimated laser illumination and a phase SLM (Figure 6.14). Our lensing system is inserted between the light source and the existing SLM, and forms an approximate light distribution on a thin-film diffuser in an intermediate image plane which is then relayed onto the image plane of the amplitude SLM. The freeform lensing approach redistributes light from dark image regions to bright ones, thus increasing both contrast and local peak brightness, which is known to have a significant impact on visual realism [97].

We make use of the forward image formation model from our simulations for the light steering phase to predict the illumination profile present at the second, amplitude-only modulator. Given the phase function from the freeform lensing algorithm, the light distribution on the image plane is predicted using the model from Equations 6.2 and 6.4. The amount of smoothness introduced at the diffuser at the intermediate image plane can be modelled using a filter kernel that approximates the PSF and the modulation pattern required for the amplitude modulator is then obtained to introduce any missing spatial information as well as additional contrast where needed. We note that careful calibration and characterization of the entire optical system is required to optimally drive the SLMs. No significant efforts beyond careful spatial registration of the two images (illumination profile caused by phase retardation and amplitude modulation on the SLM) and calibration to linear increments in light intensity were performed for this work.

6.6.2 Results

Figure 6.13 shows the phase patterns computed by Algorithm 1 as applied to the phase modulator with black corresponding to no phase retardation

Figure 6.13: *From left to right correlating to positions A to C in Figure 6.14: A: phase pattern present at phase-only LCoS modulator, B: direct image produced by lens in intermediary image plane (prior to diffuser) and C: intensity distribution present at amplitude LCoS modulator after having passed through a thin-film diffuser.*

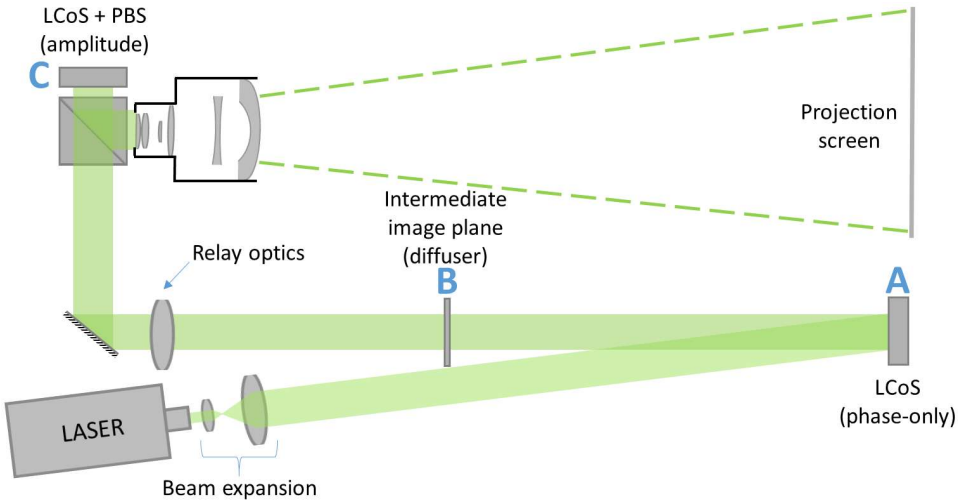


Figure 6.14: *System diagram of the proposed and prototyped high brightness, HDR projector: light from an expanded and collimated laser beam is reflected off a phase-only modulator. The per-pixel amount of phase retardation resembles the height field of the dynamic lens calculated with our algorithm. The effective focal plane of this freeform lens is in-plane with an off-the-shelf, reflective projection head consisting of the polarizing beam splitter together with an LCoS microdisplay and a projection lens. Light from dark parts of the image can be used to create high luminance features, and simultaneously reduce the black level.*

and white corresponding to a retardation of 2π. We illustrate how phase patterns with maximum phase retardation larger than 2π can be wrapped to the maximum phase retardation of the modulator, resulting in a pattern similar to a Fresnel lens. The resulting light profile resembles the target image closely, but also contains a small amount of local, high spatial frequency noise. We make use of a patterned diffuser (0.5 degree half-angle) to integrate over these local intensity variations. The resulting light profile at the diffuser is locally smooth and still provides sufficient contrast to enhance peak luminance and lower black level.

Figure 6.15 shows a selection of experimental results for our method. The first row of Figure 6.15 shows the phase pattern addressed onto the phase SLM. In the forth row of Figure 6.15 we show photos of the light steering high brightness projector and compare them to what a traditional projector with the same lumen rating out of lens would look like (second row). For the latter case we address a flat phase across the phase SLM. Rows three and five show false-colour logarithmic luminance plots on the matching scales for the traditional and light steering projector systems. All photos were captured with identical camera settings and show that our method not only recovers better black levels but also allows for significantly elevated peak luminance in highlights by redistributing light from dark regions of the image to lighter regions by making better use of available light source power. This enables high brightness high-dynamic range projection with drastically reduced power consumption when compared to dual amplitude modulation approaches.

Table 6.2: *Luminance measurements of the results depicted in Figure 6.15. Multiple exposures at varying exposure times were capture (8s, 4s, 2s, 1s, 1/2s, 1/4s, 1/8s, 1/15s, 1/30s, 1/60s, 1/125s) and combined into one linear HDR file, which was then calibrated to represent actual luminance values using a luminance spot meter (Minolta LS100). The lowest accurate measurement using the Minolta LS100 is* 0.001 cd/m^2*. We note that the relative power P_{rel} of the test patterns is significantly higher than what might be expected in theatrical high brightness HDR content and thus the gain in L_{peak} and in contrast could be even higher.*

Name (P_{rel})	HDR L_{peak} $[cd/m^2]$	HDR L_{black} $[cd/m^2]$	HDR contrast	SDR L_{peak} $[cd/m^2]$	SDR L_{black} $[cd/m^2]$	SDR contrast	**L_{peak} gain**	**Contrast gain**
SG logo (7%)	701	0.001	700900 : 1	46	0.01	4,272 : 1	**15X**	**173X**
Lena (48%)	121	0.03	4053 : 1	42	0.83	50 : 1	**3X**	**80X**
Marilyn (25%)	407	0.03	13008 : 1	41	0.63	64 : 1	**10X**	**203X**
Align (20%)	180	0.01	29677 : 1	45	0.44	101 : 1	**4X**	**292X**
Einstein (15%)	348	0.001	347700 : 1	44	0.01	2,996 : 1	**8X**	**122X**

Figure 6.15: *Result photos and measurements of the HDR prototype projector. Top to bottom: Phase Function of Lens - the phase pattern as computed by our algorithm. SDR projector for comparison - projector with identical lamp power (out of lens) used in a traditional, light attenuating mode: a uniform light field (flat phase field) is provided to the amplitude SLM which forms the image by blocking light. SDR luminance profile on a logarithmic scale. HDR projector - photograph of our lensing approach used to redistribute light from dark regions to bright regions, resulting in improved black levels and significantly increased highlight intensity. HDR luminance profile on a logarithmic scale.*

6.6.3 Real-time Freeform Lensing

While pre-processing of video frames at non-real-time frame rates is acceptable for some applications such as cinema and fixed installations using projectors, a real-time implementation is ultimately desired.

Optimization Using Fourier-Domain Solves. The key insight is that by mirror padding the input image the system arising from the discretization of $(\nabla^2 p)^2$ results in periodic boundary conditions with pure-Neumann boundary conditions at the nominal image edge. This is illustrated in Figure 6.16 and was also observed in earlier work by Ng et al. [80] for deblurring images, but has not been exploited for lensing. The modification allows the product $(\nabla^2 p)^2$ in the objective function, Equation 6.7, to be expressed as a convolution via the Fourier convolution theorem since the system matrix resulting from discretizing Equation 6.7 is circulant. This enables the use of faster Fourier-domain solves (operations) in place of slower general purpose iterative linear solvers.

We build upon the method summarized in Section 6.4 and note that for periodic boundary conditions, this problem can be solved very efficiently in Fourier-space by using *proximal operators* [85]. Proximal methods from sparse optimization allow for regularization to be imposed without destroying the structure of the system.

The specific proximal method that we use is a non-linear variant (Algorithm 2) of the well-known *proximal point method*. The proximal point method is a simple fixed-point iteration defined by Equation 6.11, that is expressed in terms of the *proximal operator*, $\text{prox}_{\gamma F}(p(\mathbf{x}))$, of the objective function $F(p(\mathbf{x}))$.

$$p^{k+1}(\mathbf{x}) \leftarrow \text{prox}_{\gamma F}(p^k(\mathbf{x})) \tag{6.11}$$

For an arbitrary convex function, $F(q(\mathbf{x}))$, the proximal operator, $\text{prox}_{\gamma F}$, (defined in Equation 6.12) acts like a single step of a trust region optimization in which a value of $p(\mathbf{x})$ is sought that reduces F but does not stray too far from the input argument $q(\mathbf{x})$:

$$\text{prox}_{\gamma F}(\mathbf{q}) = \arg\min_{\mathbf{p}} F(\mathbf{p}) + \frac{\gamma}{2}\|\mathbf{p} - \mathbf{q}\|_2^2. \tag{6.12}$$

To simplify notation, we use bold lower-case letters to refer to raster images, i.e. $\mathbf{p} = p(\mathbf{x})$, noting that there is an implied discretization step. The parameter γ serves to trade off the competing objectives of minimizing F while remaining close (proximal) to $\mathbf{q}$ but for strictly convex objectives does

not affect the final solution, only the number of iterations required to reach it.

For a least-squares objective $F(\mathbf{p}) = \frac{1}{2}\|\mathbf{Ap} - \mathbf{b}\|_2^2$, the resulting proximal operator [85] is found by expanding the resulting right hand side from Equation 6.12 and setting the gradient of the minimization term to zero. This results in Equation 6.13:

$$\text{prox}_{\gamma F}(\mathbf{q}) = \left(\gamma + \mathbf{A}^T\mathbf{A}\right)^{-1}\left(\gamma\mathbf{q} + \mathbf{A}^T\mathbf{b}\right). \tag{6.13}$$

In our case, the function F is simply the integral term from Equation 6.7. We form the proximal operator by discretizing the integral with sums over image pixels and defining: $\mathbf{A} = f\nabla^2$ and $\mathbf{b} = 1 - i_p(\mathbf{x})$.

Since proximal operators contain a strictly convex regularization term, $\|\mathbf{p} - \mathbf{q}\|_2^2$, the whole operator is a strictly convex function even if F is only weakly convex, as is the case for our problem. The proximal regularization improves the conditioning of our problem and can be interpreted as *disappearing Tikhonov regularization* [85], i.e. regularization whose effect diminishes to zero as the algorithm converges. This is helpful since the added regularization does not distort the solution.

Another benefit is that the proximal regularization does not change the structure of our problem since it only adds an identity term. This, coupled with the mirrored padding periodic boundary conditions, means that all terms in Equation 6.7 can be expressed as convolutions and the proximal operator solved in the Fourier domain. This is vastly more efficient than solving the optimization implied by the proximal operator in the spatial domain.

By denoting the forward and inverse Fourier transforms as $\mathcal{F}()$ & $\mathcal{F}^{-1}()$ respectively, complex conjugation by $*$ and performing multiplication and division point-wise, the proximal operator for Equation 6.13 can be re-expressed in the Fourier domain as Equation 6.14 for circulant matrices $\mathbf{A}$, as reported in [11] who used it to solve deconvolution problems.

$$\text{prox}_{\gamma F}(\mathbf{q}) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\mathbf{b})\mathcal{F}(\mathbf{A})^* + \gamma\mathcal{F}(\mathbf{q})}{\mathcal{F}(\mathbf{A})^2 + \gamma}\right) \tag{6.14}$$

In practice, we modify Equation 6.14 slightly by the addition of a regularization parameter α. The modified proximal operator is shown in Equation 6.15.

$$\text{prox}_{\gamma F}(\mathbf{q}) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\mathbf{b})\mathcal{F}(\mathbf{A})^* + \gamma\mathcal{F}(\mathbf{q})}{(1+\alpha)\mathcal{F}(\mathbf{A})^2 + \gamma}\right) \tag{6.15}$$

The constant $\alpha \geq 0$ regularizes the solution by favoring results with low curvature. This corresponds to solving a modified form of Equation 6.7 that imposes a penalty of $\frac{\alpha}{2}\|\nabla^2 p(x)\|^2$ once discretized (the second term of Equation 6.16 in the continuous case).

$$\begin{aligned} \hat{p}(\mathbf{x}) &= \arg\min_{p(\mathbf{x})} \int_{\mathbf{x}} \left(i_p(\mathbf{x}) - 1 + f \cdot \nabla^2 p(\mathbf{x})\right)^2 d\mathbf{x} \\ &+ \alpha \int_{\mathbf{x}} \left(\nabla^2 p(\mathbf{x})\right)^2 d\mathbf{x}, \end{aligned} \tag{6.16}$$

The effect of the parameter α is to favour smoother solutions than can otherwise be found. This helps to prevent the method from producing undesirable caustics in an attempt to achieve very bright highlights at the expense of image quality in darker regions. The effect of the α parameter is shown in Figure 6.17 for lens simulations.

