ADVANCED METHODS FOR CONTROLLING DUAL MODULATION DISPLAY SYSTEMS

by

Robin Atkins

B.ASc., Queen’s University, 2002

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

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2012

© Robin Atkins, 2012
Abstract

This thesis presents a novel method for controlling a dual-modulation display, also commonly known as a local dimming display. Dual modulation is a technology that improves the contrast and power efficiency of liquid crystal displays (LCDs) by dimming the backlight in image regions that need less light. This is an important improvement as although LCD technology accounts for nearly 90% of today’s displays, it has relatively poor performance in contrast and efficiency.

A critical component of a dual modulation display is the control algorithm. Present control algorithms cause an image artifact termed LCD clipping that affects the high spatial frequencies and is highly objectionable to many viewers. In this thesis we introduce an image metric designed to measure this artifact, as we found that existing metrics were not sufficient. The main contribution of the thesis is a new control algorithm for dual-modulation displays that eliminates the LCD clipping artifact, with minimal other tradeoffs in image quality and power efficiency. The new control algorithm requires less computational resources than previous algorithms and no change to display hardware, making it a relatively straightforward upgrade for today’s dual modulation displays.
Table of Contents

Abstract .................................................................................................................................... ii
Table of Contents ................................................................................................................... iii
List of Tables ........................................................................................................................... v
List of Figures ........................................................................................................................... vi
Dedication ............................................................................................................................. viii

Chapter 1: Introduction ........................................................................................................ 1
  1.1 The importance of contrast in displays ..................................................................... 2
  1.2 How much contrast can be perceived by the visual system? ................................. 3
  1.3 High contrast display technologies ....................................................................... 4

Chapter 2: Dual Modulation Display Technology .............................................................. 6
  2.1 Introduction ............................................................................................................... 6
  2.2 Contrast limitations in LCD panels ........................................................................... 7
  2.3 Contrast in spatially modulated backlights ............................................................... 8
  2.4 Power efficiency for spatially modulated backlights .............................................. 10
  2.5 Spatially modulated backlight algorithms .............................................................. 12
    2.5.1 Brightside’s approach to dual modulation displays ........................................... 14
    2.5.2 Dolby’s approach to dual modulation displays ................................................ 19
    2.5.3 Other relevant work on dual modulation displays ............................................ 21

Chapter 3: Drawbacks of Dual Modulation Displays .......................................................... 24
  3.1 Introduction ............................................................................................................... 24
  3.2 LCD clipping artifact ............................................................................................... 24
List of Tables

Table 1: Algorithm comparison results................................................................................... 62
List of Figures

Figure 1: Previous dual modulation algorithm ................................................................. 13
Figure 2: Reference image showing region of interest and cross section ......................... 27
Figure 3: Cross section at y-coordinate=540 of reference image illustrating LCD clipping. 28
Figure 4: An enlarged region of Figure 2 showing loss of texture in LCD correction signal 29
Figure 5: An enlarged region of Figure 2 showing the texture of the original image ........... 29
Figure 6: Local texture $\sigma_{\text{ref}}(x,y)$ of reference image ............................................. 35
Figure 7: Local texture $\sigma_{\text{test}}(x,y)$ of LCD signal ....................................................... 36
Figure 8: Ratio of texture $\sigma_{\text{test}}(x,y) / \sigma_{\text{ref}}(x,y)$ in test image to reference image ........ 36
Figure 9: Image regions where texture ratio $\sigma_{\text{test}}(x,y) / \sigma_{\text{ref}}(x,y)$ is reduced to less than 10% 37
Figure 10: Reference image showing cross section .......................................................... 38
Figure 11: Output of SSIM error metric for clipped regions ............................................. 39
Figure 12: Output of LCD clipping metric for clipped regions ......................................... 39
Figure 13: Cross-section illustration of SSIM and LCD clipping error metrics .................... 40
Figure 14: New dual modulation algorithm ....................................................................... 42
Figure 15: Target image (left) and monochrome low frequency base layer (right) .......... 45
Figure 16: Detail layer (left) and tone-mapped detail layer (right) ..................................... 47
Figure 17: Tone curve calculated for flower image detail layer ........................................ 47
Figure 18: New base layer ............................................................................................... 48
Figure 19: Backlight drive values .................................................................................. 48
Figure 20: Display model used for algorithm evaluation ................................................ 49
Figure 21: Metrics used for algorithm evaluation ............................................................ 50
Figure 22: Backlight design example ............................................................................... 51
Figure 23: Global backlight algorithm with greyscale image

Figure 24: Independent LCD/BLU algorithm with greyscale image

Figure 25: Independent square root LCD/BLU algorithm with greyscale image

Figure 26: Brightside algorithm with greyscale image

Figure 27: Independent square root algorithm with complex colors image

Figure 28: Brightside algorithm with complex colors image

Figure 29: LCD image from Brightside algorithm showing clipped regions

Figure 30: Iterative refinement with complex colors image

Figure 31: Iterative refinement with grey ramp image

Figure 32: Proposed algorithm with grey ramp image

Figure 33: Proposed algorithm with complex colors image
For my lovely wife Laura,

who believed in me and helped me though this,

and who is carrying our two boys

whom I can’t wait to meet
Chapter 1:

Introduction

In this thesis we present a novel method of controlling a dual-modulation display, also commonly known as a local dimming display. Our method improves image quality by minimizing an artifact called LCD clipping, which is a by-product of present control methods of dual modulation displays. The LCD clipping artifact is a highly noticeable loss of image detail at high spatial frequencies. Our method eliminates this artifact but increases the error in the low spatial frequencies. For some applications, this will result in improved image quality as the human visual system is less sensitive to errors at very low spatial frequencies. Our method also requires less computational resources than previous approaches, making it applicable for low-end display devices and high-end displays alike.

The thesis is organized as follows:

- We start by reviewing how much contrast can be perceived by the human visual system and the importance this plays in our perception of our world. This lays the groundwork for the motivation for the dual-modulation display technology.

- Chapter 2 describes the cause of the limited contrast in conventional LCD technology. It also describes how dual-modulation displays are able to bridge the gap between human vision and displays, by providing additional contrast (while increasing power efficiency). We review some exemplary previous work in control algorithms for dual-modulation algorithms.
• In Chapter 3, we introduce the two primary drawbacks of present control algorithms for dual modulation displays. The first is an image artifact called LCD clipping and the second is the computational resources required by the control algorithm.

• Chapter 4 introduces the image quality metrics that we will use to evaluate our results. We introduce a customized metric specifically tailored for measuring the LCD clipping artifact discussed in the previous chapter.

• Next, in Chapter 5 we introduce a new method for controlling a dual-modulation display that avoids LCD clipping. We discuss the evaluation strategy and describe our simulated display configuration.

• Chapter 6 presents the results of our simulations and the comparisons between present methods and our new method, for both test patterns and real images.

• Finally, in Chapter 7 we conclude with a discussion of the results and implications, and suggest some directions for future work.

1.1 The importance of contrast in displays

Liquid Crystal Display (LCD) devices have become ubiquitous amongst consumer as well as professional markets. They are capable of spatial resolutions approaching that of human vision and very high frame rates. After years of development, they have acquired excellent lifetime characteristics and have become relatively inexpensive to manufacture. Despite these advantages over competing technologies, they suffer one major disadvantage – that of contrast.
Contrast is the ratio between bright and dark in an image, and can also be termed dynamic range. Contrast is essential to human vision. The visual system initially processes scenes according to the spatial differences in illumination [1]. These spatial differences are the result of contrast in the scene. The absolute light levels are largely discarded from further visual processing in favor of the contrast information [2]. This visual dependence on contrast means that one of the most important features of a display is an excellent contrast ratio. State-of-the-art LCD panels are able to reproduce contrast ratios of ~2000:1, meaning the darkest “black” is 0.05% as bright as the brightest “white”. While this is well within the range of producing a comprehensible image, it is still well below the capabilities of human vision.

1.2 How much contrast can be perceived by the visual system?

Several studies have attempted to measure the contrast that a human being is able to perceive. The most relevant study was performed by Kunkel et al. who determined that we can simultaneously perceive 3.7 log orders of magnitude – i.e. contrast ratios near 5,000:1 [3]. However, this simultaneous contrast ratio assumes an observer with a fixed adaptation, which is not realistic for many viewing situations in which we do adapt.

The human visual system adapts to local regions within our field of view. We have both a mechanical mechanism to adapt (that of pupil constriction and dilation) as well as a chemical mechanism, which can be thought of as a “gain adjustment” on photoreceptors in the eye. These mechanisms allow us to perceive a much higher dynamic range (contrast) than that found by Kunkel et al [4]. For instance, a typical outdoor scene may easily have a dynamic range of 6 orders of magnitude, or 1,000,000:1 contrast. This is far in excess of the
capabilities of state of the art LCD panels, but can be easily scanned with human vision. In order to produce a visual scene with the dynamic range of human vision, a display device should have at least 5,000:1 contrast for small, local regions, and 1,000,000:1 or greater contrast for large spatial regions.

