Chapter 1

Introduction

Digital displays are ubiquitous in today's world, from huge projection screens in theaters to a tiny watch face on a person's wrist. An increasing number of digital displays are appearing around us at work, at home, and even in cars. Such growth, driven by the desire to exchange digital content generated in this world, is a signal that digital displays are becoming the de facto interface for communication.

This dissertation introduces four novel projects that demonstrate how we can augment digital displays using computational resources. Our display algorithms jointly considers the characteristics of a fixed set of display hardware and the human visual system on the other side, in order to improve the display quality.

First, we propose a software-based approach to driving multiview autostereoscopic displays. Our display algorithm can dynamically assign views to hardware display zones based on multiple observers' current head positions, substantially reducing crosstalk and stereo inversion.

Second, we present a dense projector array that creates a seamless 3D viewing experience for multiple viewers. We smoothly interpolate the set of viewer heights and distances on a per-vertex basis across the arrays field of view, reducing image distortion, crosstalk, and artifacts from tracking errors.

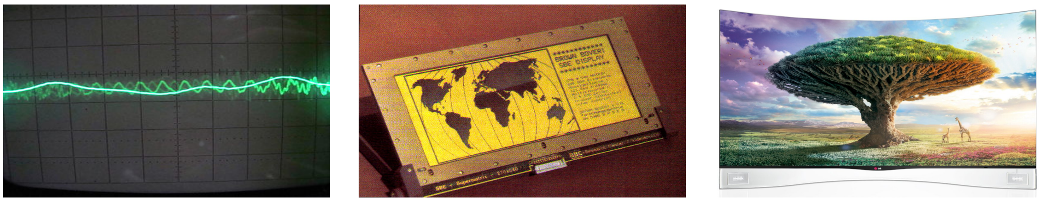
Third, we propose a method for high dynamic range display calibration that takes into account the variation of the chrominance error over luminance. We propose a data structure for enabling efficient representation and querying of the calibration function, which also allows user-guided balancing between memory consumption and the amount of computation.

Fourth, we present user studies that demonstrate that the $\sim 60Hz$ critical flicker fusion rate for traditional displays is not enough for some computational displays that show complex image patterns. The study focuses on displays with hidden channels, and their application to 3D+2D TV.

1.1 Background

1.1.1 Display Evolution

The evolution of digital displays has come a long way. Figure 1.1 presents a few milestones in the digital display evolution history: from a monochrome cathode ray tube (CRT) that is mainly used for displaying simple graph and texts, to a flat liquid crystal display (LCD) panel, to today's high definition curved organic light-emitting diode (OLED) display. It is quite surprising that it has been less than a hundred years since the earliest CRT displays were brought to the market. (Figure 1.1 and captions are partially cited from [108].)



(a) Monochrome CRT in 1922 (b) LCD Panel in 1984 (c) Curved OLED in 2013

Figure 1.1: (a) The earliest CRTs were monochrome and were primarily used in oscilloscopes and black and white televisions; (b) The first time a high resolution (540x270) LCD panel was made possible by Brown Boveri Research Center; (c) A large-sized (55 inch) curved OLED television introduced by LG into the market.

However, if we examine state-of-the-art digital displays in a more careful way, we will come to realize that there is still a long way to go before the "ultimate" digital display can exist. The digital display evolution is probably still at its infant stage.

1.1.2 The Ultimate Display

During the pursuit of the ultimate display, we first need to know what qualifies as an ultimate display. Similar to the famous Turing test for artificial intelligence, a simple experiment could be set up to test a digital display's ability to faithfully reproduce the real world scenes (as shown in Figure 1.2). In this experiment, an opaque wall would separate human judges in a room from seeing the outdoor scenes. A window and a display of the same size would be placed side by side on the wall. An ultimate digital display should be indistinguishable from a window. Human judges would be allowed to walk around in front of the wall to evaluate the two. If the judge(s) cannot reliably tell the digital display from a real window, then the display would be said to have passed the test.

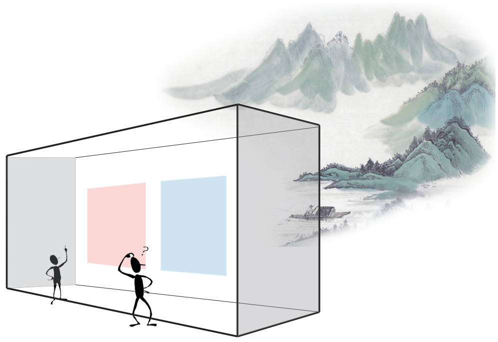


Figure 1.2: Turing test for a display. Observers inside the room are given the task to determine which frame on the wall is a display and which is a window. Observers can move around but are limited to using only visual perceptions to make the determination.