Our final algorithm is shown in Algorithm 2 and is identical to the proximal point method except that the $\mathbf{b}$ image used by the proximal operator is updated at every iteration using the warping procedure from our previous work in [18]. After precomputing the Fourier transforms of $f\nabla^2$, each iteration of the algorithm can be implemented with an image warping, some component-wise operations and a forward/inverse Fourier transform.

Implementation. The re-formulation of the algorithm results in orders of magnitude speedup when executed on a Central Processing Unit (CPU) using Fast Fourier Transform (FFT) based solvers over the Quasi-Minimal Residual Method (QMR) solver that was previously used. Typical per-frame computation times were previously on the order of 20 minutes or more [18], while the Fourier version in Algorithm 2 takes approximately 0.6 seconds at the same resolution (256×128) on a Core i5 desktop computer, a speedup of approximately 2,000 times. The conversion to Fourier domain solves also results in operations that are more friendly for parallel Graphics Processing Unit (GPU) implementation. We have implemented the algorithm both in C++ and in NVIDIA's Compute Unified Device Architecture (CUDA) using NVIDIA's CUDA Fast Fourier Transform Library (cuFFT) for the forward and inverse Fourier transforms [81]. The CUDA & cuFFT version of the code yields nearly a 150 times speedup over the single-threaded CPU version when run on a GeForce 770 GPU, resulting in roughly a 300,000 fold speedup over the naive CPU version implemented using QMR. To our knowledge this makes the algorithm the first freeform lensing method capable of operating

Algorithm 2 Paraxial caustics in Fourier space

// Initialize phase surface as a constant value
$p^0(\mathbf{x}) \leftarrow 0$
// Initialize iteration counter and constant parameters
$\mathbf{A} \leftarrow f\nabla^2$
$k \leftarrow 0$
while $k < k_{\max}$ **do**
  // Warp target image by current solution
  $i_p^k(\mathbf{x}) \leftarrow i(\mathbf{x} + f\nabla p^k(\mathbf{x}))$
  // initialize right hand side of least-squares problem
  $\mathbf{b} \leftarrow 1 - i_p^k(\mathbf{x})$
  // Update the current solution by evaluating
  // the proximal operator in Equation 6.15
  $p^{k+1}(\mathbf{x}) \leftarrow \text{prox}_{\gamma F}(p^k(\mathbf{x}))$
  // update iteration index
  $k \leftarrow k + 1$
end while
// RETURN computed mapping
return $p^{k_{\max}}(\mathbf{x})$

in real-time, see Table 6.3. This is in contrast to methods such as [100], which produce very high quality results, but have runtimes roughly five orders of magnitude higher than our GPU algorithm.

Table 6.3: *Runtimes of the FFT based algorithm run on a CPU and on a GPU (implemeted using cuFFT) for various resolution inputs with 10 iterations of Algorithm 2.*

Algorithm	Resolution	Runtime (ms)	FPS
CPU	256×128	600ms	1.7
GPU	256×128	4ms	250
GPU	480×270	14ms	71
GPU	960×540	52ms	19
GPU	1920×1080	212ms	4.7

The algorithm is well suited to hardware implementation on devices such as GPUs, Field Programmable Gate Arrays (FPGAs) or ASICs due to its use of highly parallel FFTs and component-wise operations. We run Algorithm 2 for a fixed number of iterations (typically 10 or less). Convergence to a

solution is rapid, requiring well fewer than 10 iterations, however for many hardware implementations it is desirable to have computation times that are independent of frame content.

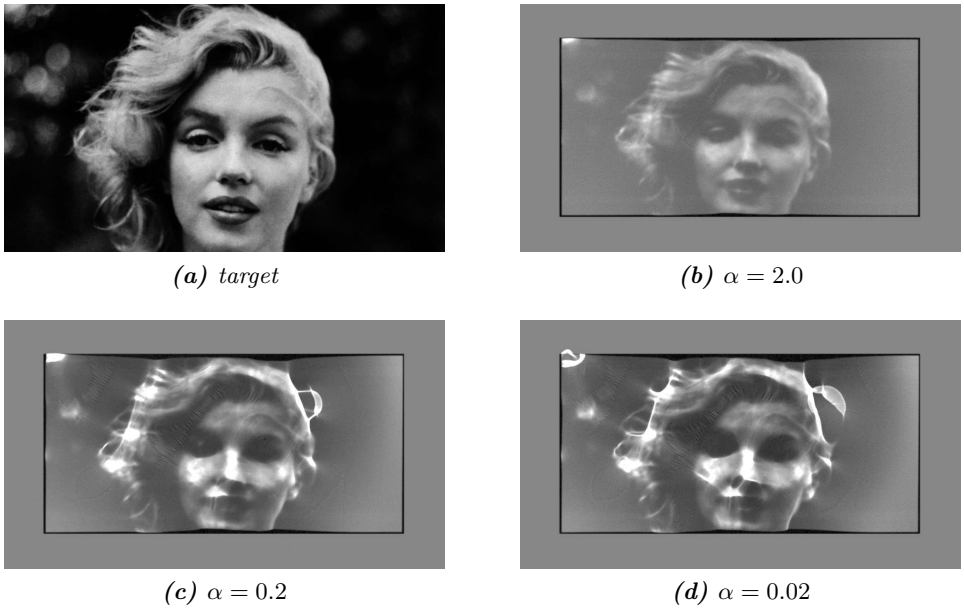
(a) Padded target

(b) Without padding

(c) Mirrored padded

Figure 6.16: *By mirror-padding the input image, pure-Neumann boundary conditions at the image edge can be achieved while retaining a Toeplitz matrix structure. This prevents distortions of the image boundary. Results were simulated with LuxRender.*

Simulation Results. Using the equivalence between physical lenses and phase functions allows solid lens models to be generated for testing via geometric optics simulation (we use Blender+LuxRender). Examples are shown in Figure 6.16 and 6.17 which illustrate the effect of mirror padding and the choice of α respectively.



(a) *target* ***(b)*** $\alpha = 2.0$

(c) $\alpha = 0.2$ ***(d)*** $\alpha = 0.02$

Figure 6.17: *LuxRender raytracing simulations: the smoothness parameter α penalizes strong caustics in the image that achieve high-brightness but poor image quality.*

6.6.4 Limitations

The architecture presented in this chapter and the resulting image quality improvements in contrast and peak luminance that can be achieved with it demonstrate the feasibility of the concept. We list some of the obvious and less obvious limitations of the implementation here that will be addressed in Chapter 7.

- The test system was built with a single, monochromatic (green) light source. For a full colour projector, at least two additional colour channels will need to be added to the system in either a parallel or a time sequential fashion. Either approach presents its own (solved) challenges. The former with respect to alignment of red, green and blue components such as SLM and dichroic mirrors and the latter with respect to synchronization/timing and thermal limitations.
- As with any display based on narrow band or monochromatic light sources (such as LEDs or lasers) care needs to be taken to manage undesirable properties such as observer metamerism and speckle.

- The phase SLM and the amplitude SLM need to be synchronized, ideally at the frame or subframe level. The amplitude modulator in the prototype was borrowed from a consumer projector which introduced an undesired latency in one of the modulation stages.
- In a full colour system (RGB), characterization and accurate modeling of the optical system including the PSF is required to ensure a colourimetrically accurate projector.
- None of the relay optics or other elements were custom designed for the prototype, which leads to light losses. Even with a more optimized light path, the addition of a phase SLM can reduce the overall light throughput. We estimate that this loss can be as high as 40-60% for the components used in the prototype. While this might seem high we note that even for bright images (ALL of 50% relative to the peak luminance) the gain in peak luminance exceeds what could be achieved in a traditional projector. Better suitable SLMs can further reduce the associated light losses.
- Careful alignment of a number of elements in the light path is required to achieve a uniform and predictable light profile on the phase modulator. In our experiment, the reflective nature of the phase SLM required off-axis illumination that was not accounted for in the simulations and algorithm and which in turn leads to errors in the resulting luminance profiles. While these errors were not clearly visible in the images projected onto the screen, the logarithmic luminance representation in Figure 6.15 reveal this non-uniformity. It can be accounted for in the lens pattern.
- Finally, the dynamic nature of the projection system with respect to peak luminance and feature size may present a challenge when colour grading content for the display. The notion of a limited light budget and a peak luminance that exceeds that of full screen white might makes sense from an HDR image statistics point of view, but would require a re-thinking in existing movie production processes.

6.6.5 Discussion

We have made two technical contributions: a simple but fast and effective new optimization method for freeform lenses (goal-based caustics), and a new dual-modulation design for projection displays, which uses a phase-only spatial light modulator as a programmable freeform lens for HDR projection.

The new freeform lens optimization approach is based on first-order (paraxial) approximations, which hold for long focal lengths and are widely used in optics. Under this linear model, the local deflection of light is proportional to the gradient of a phase modulation function, while the intensity is proportional to the Laplacian. We combine this insight with a new parameterization of the optimization problem *in the lens plane* instead of the image plane to arrive at a simple to implement method that optimizes directly for the phase function without any additional integration steps. Solved in the Fourier domain, this is the first algorithmic approach for freeform lensing that is efficient enough for on-the fly computation of video sequences.

Our new dual-modulation HDR projector design finally allows us to achieve perceptually meaningful gains in peak luminance on large cinema screens while simultaneously improving black level performance and maintaining a manageable power, light and cost budget. As such we believe that to date the approach presents one of the most sensible proposals for commercial high contrast HDR projection systems and one of the most practical ways to achieve high brightness HDR imagery in cinema.

Chapter 7

Improved RGB Projector Prototype

This chapter describes the work related to the final prototype that was developed during the PhD term in collaboration with our industry partner MTT. We discuss a number of limitations of the earlier proof-of-concept work in Chapter 6 and based on this the design decisions that led to an improved full colour RGB projector prototype.

Table 7.1 provides an overview of the new features that were incorporated into the RGB prototype introduced in this section.

Table 7.1: *RGB High Power Prototype Features*

Feature	Comment
Full colour	R, G, and B primaries at 462nm, 520nm and 638nm
Higher power	Arrays of individual diodes independently collimated
Intensity control	Source intensity adjustable per frame
Phase SLM	Custom Holoeye PLUTO device
Amplitude SLM	LCoS based (higher native contrast compared to DMD)
Synchronization	Theoretical analysis of timing for DMD implementations
Algorithm	Calibrated forward model
PSF	Measured, synthesized and incorporated into forward model
HDR Content	Tone-mapped/graded for theatrical HDR systems

7.1 Architecture of the RGB Projector Prototype

In the following section we address the most significant proof-of-concept limitations discussed in Chapter 6 and discuss improvements and design decisions for a new full colour prototype implementation which can be seen in Figure 7.1.

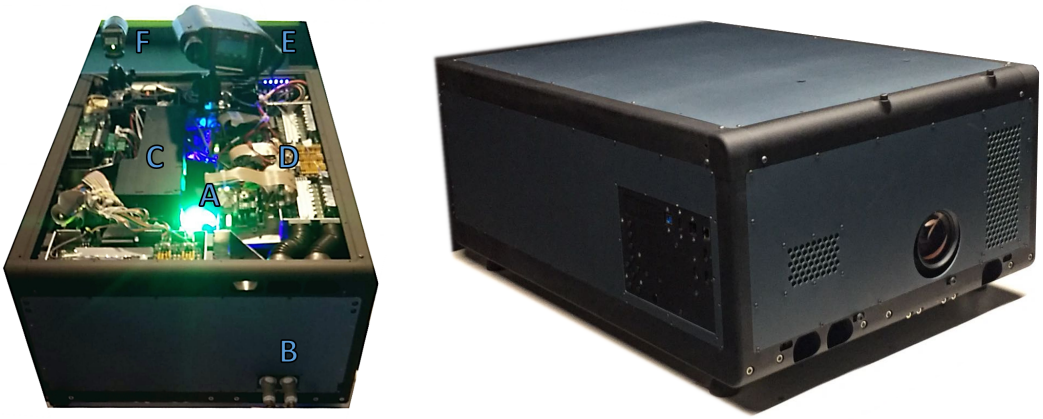


Figure 7.1: *The photos show the completed RGB light steering prototype. In the left photo the lid is removed and some of the core components are visible including the green and blue channel light sources (A), the waterline connections for a chiller to maintain the laser light sources and phase modulators at a controlled operating temperature (B), the combination optics block (C), control electronics and laser drivers (D), a top mounted spot-spectroradiometer for colour calibration (E) and a machine vision camera for spatial uniformity calibration.*

Colour. Instead of the monochromatic ($532nm$, green Diode-Pumped Solid-State (DPSS) laser) a light source consisting of red, green and blue laser diodes enables us to design a full colour projector. Switching from a DPSS-based laser to native laser diodes at wavelengths of 462nm, 520nm and 638nm simplified the light source (much fewer optical components are required for a diode laser compared to a DPSS laser). While a field-sequential system in which each of a red, green and blue light field are presented in sequence is in principle possible, we choose instead a parallel architecture in which three monochromatic light paths, including individual light sources,

phase modulators and optical components (compare Figure 6.14 up until the diffuser) are combined into a white beam with dichroic colour filters and then relayed into an amplitude modulating projection head. The range of achievable chromaticities can be seen in the diagram in Figure 7.2 relative to other common display and cinema primaries.