1.3 High contrast display technologies

In this thesis we focus on improvements to liquid crystal displays (LCD), due to their substantial dominance in the market. LCD technology is forecast to account for more than 88% of all flat panel displays in 2012, and more than 95% by 2014 [5]. Of these LCD displays, 67% are forecast to employ LED (Light Emitting Diode) backlights in 2012. The competition to LCD is mostly focused around increased contrast and improved power efficiency. The dual-modulation display technology discussed in Chapter 2 addresses these two limitations of LCD technology by improving the contrast and power efficiency of LCD displays.

After LCD, plasma displays have the second largest market share. In a plasma display, each pixel is a tiny fluorescent lamp with controllable intensity. Theoretically, these displays are capable of very high contrast ratios, as they produce light only where needed in an image. However, due to their high screen reflectivity and control stability, the practical contrast is limited to well below the capabilities of a modern LCD, for most viewing conditions. Plasma technology is expected to decline to below 4% of the market in 2014.
Organic light-emitting diode (OLED) displays also have the potential for extremely good contrast and efficiency. OLED’s are just LED’s where the substrate is an organic compound rather than silicone, simplifying the manufacturing process. Each pixel is a tiny LED light source. Like plasma, this leads to very high *theoretical* contrast, but poor *practical* contrast for the same reasons – high screen reflectivity and poor thermal stability. OLED technologies are expected to grow slowly to 2.5% of the market by 2015. Of primary interest to display makers is the potential for low power and high contrast performance.
Chapter 2:
Dual Modulation Display Technology

2.1 Introduction

Liquid crystal displays create images by transmitting a portion of the light produced by a light source behind the panel. This light source is termed the backlight, and is typically produced by a cold-cathode fluorescent light (CCFL) or more recently light emitting diodes (LEDs) [6] [7] [8]. The light produced contains considerable energy in each of the red, green, and blue portions of the visible spectrum and as little energy as possible in the ultraviolet and infrared regions. The light is diffused spatially to obtain a light field with uniform intensity. The diffused light then passes through a polarizer. This ideally transmits only one polarization of the incident light, for example the transversal component. This comprises roughly 50% of the incident light. Many displays use a reflective polarizer which reflects the undesired polarization. The polarization of the reflected light is then randomized by the diffuser and directed back through the polarizer a second time. This mechanism is termed light “recycling”, and is used to increase the efficiency of the system.

Next, the light passes through an array of color filters. Most panels use a repeating pattern of red, green, and blue band-pass filters which pass a spectral band according to their central wavelength. Light outside of the spectral band is absorbed. Thus, the red filter passes red light and absorbs blue and green. Each group of red, green, and blue filters comprises a pixel
on the LCD. For high definition television (HDTV) there are typically 1920x1080 pixels. The individual red, green, and blue color components are termed sub-pixels.

The polarized and colored light of each sub-pixel then passes through a liquid crystal (LC) layer [9]. This crystal rotates the polarization of light according to an electrical signal. With no electrical signal, the transversal component passes through the layer unchanged. With a “full” electrical signal the polarization is rotated by 90 degrees so that it is now longitudinal. The change from zero rotation to 90 degrees occurs very quickly, on the order of 2-4ms for modern panels [10]. Finally, the light passes through a second polarizer termed the analyzer. The analyzer’s function is to transmit only the longitudinal component of the polarized light and absorb the transversal component.

Thus, transversal light that is not rotated by the liquid crystal is absorbed by the analyzer and light that is completely rotated by the LC to longitudinal light is transmitted through the analyzer. The LC layer can partially rotate the polarization so that a portion of the light is transmitted through the analyzer. The precision of the rotation is determined by the digital drive electronics, which are typically able to rotate the LC by 8bit (256 steps) and up to 10bit (1024 steps) for modern panels. Each pixel in an 8bit panel can therefore produce 256 discrete intensities of red, green, and blue light for a total of ~17 million combinations.

2.2 Contrast limitations in LCD panels

The system described above is an ideal LCD, where each sub-pixel can completely transmit or completely block light. Since contrast is measured by dividing maximum luminance by
minimum luminance, and the minimum luminance of an ideal LCD is zero, the contrast of an ideal LCD would be infinite. An infinite contrast ratio is well beyond the capabilities of human perception. However, there are many factors that contribute to an LCD being a non-ideal system. The major contributing factor is that the polarizers are imperfect, largely due to cost and manufacturing constraints. Instead of transmitting only a single polarization and reflecting or absorbing the other, the polarizers instead also transmit a portion of the undesired polarization. If the first polarizer transmits even 0.5% of the undesired polarization, this is largely transmitted through the analyzer, and thus raises the minimum luminance from zero to a small non-zero number. This in turn results in less contrast. The amount of undesired light that passes through the LCD is termed “light leakage”. As described earlier, a typical LCD panel has 0.05% light leakage, resulting in a contrast ratio of 2000:1 [9].

2.3 Contrast in spatially modulated backlights

Light leakage can be controlled by the quality of the polarizers and optics, but also by the light produced by the backlight. For a given LCD, the less light that is produced in a region of the backlight, the less light leakage through the LCD in that region, thus the contrast or dynamic range of the display can be greatly enhanced. This is the approach taken for displays with a spatially modulated backlight, termed “local dimming” by the industry. Such displays were first developed by Brightside Technologies, a spinoff from the University of British Columbia, in 2004 [11]. This design used an array of thousands of LED’s each with individual control to comprise the backlight. In bright regions of the image the LEDs are turned on to full intensity, and in dark areas of the image the LEDs can be turned completely
off. This results in much less light leakage in the dark regions of the backlight, producing very little undesired light and thus very high contrast ratios approaching the capability of human vision. Since the introduction of the Brightside display, many of the major television manufacturers have adopted the technology and there are now dozens of local dimming displays from brands like Sony, Toshiba, LG, and Samsung.

The LED array used as the backlight in a dual modulation display has a much lower resolution than the LCD pixels, due to the high cost of LEDs and manufacturing constraints. To achieve uniform brightness across the screen, the optical design of the backlight ensures that the light produced by each LED is smooth and continuous by spreading it laterally. The lateral spread is termed the point spread function (PSF). The net result of the optically smooth, low resolution backlight is that the backlight for neighboring LCD pixels is nearly uniform. Thus the contrast ratio of adjacent pixels is limited by the native contrast of the LCD. Some very high end panels are able to achieve native contrast ratios of near 4000:1, approaching the limits of human vision found by Kunkel et al. Across larger regions, however, the backlight can be dimmed to completely off. This allows for nearly complete black, resulting in image contrast approaching human perception. Such displays can produce higher dynamic range than can be perceived simultaneously, but as a viewer’s eyes roam around the image they are able to adapt to local regions just as they do in real life, creating a viewing experience mimicking that of reality [12].

However, there is a drawback to the low spatial resolution of the LEDs comprising the backlight. The backlight is not always able to produce the required amount of light for a
region of a given image. This is especially true near high contrast edges. Consider, for example, a boundary between full white and full black. On the white side of the boundary the backlight is turned fully on to produce the desired light level. On the black side of the boundary the backlight is turned fully off to produce black. However, right on the boundary there is a gradient in the light field produced by the diffuse backlight (from full white to full black). Depending on the design of the backlight this gradient can span from tens to hundreds of LCD pixels. The LCD signal is simply full white on the white side of the border and full black on the black side in order to maximize the contrast ratio. On the white side of the border there may be slightly less light than is required to produce the target image, and on the black side of the border there may be slightly more light than can be blocked by the LCD. The resulting image thus shows a dark halo on the white side of the boundary and a bright halo on the black side of the boundary. A control algorithm is included in the display that determines the corresponding signals to control both the backlight and the LCD in order to best reproduce the target image. Although the physical hardware is approaching the capabilities of human vision, the algorithm can typically introduce unnecessary artifacts that degrade the image quality. These artifacts and their mitigation techniques will be discussed further in the remainder of the thesis.

2.4 Power efficiency for spatially modulated backlights

Besides contrast, power efficiency is a major concern with conventional LCD technology. With new power regulations being introduced in the US, Europe, and Asia to limit the power consumption of home televisions [13], [14], it could be the major driving factor behind mass adoption of local dimming displays or alternate display technologies. Thus it is critical that
any new algorithm for controlling a dual modulation display uses the same or less power than previous methods.

A television is usually considered an emissive device, since the total light from each pixel is a sum of the red, green, and blue sub pixels. However, to evaluate power efficiency, we will instead consider the LCD panel component alone as a subtractive device, like negative film. Light is subtracted from the backlight to create the desired image. The efficiency of the system is calculated from the ratio of light produced by the backlight to how much desired light is transmitted through the system. For a conventional LCD display, the first polarizer can transmit approximately 70% of the unpolarized light, assuming a high efficiency of light recycling. The color filters can transmit roughly 25% of the light, and the LC layer 50%. The analyzer transmits 80% of the desired light. The resulting system efficiency is the product of each of the layer efficiencies, in this example ~7%.

Local dimming displays have the potential to greatly increase the light efficiency of a display device. By producing light only where needed in the backlight, the amount of light that the LCD is required to block (absorb) can be minimized. Since the LCD is most efficient when it is not blocking any light, we can say that a design goal for an efficient local dimming display is to drive the LCD towards fully open whenever possible.

To measure the efficiency of a dual modulation display, we compare the power consumption to the power required for a global backlight. The ratio indicates how much of an improvement has been made. For example, for a certain image, the average backlight drive
value for a local dimming display is 0.25 (out of 1). The dual modulation display is only using 25% of the power of a global backlight, meaning it has 4x higher efficiency than a conventional display.

