SINGLE SHOT HIGH DYNAMIC RANGE AND
MULTISPECTRAL IMAGING BASED ON PROPERTIES
OF COLOR FILTER ARRAYS

Thesis
Submitted to
The School of Engineering of the
UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for
The Degree
Master of Science in Electrical and Computer Engineering

by
Paul M. Simon

UNIVERSITY OF DAYTON
Dayton, Ohio
May, 2011
SINGLE SHOT HIGH DYNAMIC RANGE AND MULTISPECTRAL
IMAGING BASED ON PROPERTIES OF COLOR FILTER ARRAYS

NAME: Simon, Paul M.

APPROVED BY:

Keigo Hirakawa, Ph.D.
Advisory Committee Chairman
Assistant Professor
Electrical and Computer Engineering

Eric Balster, Ph.D.
Committee Member
Assistant Professor
Electrical and Computer Engineering

Kenneth J. Barnard, Ph.D.
Committee Member
Senior Electronics Engineer
Air Force Research Laboratory (AFRL)
Wright-Patterson AFB

John G. Weber, Ph.D.
Associate Dean
School of Engineering

Tony E. Saliba, Ph.D.
Wilke Distinguished Professor & Dean
School of Engineering
©Copyright by

Paul M. Simon

All rights reserved

2011
This paper addresses the difficulty of generating High Dynamic Range (HDR) images using current Low Dynamic Range (LDR) camera technology. Typically, several LDR images must be acquired using various camera f-stops and then the images must be blended using one of several exposure bracketing techniques to generate HDR images. Based on Fourier analysis of typical Color Filter Array (CFA) sampled images, we demonstrate that the existing CFA sampled images provide information that is currently underutilized. This thesis presents an approach to generating HDR images that uses only one input image while exploiting that underutilized CFA data. We propose that information stored in unsaturated color channels is used to enhance or estimate details lost in saturated regions. A demonstration of experimental data is presented for an optimized combination of demosaicing, color balance, gamma correction, and HDR image generation.

As an extension of the HDR imaging and CFA research presented herein, multispectral imaging is also explored. We develop a theory for acquiring multispectral image data based on the Fourier analysis of the relationships in the spatial-spectral signals which then leads to the design of a multispectral CFA.
ACKNOWLEDGMENTS

Above all else, I offer my sincerest gratitude to my advisor, Dr. Keigo Hirakawa, for listening to my stories, walking me through elementary algebra, and supporting me throughout my thesis with his brilliance, grace, and excellence. Without him, this thesis would still be a jumble of nasty equations and ones and zeros.

I also need to express my sincerest appreciation to my wife Eileen, whose never-ending energy has helped me turn all my dreams into reality. She truly is a pillar of my life and my best friend.

Last but not least, I thank my parents Paul W. and Nancy, for instilling in me the need to constantly work to achieve a goal. Yes, Uncle Wasyl would be pleased, too.

The research described herein is sponsored in part by the Air Force Research Laboratory (AFRL).
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td></td>
<td>vii</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>BACKGROUND IN COLOR DIGITAL IMAGE CAPTURE</td>
<td>3</td>
</tr>
<tr>
<td>2.1</td>
<td>Image Capture</td>
<td>3</td>
</tr>
<tr>
<td>2.2</td>
<td>Signal Analysis</td>
<td>5</td>
</tr>
<tr>
<td>2.3</td>
<td>Theoretical Approach to Demosaicing</td>
<td>9</td>
</tr>
<tr>
<td>2.4</td>
<td>Theoretical Approach to Color Balance and Gamma Correction</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>HIGH DYNAMIC RANGE IMAGING</td>
<td>18</td>
</tr>
<tr>
<td>3.1</td>
<td>Background and Existing HDR Methods</td>
<td>18</td>
</tr>
<tr>
<td>3.2</td>
<td>HDR Image Analysis in CFA Sampled Images</td>
<td>19</td>
</tr>
<tr>
<td>3.3</td>
<td>Theoretical Approach to HDR Imaging</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>MULTISPECTRAL IMAGING</td>
<td>30</td>
</tr>
<tr>
<td>4.1</td>
<td>Spectral Imaging Background</td>
<td>30</td>
</tr>
<tr>
<td>4.2</td>
<td>Spectral Image Analysis</td>
<td>31</td>
</tr>
<tr>
<td>4.3</td>
<td>Theoretical Approach to Spectral Imaging</td>
<td>34</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