Optical Power. The optical power of the proof-of-concept projector was measured as ***10 lumens*** out of lens. While the overall light levels, the peak luminance and the screen size that could be illuminated using the light steering concept were impressive, a higher power system is required to show that eventually cinema screens can be illuminated using a light steering projector. ***100 to 200 lumens*** of steered light as a goal for the new prototype present a meaningful stepping stone. However the required light source power can not easily be achieved with existing individual laser diodes. Three laser diode properties are of critical interest: the total power of the laser diode as well as the emitter dimensions and the divergence of light as it is emitted. Beam expansion, collimation optics and tiling of multiple laser diodes at the light source require a mechanical design in which the individual components can be adjusted, some at 6 degrees of freedom.

Speckle. As with any display based on narrow band or monochromatic light sources (such as LEDs or lasers) care needs to be taken to manage undesirable properties such as inter-observer metamerism variations and speckle. There are typically three measures that can be taken to reduce the visible speckle contrast to the observer: randomization of polarization, increasing angular diversity of light in the optical path and broadening of the light source spectrum. Elements of all three methods can be applied to our method. While the projection head utilized in our RGB prototype is based on LCoS technology which requires linearly polarized input (see Chapter 2 for reference), it is in principle possible to randomize this polarization after the final image has been formed either before or after the projection lens. Similarly, if the light steering method is coupled with a DMD-based projector head (see Chapters 2 and 7 for reference), then polarization can be randomized within the projector light path following phase modulation and prior to the DMD amplitude modulation, since the DMD does not require linearly polarized light at the input. Light within our light steering system is ideally well collimated. This translates to a high f-number (f#) optical system and with that limited angular diversity. While it is important to preserve this high degree of collimation for the light steering part of the system,

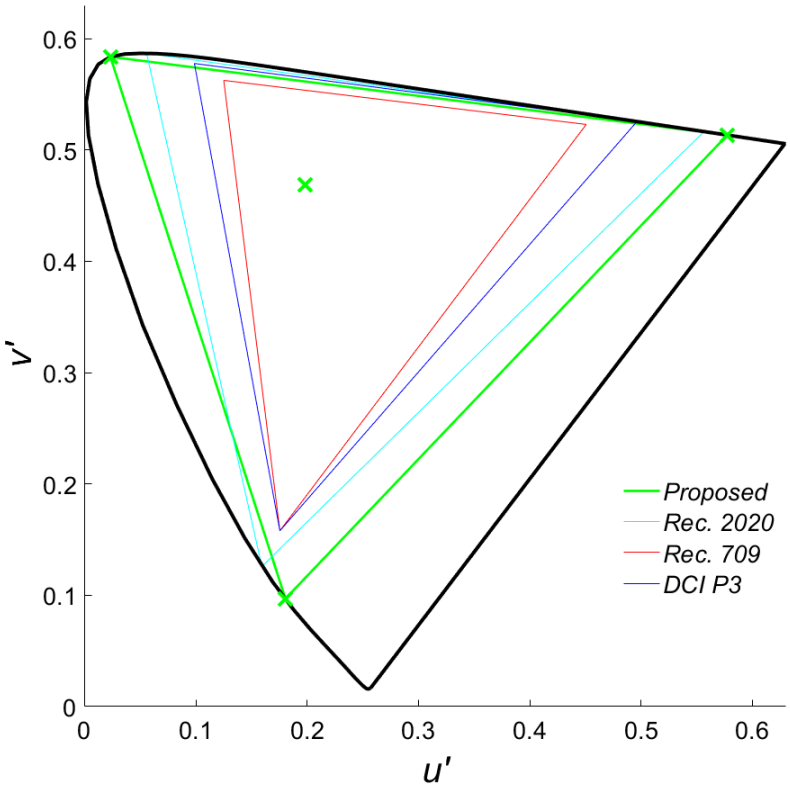


Figure 7.2: *Chromaticity diagram of the proposed projector primaries compared to a number of common colour standards. The CIE 1976 Uniform Chromaticity Scale (UCS) using the u' and $'v'$ as chromaticity coordinates provides a perceptual mostly uniform relation between individual colours (i.e. the perceived difference between colours of equal distance from each other is comparable). The green data points represent the D65 white point which is common in cinema and home display systems as well as the projector laser primaries. The chromaticity space that can be reproduced with the prototype is significantly larger and includes both the DCI P3 space as well as the Rec. 709 space. It almost entirely encompasses the Rec. 2020 colour space and exceeds it in overall area.*

this property of light is no longer required after an intermediate image has been formed (compare Figure 6.14, diffuser). The f# can then be reduced, for example with a light shaping diffuser, to match the input acceptance angles of the following optics resulting in higher angular diversity of the beam

and with it less visible speckle. Moving the diffuser inside the projector randomizes the angular spread of light over time and further reduces visible speckle contrast. This can be achieved for example by rotating a diffuser disc within the projector or by linearly displacing optical elements at or near the diffuser. A second and effective method to reduce speckle contrast is the introduction of movement to the projection screen. A slight continuous displacement of the screen surface has the effect of averaging over many angles as light from the projection lens reflects of the (non-flat) surface of the screen. Equally effective, but much less practical in a cinema setting, is the movement of the observer. Finally, by employing binned and calibrated laser diodes that, for each colour, consist of different center wavelengths, the effective spectral band of the light source can be broadened, which reduces visible speckle contrast. The broadening of the light source will result in a superimposed set of slightly magnified and demagnified images for each colour channel after steering the light, which can be modelled by a small blur kernel and is not necessarily undesirable (compare also Figure 6.6).

Synchronization. The phase SLM and the amplitude SLM, as well as a Pulse Width Modulation (PWM) dimmable laser light source need to be synchronized, ideally at the frame or subframe level. The amplitude modulator in the prototype stems from a consumer projector and exhibits undesired latency of multiple frames within its built-in image processing block that we account for. In cinema, binary DMDs are used predominantly as amplitude SLMs in order to handle large projector light output and to comply with certain standards (e.g. the DCI). Section 7.2 discusses new drive schemes (timing) that aim at mitigating anticipated temporal artifacts (flicker).

Calibration. In a full colour system colourimetric calibration requires characterizing and accurately modeling the system, including the PSF, which depending on the light source and optical path could potentially be dependent on location or even image feature sizes. We discuss our approach in Section 7.3.

7.2 Temporal Considerations in HDR Projection Displays

While parts of the prototype were built based on commodity consumer hardware, the use of development kits, together with customized light source

control electronics allows for better synchronization of pulses from the light source (laser intensity modulation), digital pulses codes required to address the phase modulator (bit planes within a subframe), and for example a binary primary amplitude modulator in the projector head (DMD mirror states). While prototyping a fully integrated solution is outside the scope of this work, we explore different theoretical solutions for synchronized drive schemes. This section discusses synchronization options for DMD based projector architectures. As part of the prototyping work in Sections 6 and 5.1 we gained a better understanding of the temporal properties of the phase SLM hardware and the PWM intensity controlled light source, and updated the prototype and the models in this section based on the findings.

LCoS Based Phase Modulator. The phase modulator we selected for the light steering projector implementation currently uses an LCoS micro display with a digital back plane and no input or output polarizing filter. The backplane updates in a vertical, top to bottom scrolling fashion. Either one or two lines are updated at a time. The relatively slow response of the LC material relative to the fast pulse codes of the back plane make it possible to achieve effectively a near-analogue phase response of the display. This is advantageous as synchronization to the fast, binary states of a micro mirror based projection head is less of an issue compared to truly binary, fast LCoS devices (e.g. ferroelectric devices).

Important for our application are:

- The overall refresh rate of the phase modulator (the refresh rate needs to be in line with the required video frame rate of the overall projector).
- The phase accuracy over the period of a frame (how reliable does the phase modulator reproduce a given phase value based on the corresponding drive level).
- The phase stability within a frame (to what extent are individual digital pulses from the back plane measurable in the overall phase response (*phase flicker*) and how does this affect light steering).
- Phase drift within a sub-frame between each line update.

In a calibrated LCoS device, per pixel, crystals will drift towards 2π or 0π respectively when an electric field is applied or removed. The *response time*

can be thought of as the time from when the electric field is applied continuously to the time when the crystal produces a 2π phase shift. Response time can be tuned with the chemical formulation of the crystal. Figure 7.3 shows an example of the rise and fall characteristics of a phase-only micro-display. If the target phase response is between 0 and 2π, voltage can be periodically applied and removed to achieve the intermediate phase retardation.

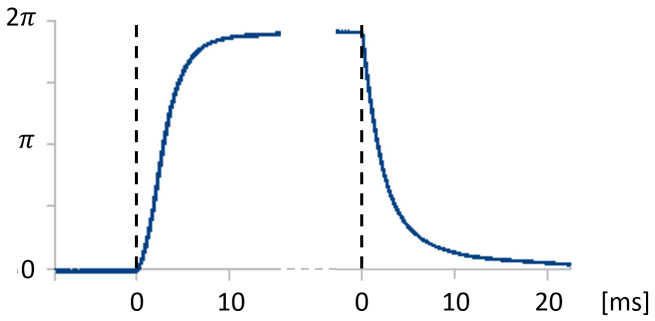


Figure 7.3: *Typical relative rise (left) and fall (right) times for the phase modulator between 0 and* 2π*.*

A circuit generating the periodic voltages has a frequency governed by the spatial resolution of the micro display and the bit precision of phase control. The crystal response time is tuned to give the most stable image given the update frequency from the driving circuitry.

The driving circuit used in the current prototype is an FPGA, it can change the state of the driving on/off voltage at around $7kHz$, in which case the pixel array is updated line by line starting at the top. The update of one *bitplane* across the entire frame therefore takes $1/7,000s$ (or around $145,000ns$). An example of variations in phase stability and phase drift of the current system can be seen in the (preliminary) measurements in Figure 7.4 in which the crystal orientation is balanced in a half turned state by a periodic square wave voltage being applied.

The total rise response time is $8.7ms$ and the total fall response time is currently $21ms$. The cell thickness for the particular panel under test was not customized for the application and provides sufficient phase retardation all the way into the Infrared (IR) part of the light spectrum. The cell thickness is hence thicker than it needs to be. For example with this particular panel the maximum possible phase retardation for a blue laser diode can be up to 6π and for a red laser diode can be close to 3π. A faster response time can be achieved by reducing the cell thickness to provide no more than

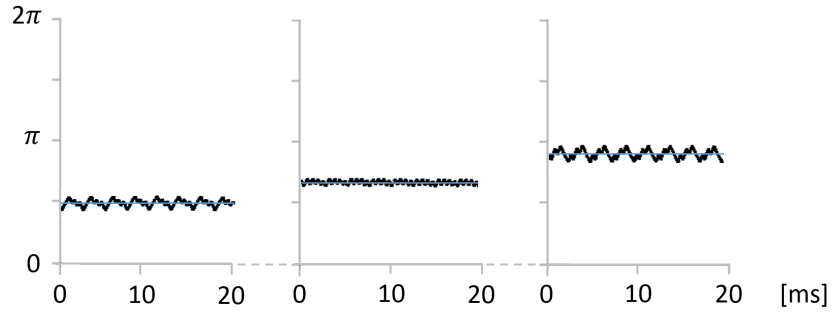


Figure 7.4: *Sub-frame phase response at drive levels 65 (left), 75 (center) and 100 (right) out of 255. The small scale temporal ripple is what we refer to as phase flicker, the deviation around the target phase value (blue line) indicates the phase stability.*

the maximum required phase retardation per wavelength. Furthermore the lensing (steering) function can and should account for the effective refresh rate of the phase modulator via a simple model. As a new frame arrives, the driving state of all pixels is updated on the next refresh cycle.

Fast response time and high phase stability are somewhat opposing goals along one shared dimension of temporal control. This is because within the duration of a video frame, applying a continuous electric field early on will allow the crystals to move into the correct position quickly (after which the electric field should be removed or only pulsed *on* occasionally), whereas phase stability throughout the entire video frame is best achieved by having the *on* and *off* states of the electric field spread relatively evenly throughout the duration of the video frame. Both goals could eventually be accounted for in the lensing algorithms as well as in the underlying digital update scheme of Pulse Code Modulation (PCM) used to map phase code words to the optical phase response.