2.5 Spatially modulated backlight algorithms

The first work that was published regarding dual modulation displays was by H. Seetzen et al in 2003 [6]. They proposed the fundamental dual modulation algorithm described below. This work then went on to inspire the backlight determination algorithm used by Brightside technologies in their local dimming displays circa 2004 [15]. In 2007, Dolby Labs purchased Brightside and a few improvements were made to the backlight determination algorithm in the following years. These are mostly confidential, but we briefly review a few changes that have been published externally. Much other work has presumably been done by display makers, but the details of their techniques are proprietary and too recent to have been made available publicly through the patent process. More recent work by Guarnieri et al also follows the same approach [16].
A spatially modulated backlight algorithm generally proceeds as follows, shown in Figure 1:

1) Determine the backlight drive values from the target input image
2) Simulate the light field produced by the backlight
3) Determine the LCD drive values to compensate for the difference between the backlight light field and the input image

![Figure 1: Previous dual modulation algorithm](image)

What works well with this approach is the LCD image is essentially a high-resolution error correction signal for the backlight. The closer the backlight is to the target image, the less the LCD image has to do. If the simulation in Step 2 is very accurate, then many types of target images can be reproduced nearly perfectly. The remainder of this chapter reviews previous dual-modulation display algorithms in more depth. Although dual modulation displays can result in a significant improvement to contrast, they typically come with two drawbacks, which will be discussed further in Chapter 3.
2.5.1 Brightside’s approach to dual modulation displays

Mathew Trentacoste proposed a method for determining backlight drive values for the Brightside display [15], building upon the prior work done by Seetzen [6], [17]. Trentacoste’s research focused on how the backlight determination could be optimized to improve the resulting visual image quality. He started by describing a method using nonlinear optimization to determine the backlight drive values for a given image. This work has been extended since by others [18]. For each iteration the backlight light field is simulated by convolving the backlight point spread function with the drive values for each element and then multiplying by the LCD pixel intensity:

\[ I(p, \delta) = p(PSF_D \ast \delta_D) \]  

where \( \delta_D \) is the drive value for the backlight

\( PSF_D \) is the point spread function (PSF) for each backlight element

\( p \) is the LCD pixel intensity

Trentacoste then computes the error between the reproduced image \( I(p, \delta) \) with the original \( I \), and uses this to refine each iteration. To improve results, he first filtered both original image \( I \) and reproduced image \( I(p, \delta) \) using a model of human vision that accounts for our limited contrast sensitivity at sharp edges, as well as our non-linear sensitivity to lightness.

\[ \varphi(I) = L(PSF_e(Y_{avg}) \ast I) \]  

where \( PSF_e(Y_{avg}) \ast I \) is a model of blur in the human eye caused by optical scatter

\( L(I) \) models the non-linear response of the human eye

For each iteration, a new set of backlight drive values \( \delta \) are computed from the previous values and the error from the previous iteration. The backlight is then again simulated by the
convolution of the backlight PSF with the backlight drive values. The LCD pixel values are computed by dividing the original image by the simulated backlight in that area.

\[ p = \frac{I}{B} = \frac{I}{(PSF_D \ast \delta_D)} \]  

(3)

where \( I \) is the original target image

\( B \) is the simulated backlight \((PSF_D \ast \delta_D)\)

Substituting equation 3 into equation 1 results in \( I(p, \delta) = I \), which is the desired outcome. However, the solution is not that simple because \( p \) is constrained to possible values, namely clamped to [0,1] and then quantized to 8 bits, representing fully closed or fully open LCD pixels. This allows the iterative solver to correctly model areas where the backlight produces too little light \((B < I)\), which results in values of \( I/B \) greater than one which are then clipped, introducing errors in the final image. This model can also account for quantization of LCD pixels, for example when very small values are rounded to zero or one. This occurs when the backlight produces much too much light \((I >> B)\), which results in very small values of \( p \) which are rounded, again introducing errors in the final image.

This model does not account for errors introduced by light leakage in the display. This is largely due to the assumption that any light leakage is not visible since its magnitude is below the magnitude of light scattering in the human eye, termed veiling glare, and modeled by equation 2. While this is a valid assumption for some images on the Brightside display, it does not hold true for other patterns. One such example is a single white pixel on a black background. The single white pixel requires a significant number of backlight elements to be active in order to generate the desired level of illumination on the backlight. Since the backlight has a low spatial resolution, there is a significant amount of light leakage through
the black LCD pixels which is clearly visible as a halo surrounding the white pixel. Veiling glare does not hide the surrounding halo because the average intensity $Y_{avg}$ caused by the single pixel is very small, resulting in minimal light scatter in the eye.

To improve the performance of the iterative solver, Trentacoste proposed to first condition the input image to one more easily achievable by the low spatial resolution backlight. Thus, instead of using the input image $I$ as the target, he instead uses the preconditioned image $\tilde{B}$, the target backlight. The preconditioning involves three steps.

First the color is removed from the image by taking the maximum of the RGB color components. Using the maximum rather than some weighted average ensures that there is enough light produced by the backlight for each of the color channels.

The second step is to take the square root (or some other exponent) of the intensity image. This is done to allocate a portion of the dynamic range to each of the two modulators. A square root allocates the dynamic range evenly between the backlight and the LCD. A power of one allocates all of the dynamic range to the backlight. Using a power less than one reduces the dynamic range of the target image, supposedly to something more achievable by the backlight. However, this step has the effect of raising the average intensity of the target image, which increases the power used to produce the image. It also puts some responsibility for the dynamic range of the image on the LCD, which limits its ability to correct for imperfections in the backlight. The result is that the backlight is more likely to produce more light than necessary, and more light than the LCD is able to block completely. A simple
example is a full-screen gray image. The square root results in a target backlight of 71% intensity, rather than the 50% intensity that is actually needed. The LCD is responsible for blocking light to reduce it to the desired intensity.

The third step is to down-sample the image to a lower resolution using a low-pass filter. This takes advantage of the low spatial resolution of the LEDs, which are not able to produce a high frequency signal. However, this step further reduces the dynamic range of the target image. The intensity of small bright features is reduced, and the intensity of small dark features is raised. This again places an increased responsibility on the LCD for producing the image dynamic range. It also increases the likelihood that the backlight will produce too much or too little light for small features. For the example used earlier of the single white pixel on a black background, the low pass down-sample greatly reduces the intensity of the single pixel for the target backlight image. The resulting backlight produces too little light, and the LCD pixel must be clipped to fully open to compensate. The resulting image is both too dim and has lost any detailed texture.

Performing the iterative non-linear optimizing routine is too complex to perform in real-time within a consumer electronics device, largely due to the requirements of the 2D convolution of backlight drive values with the backlight PSF. This operation cannot be performed more than once per frame in current hardware, and there is a desire to skip it altogether. However, without the simulation of the backlight the LCD correction signal cannot be determined accurately. Instead of relying upon the iterative optimizer as the algorithm, instead we need
to consider this the optimal solution for evaluating approximations that can be done in real-time, requiring a maximum of one simulation of the backlight light field per frame.

Trentacoste proposes a simple solution to this, which accounts for the predicted drive values for surrounding backlight elements. This is done by subtracting the predicted contribution of light from the surrounding target backlight region by the backlight PSF for each backlight element.

\[
\delta_j = B_j - \sum_{i} N(s_j) w_{ji} B_i
\]

\[ w_{ji} \]

where  
\( B_j \) is the target backlight in the region of the backlight element \( j \)  
\( w_{ji} \) is the point spread function of the backlight element \( j \)  
\( B_i \) is the target backlight in the region of each neighboring backlight element \( i \)  
\( w_{ji} \) is the point spread function of each neighboring backlight element \( i \)

The net result of this solver is a sharpening of the backlight drive values. This helps to counter for the reduction of dynamic range introduced in previous steps. If a neighboring region has a high intensity, the solution predicts that the corresponding backlight element will produce a lot of light, some of which will spread into the immediate region, thus requiring a lower drive value than otherwise needed. Conversely, if a neighboring region is dark, relatively more light will be needed from the immediate region to make up for the difference. The advantage of this solution is that it is performed in a single step without iterating, thus can be implemented without great difficulty in real-time. However, the cost is an imperfect solution, as for many cases the solution will be a negative or greater than one drive value. Like the LCD drive values, the resulting backlight drive value \( d_j \) are constrained
to [0,1] to account for this. Instead of clamping the immediate backlight drive value, it would be optimal to propagate the difference to surrounding backlight elements in some stable manner.

The sharpening filter also has a negative impact on the stability of the backlight for video sequences. Continuing our previous example, as a bright single pixel moves across a black screen, this solution causes only the backlight element in the immediate vicinity to be active, and the surrounding elements to be off. The object can move smoothly by single pixel increments. However, the backlight is constrained by its lower resolution, so cannot move smoothly to illuminate the single pixel. Instead, as the single pixel transitions from one backlight region to another, the backlight elements across the transition switch from full off to full on in the space of a single image frame. The light from the backlight that leaks through the LCD also transitions sharply in this frame, so the halo that was centered on one backlight element in a first frame is now centered on the second backlight element – not the object itself. This discontinuous motion of the halo is a highly objectionable visible artifact, especially as the center of the halo is not centered on the single pixel. It would be more desirable to have a low-pass spatial filter on the backlight than a high-pass filter, to add some measure of temporal stability.