This test involves several key factors regarding what an ultimate display should be capable of doing. One of the most challenging ones is the ability to reproduce a 3D scene with depth perception, without the assistance of special eyewear. As human evaluators walk around the display, they should be able to observe a continuous motion parallax of the outside world, and each individual's eyes could focus at objects from different distances with appropriate accommodation and vergence. This would require the display to have high spatial and angular resolution. Another important factor would be the dynamic range of the display, which determines how bright and how much contrast human observers can perceive in the reproduced scene. Display factors like these and how they affect human visual system will be introduced with more details in

Section 1.1.4.

1.1.3 Display Models

The core element of a digital display is *the pixel*. As shown in Figure 1.3, one simple way to model a display is using a 2D grid of pixels. All pixels on the grid can be controlled independently with colors specified to form an image. Normally a pixel sends out uniform light rays in all directions, which means that the color of the pixel is independent of viewing direction. In applications where angle-dependent views are desired, direction-dependent pixels are used to model the display. These pixels have different colors when viewed from different angles.

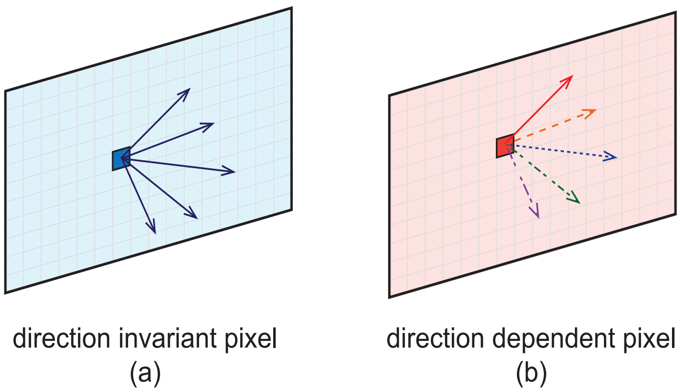


Figure 1.3: Display pixels: (a) direction-invariant pixel sends out uniform light rays in all directions; (b) direction-dependent pixel sends out distinguished light rays in different directions.

When we form a 2D grid of direction-dependent pixels, a light field display model is obtained. The advantage of a light field display model is its capability to completely characterize the flow of light through a fixed region in 3D space. It represents the radiance as a function of not only position but also direction [61].

1.1.4 Display Factors and How They Affect Human Visual System

There are many factors that could be used for evaluating the quality of digital displays. Quite interestingly, we can often find fairly direct mappings between each display property and the corresponding human visual perception that is affected (see Figure 1.4). Therefore, in this dissertation, these correspondences are chosen to establish a framework, to categorize the advances in display technologies.

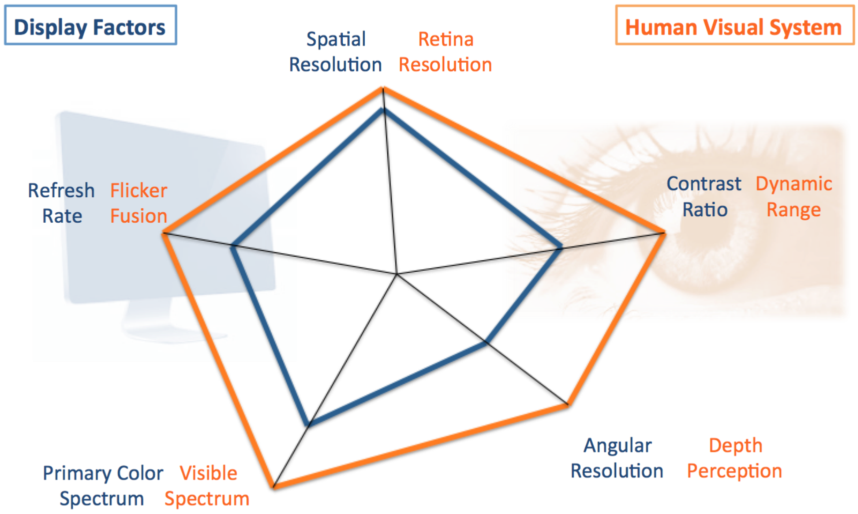


Figure 1.4: This graph establishes a framework to categorize the corresponding relationship among different display factors and modes of human visual perception. Each corner represents a pairing of display factor and human perception.