In the future a dedicated ASIC could be developed instead of an FPGA to control the phase modulator would allow for faster update rates and lower part costs for volume production. In the *phase stability* plots in Figure 7.4 one would then see more peaks and valleys during the same time period.

Additionally there are options to modify or re-formulate the birefringence properties of the LC material used in the current phase modulator to enable a faster response time (for example 2× improvements).

DLP Technology Characteristics. DLP technology makes use of a binary modulator which flips per pixel micro mirrors back and forth across their diagonals. Each mirror at any time can be either in the *on* state in which it directs light rays to the screen or an *off* state in which it directs light rays to an off-screen location, the so-called *light dump* area. A mirror creates grey scales by flipping back and forth rapidly. For example over the course of a video frame it would spend more time in an *on* state to render a brighter pixel or more time in an *off* state for a darker pixel. Typically, each pixel requires an 8 bit (or more) grey scale drive value per frame of video (usually 60 Frames per Second (fps)). Figure 7.5 shows how these grey scales are translated into mirror flips.

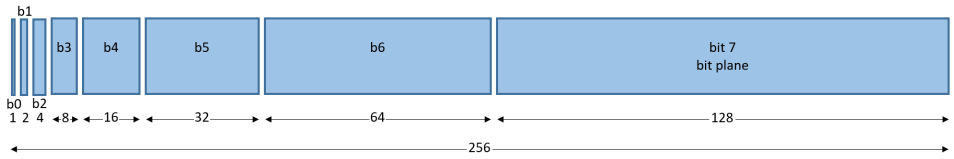


Figure 7.5: *DMD timing: high level working principle (reproduced from TI DLP documentation depicting the conceptual bit partition in a frame for an 8-Bit colour).*

Whether a bit is set to 0 or 1 determines whether the mirror is flipped to the *on* or to the *off* position. The bit position determines the relative duration of time that the mirror remains in the state. The maximum number of flips per seconds that the mirrors can typically achieve is just under $10kHz$ (see 7.5; up to $80kHz$ have been reported for professional applications), thus for this estimation we set the shortest period of a mirror state to $0.1ms$ or $100,000ns$). We will refer to this as the *mirror flip* period equal to the period of b0 in the diagram in Figure 7.5.

Asynchronous Light Pulses. If a pulsed light source is used (for example to produce light at 50 percent of the maximum level), flickering will occur if the *off* and *on* pulses are asynchronous to the mirror flipping and the periods of *off* and *on* significantly differ from frame to frame on a static image due to for example a low pulse frequency of the light state.

In Figure 7.6 note how in frame 1 the light is on for 2/5 of the time and in frame 2 the light is on 3/5 of the time due to the fact the signals are asynchronous. If the *off* and *on* light source periods are short relative to the *mirror flip* period, the difference between *off* and *on* periods between static frames should be drastically reduced and be imperceptible to the human

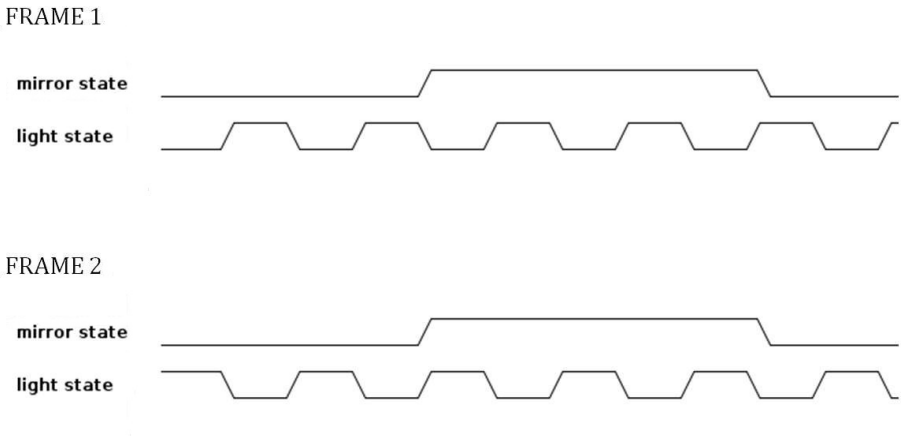


Figure 7.6: *Slow, asynchronous light pulses.*

eye. Figure 7.7 shows an example in which the light source is modulated significantly faster than the shortest possible mirror flip period.

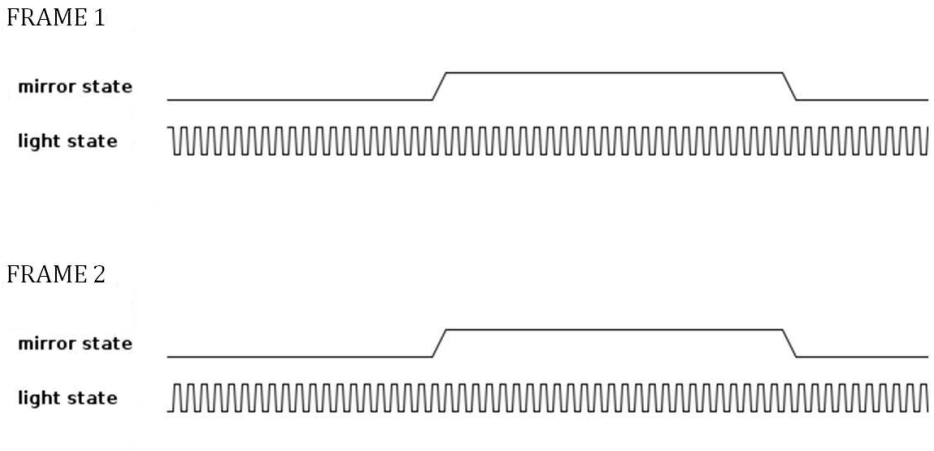


Figure 7.7: *Fast, asynchronous light pulses.*

In Figure 7.7 note that only a single minimum width mirror flip is shown with a drive value of 1, and the light state is analogous to the *pwm clock* described below. Also note that in this example the light is on 27/54 in frame 1 and 28/54 of frame 2.

Synchronous Light Pulses. Provided the light source *off* and *on* periods are synchronous to mirror flips (such as in Figure 7.8), there should be practically no intensity difference between static frames and the light source pulse generator need only run at the period of the mirror flips, which in turn can drastically reduce the requirements of the control solution and impact Electromagnetic Interference (EMI) considerations.

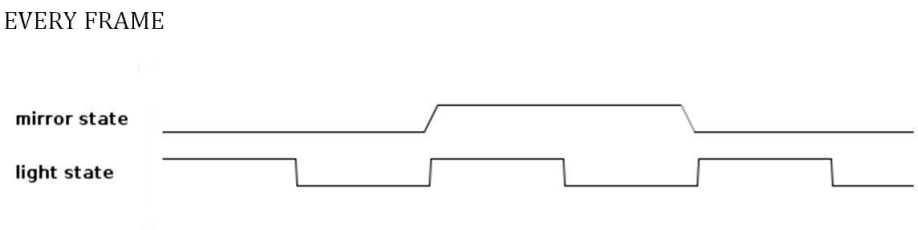


Figure 7.8: *Slow, synchronous light pulses.*

When a new frame arrives, the mirror flip logic for all pixels can be updated simultaneously via a double buffering scheme (or in blocks from top to bottom if desired).

Laser Control Solution. After a first review of available laser driver solutions we selected the iC-Haus iC-HG [49] device to directly drive laser diodes at high current and the option to pulse at very high frequency. The device has up to $200MHz$ switching capability from a differential pair input. Synchronized switching of up to $100MHz$ of a $500mW$ (optical power or $650mA$ and $2.2V$ electrical power per diode) $638nm$ laser diode array was confirmed during a previous project in the UBC CS PSM lab [40]. A constant voltage input to the iC-HG sets the current limit in the *on* state. We drive the constant voltage current input with a high speed Digital-to-Analog Converter (DAC).

Combining LCoS Phase Modulator, Binary DMD-based SLM, and Pulsed Lasers. Figure 7.9 shows an example of the timing for a combined system. The following assumptions were used for a first estimate of the overall system temporal response:

- For the shortest mirror flip duration, there are about 100 light state *PWM clock* pulses (fewer shown in the diagram below for clarity) and the *LCoS* phase error will drift between, in this example, 0.1π and -0.1π an average of 1.5 times.

- A 2× faster (compared to the prototype) phase LCoS SLM and faster controller chip (ASIC) are used in these visualizations along with the fast asynchronous *PWM clock* drive scheme introduced above.

Table 7.2 shows an exemplar update speed and the resulting pulse duration for the different modulation elements within the projector.

Table 7.2: *Pulse durations of the different light modulation stages within the projector.*

Component	updates per second	shortest period [ms]	frame period [ms]	pulses per frame	pulses per mirror flip
DMD	$10,000$	0.1	$16.\overline{66}$	$166.\overline{66}$	1
LCoS	$15,000$	$0.0\overline{66}$	$16.\overline{66}$	250	1.5
Laser	$1,000,000$	0.001	$16.\overline{66}$	$16666.\overline{66}$	100

The 0.1π error (drift) in phase modulation is relatively low compared to the maximum amount of possible phase retardation of $> 2\pi$ and the fast laser light source *washes over* even the fastest DMD mirror flips. Many alternative implementations are possible, including a Continuous Wave (CW) or *constant-on* laser driver, in which excess light is steered away from the active image area.

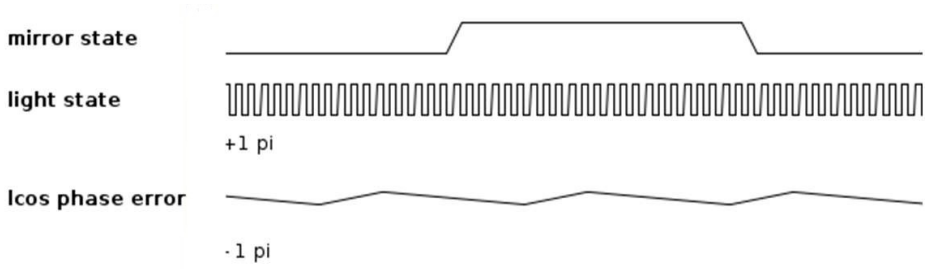


Figure 7.9: *Relative timing diagram for phase LCoS, DMD and laser pulse combination: DMD mirror flips (shortest possible period), light source pulses (fewer than actual for illustration purposes), and anticipated phase drift of the LCoS-based phase modulator).*

DMD-Based Experimental Prototype. An experimental prototype that combines the LCoS based phase modulator, a PWM-dimmed (order

of kHz) laser light source and a DMD according to the asynchronous drive scheme, was built for demonstration purposes, and while direct timing measurements were not taken, visible temporal artifacts were not dominant enough to be noticeable. The relatively low magnitude of the phase error and the low amplitude of the phase flicker relative to its maximum of 2π aided in masking possibly present temporal artifacts. Figure 7.10 shows a DMD-based system.

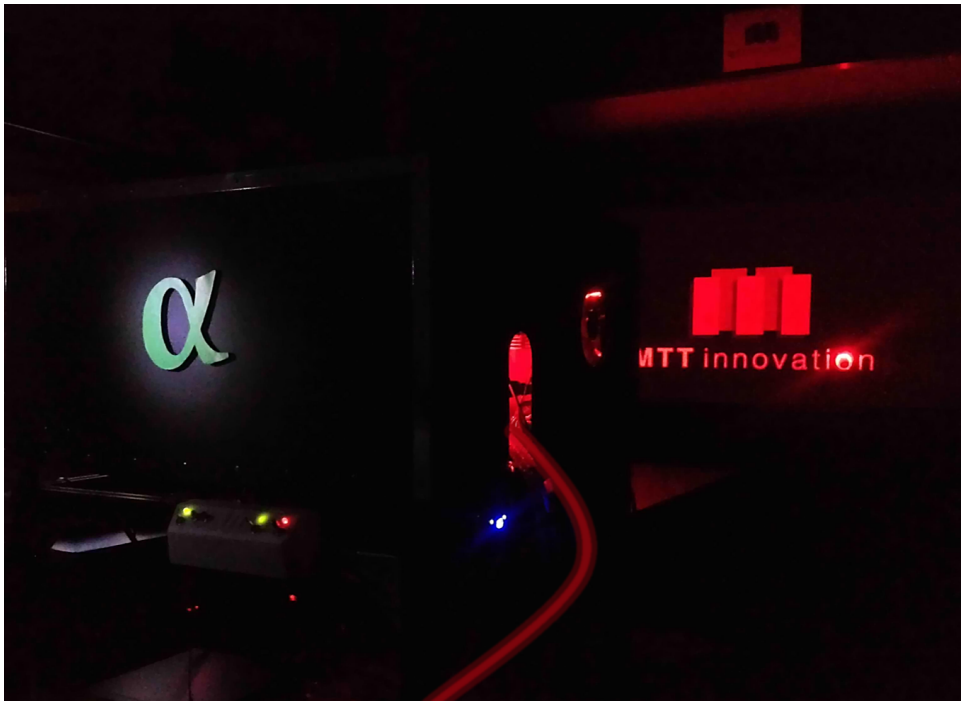


Figure 7.10: *Prototype using a DMD amplitude modulator. The laser is coupled into the system via a fiber (red, right side of the laser safety enclosure). Its intensity can be adjusted via a PWM drive scheme.*

7.3 Colourimetric Calibration of the Projector

The system being targeted in this subsection is a two-projector system consisting of a traditional non-steering projector and a light steering projector. The traditional projector is an off-the-shelf projector that uses amplitude modulation while the light steering projector is a phase/amplitude modulation design.