1. Block Diagram of the Camera Processing Pipeline .......................... 4
2. Color Digital Image Processing .................................................. 6
3. Actual, Idealized, and Aliased Representations of Fourier Transform of an RGB Image .................................................. 8
4. Using a Modulated Low Pass Filter to Isolate the Chrominance 1 Portion of the CFA Image .................................................. 10
5. Using a Directional Modulated Filtering to Isolate the Chrominance 2 Portion of the CFA Image .................................................. 11
6. Directionally Sensitive Emphasis and Deemphasis to Avoid Aliasing in the Chrominance Reconstruction .................................. 12
7. Reconstruction of an Image: Chrominance 1 ($f_{C1}$), Chrominance 2 ($f_{C2}$), and Luminance ($f_{L}$) .................................................. 13
8. Creation of an Unamplified Representative Image: Images at F-Stop 0, F-Stop -2, and F-Stop -1 .................................................. 25
9. 2D Fourier Transform of CFA After White Balance .......................... 27
10. Comparison of Original, Unaltered Image and Image Created Using Single-Shot HDR Algorithm .................................................. 28
11. 3D Spectral Content ................................................................. 31
12. Periodic Sampling of the 3D Spectral Contour ................................ 34
13. Multispectral Image Sampling ...................................................... 36
14. Multispectral Image Processing Step-By-Step .................................. 37
15. Example of a 2D Idealized Fourier Transform of a Multispectral Image 38
1 INTRODUCTION

The current state of the art in digital imagery and digital image processing provides more flexibility and the “heavy lifting” of data processing is becoming more transparent. Improvements to digital image and signal processing and image sensors based on Charge Coupled Devices (CCD) and Complementary Metal–Oxide Semiconductor (CMOS) technologies are pushing the speed of cameras and the number of mega-pixels per image upward every year. Despite that, there are still improvements that can be made to the output images or the camera pipeline. Color Filter Arrays (CFA) effect how the image is sampled and reconstituted. The CFA is a set of color filters overlaid upon the CCD or CMOS sensor. The CCD is the silicone photo-sensor that detects light intensity, thus the CFA filters the incoming light to provide specific wavelength selectivity to each pixel location on the sensor.

demosaicing is the interpolative reconstruction process used to generate a full color image from the incomplete color samples that were collected by combining image sensor and CFA. It is the process that renders the raw image data into a viewable format. The demosaicing process effects the quality of the overall image and the speed of reconstruction. White balance, color correction, and gamma correction further adjust how the image appears on a color monitor or when printed. Further, the end-user features such as smile recognition and facial recognition aid the photographer by triggering the shutter based on a smile or the placement of the target's face. Post processing techniques, such the red-eye correction, further improve the overall quality of captured images. Consumer cameras are becoming extremely feature rich and user friendly. This has the effect of shadowing the underlying function of demosaicing, or the needs for proper lighting, and still makes those cameras somewhat inflexible since they are designed to respond to every-
day types of images without giving the user access to the internal functions. Yet, new features are always being rolled out, the best providing a unique or complex capability in an easy-to-use function.

In this thesis, novel features of cameras based on the Fourier analysis of the CFA sampled image are presented. The first feature is High Dynamic Range (HDR) imaging, which extends the dynamic range of images, eliminating saturation that occurs in brighter regions of images and reduces noise in darker regions of images. HDR imaging requires special processing to arbitrate between brighter, possibly saturated, areas and darker, possibly noisier, areas within a field of view. This document presents a unique algorithm for reproducing HDR images with existing LDR camera hardware and only one input source image. The algorithm that we developed is based on Fourier analysis of the CFA sampled image and takes advantage of details that heretofore been ignored.