2.5.2 Dolby’s approach to dual modulation displays

Dolby’s approach to the local dimming algorithm builds upon Trentacoste’s work, addressing several of the issues identified above [19]. Two of the changes were in the way that the
original image $I$ is preconditioned to the target backlight $\bar{B}$ in advance of the backlight determination.

The first was to replace the square root in the image preconditioning with a power of one. This shifts the full responsibility for image dynamic range to the backlight, while reducing the power required by the backlight. The full range of the LCD panel becomes available to correct for backlight errors, rather than having to correct for errors as well as add dynamic range to the image.

The second was to use a weighted combination of maximum and mean down-sample filters rather than just a mean down-sample. This better preserves the intensities of small bright features such as the single pixel case, helping to ensure that the backlight produces enough light for small objects. However, the effect of this is to bias the target backlight even brighter, reducing the dynamic range of the backlight and increasing power consumption. Since the maximum is only partially weighted, this approach does not ensure that enough light is made available for small bright regions, it just improves the level over using the average alone.

The Dolby approach also eliminated the sharpening solution to the backlight drive values, instead replacing it with a low pass filter to smooth the backlight light field, which helps to mitigate the highly objectionable temporal instability for moving objects. However, this also reduces the dynamic range of the backlight, as peaks are brought down and darks brought up. This reduction partially counters the increased dynamic range from removing the square root.
division of dynamic range. Adjusting the values in this way also increases the chances of there being insufficient light produced for bright image regions and too much light produced for dark regions. The Dolby approach still does not account for light leakage through the display caused by too much light in the backlight, and it does not attempt to propagate a difference in backlight intensities to neighboring elements. The result is that the backlight can still produce too much or too little in any area of the LCD which cannot be compensated for by the LCD, causing clipping and loss of texture detail.

### 2.5.3 Other relevant work on dual modulation displays

More recently, Guarnieri et al investigated how to determine a backlight control signal with the additional constraint of minimizing parallax error [20], [21], [16]. Parallax occurs when two objects at different depth are displaced from each other, and is the basis for stereoscopic vision. Parallax error occurs when two objects which are supposed to be at the same location appear displaced. The parallax error that Gaurnieri et al were investigating occurs when viewing a dual-modulation display off-axis, causing the images created by the two modulators to be misaligned, since they are separated by a non-zero depth. When viewing off-axis, the backlight appears spatially displaced, so that the light field is incorrectly positioned with respect to the front modulator. This was of particular importance since the backlight in his display had very high spatial resolution. For low resolution backlights, slight spatial displacements do not have a noticeable effect, which greatly mitigates the problem of parallax. To account for parallax, they found that he had to digitally blur the high-resolution backlight, mimicking the optical blur that occurs with a low resolution backlight. The
algorithm he uses is based on the previous work by Trentacoste, and is identical to the block diagram of Figure 1.

Guarnieri et al [21] introduced a new constraint to the optimization objective function, improving upon that of previous work [6]. Their solution accounts for the limited dynamic range of each modulator. Rather than constraining each modulator to [0,1] in the model, this approach constrains them to \([1/d, 1]\), where \(d\) is the dynamic range or contrast of the modulator, and \(1/d\) is the light leakage.

\[
\frac{1}{d_b} \leq L_b(x, y) \leq 1
\]  
(5)

where \(L_b\) is the signal to the rear modulator (backlight)

\(d_b\) is the contrast of the rear modulator

Similarly, for the front modulator (LCD)

\[
\frac{1}{d_f} \leq L_f(x, y) \leq 1
\]  
(6)

where \(L_f\) is the signal to the front modulator (LCD)

\(d_f\) is the contrast of the front modulator

As in previous approaches, Guarnieri et al calculate the front modulator signal by dividing the target image \(I\) by the backlight image:

\[
L_f(x, y) = \frac{I(x, y)}{L_b(x, y)}
\]  
(7)

By substituting equation 7 into equation 6, we can re-write the constraints as:

\[
\frac{1}{d_f} \leq \frac{I(x, y)}{L_b(x, y)} \leq 1, \quad \Rightarrow \quad I(x, y) \leq L_b(x, y) \leq d_f I(x, y)
\]  
(8)
Combining the two constraints on our backlight from equations 5 and 8 yields the combined constraint for the backlight drive values in order to avoid clipping:

$$\max \left\{ I(x, y), \frac{1}{d_b} \right\} \leq L_b(x, y) \leq \min \{d_f I(x, y), 1\}$$  (9)

To minimize parallax, he then uses an optimization procedure to produce a backlight that is as smooth as possible while meeting the constraints above. This results in a constant backlight anywhere where the constraints are not met. Note that the constraints in equation 9 are not compatible when $I(x, y) < \frac{1}{d_b d_f}$. Guarnieri suggests filtering the input image to remove exceedingly dark pixels, or relaxing the constraints as needed to ensure a solution.

Unlike the display of Guarnieri et al, our dual modulation display does not have a high resolution backlight. Hence, the parallax issues do not play a role in our backlight solution, since our backlight is optically constrained to a smooth solution. We must also account for light contribution from neighbouring elements due to our low resolution backlight. Thus $1/d_b$ is very small for small spatial frequencies, but infinitely high for large spatial frequencies. Modifying this constraint to account for the light in the local region introduces a recursive element into the optimization, increasing the complexity to that of the method proposed by [15].
Chapter 3: Drawbacks of Dual Modulation Displays

3.1 Introduction

Previous control algorithms for dual modulation displays have two main drawbacks, both of which are presented in more detail in this chapter.

1) LCD clipping artifact: Dual modulation displays are not able to reconstruct high frequency and high contrast edges in the target image. For present control algorithms this results in clipped pixels, which are groups of pixels which have lost all texture or contrast information. This can make an image look “plastic” or “flat”.

2) Computational complexity: Simulating the light field (step 2) has to be done accurately, which requires heavy computation and memory requirements in addition to accurate models and measurements of the display.

3.2 LCD clipping artifact

The dual modulation algorithms described in the previous chapter each determine the LCD drive values for a given target image and simulated backlight light field. As we saw repeatedly in the previous chapter, this is typically done by dividing the target image by the light field for each pixel:

\[
\text{LCD} = \frac{\text{Target Image}}{\text{Light Field}} \quad (10)
\]
When the light field matches the target image, the LCD is fully open (=1). When the target image is slightly less than the light field (such as $\frac{1}{2}$), the LCD is controlled so that the light transmitted is reduced by 50% to correct for the difference.

There are two situations when the LCD is unable to correct for a difference between the light field and the target image. The first situation is if there is less light produced by the light field than that required by the target image ($\text{Light Field} < \text{Target Image}$). In this case, the resulting LCD signal is greater than one. As the LCD is not physically able to produce light (this would be $>1$), the signal must be limited to less than or equal to one – termed clipping. Clipping removes contrast in an image. Consider the example of a car headlight at night that is bright and contains detailed structure of the lens. If there is not enough light in this region of the backlight the resulting image will be too dim and will lose the structure detail of the lens. Considering that contrast is essential to the human visual system, this loss of contrast is a very undesirable artifact. In addition to losing contrast information, image regions that are clipped will also be less bright than intended, since the desired LCD signal was actually an impossible light-producing signal to correct for insufficient light in the backlight. If this corresponds to a dark, barely visible portion of the image, this may render that portion invisible. If this corresponds to a bright portion of the image such as a sun, it may appear dimmer than desired, reducing the visual impact of the image.

The second situation where the LCD is not able to correct for the light field is where there is too much light produced by the backlight ($\text{Light Field} >> \text{Target Image}$). Due to light leakage through the LCD it may not be possible to adequately block the light produced by the
light field. In this case the desired LCD intensity to correct for the backlight can be zero or below the minimum digital drive value. Consider a dark area of an image, with target light level of only 0.1% of the maximum. If the light field in that region is very bright, for example 50% of maximum, then the LCD drive signal for that region will be 0.1/50, or 0.002%. For an 8bit LCD panel, this is less than the first code value. The result is that in this region, detailed texture information will be clipped to a black LCD value. It will also appear brighter than desired. Consider as an example a dark bird flying in front of a bright sky. In the target image the dark bird has detailed texture information for its feathers. However, since it is a small dark feature in front of the bright sky, the backlight light field is very bright in that region. As the LCD attempts to block the excess light produced by the backlight, all the detail in the bird’s feathers is clipped, and it appears as a dark grey due to light leakage.

In general, clipping of the LCD can be avoided by ensuring that the backlight light field is as close as possible to the target image, erring always on the “too bright” side. However, due to the low spatial resolution of the backlight, this is not possible for small image features and high contrast boundaries. Also, because the light at each pixel is the sum of the contribution of many surrounding backlight elements it is not always possible to balance between producing too much light for one region and not enough light for another region.
To illustrate the LCD clipping artifact we analyze the result of applying the Brightside dual modulation algorithm to an image shown in Figure 2 below. A grey line was inserted to illustrate the location of a cross section described below.

Figure 2: Reference image showing region of interest and cross section
We look at the cross section of the pixel values for the corresponding target image, light field simulation, and LCD correction. We can see from Figure 3 below that the target image signal (red curve) does not extend beyond 90% of maximum for this cross section. This brighter region is surrounded by very dark pixel values. The backlight in this region (green) is calculated to hover around 75%. The problem occurs when we try to compensate the LCD signal for the low backlight intensity in this brighter region. In this region the LCD signal (blue) is clipped to one. This region has lost all contrast information, greatly degrading the image.