7.3.1 Light Steering Image Formation Model

Working from light source to screen, collimated laser light is relayed from a light-source to a reflective phase modulator by a series of optics. Light leaving the phase modulator for a given colour channel is combined with light from the remaining channels and relayed through a diffuser to a Philips or RGB prism. The prism splits the light into its component colours which are each modulated by an amplitude modulator, recombined within the prism and directed into the main projection lens. Figure 7.11 shows the high level blocks described herein.

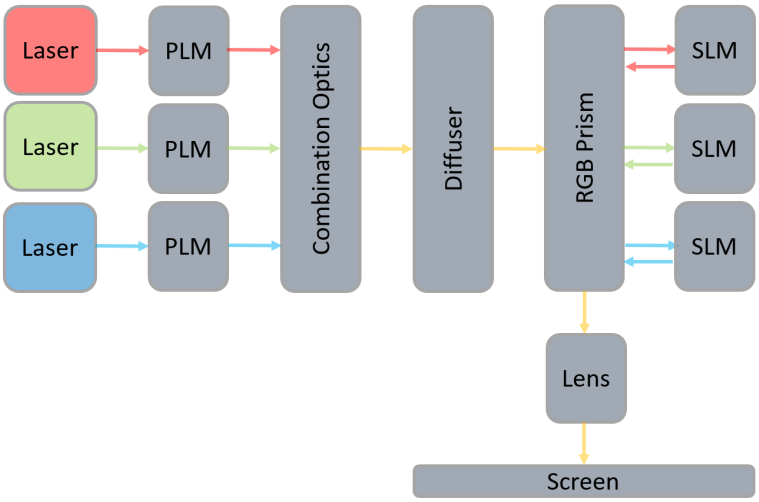


Figure 7.11: *Diagram depicting the high level optical blocks within the light path of the prototype projector from laser light source (left) to projection lens (right). The combination of RGB light into white light was chosen for convenience and to accommodate commercially available projection hardware with pre-aligned SLMs. Better contrast performance can be expected from a discrete RGB light path.*

This design is able to produce local intensities well above the typical full-screen intensity by virtue of the phase modulator, which is able to introduce phase variation to the incident wavefront. This allows it to function as a programmable lens in response to a software-driven phase pattern. The phase pattern is computed and attempts to redistribute light from the input illumination profile to target light profile, chosen ideally to approximate an upper-envelope of intensities in the underlying target image. This redistributes light from dark areas to bright regions which, due to properties of

typical images, tend to have large dim regions and small bright highlights resulting in considerable focusing of spare light which can reach levels more than 10× of what could be achieved by a conventional projectors.

A diffuser is incorporated into the optical design to reduce speckle and also optionally acts as a low-pass filter over the light field, since the phase modulator can introduce a number of artifacts. These include fixed texture, a component of unsteered illumination and diffraction artifacts. The diffuser is effective at removing diffraction artifacts and fixed texture but generally cannot compensate for the unsteered light(see Figure 7.12).

Figure 7.12: *Unsteered Component. Laser light that is not steered by the phase modulator.*

Typically around 10% of the light ends up in the unsteered component and is related to the illumination incident on the phase modulator after filtering by the diffuser. Consequently the unsteered component is an important contributor to the subsequent image formation. Measurements of the unsteered component can be obtained by designing a phase pattern to steer all available light off-screen; what remains is the unsteered component (Figure 7.12).

For laser diode-based systems such as in the RGB prototype the unsteered component shows individual diode beams reflected from the phase modulator, in Figure 7.12 these show as vertical (red) and horizontal (green, blue) stripes. The difference in orientation is due to differing polarization orientation of the diodes. Beams from the light source are polarized in the same direction and so all stripes are oriented similarly, by design. Figure 7.13 shows the full-screen white pattern once a corrective lens pattern has been applied to the illumination profile in Figure 7.12. In case of fiber-coupled laser, the unsteered component would be significantly more uniform.

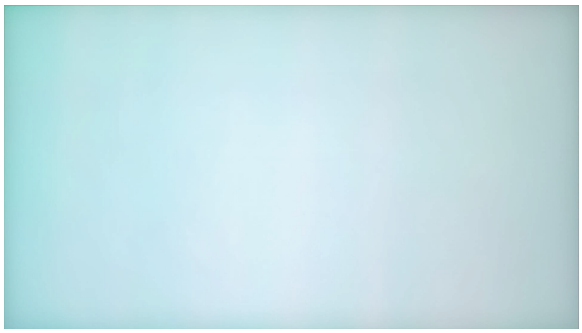


Figure 7.13: *Normalized full-screen white pattern after a correction phase pattern has been applied to the illumination profile.*

7.3.2 Optical Model

This section describes the algorithms that are used to drive the system, beginning with high-level algorithmic blocks for the overall algorithm. Key blocks are further described in dedicated subsections. Each rectangular block corresponds to a set of operations, parallelograms indicate data passed between the blocks. Solid arrows indicate known interactions between blocks while dashed arrows indicate optional connections. Figure 7.14 shows a high level overview of the functional algorithm blocks.

7.3.3 High-Level Algorithm

The high-level view of the algorithm takes in the input image. In the RGB prototpye system, this input is the CIE 1931 XYZ colour space with PQ encoding of each channel.

In the input transformation block, image data content is linearized and converted to the working colour space of the system. The output of input transformation is a linear image expressed currently in linear RGB laser primaries.

In the content mapping block the linear image is split (or distributed) between the light steering projector and the non-steering projector with a distinct amplitude pattern each in the case of two projectors or in case of an integrated system a shared amplitude pattern (one projection head with a light steering and a non-light steering light source). The output of the content mapping block is a target light field image, a target (full) image (per projector) and a power control signal. The power-control signals and target image are inter-related depending on the power-control approach that

is taken.

To physically redirect light, the target light field is used as input to the phase pattern generation algorithmic block. This computes the drive parameters needed to affect light-redistribution by the phase modulators. In addition, the target light field is also used by the forward model algorithmic block, which implements a feed-forward simulation of the light steering image-formation model since, in practice, the phase modulator and subsequent optical path is unable to reproduce arbitrary target light fields exactly. The forward model produces a predicted light field image that, in combination with the target image, is used by the amplitude pattern generation block to determine the necessary amplitude patterns for both the light steering and the non-light steering block.

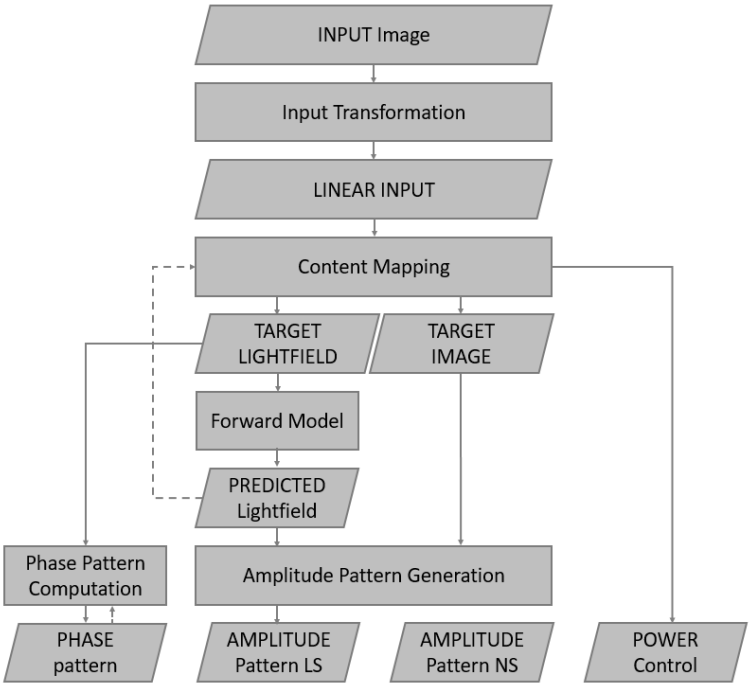


Figure 7.14: *High-level algorithm blocks.*

7.3.4 Input Transformation

The input transformation block functions primarily to transform the input image from the input PQ encoded images in the XYZ colour space to the colour space defined by the laser primaries.

Table 7.3: *Projector Chromaticity Coordinates*

Colour	Wavelength(s) (nm)	Chromaticity x	Chromaticity y
Red	638	0.71679	0.28317
Green	520	0.07430	0.83380
Blue	462	0.14075	0.03357
White	638 / 520 / 462	0.31271	0.32902

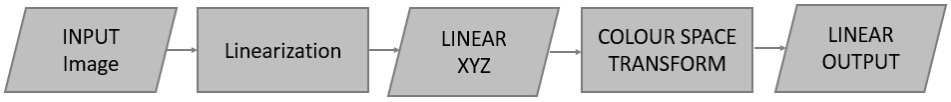


Figure 7.15: *Input transformation block takes PQ-XYZ inputs, linearizes them and transforms them to the light steering projector colour space.*

The forward and inverse transformations for the PQ encoding are given in the following equations:

$$L = \left(\frac{P^{\frac{1}{m_2}} - c_1}{c_2 - c_3 P^{\frac{1}{m_2}}} \right)^{\frac{1}{m_1}} \tag{7.1}$$

$$P = \left(\frac{c_1 + c_2 L^{m_1}}{1 + c_3 L^{m_1}} \right)^{m_2} \tag{7.2}$$

where P and L represent PQ and linear values mapped to the range $[0, 1]$. These ranges should be adjusted to the nominal working range, e.g. $[0, 2^{10} - 1]$ for 10 bit PQ and $[0, 10000] cd/m^2$ for L. The transformations can be implemented as a 1D Look Up Table (LUT), however care over the sampling rate is important to resolve all regions of the curve.

For the colour transformation, the RGB projector primaries and white point (D65) are shown in Table 7.3.

To obtain RGB images in laser primaries from these, it is necessary to convert to the RGB projector primaries *MTTP3*. This transformation is chosen to preserve the luminance of each channel, leading to the following relationship between XYZ images:

$$\begin{bmatrix} \frac{Y_w x_w}{y_w} \\ Y_w \\ \frac{Y_w (1 - x_w - y_w)}{y_w} \end{bmatrix} = \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = \begin{bmatrix} \frac{x_r}{y_r} & \frac{x_g}{y_g} & \frac{x_b}{y_b} \\ 1 & 1 & 1 \\ \frac{(1 - x_r - y_r)}{y_r} & \frac{(1 - x_g - y_g)}{y_g} & \frac{(1 - x_b - y_b)}{y_b} \end{bmatrix} \begin{bmatrix} Y_r \\ Y_g \\ Y_b \end{bmatrix}, \tag{7.3}$$

where $[X_w, Y_w, Z_w]^T$ is the luminance of the combined image and $[Y_r, Y_g, Y_b]^T$ is the luminance of each channel treated independently under the constraint that $Y_w = Y_r + Y_g + Y_b$. The per-channel luminances $[Y_r, Y_g, Y_b]^T$ corresponding to $Y_w = 1$ can then be found by solving the system above.

The per-channel luminance values are used to define the transformation $\mathbf{M}$ from *MTTP3* to XYZ. This transformation can be defined as follows:

$$\begin{bmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{bmatrix} = \mathbf{M} \begin{bmatrix} Y_r & 0 & 0 \\ 0 & Y_g & 0 \\ 0 & 0 & Y_b \end{bmatrix}, \tag{7.4}$$

meaning that input images in which each channel stores the luminance of its corresponding primary should map to the chromaticity of the primary at the luminance stored in the image. The transformation can be found using

$$\mathbf{M} = \begin{bmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{bmatrix} \begin{bmatrix} \frac{1}{Y_r} & 0 & 0 \\ 0 & \frac{1}{Y_g} & 0 \\ 0 & 0 & \frac{1}{Y_b} \end{bmatrix}. \tag{7.5}$$

For the chromaticities and white point listed above, this gives the following result for $\mathbf{M}$:

$$\mathbf{M} = \begin{bmatrix} 2.5313 & 0.0891 & 4.1927 \\ 1.0000 & 1.0000 & 1.0000 \\ 0.0001 & 0.1102 & 24.5958 \end{bmatrix} \tag{7.6}$$

Similarly, the inverse mapping from XYZ to *MTTP3*:

$$\mathbf{M}^{-1} = \begin{bmatrix} 0.4064 & -0.0287 & -0.0681 \\ -0.4082 & 1.0333 & 0.0276 \\ 0.0018 & -0.0046 & 0.0405 \end{bmatrix} \tag{7.7}$$

7.3.5 Content Mapping

The content mapping block takes as input the linear input image and determines the split between the light steering and the non-light steering projectors as well as the power levels required.