Second, this thesis proposes the design of a multispectral CFA based on the Fourier analysis of spatial-spectral data. Multispectral imaging, instead of just three wide-bandwidth wavelength color channels, has multiple narrow-bandwidth color channels. Multispectral imaging can feasibly use existing sensor technology, however existing CFA cannot support the additional wavelengths and as such the design methodology of the CFA and demosaicing approaches need to be refined. This document provides the method for efficient single-shot capture and demosaicing of multispectral images by developing a scalable CFA based on the lessons learned from HDR image processing and analysis. General cases are presented and explored.
2 BACKGROUND IN COLOR DIGITAL IMAGE CAPTURE

Color digital image capture technology encapsulates the broad array of hardware, firmware, software, and theoretical processes used to generate digital images. To fully grasp the complexity of digital imagery, the underlying processing of digital cameras is reviewed first. Most importantly, the Fourier analysis of CFA sampled images is shown to provide clear evidence and support for demosaicing and image generation processes. The analysis also provides a path towards HDR and multispectral imaging.

2.1 Image Capture

The series of signal and image processing steps that are taken to generate a digital image in a digital camera is called the camera pipeline. Figure 1 shows a block diagram of the camera pipeline, with the specific steps taken from light entering a camera to the two possible output, a raw data format and a compressed data format image. As can be seen in Figure 1, the camera pipeline is broken up into 3 main sections: the processes that occur in the camera body, on the sensor chip, and in dedicated digital signal processing (DSP). The importance of the camera pipeline and the order of the processes is a critical detail that will be addressed later in this document.

The optics and shutter mechanism of a digital camera is identical to the film camera. Light, or more specifically, photons enter the camera through a lens, generally referred to as the camera’s optics. Digital cameras have a method for collecting photons over a time period similar to an optical shutter or by using an actual shutter mechanism. The optics focus the photons through an aperture so that when the shutter mechanism opens, the photons are directed into the camera.
body onto the photosensitive elements inside the camera, be it film or a CCD or CMOS. The amount of time that the shutter is open determines exposure time, or how much light enters into the camera. The end-user is able to select a slower shutter speed, with the end result of more light entering the camera body, thus higher exposure time. A faster shutter speed allows less light to enter the camera body, thus lower exposure time. The f-number, or f-stop, is the focal length of the optics divided by the effective aperture diameter. This number will be referenced later in the document.

Next, the role of that the sensor chip plays is analogous to the role of film in a traditional camera. The CCD is a silicon-based integrated circuit that acts as a photo-sensor. When the photons that are allowed to enter the camera body through the lens and aperture penetrate the Color Filter Array and the CCD/CMOS, the intensity of the light is measured. Lastly, the analog signal that is generated by the CCD or CMOS is converted to a digital signal in an analog-to-digital (A/D) process. From this point, the raw image can be output for independent processing.
or it can be sent on in the camera pipeline for additional in-camera processing.

The final stage in the digital camera pipeline are the set of processing of the digital image data itself. Taken together, they are referred to as a digital camera processing pipeline, and may include other processes that are outside the scope of this document. Contained in the pipeline are white balance and color correction, demosaicing, gamma correction, and compression. White balance and color correction deal with matching the digital data with the appearance of physical world. Demosaicing is the process of converting from a sampled image to an actual, tristimulus image. Gamma correction is another method of adjusting the color content in an image to appear more realistic on paper or on a display. Lastly, compression is the process of data reduction to allow for efficient transfer and storage of the output image. With the exception of compression, all of these steps will be covered in this document to some degree, and all have a significant impact on the quality of the output image.