Figure 3: Cross section at y-coordinate=540 of reference image illustrating LCD clipping.
Figure 4 below shows an enlargement of the clipped region in the LCD correction signal. Notice how all texture and contrast information has been lost. For comparison, Figure 5 shows the target image which has the detailed texture information.

![Figure 4: An enlarged region of Figure 2 showing loss of texture in LCD correction signal](image)

![Figure 5: An enlarged region of Figure 2 showing the texture of the original image](image)
3.3 Computational complexity

The light field simulation of step 2 from Figure 1 models the spread of light through the display in order to predict the physical light field that will be produced for a given set of backlight control values. The spread of light from each backlight element through the display optics is termed the point spread function (PSF). The simulation should be done as accurately as possible so that the LCD correction in step 3 can be performed accurately to achieve the target image. The approach generally taken for this is to perform a convolution of the backlight drive values with the measured or modeled PSF from each backlight element. This can be implemented in a number of ways within hardware or software by accumulating the results of multiple separable filters.

For most displays with locally modulated backlight, this step requires the majority of the computational resources due to the large size of the spatial filters required to model the small contribution of light from very distant backlight elements. The large spatial filters require buffering significant portions of the image into memory (nearly an entire frame) in addition to large footprints in FPGA architectures. This results in high cost for the chip and memory, as well as increased power and system latency. As the number of backlight elements or desired frame rates increase, this step requires greater and greater resources.

In addition to the computational cost, this approach requires very accurate models of the display optics and calibration. Small errors in the model or measurements can lead to large over or under-estimations of the light field, which then propagate to large errors in the LCD.
compensation. For example, a target image has a luminance intensity of light in a dark region of 0.1 cd/m². The light produced by the backlight in the region is measured at 0.11 cd/m², a perfectly good value for the region, as the LCD can block the extra 10% of luminance. However, due to inaccuracies of the light field simulation, the system models the light in the region as 0.09 cd/m². This results in the case where the \((\text{light field} < \text{target image})\), and causes the LCD to be fully open in the region, rather than partially closed. This results in the region being a) too bright, and b) clipped of any texture detail. The only way to avoid this situation is with highly accurate measurements, models, and computations. On a factory floor for mass production of displays, such measurements can be prohibitively time-consuming.
Chapter 4:

Image Quality Metrics

4.1 Introduction

Many image quality metrics are available for evaluating visibility of distortions to images. Referenced metrics compare an original image against a distorted version of the image. For our comparisons we will compare the target input image against the simulated image produced by a model of our local dimming display. We are interested in the particular distortions caused by dual modulation control algorithms: high-frequency distortions caused by LCD clipping as well as differences in brightness of image regions on the order of the LED backlight.

The most commonly used metrics in evaluating image processing are the mean-square error (MSE) and related peak signal to noise ratio (PSNR) [22]. These are widely used due to being common standard error metrics as well as being easy and fast to compute. However, over the last decade they have been shown to not match well with perceived visibility of image artifacts [22], and other metrics have been developed to improve performance.

A popular metric that has often been used for evaluating dual-modulation algorithms is Daly’s Visual Difference Predictor (VDP) [23], and later Mantiuk’s extension to High Dynamic Range, the HDR-VDP [24]. The VPD models low-level processes in the human visual system in order to determine visibility of artifacts. It operates on multiple spatial
frequencies, mimicking the human contrast sensitivity function (CSV). For a defined set of viewing conditions (distance, surround environment, etc), this can be a very accurate approach. However, it is slow (10-20 seconds per image), and is only valid for the viewing conditions defined. This includes the absolute light levels of the display and the viewing environment. For many applications, these conditions are not known precisely or are variable within a wide range.

Another suitable metric is the Structured Similarity Metric (SSIM) [25]. This follows similar reasoning as the VDP but is greatly simplified. It combines estimated distortions due to changes in luminance, contrast, and image structure (texture, etc) to generate a single number indicating the similarity between two input images. One indicates a perfect match, and less than one indicates some differences. It does not directly try to estimate the visibility of the differences, as this is dependent on viewing conditions. Instead, it is useful as a relative metric for comparing different image distortions. Like the VDP, this also includes a spatial component in the form of the contrast sensitivity function and has been extended to operate on multiple spatial scales via the extension MS-SSIM [26]. The advantages of this metric over the VDP is much faster processing time (<1s per frame in Matlab), and independence from the viewing environment. For our evaluations we will compare image performance using (1-mSSIM), or the mean of the SSIM error predictions scaled such that 0 is no error, and higher numbers indicate more error.
4.2 LCD clipping metric

In our tests we found that both the VDP and SSIM can detect the LCD clipping artifact. However, both of these metrics pool this texture loss with other fidelity measures, diluting the impact of small regions of clipped pixels. Since we are interested in specifically measuring the loss of high-frequency spatial detail associated with LCD clipping, we instead developed a metric specifically suited for this task – the LCD Clipping metric. The metric measures areas in an image where the texture has been reduced to below some threshold. This is not so much an entirely new metric as a modification to the existing SSIM metric to place more weight on the image structure component. It is also faster to compute than SSIM so is more suitable for finding iterative solutions.

The metric compares spatial texture between a reference and test image, and reports regions where detail has been lost in the test image. We use the standard deviation of a local 3x3 pixel region surrounding each pixel to measure image texture. This is executed in Matlab very quickly using the stdfilt command in the Image Processing Toolbox. This returns a 2D map corresponding to the local texture at each pixel of the image.

We evaluate the loss of texture by counting any pixels in the test image where the texture is reduced to below ten percent of the reference image. We divide the number of clipped pixels by the total number of image pixels to compute an average number of clipped pixels for an image. A more sophisticated approach could measure the size of clipped regions, but this was not explored in this thesis as we found the simple averaging to work well.
The procedure for executing the metric is as follows:

a) Compute local standard deviation \( \sigma_{\text{ref}}(x,y) \) of a 3x3 region of the reference image

b) Compute local standard deviation \( \sigma_{\text{test}}(x,y) \) of a 3x3 region of the test image

c) Count pixels where \( \sigma_{\text{test}}(x,y) < \sigma_{\text{ref}}(x,y) / X \)

Where \( X \) is the allowable loss of contrast, \( X=10 \) in our thesis

d) Divide count by total number of pixels in reference image

The first step is executed using the Matlab \texttt{stdfilt} command. The result is shown in Figure 6 below for an example reference image. Edges are clearly visible as expected from a local standard deviation.

![Figure 6: Local texture \( \sigma_{\text{ref}}(x,y) \) of reference image](image)
The second step is computed the same way for the test signal. Notice the change in texture on the circled area below in Figure 7.

![Image of local texture](image1.png)

**Figure 7: Local texture $\sigma_{test}(x,y)$ of LCD signal**

In the third step, by comparing the ratio of test and reference images we can find areas where the texture has been reduced beyond our desired threshold. Figure 8 below shows the ratio of texture between test and reference images. A low ratio indicates that texture has been decreased significantly.

![Image of ratio of texture](image2.png)

**Figure 8: Ratio of texture $\sigma_{test}(x,y)/\sigma_{ref}(x,y)$ in test image to reference image**
If we threshold the result in Figure 8 to regions where the ratio is less than 10% we obtain the resulting image in Figure 9.

![Image](image.png)

**Figure 9: Image regions where texture ratio $\sigma_{\text{test}}(x,y)/\sigma_{\text{ref}}(x,y)$ is reduced to less than 10%**

The final steps are to count the pixels where the texture has been reduced to below threshold, and divide by the total number of image pixels. This is executed in Matlab by computing the mean of the binary image. For the example used above in Figure 9 this works out to 1.6%. 
4.3 Evaluation of clipping metric against SSIM

We compare the SSIM image quality metric for measuring clipped pixels against our modified clipping metric for the image shown in Figure 2, repeated below in Figure 10 for convenience.

Figure 10: Reference image showing cross section
Figure 11 and Figure 12 show the error estimates from both the SSIM and modified metric respectively for the highlighted region in Figure 10. For both metrics, the intensity represents the magnitude of the error. We can see that while the SSIM metric detects the error corresponding to the clipped pixels, the LCD clipping metric detects these errors even more clearly.

Figure 11: Output of SSIM error metric for clipped regions

Figure 12: Output of LCD clipping metric for clipped regions
In Figure 13, below we show a cross-section of the output of each of the two error metrics. The SSIM metric (cyan) is only slightly higher in each of the clipped regions, while the LCD clipping metric (magenta) shows a definitive change. The reason for this is simple: the SSIM metric is designed to detect loss of contrast, but it pools this error with other errors as well, thus diluting the measurement of the clipped regions. This works very well for most image applications but is not sensitive enough for specifically measuring LCD clipped pixels in our application. Using the customized metric allows us to directly measure the loss of high-frequency texture detail caused by LCD clipping.

Figure 13: Cross-section illustration of SSIM and LCD clipping error metrics
Chapter 5:  
Method for Improving Dual Modulation Displays

5.1 Introduction

We have seen that the previous approaches to dual modulation algorithms reviewed in Chapter 2 share the same underlying approach of first determining backlight drive values and then using the LCD to correct for errors. This approach has two fundamental limitations that limit performance.