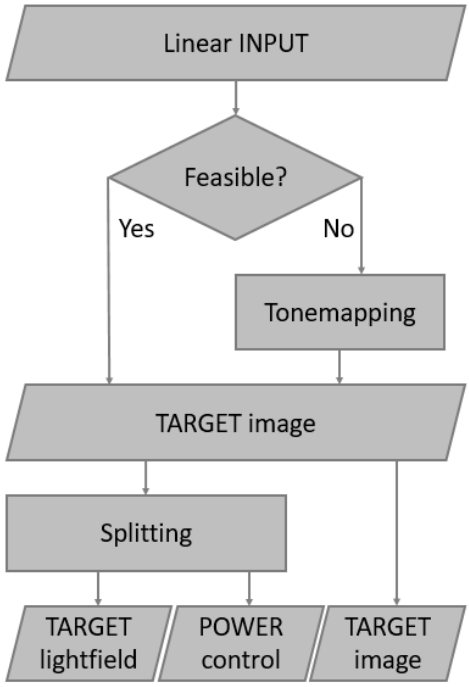


Figure 7.16: *Content Mapping Algorithm Block*

The algorithm first checks if the input image is feasible given the system power budget. This is currently done using a power heuristic. If not, the input is tone-mapped. The resulting image (either passed through or tone-mapped) is then the target light field for subsequent stages. For feasible input content, the linear input and target images are identical. The target image is then used to generate a target light field image.

The split between steering and non-steering can be achieved by a number of methods. A simple one is raising the target image to an exponent ($\gamma > 1$) to determine the light steering image. More sophisticated schemes have been proposed earlier. This new image does not accurately reproduce light steering projector data since it distorts the content to emphasize highlights. It is expected that a more accurate (and light efficient) splitting can be obtained by using the non-steering projector for most of the image formation up to 48 cd/m^2, gradually phasing in the light steering projector as depicted in Figure 7.17.

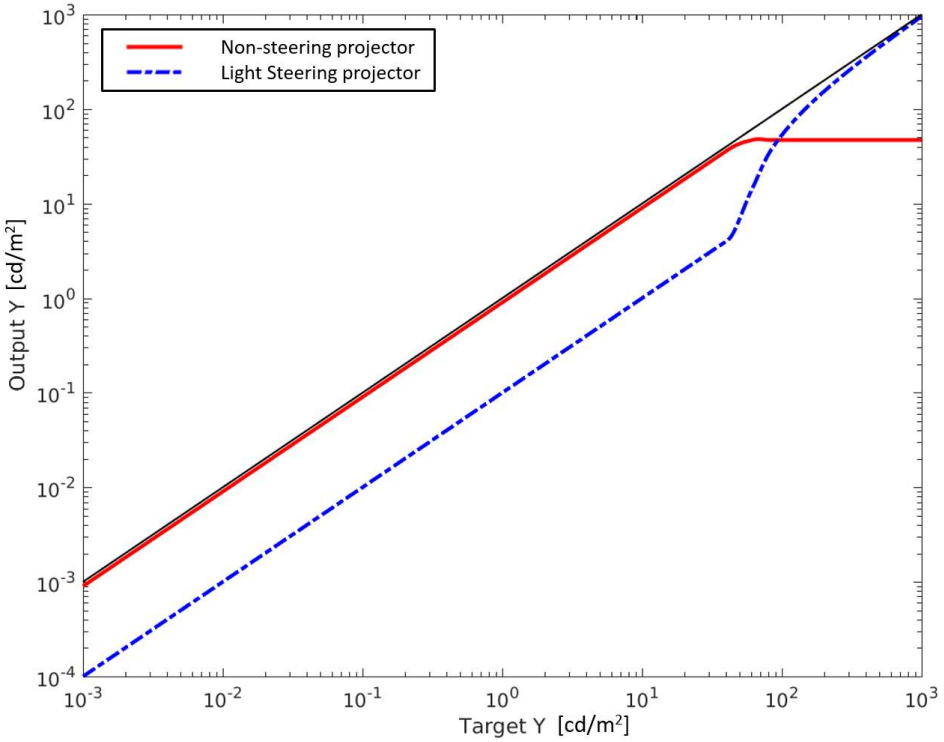


Figure 7.17: *Steering and Non-Steering splitting as a function of target luminance in cd/m^2. Note the log-log scale.*

The split function attempts to utilize the non-steering projector for 90% of image formation up to 47 cd/m^2, at which point the steering projector begins to take over. It is desirable to use the steering projector for a portion of the image at every pixel in order to avoid bright image features having a *painted-on* appearance. The splitting is 1D and could be implemented as a function or as a LUT.

7.3.6 Forward Model

The forward model block takes as input the target light field from the content mapping block and uses it to predict the output of the optical system, referred to as the predicted light field. This is necessary since not all target light fields are achievable and, to compute the correct amplitude patterns, it is also necessary to know how the actual light field differs from the target light field. This consists mainly of applying the calibrated system PSF and

the unsteered component to the output image, taking care to account for overall power-levels.

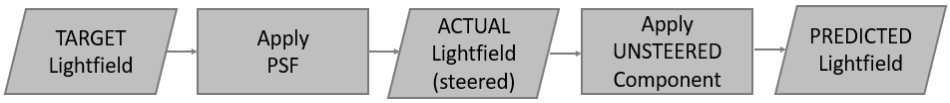


Figure 7.18: *Forward model algorithmic block*

The forward model takes the target light field as input and applies the system PSF to it to predict the result of the actual light field after blurring by the diffuser. An example of the PSF of the RGB projector system is shown below, tiled into a 4×4 pattern:

Figure 7.19: *Point Spread Function applied to a test patters. Note the different size and shape for red, green and blue colour channels.*

The resulting light field then has the effect of the unsteered component added. This is added after blurring since measurement of this image can only be accomplished after passing through the diffuser. In the current system, the fixed pattern is highly non-uniform. In a fiber-coupled system it would approximate a Gaussian profile.

7.3.7 Phase Pattern Computation

The phase pattern generation block calculates the phase patterns required to achieve a target light field.

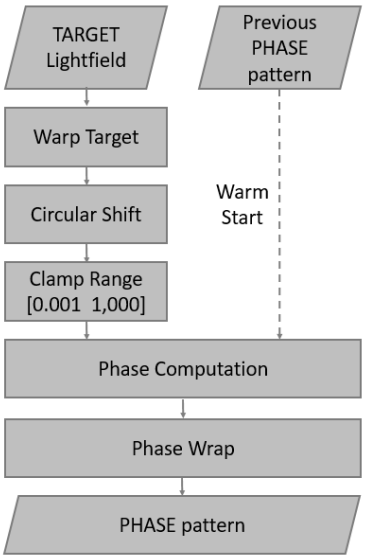


Figure 7.20: *Phase Pattern Computation Block*

In order to frame the image correctly on the amplitude modulator and separate out higher diffraction orders it is necessary to pre-process the target light field. This involves warping by a calibrated distortion intended to align the three channels. Currently, each point in the target image, $[x, y]$, of dimensions $W \times H$ is mapped to a point in the source image, $[x_m, y_m]$, by a 2D cubic polynomial:

$$x_n = \frac{x}{W} \tag{7.8}$$

$$y_n = \frac{y}{H} \tag{7.9}$$

$$b = \left[1, x_n, y, x_n^2, x_n y_n, y_n^2, x_n^3, x_n^2 y_n, x_n y_n^2, y_n^3\right]^T \tag{7.10}$$

$$x_m = b^T \beta_x \tag{7.11}$$

$$y_m = b^T \beta_y \tag{7.12}$$

The source image is then linearly sampled at the $[x_m, y_m]$ corresponding to each destination pixel $[x, y]$. Normalization of the target coordinates (the $[x_n, y_n]$ coordinates), allows the mapping to be computed even for resolution mismatches between source and target images. The 10×1 fit parameter vectors β_x and β_y are obtained from calibration. Once warped and re-sampled, the resulting image is circularly shifted (depending on the optical

configuration), clamped to the range $[0.001, 1000.0]$. At this point the phase computation algorithm is applied. The final phase pattern is then mapped to the 8 bit output range of the phase panels.

7.3.8 Amplitude Pattern Generation

The amplitude pattern generation block determines the amplitude pattern for the steering projector and the non-steering projector using the target image and the predicted light field as input.

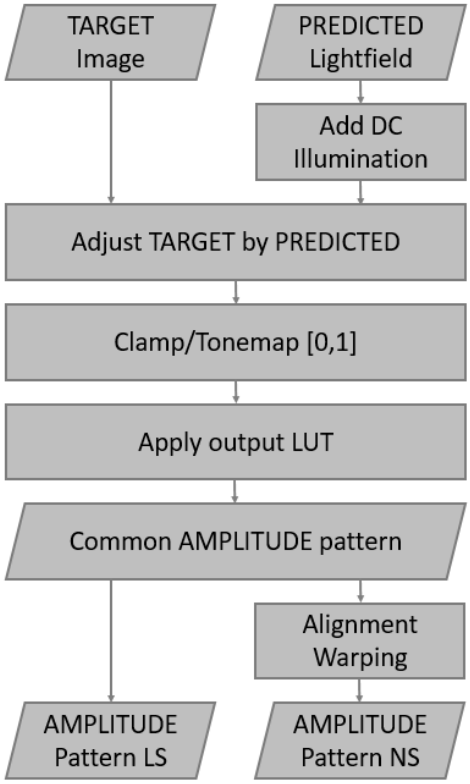


Figure 7.21: *Amplitude Pattern Generation Block*

The algorithm first adds the non-steering illumination to the predicted light field. This is the total light available on-screen. A common amplitude pattern is then computed for both the light steering and the non-light steering amplitude modulator. The resulting image is clamped to a valid range of transmission factors of $[0, 1]$, (or could be tone-mapped in order to preserve texture in out-of-range regions). Any necessary LUTs are then applied to

account for the response of the amplitude SLMs and the pattern is then directly send to the projection head. In order to spatially align the steering and the non-steering projectors, a warping is used based on calibrated pixel correspondences which uses the same cubic warping function as in the phase pattern generation block.

7.4 Results

Figure 7.22 shows photos comparing a cinema projector and our light steering projector side-by-side, both with the same optical power out of lens, playing a video processed using the algorithm framework introduced in this section. The contrast and peak luminance performance was confirmed to be comparable (slightly higher due to better light management within the light path) to the proof-of-concept work in Chapter 6.

Figure 7.22: *Two scenes from the movie Avatar by 20th Century FOX (top and bottom) displayed on the light steering prototype (left) and on a traditional projector with same power out of lens (right). The light steering projector (left) exceeds the comparison projector contrast and peak luminance significantly by about a factor of 20 (light steering projector: 1,000 cd/m^2; right projector: 48 cd/m^2).*

Chapter 8

Discussion and Conclusion

In this thesis we have taken a critical look at the current HDR cinema pipeline from a perceptual point of view and have identified the most significant bottlenecks. We explored a variety of approaches to address these limitations with new optical system designs and computational processing.

8.1 Discussion

In this section we reiterate over the contributions of the work presented in this thesis and briefly discuss each approach in light of computational display and visual perception. A more detailed discussion of each individual topic is included in the respective chapters.

The light steering projector and the associated algorithms introduced in Chapter 6 provide a practical solution to the unsolved challenge of achieving perceptually meaningful contrast and peak luminance in large screen cinema environments where today light source power, system cost, mechanical dimensions and thermal management limit the performance. Compared to traditional cinema projectors our approach provides a visually more appealing image with 20 times or more of the peak luminance capability and with orders of magnitude darker black levels. Compared to the new class of high contrast laser projectors that has recently been introduced into the market as a PLF offering our approach is attractive because it requires a significantly lower power light source and with that also enables a lower cost projector alternative that exceeds the peak luminance of current *brute-force* laser projector offerings by about 10 times. Other high contrast high brightness technologies for large screens include tiled LCD TVs and direct view LED walls. Both achieve great black level and peak luminance comparable to our approach. The difference lies in cost (every pixel of an LED wall or an LCD display has to have the capability to achieve the peak luminance, which results in high local power and with that high cost for LEDs and associated driver electronics). Additionally to date it is not possible to manufacture tiled large displays without visible seams. While the gap size of seams has

improved over the last years, the HVS remains extremely sensitive to even the smallest vertical and horizontal discontinuities in an image. Finally projection screens, for several reasons including vandalism, existing building infrastructure (power and space) and speaker placement behind perforated screens, are currently preferred over direct view displays in cinema. Having said that, direct view LED and OLED displays will in the future provide a feasible, and from visual quality perspective, very compelling alternative to projectors including the systems introduced in this thesis.