### 2.2 Signal Analysis

The process of demosaicing has been researched extensively [1, 2, 3, 4, 5, 7], and is still an active field of research. Demosaicing, to define it precisely, is the interpolative process of transforming the incomplete image output of a Bayer Color Filter Array (CFA) / sensor pair, also known as a Bayer image, into a standard RGB image. Color filters in a Bayer CFA are arranged as

\[
\begin{bmatrix}
G & R & G & R & \ldots \\
B & G & B & G & \ldots \\
G & R & G & R & \ldots \\
B & G & B & G & \ldots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\]
where R, G, and B represent the locations of the red, green, or blue filters, respectively. The process of demosaicing begins with the raw intensities at each pixel location. These intensities are the direct measurements of light as detected by the sensor. The CFA filters the light that enters the camera and provides wavelength selectivity, thus each pixel represents a specific range of wavelengths, or color. The Bayer image can be viewed directly, however it appears as a highly modulated gray scale image since the raw pixel values have no color information associated with them and they are merely digital numbers corresponding to the amount of light transmitted by the different filters. A comparison can easily be observed in Figure 2, where Figure 2a is the original color image, Figure 2b is the typical Bayer CFA that is overlaid onto the sensor, which then generates the Bayer image shown in Figure 2c. The raw output of most modern SLR cameras is this very Bayer image. It should also be noted that the green color content of any arbitrary image is down sampled by a factor of 2 when it is recorded and the image’s red and blue color content are down sampled by a factor of 4. Reconstructing a typical RGB image from a Bayer CFA thus requires interpolation to fill in the missing pixels and
to generate the three individual color layers, or color channels, of the RGB image. Arguably, the green component contains more spatial information than either the red or blue color content if no saturation has occurred. Saturation occurs when any of the color channels reach the maximum intensity possible and can be regarded as potential loss of information.

Luminance refers to the brightness of linear RGB components. Luminance represents the achromatic component, or brightness, of light. Chrominance refers to the color content of linear RGB images. Because human vision has finer spatial sensitivity to luminance differences than chromatic differences, imaging systems typically store chromatic information at lower resolution. Because the luminance can be approximated to the color green, it is the green color content of the CFA sampled image that provides the spatial resolution. This reasoning justifies the higher sampling rate of the green component compared to red or blue in a Bayer CFA. It is the red and blue color content of the CFA sampled image that provides the chrominance components. Therefore, the chrominance component is further split into the chrominance 1 and chrominance 2 components because of chrominance sub-sampling, representative of the red pixel locations and the blue pixel locations, respectively.

The Fourier analysis is an important tool in image processing. The Fourier transform is used to decompose the image from the input spatial domain to the frequency domain. For a MxN image, the discrete Fourier transform (DFT) is defined as

$$X(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n) e^{-j2\pi(km + ln)}$$  \hspace{1cm} (1)

where $x(m, n)$ is the image in the spatial domain and the exponential term is the basis function corresponding to each point in the Fourier domain, $X(k, l)$. The basis functions are sine and cosine waves with increasing frequencies.
Fourier analysis of an image provides insight into the spectral content of the image. Figure 3a is a plot of the discrete Fourier transform of a CFA sampled image as generated by performing a discrete Fourier Transform on a CFA image. Figure 3b is an idealized 2D representation of the output of a Fourier transform of a CFA image. Typically, the luminance, which is located at the center of the 2D representation, will have a higher energy and spatial bandwidth than the chrominance components that are located at the corners and on the edges of the 2D representation. The red, green, and blue shading in Figure 3b also presents the frequency support contained in the luminance and chrominance, respectively. Separating out the luminance, chrominance 1, and chrominance 2 components from the total spectral content in a disciplined manner leads to a successful demosaicing process.

Care must be taken when manipulating the images to prevent spectral overlap when multiplexing the image components. The spectral overlap is also known as aliasing, which is the interference or crosstalk that can occur between spectral components that contaminates the spectral content. Aliasing is strongest between the luminance component and the modulated chrominance 2 component, as can be seen in Figure 3c. Aliasing tends to be most problematic with high spacial frequency, and leads to a zippering artifact in the reconstructed image.
From the power spectral density shown in Figure 3a, it appears that the luminance component is recoverable from the CFA signal using a low pass filter. The two chrominance components located in the corners and along the edges also appear to be recoverable from the CFA using appropriate band pass filters. Thus the CFA sampled image is separated into the component parts.