The first limitation is that with a low spatial resolution backlight, there is no correct solution for the backlight near areas of high contrast. Instead, the solution must compromise between propagating error to the light or dark side of the boundary. All solutions to the backlight in these regions cause clipping errors on the LCD. Since human vision is much more sensitive to loss of high frequency texture detail than to low frequency errors [27], this approach tends to introduce the most visible types of errors.

The second limitation is that each of the previous methods requires first simulating the light field produced by the backlight in order to calculate the LCD. This step is the most computationally expensive component of the present dual modulation algorithms, and the accuracy directly influences the quality of the final image. In order to function in real time, most implementations make simplifying assumptions to the shape of the PSF. These assumptions lead to inaccuracies in the light field simulation which propagate into the LCD
calculation and affect the final image. As the number of backlight elements and the desired frame rate increases, this step becomes more and more of a bottleneck in the image processing pipeline.

Our new algorithm described in this chapter proposes to re-arrange the order of the steps in previous approaches by first determining an optimal solution to the LCD, and then compensating with the backlight as well as possible. This approach results in errors only in low-frequency portions of the signal where they are less noticeable.

In our approach, we first determine an optimal LCD image that is within the dynamic range limitations of the hardware, and then determine backlight drive values to increase the dynamic range according to the desired input signal. This method has the advantage that any errors introduced in the second and third steps are low frequency and masked by the high frequency detail that is displayed correctly on the LCD. It also has the advantage of avoiding the light field simulation, which was the most computationally expensive portion of the previous algorithms.

Figure 14: New dual modulation algorithm
Several authors cited in the previous section observed that the resulting LCD image shared some similar attributes to a tone-mapped image [6], [15]. In our approach we take this observation a step further and propose that we can generate the LCD image directly by appropriately tone-mapping the target image for the goals of our system. As described previously, these goals are:

1) Preserve the high-spatial frequency detail of the target image (avoid clipping)
2) Preserve the overall luminance and colors of the target image
3) Minimize power consumption by driving the LCD as open as possible

From our tone-mapped image we can then compute a target backlight by subtracting it from the target image. The final step is to generate backlight drive signals to achieve the target backlight as closely as possible.

Our new approach eliminates both the LCD clipping artifact as well as the light field simulation. However, our new approach is not able to achieve the same perfect reconstruction as previous approaches, as it uses the low resolution backlight as the error correction signal instead of the high-resolution LCD. For some display application, these advantages outweigh the drawback of imperfect image reconstruction for low spatial frequencies.
5.2 Detailed description of proposed algorithm

The key to our algorithm is in the first step, where we determine the LCD image from the target image. We start by splitting the target image into a base layer and a detail layer. The base layer contains the low spatial frequency information that is appropriate for the backlight. The remaining high spatial frequency detail is then allocated to the LCD. This is similar in practice to multi-scale tone mapping, which is commonly used to reduce the dynamic range of an image. This technique exploits the property that human vision is less sensitive to inaccuracies over large spatial frequencies so long as local spatial detail is preserved.

To generate a base layer that is appropriate for our backlight, we could convolve the image by the point spread function of each LED. The resulting base layer image would be within the spatial capabilities of our backlight.

\[ Base = PSF \ast Target \quad (11) \]

\[ \text{where } \ast \text{ denotes 2D convolution} \]

However, this amounts to a 2D convolution with a large spatial kernel, which is one of the steps of previous algorithms that we wanted to avoid due to the large memory and computational requirements. On our test machine, for example, this step takes 16 seconds. However, unlike previous approaches, in our method this step does not simulate the backlight, which requires high accuracy and precision, but instead just generates a low-pass filtered version of our target image. Eliminating the need for high accuracy allows us to make many simplifications to this step without introducing image artifacts.
Two efficient approaches for simplifying the low pass filter step are:

A) Translate the input image into the frequency domain using a fast Fourier transform, perform a multiplication, and then return to the spatial domain (0.20s or 80x faster). This method preserves the same accuracy as the convolution in the spatial domain but is much faster for certain hardware configurations.

B) Down-sample the input image to a lower resolution, convolve with a low pass filter, then up-sample (0.06s or 270x faster). This method is the fastest and can be easily implemented on most hardware configurations.

There are many other possible approaches, as optimizing image filtering with large spatial kernels has been extensively researched by others [28]. Since our backlight does not have color, we can also simplify this step by performing it on a single monochrome channel. Figure 15 shows an illustration of a target image and the resulting base layer.

![Target image and monochrome low frequency base layer](image)

**Figure 15: Target image (left) and monochrome low frequency base layer (right)**

---

1 Copyright Dolby Laboratories, 2010, by permission
Once we have the low frequency base layer, we can compute the high frequency detail layer required to regenerate the target image:

\[ \text{Detail} = \frac{\text{Target}}{\text{Base}} \]  

(12)

In regions with low spatial frequencies, the base layer will closely match the target image and the detail layer will be one. This is the result that we want to ensure that the LCD is fully open whenever possible. If the target is darker than the low frequency base layer then the detail layer can correct for this by dropping below one. Where the target is brighter than the base layer the detail layer will extend to a value greater than one. This will occur for small highlights that are brighter than the local average. This can be seen in some of the stems and leaves in Figure 16 below. If the detail layer were used directly as the LCD signal, the details in the stems in the detail layer (right hand side) would be clipped to one, eliminating all the fine detail.

A more sophisticated approach to dealing with values greater than one is to tone map the detail layer into the valid range (0:1) of the LCD. Any global or local tone mapper could be used for this step. In our implementation we apply a global tone curve to each of the R,G,B channels, illustrated in Figure 17. This is a form of the Nakka-Rushtan model for the response of cone photoreceptors [29]. The tone curve reduces the contrast of increasingly bright pixels without clipping detail.

\[ \text{Detail}_{TM} = \frac{c_1 \cdot \text{Detail}}{1 + c_2 \cdot \text{Detail}} \]  

(13)

The tone curve parameters \( c_1, c_2 \) can be calculated for each frame based on the minimum and maximum values in the detail layer, or can be defined as constant for all images. The
resulting image is used as the signal to the LCD. This LCD signal is as open as possible in regions of low spatial frequency and contains all the detailed information of the original image.

Figure 16: Detail layer (left) and tone-mapped detail layer (right)

Now that we have established the desired LCD signal, we need to determine the backlight drive values that will restore as much as possible of the original intensities and dynamic range.

Figure 17: Tone curve calculated for flower image detail layer

---

2 Copyright Dolby Laboratories, 2010, by permission
range of the input image. We cannot use the base layer that we generated in Equation 11 directly; as we have since altered some of the pixel values when we tone mapped the detail layer. Instead, we re-compute a base layer from the target image and our LCD layer from the previous step. This is opposite to how previous algorithms function, where the LCD layer is computed last by dividing the light field simulation. We will also account for light leakage through the LCD due to the limited contrast ratio $CR$.

$$Base = \frac{Target}{LCD \cdot (1-CR)+CR}$$  \hspace{1cm} (14)

Figure 18: New base layer

Finally, the base layer image is used to calculate backlight drive values. For this step we can borrow the same solution used in previous approaches, as there is no substantial difference in algorithm goals. The result of applying the backlight determination from the Brightside approach for our simulated hardware is shown in Figure 19.

Figure 19: Backlight drive values

---

3 Copyright Dolby Laboratories, 2010, by permission
5.3 Evaluation strategy

We compare the performance of our proposed algorithms against the existing approaches using a model of the dual modulation display and a combination of metrics. The model is implemented in Matlab and follows the process shown in Figure 20 below.

![Figure 20: Display model used for algorithm evaluation](image)

The model takes as input the backlight drive values and LCD image produced by the algorithm under test, as well as two parameters describing the physical system. The first is the Point Spread Function, which is the full model of how light propagates from each backlight element through to the LCD layer. The second is the contrast ratio of the LCD, which is how much light the LCD is able to block when completely closed. The Light Field Simulation step is a simple 2D convolution of the intensity of each backlight element by its intensity, and the Image Reconstruction step models how the light field propagates through
the LCD, taking into account desired light as well as undesired light leakage. The final Reconstructed Image is then ready for comparison against the original Target image using the appropriate image metrics.

To evaluate high-frequency clipping on the LCD layer, we use the LCD Clipping metric introduced in Chapter 4. In addition to the clipping metric, we also use SSIM to evaluate the overall difference between the target and reproduced image. This accounts for overall luminance and color shifts introduced by the algorithm and by undesired light leakage through the LCD. Finally, we are interested in evaluating the power requirements for each approach. This is determined simply from the sum of backlight drive values. For the same image quality we prefer solutions that require less power. Figure 21 below shows the metrics used to evaluate the algorithms.

Figure 21: Metrics used for algorithm evaluation
5.4 Design example

For purposes of illustration and comparison of various approaches, we describe a particular design example shown below in Figure 22. This hypothetical display has 325 backlight elements arranged in a rectangular grid of 25x13. The intensity of each backlight element can be controlled with 256 discrete values. The backlight array illuminates an LCD having a contrast of 100:1 and a resolution of 1920x1080 pixels, each with 3 color channels controllable to any one of 256 discrete values. Each backlight element has a PSF that significantly overlaps a 7x5 pattern of neighboring backlight elements, corresponding to an LCD area of ~540x420 pixels.