An appearance reproduction algorithm has been introduced in Chapter 3 which was inspired from the fields of colour appearance modelling and tone reproduction research. While traditionally the difference in peak luminance, contrast and colour performance as well as screen size between different available display devices and cameras was comparable and hence one unified video signal was sufficient to represent content for *all* devices, today an ever increasing performance gap between the different types of displays surrounding the user in their daily lives (mobile phones, tablets, computer screens, TVs, cinema, advertising displays and even Virtual Reality (VR) and Augmented Reality (AR) devices) also requires that video signals account for the different properties of these displays as well as the different viewing environments that these displays are being viewed in (e.g. dark in cinema, bright for mobile screens). Our model takes into account the characteristics of the display device as well as parameters describing the viewing environment and leads to precise appearance reproduction. Where these variables are unknown we have proposed a robust method to directly approximate the scene-referred parameters leading to plausible reproduction of the image content. Both the field of colour appearance modelling and tone reproduction operators have recently received a large amount of attention from the computer graphics research community. We have aimed to bring the two fields closer together with a new model that can handle a large dynamic range *and* reproduce colours accurately. Our research and findings related to the HVS's lightness perception have been instrumental in prototyping bit efficient transmission of video data over existing interfaces and in preparing both SDR and HDR content for viewing on our light steering projector.

In Chapter 5, based on the initial proof-of-concept projector performance we have analysed typical image statistics to understand the implications for a meaningful hardware design and propose two projector architectures that can enable scaling to larger screens with available electro-optical components. These findings have been taken into account to develop a new full featured projector prototype introduced in Chapter 7. Here we have tried

to address the major practical limitations of the proof-of-concept work in Chapter 6, discussed temporal considerations as well as detailed characterization and optical modelling of the hardware and demonstrated that the methods presented in this thesis can go beyond the experimental and theoretical research work in the lab and possibly make a meaningful impact in the cinema in the future.

8.2 Future Work

Throughout the work presented in this thesis we have attempted to put value on *perceptual engineering* by thinking about perceptually meaningful targets first, and only then finding the means to achieve them. This methodology might be risky in the sense that the solutions found present a number of new challenges that require solving, the potential gains however can be large. The methods and systems developed in this thesis only present a starting point for research in the new fields of computational projection display and accurate colour appearance reproduction. Taking into account the recent advances in display technologies we look forward to further research into displays that combine optical modelling and computational processing to achieve perceptually meaningful gains. Contrast (both global and local) as well as peak luminance levels comparable to real world scenes are important cues in depth perception of the HVS. Stereoscopic and light field displays are emerging, but need to improve in spatial resolution, contrast, luminance and practicality of their implementation. Further work to close the gap between these new type of displays and our work will provide even more realistic image appearance. While we have attempted to address the major practical and engineering challenges in this research work, there remains a long path to a product that is viable in the market and robust enough to be utilized in professional environments such as cinema.

8.3 Conclusion

We have conceived a new computational projection display architecture that for the first time allows an increase in both the contrast and more importantly the peak luminance in a perceptually meaningful fashion - by orders of magnitude. We have introduced new algorithmic approaches to compute dynamic freeform lensing phase patterns efficiently for this computational projection display. An initial monochromatic proof-of-concept projector has demonstrated that the approach shows promise, but is limited in power

(screen size), colour (achievable chromaticity) and overall image quality. To understand these limitations as well as requirements for cinema in more detail we have developed a colour appearance model and tone mapping operator that operates in an absolute, calibrated colour space over a wide range of luminance values and takes into account viewing environment and the HVS's adaptation to display and environment. We have explored new and existing tools to transmit high bit depth data between devices and we have touched upon perceptual aspects of image signal discretization that supports high brightness HDR data. Inspired by projector power requirements to reproduce HDR image content faithfully as well as based on the overall efficiency of optical components within the projector we have proposed a new hybrid architecture that balances the overall system light throughput and peak luminance. Finally we have prototyped a full featured projector to address the remaining limitations of the proof-of-concept projector and developed an optical model that allows colourimetric accurate image reproduction of high brightness HDR content in a cinema environment (see Figure 7.22) with an up to now unachievable peak luminance of up to 1,000 cd/m^2 and with in-scene contrast of up to $1,000,000:1$.

Bibliography

[1] A. Adams. *The print*. The Ansel Adams Photography series. Little, Brown and Company, 1983.

[2] D. Armitage, I. Underwood, and S.-T. Wu. *Introduction to microdisplays*, volume 11. John Wiley & Sons, 2006.

[3] T. Aydin, R. Mantiuk, K. Myszkowski, and H. Seidel. Dynamic range independent image quality assessment. *ACM Transactions on Graphics*, 27:69(1)–69(10), 2008.

[4] A. Ballestad and A. Kostin. Method and apparatus for image data transformation, November 26 2013. US Patent 8,593,480.

[5] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers. *Advanced High Dynamic Range Imaging: Theory and Practice*. AK Peters / CRC Press, Natick, MA, 2011.

[6] M. Berry. Oriental magic mirrors and the laplacian image. *European journal of physics*, 27(1):109, 2006.

[7] O. Bimber and D. Iwai. Superimposing dynamic range. *ACM Trans. Graph.*, 27(5):150, 2008.

[8] G. Blackham and A. Neale. Image display apparatus, March 18 1998. EP Patent App. EP19,970,306,624.

[9] W. P. Bleha and L. A. Lei. Advances in liquid crystal on silicon (LCOS) spatial light modulator technology. In *SPIE Defense, Security, and Sensing*, pages 87360A–87360A. International Society for Optics and Photonics, 2013.

[10] E. Buckley. 70.2: Invited paper: holographic laser projection technology. In *Proc. SID*, volume 39, pages 1074–1079, 2008.

[11] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[12] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman. Spatially nonuniform scaling functions for high contrast images. In *Proceedings of Graphics Interface '93*, pages 245–253, May 1993.

[13] S. Daly, T. Kunkel, X. Sun, S. Farrell, and P. Crum. 41.1: Distinguished paper: Viewer preferences for shadow, diffuse, specular, and emissive luminance limits of high dynamic range displays. In *SID Symposium Digest of Technical Papers*, volume 44, pages 563–566. Wiley Online Library, 2013.

[14] S. Daly, T. Kunkel, X. Sun, S. Farrell, and P. Crum. Preference limits of the visual dynamic range for ultra high quality and aesthetic conveyance. In *Proc. SPIE*, volume 8651, page 86510J, 2013.

[15] G. Damberg, A. Ballestad, E. Kozak, R. Kumaran, and J. Minor. Efficient, high brightness, high dynamic range projection, 2014.

[16] G. Damberg, J. Gregson, A. Ballestad, E. Kozak, J. Minor, R. Kumaran, and W. Heidrich. High-brightness HDR projection using dynamic-phase modulation, 2015.

[17] G. Damberg, J. Gregson, and W. Heidrich. High brightness HDR projection using dynamic freeform lensing. *ACM Transactions on Graphics (TOG)*, 35(3):24, 2016.

[18] G. Damberg and W. Heidrich. Efficient freeform lens optimization for computational caustic displays. *Optics Express*, 23(8):10224–10232, 2015.

[19] G. Damberg, H. Seetzen, G. Ward, W. Heidrich, and L. Whitehead. 3.2: High dynamic range projection systems. In *Proc. SID*, volume 38, pages 4–7. Wiley Online Library, 2007.

[20] G. Damberg, H. Seetzen, G. Ward, M. Kang, P. Longhurst, W. Heidrich, and L. Whitehead. High dynamic range projector, 2007.

[21] P. Dasgupta, M. K. Das, and B. Das. Physical properties of three liquid crystals with negative dielectric anisotropy from x-ray diffraction and optical birefringence measurements. *Molecular Crystals and Liquid Crystals*, 540(1):154–161, 2011.

[22] P. Debevec. A median cut algorithm for light probe sampling. In *ACM SIGGRAPH 2005 Posters*, SIGGRAPH '05, New York, NY, USA, 2005. ACM.

[23] L. Digital Cinemea Initiatives. DCI specifications documentation. `http://www.dcimovies.com/`, visited on 01 February 2017.

[24] F. Drago, K. Myszkowski, T. Annen, and N. Chiba. Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum*, 22(3):419–426, 2003.

[25] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics*, 21(3):257–266, 2002.

[26] M. Fairchild. The HDR photographic survey. In *Proceedings of the Fifteenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications*, volume 15, pages 233–238, 2007.

[27] M. D. Fairchild. *Color appearance models*. Addison-Wesley, Reading, MA, 2nd edition edition, 2005.

[28] M. D. Fairchild. The HDR photographic survey. In *Color and Imaging Conference*, pages 233–238. Society for Imaging Science and Technology, 2007.

[29] M. D. Fairchild and G. M. Johnson. Meet iCAM: an image color appearance model. In *IS&T/SID* 10^{th} *Color Imaging Conference*, pages 33–38, Scottsdale, 2002.

[30] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. *ACM Transactions on Graphics*, 21(3):249–256, 2002.

[31] E. J. Fernández, P. M. Prieto, and P. Artal. Wave-aberration control with a liquid crystal on silicon (LCOS) spatial phase modulator. *Optics express*, 17(13):11013–11025, 2009.

[32] J. A. Ferwerda, S. Pattanaik, P. Shirley, and D. P. Greenberg. A model of visual adaptation for realistic image synthesis. In *SIGGRAPH 96 Conference Proceedings*, pages 249–258, 1996.

[33] M. Finckh, H. Dammertz, and H. P. Lensch. Geometry construction from caustic images. In *Proc. ECCV*, pages 464–477, 2010.

[34] G. A. Fishman and S. Sokol. *Electrophysiological Testing in Disorders of the Retina, Optic Nerve, and Visual Pathway*. American Academy of Ophthalmology, San Francisco, 1990.

[35] J. Froehlich, T. Kunkel, R. Atkins, J. Pytlarz, S. Daly, A. Schilling, and B. Eberhardt. Encoding color difference signals for high dynamic range and wide gamut imagery. In *Color and Imaging Conference*, volume 2015, pages 240–247. Society for Imaging Science and Technology, 2015.

[36] A. Gilchrist, C. Kossyfidis, F. Bonato, T. Agostini, J. Cataliotti, X. Li, B. Spehar, V. Annan, and E. Economou. An anchoring theory of lightness perception. *Psychol Rev.*, 106(4):795–834, Oct 1999.

[37] A. Gilchrist and A. Radonjić. Frameworks of illuination revealed by probe disk technique. *Journal of Vision*, 10(5):1–12, 2010.

[38] R. R. Hainich and O. Bimber. *Displays: fundamentals & applications*. CRC press, 2016.

[39] P. R. Haugen, H. Bartelt, and S. K. Case. Image formation by multifacet holograms. *Applied optics*, 22(18):2822–2829, 1983.

[40] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM Transactions on Graphics (ToG)*, 32(4):45, 2013.

[41] A. Hermerschmidt, S. Osten, S. Krüger, and T. Blümel. Wave front generation using a phase-only modulating liquid-crystal-based microdisplay with HDTV resolution. In *International Congress on Optics and Optoelectronics*, pages 65840E–65840E. International Society for Optics and Photonics, 2007.

[42] Holoeye Photonics AG. Holoeye corporation. `http://www.holoeye.com`, visited on 01 February 2017.

[43] D. C. Hood, T. Ilves, E. Maurer, B. Wandell, and E. Buckingham. Human cone saturation as a function of ambient intensity: A test of models of shifts in the dynamic range. *Vision Research*, 18(8):983–993, 1978.

[44] R. Hoskinson, S. Hampl, and B. Stoeber. Arrays of large-area, tip/tilt micromirrors for use in a high-contrast projector. *Sensors and Actuators A: Physical*, 173(1):172–179, 2012.

[45] R. Hoskinson and B. Stoeber. Increasing projector contrast and brightness through light redirection. *Optical Imaging Devices: New Technologies and Applications*, 44:49, 2015.

[46] R. Hoskinson, B. Stoeber, W. Heidrich, and S. Fels. Light reallocation for high contrast projection using an analog micromirror array. *ACM Transactions on Graphics (TOG)*, 29(6):165, 2010.

[47] M. B. Hullin, I. Ihrke, W. Heidrich, T. Weyrich, G. Damberg, and M. Fuchs. State of the art in computational fabrication and display of material appearance. In *Eurographics Annual Conference (STAR)*, 2013.

[48] R. W. G. Hunt. *The reproduction of color*. Fountain Press, England, 1996. Fifth edition.

[49] iC-Haus GmbH. iC-HG product page. `http://ichaus.de/HG`, visited on 01 February 2017.