In this section, I will introduce several digital displays properties (*Spatial Resolution*, *Angular Resolution*, *Contrast Ratio*, *Primary Colors*, *Refresh Rate*) and their effect on relevant types of perception in human visual system.

1.1.4.1 Spatial Resolution

Spatial resolution can be defined as pixels per square inch on a display panel, which corresponds to the human eye retina resolution. The pixel spatial density is often the primary focus for display manufacturers. However, there is a limit on human visual acuity. The spatial resolution of modern displays is starting to exceed the limit that human eyes can resolve.

As shown Figure 1.5, normal visual acuity is commonly referred to as 20/20 vision [110], which means a normal human eye is able to separate contours that are approximately 1.75 mm apart at 20 feet away.

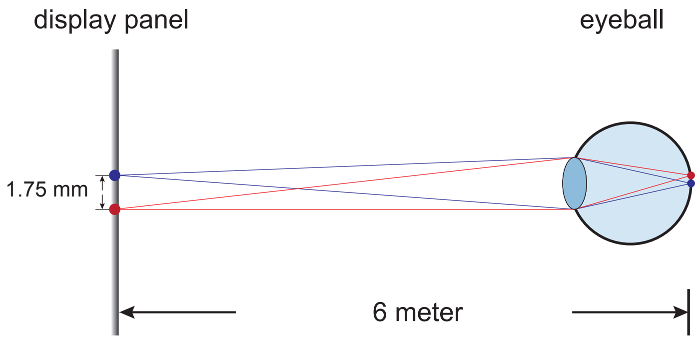
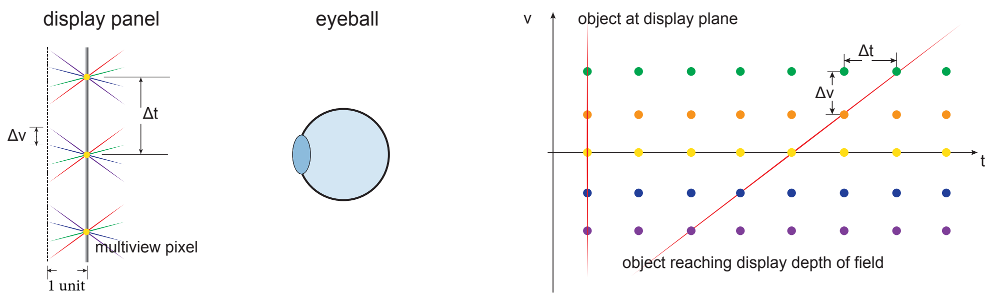


Figure 1.5: Spatial resolution and its relationship to eye retina resolution. A normal human eye is only able to separate contours that are approximately 1.75 mm apart at 6 meters away.

Normal visual acuity is equivalent to a visual angle of 0.0029 degrees of arc. Normal vision cannot tell the difference below this threshold. Emerging displays with 400 pixels per inch spatial resolution, if viewed from 10 inches away, are surpassing this threshold as two neighboring pixels form a visual angle of only 0.0025 degree of arc.

1.1.4.2 Angular Resolution

Angular resolution describes the density of distinguished views that a display can send out in the angular domain. The display angular resolution is an important factor in determining the depth of field of a glasses-free 3D display (see Figure 1.6), and mainly affects the human visual depth perception.



(a) (b)

Figure 1.6: (a) Multiview pixels allow the viewer to perceive different images from different angles. We follow the conventions in [15] that define Δt as the pitch between neighboring multiview pixels, and Δv as the pitch of view-dependent subpixel at 1 unit distance from t plane. (b) Display depth of field can then be derived from this parameterization of light rays. Objects at display plane depth correspond to a vertical line as it remains at same position from different views. Objects at greater depth from display plane correspond to a more slanted line, and become blurry after exceeding the display bandwidth at $\Delta t/\Delta v$. More details of the derivation could be found in [120].

Depth perception is achieved when two eyes receive views from slightly different angles. A typical modern autostereoscopic display can generate 3D images at 3.58 degree angular resolution (number of views over display field of view), which is only capable of showing a shallow depth of field of $\pm 32mm$. Such angular resolution in current displays is far from satisfactory because of the limited ability in rendering 3D scenes.