Figure 22: Backlight design example
Chapter 6: Results

6.1 Introduction

We begin this chapter by describing the results from benchmarking previous approaches using the metrics and evaluation procedure described previously. We then give the results from our new algorithm. We tabulate the results for ten arbitrarily chosen test images and discuss the results and implications.

6.2 Results from global backlighting algorithm

We start by analyzing the results for a global backlight, where no local dimming is applied. This is the case for the majority of LCD type displays, and suffers from light leakage and low power efficiency. Of course, there is no clipping of the LCD signal. The results for this case are shown below in Figure 23. The results show us what we expect. The backlight is at 100%, so the light field simulation is a full-white image and the LCD image is exactly the same as the target image. The reconstructed image looks close to the target, but the SSIM map shows that there is some light leakage, most significant in the dark regions. The clipping map shows no additional flat regions in the LCD image that were not present in the input image.
6.2.1 Results from “independent” dimming algorithms

As a second benchmark, we analyze what happens if the LCD and backlight signal are determined independently from the target image, without compensating one for the other. The LCD image is the same the previous example, but now the backlight approximately reproduces the input signal. The dimmed regions of the backlight reduce the power consumption to just one third of full white. However, we can see from the SSIM map that the reconstructed image is much worse. This is because we are now generating much less light in the dark regions of the image, but have failed to compensate for this. The resulting image is much too dark in dark regions. There are still no clipped pixels on the LCD signal.
As the third benchmark, we start to address the problem introduced above by splitting the dynamic range of the Target image into two parts – half for the LCD and half for the backlight. This corresponds to the first step of the original Brightside approach, which was to take the square root of the Target image as the input to the backlight determination algorithm. This results in an overall brighter backlight but enormously improves the SSIM error for the reconstructed image, still without introducing any clipped pixels. The goal for all subsequent algorithms is to improve upon this very simple approach. They will need to lower power and/or SSIM error, while only introducing very little or no additional image clipping. Beyond the scope of this research, they will also have to justify the image or power improvement against the additional computational complexity required.

![Figure 25: Independent square root LCD/BLU algorithm with greyscale image](image)

### 6.2.2 Results from Brightside algorithms

In the Brightside approach the backlight is generated in the same way as above, but this time the LCD image is computed by dividing the Target image by the Light Field Simulation as described in Chapter 2. Performing the light field simulation and division to produce the LCD image is much more complex than the previous method, but is able to reduce the SSIM error...
error somewhat. However, this approach also comes at the cost of producing a few clipped pixels in the brightest region of the image. This region of the backlight did not produce enough light to reach the target image. Thus the LCD is fully open and we have lost the texture of the smooth gradient. This might not be too visually objectionable for this particular image, but is often judged as an objectionable artifact for real imagery containing small bright features with important image detail such as clouds, lights, and reflections from water.

Figure 26: Brightside algorithm with greyscale image

To further illustrate the problem associated with the LCD clipping, we compare the independent approach (Figure 27) with the Brightside (Figure 28) approach for a more complex image. Again we see that both algorithms produce the same backlight signals. As with the greyscale image, the Brightside algorithm improves the SSIM error by compensating for the light field simulation, but introduces a massive number of clipped pixels in doing so. Most visible are the bright green and red regions which have grown in size and lost all texture in their brightest regions. These clipped pixels are detected by the SSIM metric, but are not weighed proportionally with their impact on image quality. In visual tests,
observers would much prefer the image with higher SSIM error over the image with clipped pixels, despite the better SSIM results. Figure 29 shows an enlarged view of the LCD signal to illustrate this. The clipped pixel metric specifically identifies this critical artifact of dual modulation algorithms.

Figure 27: Independent square root algorithm with complex colors image

Figure 28: Brightside algorithm with complex colors image
6.2.3 Results from iterative refinement algorithm

Before moving on to the proposed algorithm, it is useful to consider how good the above approaches could be, if the backlight were determined in such a way to prevent clipping completely while also minimizing light leakage and power. To do this, we start with the results of the Brightside method and then solve iteratively to open the LCD as much as possible without introducing clipping. In each step, we calculate how much “headroom” is in the LCD. We arbitrarily decide upon a target for the LCD to be 95% open. This leaves a little for overshoot while nearly maximizing the power efficiency and minimizing unwanted light leakage. For each region of the image, if the LCD is more than 95% open, we increase the backlight in that region slightly. If it is less than 95% we decrease it. For each step we calculate each of the metrics described above.
Figure 30 below shows the results for the complex colors image after 30 iterations. We can see that the backlight has been brightened from 62% to 79% in order to reduce the clipped pixels from 2.6% to 0.2%, and the SSIM error has been reduced by nearly 70%. The solution converged in just ten iterations. At typical video rates of 60 frames per second, this is just 17ms. A viable solution for an improved local dimming display could be to perform this iteration on subsequent video frames, which should not vary much at the backlight resolution, even for moving video. Some predictive element could also be incorporated to improve performance for quickly moving objects.

Figure 30: Iterative refinement with complex colors image
Performing the iterative method on the grey ramp image shows how after the same 30 iterations it can simultaneously reduce the number of clipped pixels by 40%, lower power by 30%, as well as decrease SSIM error by 20%. The algorithm would be able to completely eliminate the clipped pixels if allowed to proceed for additional iterations, or if more aggressive iterations methods were employed. Notice how the LCD image is nearly completely open, since this low frequency signal can be reproduced very closely by the backlight.

Figure 31: Iterative refinement with grey ramp image
6.3 Results of the proposed algorithm

Results for the proposed algorithm on the grey ramp image are shown below in Figure 32. The new algorithm nearly completely eliminates the LCD clipping artifact. Some remaining clipping is where the texture has been reduced to less than 10% by the tone curve described in Chapter 5. Complete elimination should be possible by fine tuning the multi-scale tone mapping algorithm. The new algorithm is also able to reduce the backlight power by 25% over the Brightside approach. As expected, the algorithmic simplifications taken for this approach result in a less accurate reproduction of the original image, since we are effectively using the backlight to correct for errors in the LCD image, rather than vice-versa. Since the backlight has limited spatial resolution, and is not being accurately modeled, the effect is to raise the SSIM error by over 3x for this image. As mentioned before, SSIM is an average metric for the entire image, so does not weigh the effect of a small region of clipped pixels as highly as a human viewer might. It would be up to the designer of a display device to decide if the lower power and reduced clipping artifacts offset for the lowered overall accuracy.

Since many consumer televisions are viewed in non-ideal circumstances (bright lights and open windows, un-calibrated, vivid expansion mode, etc.), the benefits of the new approach may well outweigh the overall decrease in quality.

![Figure 32: Proposed algorithm with grey ramp image](image)

Target Image  Backlight [36%]  Light Field Simulation  LCD Image

Reconstructed Image  SSIM Error: [9.7e-006]  Clipped Pixels: [0.22%]
Likewise, for the complex color image, we see that the clipping was reduced from 2.6% to 0.1% (95% reduction); while the backlight was increased by 30% and the SSIM error also increased approximately 3x.

![Figure 33: Proposed algorithm with complex colors image](image-url)
6.4 Results for real images

Finally, we depart from test patterns and evaluate the results for a selection of real images.

Table 1 below shows the tabulated metric results for ten samples of typical cinematic content found on the internet.

Table 1: Algorithm comparison results

<table>
<thead>
<tr>
<th>Image Thumbnail</th>
<th>Global Backlight</th>
<th>Brightside</th>
<th>Iterative Refinement</th>
<th>New Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Power (%)</td>
<td>SSIM x10^5</td>
<td>Clipping (%)</td>
<td>Power (%)</td>
</tr>
<tr>
<td>1 Flower</td>
<td>100</td>
<td>5.1</td>
<td>0.0</td>
<td>53</td>
</tr>
<tr>
<td>2 True Grit</td>
<td>100</td>
<td>59</td>
<td>0.0</td>
<td>32</td>
</tr>
<tr>
<td>3 Avatar</td>
<td>100</td>
<td>25</td>
<td>0.0</td>
<td>25</td>
</tr>
<tr>
<td>4 Navaron</td>
<td>100</td>
<td>150</td>
<td>0.0</td>
<td>4</td>
</tr>
<tr>
<td>5 Inception</td>
<td>100</td>
<td>10</td>
<td>0.0</td>
<td>48</td>
</tr>
<tr>
<td>6 Outlander</td>
<td>100</td>
<td>51</td>
<td>0.0</td>
<td>43</td>
</tr>
<tr>
<td>7 Pirates</td>
<td>100</td>
<td>98</td>
<td>0.0</td>
<td>12</td>
</tr>
<tr>
<td>8 Shoopah</td>
<td>100</td>
<td>110</td>
<td>0.0</td>
<td>9</td>
</tr>
<tr>
<td>9 Third Man</td>
<td>100</td>
<td>56</td>
<td>0.0</td>
<td>16</td>
</tr>
<tr>
<td>10 Old Men</td>
<td>100</td>
<td>56</td>
<td>0.0</td>
<td>37</td>
</tr>
<tr>
<td>Average</td>
<td>100</td>
<td>62</td>
<td>0.0</td>
<td>28</td>
</tr>
<tr>
<td>Worst</td>
<td>100</td>
<td>150</td>
<td>0.0</td>
<td>53</td>
</tr>
</tbody>
</table>

Average and Worst values are calculated for all images.
6.5 Discussion

This thesis establishes a benchmark for display performance for the global backlit displays that we are very familiar with (computer monitors, televisions, etc.). This type of display suffers from unwanted light leakage through the LCD panel, which raises the black level and de-saturates colors. In our evaluation conditions for ten arbitrary test images we find that the light leakage introduces a mean SSIM error of $62 \times 10^{-5}$.