2.3 Theoretical Approach to Demosaicing

demosaicing, as it was defined earlier, is an interpolative process of reconstructing an image. The most important aspect of Figure 3b is that it demonstrates the modulation of the chrominance 1 and chrominance 2 components off of the baseband as compared to the luminance component. As can be seen in Figure 3b, the chrominance 1 component is shifted off of the baseband to the corners and the chrominance 2 component is shifted to the top, bottom, and side edges. By observing the CFA arrangement of pixels, Figure 2b, it is clear that the chrominance components contain a modulation factor. Therefore, the process of demosaicing also demodulates the chrominance components in the reconstruction process.

To demosaic the Bayer image and to reconstruct a full color image from CFA sub-sampled data, a three dimensional Gaussian contour centered at zero (DC), \( h_{LP}(x,y) \), is convolved with a modulated version of the CFA to generate the chrominance 1 portion of the reconstituted image, or specifically

\[
 f_{C1}(x,y) = h_{LP}(x,y) \ast \left( f_{CFA}(x,y) \left( (-1)^x + (-1)^y \right) \right)
\] (2)

where \( f_{CFA}(x,y) \) is the CFA sampled image and \( \left( (-1)^x + (-1)^y \right) \) is the modulation factor. The amplitude response of the Gaussian is shown in Figure 4a. The CFA is multiplied by the modulation factor so that the chrominance 1 portion of the image is shifted to the baseband, as shown in Figure 4b. The Gaussian acts as a low-pass filter on the modulated version of the CFA, thus the effective pass band of
the filtering process is shown in Figure 4c. Therefore, high frequency content is eliminated and the low frequency content is stored. An example of the process for computing $f_{C1}(x, y)$ for a typical image is shown in Figure 7a.

To extract the chrominance 2 component from the CFA sampled image, the demodulation needs to occur in a directionally sensitive manner, depending on which process is being performed. A pair of Gaussian curves, $h_x(x, y)$ and $h_y(x, y)$, must be generated such that they are narrow in the opposing direction and wide in the operative direction of processing. Figures 5a and 5b shows the directional filters $h_x(x, y)$ and $h_y(x, y)$, respectively. Those Gaussian curves are convolved with horizontally and vertically demodulated versions of the CFA such that the pass band of the filtering process is shown in Figure 5c. To clarify this point, the equations are as follows:

$$f_{C2a}(x, y) = h_x(x, y) \ast (f_{CFA}(x, y)(-1)^x)$$  \hspace{1cm} (3)

and

$$f_{C2b}(x, y) = h_y(x, y) \ast (f_{CFA}(x, y)(-1)^y)$$ \hspace{1cm} (4)

where $f_{CFA}(x, y)$ is the CFA sampled image and $(-1)^x$ and $(-1)^y$ are the directional modulation factors.
An image that has high spatial frequency in the horizontal direction, for example a picket fence, may introduce aliasing in the spectral domain. An idealized 2D representation of the output of a Fourier transform of an image that contains high spatial frequency is shown in Figure 6a. Note the aliasing in the spectral domain. By deemphasizing the contribution of horizontally modulated portion of the CFA image and compensating with the vertically modulated portion of the CFA image, as shown in Figure 6b, the aliasing will be avoided in the demosaicing and reconstruction process. Therefore, an additional weighting coefficient is required to properly balance the vertically and horizontally modulated portions of the chrominance component, eliminating aliasing artifacts. The coefficient is generated by convolving modulated versions of $h_x(x, y)$ and $h_y(x, y)$ by the original CFA. This can be represented as

$$f_x(x, y) = h_x(x, y) \ast f_{CFA}(x, y)$$

(5)

and

$$f_y(x, y) = h_y(x, y) \ast f_{CFA}(x, y),$$

(6)
which is then convolved with a neighborhood moving average filter, \( f_m(x, y) \), such that the coefficient is

\[
c(x, y) = \frac{\left( f_y(x, y) \right)^2 \ast f_m(x, y)}{\left( \left( f_x(x, y) \right)^2 \ast f_m(x, y) \right) + \left( \left( f_y(x, y) \right)^2 \ast f_m(x, y) \right)}.
\]