1.1.4.3 Contrast Ratio

Contrast ratio refers to the ratio between the largest and smallest possible values of light intensity that a display can output. The corresponding term for human visual perception is *dynamic range*, which describes what a human is capable of seeing: from objects in dark starlight to very bright sunlight. Because "stops" is often used to describe the dynamic range of eyes, we provide a conversion map between contrast factors and stops in Table 1.1.

Table 1.1: Conversion Table: Stops to Factors

Factor	1	2	4	8	16	100	1000	1 Million	1 Billion
Stops	0	1	2	3	4	6.64	9.97	19.9	30

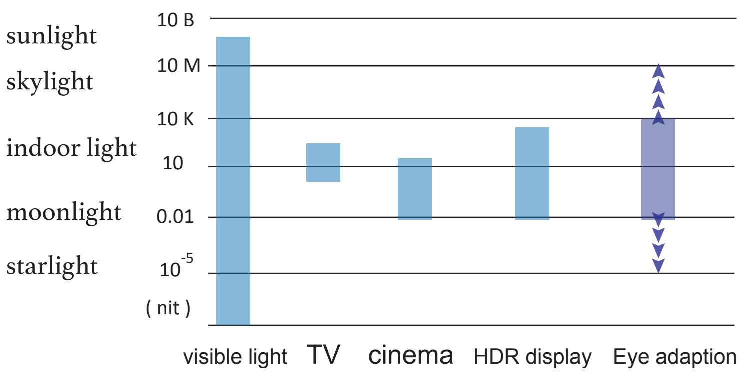
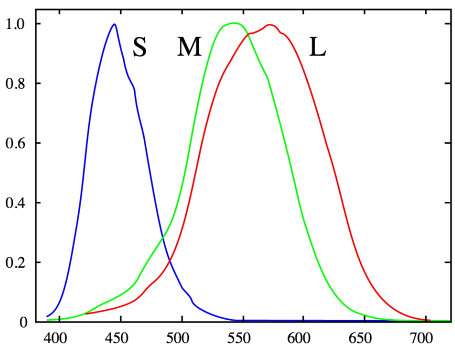


Figure 1.7: The contrast ratio of existing displays is quite limited when compared against real world scenes or human eye dynamic range. For example, good LCD screens only present a contrast ratio of ∼300:1 (8 stops); in a controlled dark environment like cinema, the screen could reach ∼10000:1 (13 stops). High dynamic range display prototypes claim to be over ∼400000:1 (18 stops), which is very promising. But notice this only reaches the eye's instantaneous dynamic range. If we consider situations where the pupils adjust openings, then it can see over a range of nearly 24 stops (as indicated by the rightmost bar and arrows). In the real world, the difference between sunlight and starlight illumination can exceed a factor of 1,000,000,000 in power (30 stops).

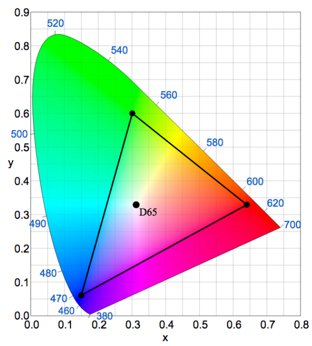
Figure 1.7 presents the relative relationship between dynamic range of human eyes and the contrast ratio of existing displays. As we can see, no conventional digital displays come even close to the dynamic range on eyes, let alone the real world environment. Emerging high dynamic range display prototypes [84, 99] lead us to a promising display technology in this domain. However, many problems still exist with these prototypes that must be resolved before they can be widely adopted.

1.1.4.4 Primary Colors

Primary colors are a set of colors that can be combined to generate a useful range of colors. A typical set of three color primaries that are standardized for high-definition television is shown in Figure 1.8(b). How we choose primary colors originates from the three types of cones in our retina, which together decide the color gamut of human perceivable colors (human cone cells response shown in Figure 1.8(a)).



(a)



(b)

Figure 1.8: (a) This plot shows the normalized responsivity spectra of three types human cone cells, which are responsible for color vision [107]. (b) Diagram of the CIE 1931 color space that shows the Rec. 709 (HDTV) color space in the triangle and the location of three primary colors [109].

Human eyes perceive a wide range of colors that no commercial displays can yet match completely. Notice that the Rec. 709 color primary triangle only covers 35.9% of the total human perceivable color space. Display manufacturers are starting to add more primary colors in order to cover the range of colors that human eye is capable of perceiving. Besides color coverage, accurate color reproduction of specified input colors is also an important criteria for digital displays.

1.1.4.5 Refresh Rate

Another dimension for evaluating display quality lies in the time domain: *refresh rate*, the number of times in a second that a display hardware updates its buffer. Notice that this is different from *frame rate*, which refers to the number of consecutive images being captured by an imaging device in a second.