Next, we compare the global backlight performance with a locally modulated backlight using a published local dimming algorithm from Brightside technologies. Having individual control over hundreds of regions of the backlight allows such a display to lower the backlight power to an average of 28%. This is an enormous savings as this saving does not only impact the display cost due to reduced power requirements, but also the ongoing cost of operating the display. In addition to saving power, such displays can improve image quality, due to having less light unwanted light leakage through the LCD. In our evaluation, we found that the mean SSIM error is reduced to $19 \times 10^{-5}$, close to a 70% improvement compared to a global backlight.

However, in order to achieve these impressive performance gains, previous control algorithms tend to introduce an artifact that we denote “LCD Clipping”. The effect of this artifact is a local reduction of spatial detail, usually for small regions of an image. Typical image metrics based on image averages such as SSIM do not to weigh such errors very highly, due to the very small size of the errors compared to the entire image. However, many
viewers find such errors disproportionately distracting compared to their size. To account for
this, we introduced a modified metric specifically designed for LCD clipping which
measures the number of pixels where the local spatial detail has been lost. Global backlights
do not introduce any clipping. The Brightside approach introduces an average of 2.7%
clipped pixels for the images we tested. Notice that these clipped pixels do not severely
impact the SSIM score. The clipping metric contributed in this thesis can be easily and
quickly applied to detect lost spatial detail in an image, providing a useful tool in the design
of dual modulation algorithms.

Next, we investigate how the performance of the Brightside approach could be improved by
refining the usual algorithm over subsequent image frames. This technique employs the
clipping metric to improve image quality on subsequent image frames by making
adjustments to the backlight control signals. Allowing the algorithm to be refined in this
manner can virtually eliminate the LCD clipping artifact (to 0.1% of pixels), as the backlight
is adjusted appropriately in clipped regions. However, there is a cost to this improvement. To
prevent the LCD clipping artifact, the backlight is on average increased from 28% to 70%,
greatly reducing the improved power efficiency that we enjoyed with the Brightside
approach. A portion of this additional light results in unwanted light leakage, in turn raising
the SSIM error slightly to $30 \times 10^{-5}$.

When we evaluate our new algorithm under the same conditions as the previous two
approaches, we find an average backlight level of 30%, very similar to the Brightside
approach and much lower than the iterative approach. We also verify that the LCD clipping
artifact is eliminated. The cost of this simpler, low power algorithm with no LCD clipping is a higher average SSIM error of $39 \times 10^{-5}$. This is one and a half times higher than the Brightside approach, but still much lower than the global backlight approach. The main difference between the image results of the new algorithm with the previous ones is that in the new algorithm the error is contained in the low spatial frequency backlight signal; whereas the Brightside approach causes the error to be propagated to the high spatial frequency LCD layer. Since human vision is less sensitive to image artifacts in low spatial frequencies, the fundamental approach of the new algorithm should outperform previous approaches despite having a higher SSIM error.
Chapter 7:

Conclusion

LCD displays are the dominant display technology in the market today and for the near future. They are relatively cheap to make, have excellent lifetime characteristics, and have excellent image quality. Their most significant weaknesses are contrast and power efficiency. To address these two weaknesses display makers are turning to dual modulation backlights. These improve contrast and efficiency by dimming the backlight in image regions that need less light. A critical component of a dual modulation display is the control algorithm, which determines how to best control the backlight and LCD together to achieve the desired performance. Existing control algorithms for dual modulation displays can introduce an image artifact termed LCD clipping. This thesis proposes a new control algorithm that eliminates this artifact, with minimal other tradeoffs in image quality and power efficiency. The new algorithm has the added advantage that it requires less computational resources than previous algorithms. Furthermore, it does not require a change to the display hardware, making it a relatively simple upgrade for today’s dual modulation displays.

In this thesis we compared our new algorithm against a previously published control algorithm for dual-modulation displays from Brightside technologies as well as non-dual-modulation display. We evaluated the performance of these algorithms using four main criteria:
1. **LCD Clipping**: The previous control algorithm caused an average of 2.7% and up to 6.8% of all LCD pixels to be clipped. These pixels had lost all texture and contrast information and are highly visible. By comparison, our proposed algorithm clips an average of 0.0% and up to 0.1% of pixels. Non dual-modulation displays also have 0.0% clipped pixels.

2. **Power Efficiency**: The previous algorithm used on average 28% of the power of a non-dual-modulation display. Our proposed algorithm uses on average 7% more power than the previous method, but still much less than a non-dual-modulation display.

3. **Image Quality**: The previous algorithm reproduced the target image with an average SSIM error of $19 \times 10^{-5}$. Our algorithm increases the error by 50% to $39 \times 10^{-5}$. The increased error is in the low spatial frequency channels of the image, where the human visual system is least sensitive. Further studies should be conducted on the actual display hardware to assess the overall visual quality. By comparison, non-dual-modulation displays have an average error of $62 \times 10^{-5}$.

4. **Computational Requirements**: Previous algorithms required highly accurate models of the display optics on a per-frame basis. We found that our approach executed 42% faster than the Brightside approach (300ms compared to 510ms in Matlab on our test machine). For displays with greater numbers of backlight elements the performance increase will be even higher. Future displays having 120Hz or 240Hz video,
thousands of backlight elements, or 3D processing are nearly if not completely infeasible with previous algorithmic approaches.

In summary, the main contribution of this work is a new algorithm which produces images without any clipped pixels, while maintaining the power efficiency of previous approaches and only slightly increasing SSIM. The increased SSIM error is constrained to low spatial frequency channels where it is less objectionable than clipped pixels in high spatial frequency channels. The new algorithm also eliminates the need for an accurate model of the display optics, reducing the computational complexity, making it a feasible solution for future high-end displays as well as current lower-end products.

7.1 Future work

There are many areas for improvement and refinement to the new algorithm described in this work. One future activity is to perform visual assessment and validation of the images on an actual dual-modulation display. This was planned to be done on a novel dual-modulation display, but had to be postponed due to availability of the hardware. However, from previous validation work we are confident that our simulated display model is a highly accurate model of the physical hardware, and the SSIM metric has been previously validated as a good metric for assessing image quality. Evaluating with moving images for temporal stability and other motion artifacts would also be an important component of this work.

Aside from the visual assessment, it is possible to do nearly endless tuning and modifications to the multi-scale tone mapping algorithm which is the central feature of the new algorithm.
This is an active area of research in the image capture community [30], although the application of these to dual modulation displays is novel to our knowledge. It is highly likely that future improvements to multi-scale or other local tone mapping operators will also improve performance for the dual modulation display application. As new tone mapping algorithms are developed they can be incorporated into the new dual modulation display algorithm relatively easily.

Secondly, there are nearly endless possibilities for improving the determination of backlight element control signals. These have been well described in previous work [15]. For consistency in this thesis we used the same backlight determination algorithm for each of the approaches that we investigated. However, there are many opportunities to fine-tune and improve the exact filtering and weighting used in this step, which would improve results for each of the Brightside, iterative, and new algorithms. This should be done in conjunction with the specific display hardware to ensure that the tuning is most applicable for the particular application. One novel approach to make improvements to this step could be to train a neural network via back-propagation from the optimal iterative solution, with the goal of achieving the same excellent image quality of the iterative solution with a single network lookup with appropriate input conditioning and weighting.

An interesting extension to this work could be to further develop the idea of an iterative algorithm for use with certain applications. The iterative refinement algorithm is a vast improvement over the global backlight approach, with a 30% savings of power, 50% improvement in SSIM, and virtually no clipped pixels. Instead of having a highly complex,
highly accurate algorithm that is computed for each frame, a simpler, low complexity algorithm could be used to refine an early guess over subsequent image frames. The guess would be refined based on the output of metrics such as SSIM and the LCD clipping metric. Such an approach could be low cost, since it would not be a requirement to do extensive image processing per-frame. It may also be high quality, at least after enough iterations have been applied. It would also be able to produce a viewable image immediately, even if it is not initially high quality. Over a few subsequent iterations the quality would improve. The major drawback would be that it is less suitable to moving imagery than a faster solution. However, for certain applications requiring excellent image quality but not necessarily fast image refresh rates this could be a very useful approach. Products such as digital artwork or picture frames are perfect examples, where the content changes slowly and predictably or not at all, and power savings and high image quality are important differentiating factors. In our evaluation we found that the iterations typically converged to an optimal solution in fewer than 10 frames. At video rates of 60Hz this corresponds to 1/6 of a second. With future displays being capable of 120 and 240Hz, this is as low as 80ms and 40ms respectively. If our iterative approach were coupled with motion vectors it may become practical to refine the backlight control signals for relatively unchanging regions of a video stream.
References


[27] P. Barten, Contrast sensitivity of the human eye and its effects on image quality, SPIE