The neighborhood moving average filter is a 5-by-5 moving window where the average of the window is stored at the center pixel location. By convolving the neighborhood moving average filter by the square of \( f_x(x, y) \) and \( f_y(x, y) \) as shown in \( (7) \), the coefficient, \( c(x, y) \), becomes a value between zero and one representing the average pixel intensity. The resultant weighting coefficient is multiplied by the directional portions and summed to generate the final estimate of the chrominance 2 portion of the image by

\[
f_{C2}(x, y) = \left( c(x, y) \ast f_{C2a}(x, y) \right) + \left( (1 - c(x, y)) \ast f_{C2b}(x, y) \right).
\]

The weighting coefficient helps balance the final directional characteristics of the image, depending on the image content. An example of this computation for \( f_{C2}(x, y) \) for the same imaged used in the calculation of \( f_{C1}(x, y) \) is shown in Figure 7b.
The estimate for the luminance component is generated by subtracting $f_{C1}(x, y)$ and $f_{C2}(x, y)$ from the original CFA, or by using the following equation

$$f_L(x, y) = f_{CFA}(x, y) - f_{C1}(x, y) - f_{C2}(x, y)((-1)^x - (-1)^y).$$  \hfill (9)

The luminance, or $f_L$, is shown in Figure 7c. Again, $((-1)^x - (-1)^y)$ is a modulation factor that is included to shift the chrominance component back off the baseband. Since $f_{C1}$ and $f_{C2}$ contain mostly low frequency content, the luminance contains the high frequency content of the image. Therefore, there is significantly more detail in the luminance than the other two image components which reinforces the Fourier analysis.

Using the three components, luminance or $f_L$, chrominance 1 or $f_{C1}$, and chrominance 2 or $f_{C2}$, the standard RGB image is generated. In matrix notation,
this is stated as
\[
\begin{bmatrix}
  f_L \\
  f_{C1} \\
  f_{C2}
\end{bmatrix}
= \begin{bmatrix}
  \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\
  -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\
  -\frac{1}{4} & 0 & \frac{1}{4}
\end{bmatrix}
\begin{bmatrix}
  f_R \\
  f_G \\
  f_B
\end{bmatrix}
\] (10)

or
\[
\begin{bmatrix}
  f_R \\
  f_G \\
  f_B
\end{bmatrix}
= \begin{bmatrix}
  1 & -1 & -2 \\
  1 & 1 & 0 \\
  1 & -1 & 2
\end{bmatrix}
\begin{bmatrix}
  f_L \\
  f_{C1} \\
  f_{C2}
\end{bmatrix},
\] (11)

where \( f_R, f_G, \) and \( f_B \) are the red, green, and blue color channel layers. Solving for the missing component via matrix multiplication generates the final, full-color RGB image.

**2.4 Theoretical Approach to Color Balance and Gamma Correction**

Despite having an accurate recreation of the full RGB image, the color may not completely match the real or expected color. Two steps that are frequently overlooked are color balance and gamma correction.

The first step after demosaicing is the process of color balance. In this process, the nominal average color of the image is matched to a predefined value. To start, let \( X \) represent the image captured by the camera sensor such that
\[
X_i = \begin{bmatrix}
  R_i \\
  G_i \\
  B_i
\end{bmatrix}.
\] (12)

\( M \) will represent the 3-by-3 color transformation matrix, and \( Z \) represents the linear RGB image, then
\[
X = MZ.
\] (13)
However, since the process is being designed to find the linear RGB image from the sensor-captured image, the equation then becomes

\[ Z = \Phi X, \]  

(14)

where \( \Phi = M^{-1} \), the inverse of the color correction matrix. Equation (13) does not take into account for any offset required to perform a zero-correction. Therefore, with \( \bar{X} \) defined as the average of \( X \) and \( \bar{Z} \) defined as the average of \( Z \), then the equation can be extended as

\[ X - \bar{X} = M(Z - \bar{Z}). \]  

(15)

And so

\[ X = MZ - (MZ - \bar{X}), \]  

(16)

where \( M\bar{Z} - \bar{X} \) is the zero calibration, or \( Z_{\text{offset}} \), of the linear RGB image. Combining equation (14) with a rearranged form of equation (16) so that the equation solves for \( Z \) yields

\[ Z = \bar{Z} + \Phi (X - \bar{X}) \]  

(17)

which is the final form of how to perform the color correction on the image \( X \).