Human eyes perceive images by continuously integrating incoming photons. Although the exact underlying mechanism is very complicated, the common understanding is that most people do not detect flicker above 75 Hz. This is usually referred to as Critical Flicker Fusion (CFF) threshold, the frequency at which an intermittent light stimulus appears to be completely steady to the average human observer.

Modern LCD displays can operate at a frequency above 200 Hz. With traditional display applications where there is very little difference between successive frames, this refresh rate is enough. However, as soon as we start considering new display applications that put unconventional content into successive frames, there will be cases where a higher refresh rate is desired and standard CFF breaks.

1.1.5 Why Computation is Important

The growing demand for consuming digital information gives birth to a variety of novel displays. However, it also presents one fundamental challenge in display technologies: How to improve display output bandwidth. For example, how to reproduce the dense light field of a 3D scene; how to deliver images in high contrast, vivid colors; how to increase display frame rate so that unprecedented viewing applications are possible. On the pathway to overcome these obstacles, computation plays an important role in connecting available resources to push forward what displays are capable of showing.

Heterogeneous display hardware is invented to increase display output bandwidth across different aspects such as field of view, angular resolution and dynamic range. Some examples are listed in Figure 1.9.

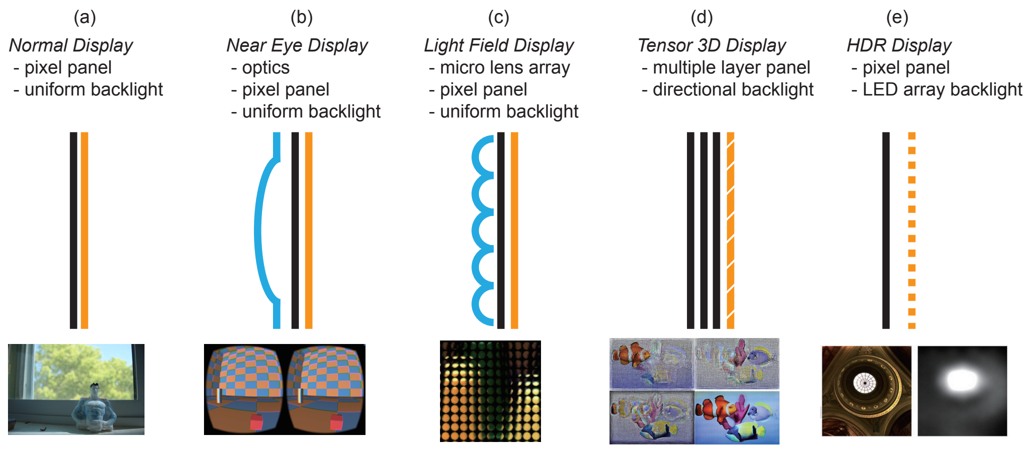


Figure 1.9: Different display hardware designs: (a) straight forward mapping on normal displays; (b) on near eye displays, images are warped to include wide field of view; (c) directional light rays are encoded for light field displays [73]; (d) a light field is decomposed into multiple frames to be shown on a tensor display [105]; (e) High Dynamic Range display utilizes modulated LEDs to achieve improvement in contrast [84].

Unlike traditional displays that share a common model of a single layer, 2D grid of pixels, emerging computational displays vary substantially in terms of hardware designs. This imposes challenges on display calibration, content encoding / decoding, and content retargeting.

Some display designs require the content to be decomposed into different layers, for example the designs in Figure 1.9 (d),(e). The processing step will be very challenging if the application requires real-time and content dependent decomposition. Even if the processing can be done off-line, a huge amount of memory space will be used up for calibration and content retargeting purposes. Computation is in strong demand for these displays to be widely adopted in real world applications.

Another way to improve display bandwidth is through viewer tracking. Figure 1.10 shows that viewer information could be included in the loop to update the content for each video frame.