[50] Y. Isomae, Y. Shibata, T. Ishinabe, and H. Fujikake. P-199l: Late-news poster: Optical phase modulation properties of 1 μm-pitch LCOS with dielectric walls for wide-viewing-angle holographic displays. In *SID Symposium Digest of Technical Papers*, volume 47, pages 1670–1673. Wiley Online Library, 2016.

[51] M. H. Kim, T. Weyrich, and J. Kautz. Modeling human color perception under extended luminance levels. *ACM Transactions on Graphics*, 28(3):27:1–9, 2009.

[52] C. Kiser, E. Reinhard, M. Tocci, and N. Tocci. Real-Time automated tone mapping system for HDR video. In *IEEE International Conference on Image Processing*, 2012.

[53] T. Kiser, M. Eigensatz, M. M. Nguyen, P. Bompas, and M. Pauly. Architectural causticscontrolling light with geometry. In *Advances in Architectural Geometry 2012*, pages 91–106. Springer, 2013.

[54] P. Kohns, J. Schirmer, A. A. Muravski, S. Y. Yakovenko, V. Bezborodov, and R. Dãbrowski. Birefringence measurements of liquid crystals and an application: An achromatic waveplate. *Liquid crystals*, 21(6):841–846, 1996.

[55] G. Krawczyk, R. Mantiuk, K. Myszkowski, and H. Seidel. Lightness perception inspired tone mappping. In *First ACM Symposium on Applied Perception in Graphics and Visualization (APGV)*, page 172, 2004.

[56] J. Kuang, G. M. Johnson, and M. D. Fairchild. iCAM06: a refined image appearance model for HDR image rendering. *Journal of Visual Communication and Image Representation*, 18(5):406 – 414, 2007.

[57] T. Kunkel and E. Reinhard. A neurophysiology-inspired steady-state color appearance model. *Journal of the Optical Society of America A*, 26:776–782, 2009.

[58] T. Kunkel, G. Damberg, and L. Johnson. A high bit depth digital imaging pipeline for vision research. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, pages 121–121. ACM, 2011.

[59] Y. Kusakabe, M. Kanazawa, Y. Nojiri, M. Furuya, and M. Yoshimura. A high-dynamic-range and high-resolution projector with dual modulation. In *IS&T/SPIE Electronic Imaging*, pages 72410Q–72410Q. International Society for Optics and Photonics, 2009.

[60] A. Lasansky. Synaptic action mediating cone responses to annular illumination in the retina of the larval tiger salamander. *Journal of Physiology*, 310:205–214, 1981.

[61] G. Lazarev, A. Hermerschmidt, S. Kruger, and S. Osten. Optical imaging and metrology: Advanced technologies, 2012.

[62] G. Lazarev, F. Gädeke, and J. Luberek. Ultrahigh-resolution phase-only LCOS spatial light modulator. In *Proc. of SPIE Vol*, volume 10125, pages 101250M–1, 2017.

[63] G. Lazarev, A. Hermerschmidt, S. Krüger, and S. Osten. LCOS spatial light modulators: trends and applications, 2012.

[64] P. Ledda, L. P. Santos, and A. Chalmers. A local model of eye adaptation for high dynamic range images. In *AFRIGRAPH '04*, pages 151–160, New York, NY, USA, 2004. ACM.

[65] L. Lesem, P. Hirsch, and J. Jordan. The kinoform: a new wavefront reconstruction device. *IBM Journal of Research and Development*, 13(2):150–155, 1969.

[66] C. Li, M. R. Luo, B. Rigg, and R. W. G. Hunt. CMC 2000 chromatic adaptation transform: CMCCAT2000. *Color Research and Application*, 27(1):49–58, 2002.

[67] Y. Li, L. Sharan, and E. Adelson. Compressing and companding high dynamic range images with subband architectures. *ACM Transactions on Graphics*, 24(3):836–844, 2005.

[68] T. Lu, F. Pu, P. Yin, T. Chen, W. Husak, J. Pytlarz, R. Atkins, J. Fröhlich, and G. Su. ITP colour space and its compression performance for high dynamic range and wide colour gamut video distribution. *ZTE Communications, Feb*, 2016.

[69] M. R. Luo, A. Clark, P. Rhodes, A. Schappo, S. Scrivner, and C. Tait. Quantifying colour appearance: Part I. LUTCHI colour appearance data. *Colour Research and Application*, 16:166–180, 1991.

[70] R. Mantiuk, R. Mantiuk, A. Tomaszewska, and W. Heidrich. Color correction for tone mapping. *Computer Graphics Forum*, 28(2):pp. 193–202, 2009.

[71] R. Mantiuk, S. Daly, and L. Kerofsky. Display adaptive tone mapping. *ACM Trans. on Graphics*, 27(3):68, 2008.

[72] J. J. McCann and A. Rizzi. *The Art and Science of HDR Imaging*. John Wiley and Sons, Chichester, 2012.

[73] S. Miller, M. Nezamabadi, and S. Daly. Perceptual signal coding for more efficient usage of bit codes. *SMPTE Motion Imaging Journal*, 122(4):52–59, 2013.

[74] J. C. Miñano, P. Benítez, and A. Santamaría. Free-form optics for illumination. *Optical Review*, 16(2):99–102, 2009.

[75] P. Moon and D. E. Spencer. Visual data applied to lighting design. *Journal of the Optical Society of America*, 34(10):605–617, 1944.

[76] N. Moroney, M. D. Fairchild, R. W. G. Hunt, C. J. Li, M. R. Luo, and T. Newman. The CIECAM02 color appearance model. In 10^{th} *Color Imaging Conference*, pages 23–27, Scottsdale, 2002.

[77] K. Myszkowski, R. Mantiuk, and G. Krawczyk. *High Dynamic Range Video*. Morgan and Claypool Publishers, San Rafael, 2008.

[78] K.-I. Naka and W. A. Rushton. S-potentials from luminosity units in the retina of fish (cyprinidae). *The Journal of physiology*, 185(3):587, 1966.

[79] M. Nezamabadi, S. Miller, S. Daly, and R. Atkins. Color signal encoding for high dynamic range and wide color gamut based on human perception. In *IS&T/SPIE Electronic Imaging*, pages 90150C–90150C. International Society for Optics and Photonics, 2014.

[80] M. K. Ng, R. H. Chan, and W.-C. Tang. A fast algorithm for deblurring models with neumann boundary conditions. *SIAM Journal on Scientific Computing*, 21(3):851–866, 1999.

[81] NVIDIA. CUDA programming guide, CUSPARSE, CUBLAS, and CUFFT library user guides.{Online}. `http://docs.nvidia.com/cuda/`, visited on 01 February 2017.

[82] Y. Ohno. Color rendering and luminous efficacy of white led spectra. In *Proc. of SPIE Vol*, volume 5530, page 89, 2004.

[83] M. Papas, T. Houit, D. Nowrouzezahrai, M. Gross, and W. Jarosz. The magic lens: refractive steganography. *ACM Transactions on Graphics (TOG)*, 31(6):186, 2012.

[84] M. Papas, W. Jarosz, W. Jakob, S. Rusinkiewicz, W. Matusik, and T. Weyrich. Goal-based caustics. *Computer Graphics Forum*, 30(2):503–511, 2011.

[85] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):123–231, 2013.

[86] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, and D. P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. In *SIGGRAPH 98 Conference Proceedings*, pages 287–298, 1998.

[87] M. Pauly and T. Kiser. Caustic art. Technical report, EPFL, 2012.

[88] T. Pouli and E. Reinhard. Progressive color transfer for images of arbitrary dynamic range. *Computers and Graphics*, 35(1):67–80, 2011.

[89] C. Poynton. *Digital video and HD: Algorithms and Interfaces*. Elsevier, 2012.

[90] E. Reinhard and K. Devlin. Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Visualization and Computer Graphics*, 11(1):13–24, 2005.

[91] E. Reinhard, E. A. Khan, A. O. Akyüz, and G. M. Johnson. *Color Imaging: Fundamentals and Applications*. A K Peters, Wellesley, 2008.

[92] E. Reinhard, T. Kunkel, Y. Marion, J. Brouillat, R. Cozot, and K. Bouatouch. Image display algorithms for high and low dynamic range display devices. *Journal of the Society for Information Display*, 15(12), 2007.

[93] E. Reinhard, T. Pouli, T. Kunkel, B. Long, A. Ballestad, and G. Damberg. Calibrated image appearance reproduction. *ACM Transactions on Graphics (TOG)*, 31(6):201, 2012.

[94] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. Photographic tone reproduction for digital images. *ACM Transactions on Graphics*, 21(3):267–276, 2002.

[95] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, W. Heidrich, and K. Myszkowski. *High dynamic range imaging: Acquisition, display and image-based lighting*. Morgan Kaufmann Publishers, San Francisco, 2nd edition edition, 2010.

[96] A. Rempel, W. Heidrich, H. Li, and R. Mantiuk. Video viewing preferences for HDR displays under varying ambient illumination. *Proc. APGV*, pages 45–52, 2009.

[97] A. Rempel, W. Heidrich, and R. Mantiuk. The role of contrast in the perceived depth of monocular imagery. *Proc. APGV*, page 115, 2011.

[98] M. D. Robinson, G. Sharp, and J. Chen. *Polarization engineering for LCD projection*, volume 4. John Wiley & Sons, 2005.

[99] C. Schlick. Quantization techniques for the visualization of high dynamic range pictures. In P. Shirley, G. Sakas, and S. Müller, editors, *Photorealistic Rendering Techniques*, pages 7–20. Springer-Verlag Berlin Heidelberg New York, 1994.

[100] Y. Schwartzburg, R. Testuz, A. Tagliasacchi, and M. Pauly. High-contrast computational caustic design. *ACM Trans. Graph. (Proc. Siggraph)*, 2014. (in print).

[101] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, Ghosh, and A. Vorozcovs. High dynamic range display systems. *ACM Trans. Graph. (Proc. SIGGRAPH)*, pages 760–768, August 2004.

[102] H. Seetzen. *High dynamic range display and projection systems*. PhD thesis, University of British Columbia, 2009.

[103] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs. High dynamic range display systems. *ACM Trans. on Graphics*, 23(3), 2004.

[104] D. Stidwill and R. Fletcher. *Normal Binocular Vision: Theory, Investigation and Practical Aspects*. Wiley-Blackwell, Chichester, 2011.

[105] Texas Instruments Inc. DLP technology documentation. `http://www.ti.com/lsds/ti/dlp/overview.page`, visited on 01 February 2017.

[106] M. D. Tisdall, G. Damberg, P. Wighton, N. Nguyen, Y. Tan, M. S. Atkins, H. Li, and H. Seetzen. Comparing signal detection between novel high-luminance HDR and standard medical LCD displays. *Display Technology, Journal of*, 4(4):398–409, 2008.

[107] M. D. Tocci, C. Kiser, N. Tocci, and P. Sen. A Versatile HDR Video Production System. *ACM Transactions on Graphics (TOG) (Proceedings of SIGGRAPH 2011)*, 30(4), 2011.

[108] H. Toyoda, T. Inoue, N. Mukozaka, T. Hara, and M. H. Wu. 39.3: Invited paper: Advances in application of liquid crystal on silicon spatial light modulator (LCOS-SLM). In *SID Symposium Digest of Technical Papers*, volume 45, pages 559–562. Wiley Online Library, 2014.

[109] J. Tumblin and H. Rushmeier. Tone reproduction for computer generated images. *IEEE Computer Graphics and Applications*, 13(6):42–48, 1993.

[110] UHD Alliance. UHD Alliance online ressources. `http://www.uhdalliance.org/`, visited on 01 February 2017.

[111] J. H. van Hateren. Encoding of high dynamic range video with a model of human cones. *ACM Transactions on Graphics*, 25(4):1380–1399, 2006.

[112] G. Ward, H. Rushmeier, and C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4):291–306, 1997.

[113] A. M. Weiner, D. E. Leaird, J. Patel, and J. R. Wullert. Programmable shaping of femtosecond optical pulses by use of 128-element liquid crystal phase modulator. *IEEE Journal of Quantum Electronics*, 28(4):908–920, 1992.

[114] S.-T. Wu, U. Efron, and L. D. Hess. Birefringence measurements of liquid crystals. *Applied optics*, 23(21):3911–3915, 1984.

[115] Y. Yue, K. Iwasaki, B.-Y. Chen, Y. Dobashi, and T. Nishita. Poisson-based continuous surface generation for goal-based caustics. *ACM Trans. Graph.*, 2014. (in print).

[116] Y. Yue, K. Iwasaki, B.-Y. Chen, Y. Dobashi, and T. Nishita. Pixel art with refracted light by rearrangeable sticks. *Computer Graphics Forum*, 31(2pt3):575–582, 2012.

[117] Z. Zhang, Z. You, and D. Chu. Fundamentals of phase-only liquid crystal on silicon (LCOS) devices. *Light: Science & Applications*, 3(10):e213, 2014.

[118] M. Zink and M. Smith. Managing HDR content production and display device capabilities. *International Broadcasting Convention 2015*, 2015.