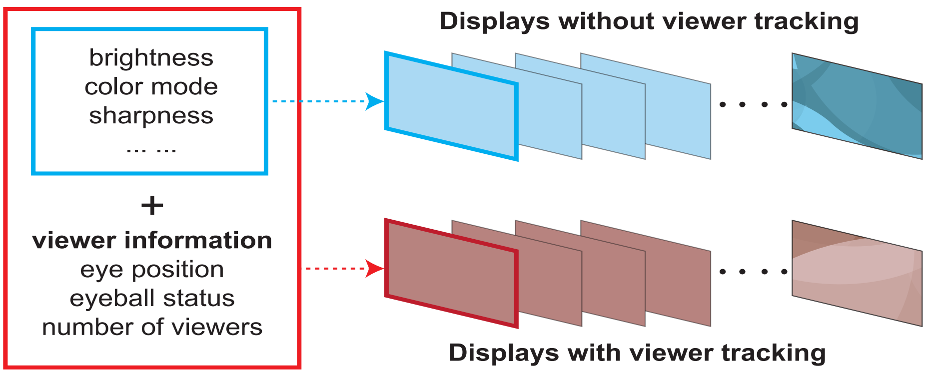


Figure 1.10: Traditional displays tend to be ignorant about viewer information when preparing the content for each frame in a stream. In emerging displays applications, each frame is often calculated based on tracked viewer information in real-time, as well as other parameters used by most displays.

Take 3D display as an example. Currently what largely limits the performance of 3D displays is the inability to produce small display pixels with high angular resolution. If the display updates its pixels as the viewer moves, it is essentially increasing the display bandwidth in angular domain. Head mounted displays are similar in the sense that displays are attached to the eyes. Fast and precise tracking of viewers' eye positions are essential for rendering convincing 3D scene in these applications.

In this dissertation, I first present two projects that utilize viewer tracking to improve glasses-free 3D display quality. In a third project, I propose a method to calibrate HDR displays while balancing the memory and computation usage. The last project studies flicker perception when image frames are created for computational display hardware designs.

1.2 Contribution

1.2.1 Dynamic Mapping for Multiview Autostereoscopic Displays

In Chapter 2, we introduce a project that addresses several image artifacts when viewing multiview autostereoscopic displays. On commercial multiview autostereoscopic displays, crosstalk between adjacent views is often severe, stereo inversion occurs at some head positions, and legacy 2-view content is difficult to display correctly. We introduce a method for driving multiview displays, dynamically assigning views to hardware display zones, based on potentially multiple observers current head positions. Rather than using a static one-to-one mapping of views to zones, the mapping is updated in real time, with some views replicated on multiple zones, and some zones left blank. Quantitative and visual evaluation demonstrates that this method substantially reduces crosstalk.

1.2.2 Projector Array with High Angular Resolution

As has been discussed in Section 1.1.4, angular resolution on most autostereoscopic displays is still far from desirable in terms of delivering convincing 3D imageries.

In Chapter 3, the "Projector Array" setup is introduced to explore the display design space that realizes high angular resolution over a large field of view. We demonstrate this technique using a dense horizontal array of pico-projectors aimed at an anisotropic vertical diffusion screen, yielding 1.5 degree angular resolution over 110 degree field of view. To create a seamless viewing experience for multiple viewers, we track the viewers and smoothly interpolate the set of viewer heights and distances on a per-vertex basis across the array's field of view, reducing image distortion, cross talk, and artifacts from tracking errors.

1.2.3 Chromatic Calibration of an HDR Display

High dynamic range (HDR) display prototypes have been built and used for scientific studies for nearly a decade, and they are now on the verge of entering the consumer market. However, problems remain regarding the accurate color reproduction capabilities on these displays. In this project, we first characterize the image reproduction capability of a state-of-the-art HDR display through a set of measurements, and present a novel calibration method that takes into account the variation of the chrominance error over HDR display's wide luminance range. Our proposed 3D octree forest data structure for representing and querying the calibration function successfully addresses the challenges in calibrating HDR displays: (1) high computational complexity due to nonlinear chromatic distortions; (2) huge storage space demand for a look-up table. We show that our method achieves high color reproduction accuracy through both objective metrics and a controlled subjective study.

1.2.4 Flicker Studies on Displays with a Hidden Channel

In Chapter 5, we investigate flicker perception on computational displays with temporally encoded hidden channels. The emergence of high frame rate computational displays has created an opportunity for viewing experiences impossible on traditional displays. These displays can create views personalized to multiple users, encode hidden messages, or even decompose and encode a targeted light field to create glasses-free 3D views.

Yet as these displays break new ground in functionality, they also bring complex display patterns that have never before appeared on traditional displays. Commonly accepted standards for traditional displays might no longer apply to these new displays. For example, under what conditions do viewers no longer perceive flicker?

We performed user studies to investigate the necessary frame rate for flicker fusion on this type of display and then built a prototype 3D+2D TV configured according to our user studies, on which users reported no flicker.