Further analysis can be performed on the transformation matrix, \( M \), and its inverse, \( \Phi \). The transformation matrix, \( M \), was found by following (17) and capturing a series of digital images of a 24-color Macbeth ColorChecker chart with simulated sunlight, comparing the captured color values versus the printed values, and then generating the pseudo-inverse matrix that equates the actual and sampled values.
The normalized transformation matrix for these experiments was found to be

\[
M = \begin{bmatrix}
1 & -0.026 & -0.181 \\
-0.055 & 1 & -0.049 \\
-0.023 & 0.039 & 1
\end{bmatrix}
\]

(18)

which has a condition number of 2.4. Since the condition number is so close to one, that indicates that the internal processing performed in the camera are nearly transparent and do no inject any unexpected color shifts.

The gamma correction process takes the linear RGB image and performs a non-linear transformation [8], thus making the image appear more lifelike in print and on monitors. The gamma correction process is loosely referenced as an exponential of 2.2. In actuality, it is a piecewise process that ranges in value from minimum gamma = 1 to a maximum gamma = 2.4. Additionally, the input value must range from zero to one, thus for standard images that range from zero to 255, those values must be divided by 255 to get the proper scaling. To convert from the linear RGB values that are calculated above to sRGB values, the equation is as follows:

\[
C_{\text{srgb}} = \begin{cases} 
12.92C_{\text{linear}} & \text{if } C_{\text{linear}} \leq 0.0031308, \\
(1 + a)C_{\text{linear}}^{1/2.4} - a & \text{if } C_{\text{linear}} > 0.0031308.
\end{cases}
\]

(19)

Conversely, to convert from sRGB to linear RGB, then

\[
C_{\text{linear}} = \begin{cases} 
\frac{C_{\text{srgb}}}{12.92} & \text{if } C_{\text{srgb}} \leq 0.04045, \\
\left(\frac{C_{\text{srgb}} + a}{1+a}\right)^{2.4} & \text{if } C_{\text{srgb}} > 0.04045.
\end{cases}
\]

(20)

The Fourier analysis provides the mathematical basis and the demosaicing
algorithm serves as the primary underpinning for the generation of HDR images. Additionally, the Fourier analysis of the CFA image provides additional opportunities to improve the image processing and harness underutilized information and image content.
3 HIGH DYNAMIC RANGE IMAGING

High Dynamic Range (HDR) imaging is an applied method of image processing to capture images that contain significant detail both in darker and brighter regions of images. Standard digital and film cameras are typically low dynamic range (LDR) devices, whereas the human visual system is HDR. HDR imaging is positioned to generate images that are closer to what is seen with the human visual system. With standard LDR images, there is a trade off between the two extremes of the lighting continuum. Images that capture detail with a short exposure have more detail in the bright regions of the image, while sacrificing darker regions to noise. Images that capture detail with long exposure sacrifice saturation in brighter regions of the image for detail in darker regions. These trade-offs effect the overall image processing and photography techniques. Additionally, amateur photographers may not be aware of the limitations of their equipment or a specific technique causing overly dark or saturated images. HDR image generation is a method of stretching the brightness / darkness dynamic range so images appear to not lose detail to the darker or brighter regions. Several methods exist to generate HDR images, notably exposure bracketing, specialized hardware, and the method that was researched for this document which demonstrates comparable results to other current methods, with the added benefit of its simplicity and ease of implementation.

3.1 Background and Existing HDR Methods

In a technique known as image bracketing [6, 9, 11], multiple images of varying exposure are taken in very quick succession. The multiple images are analyzed and merged such that the images captured with a short exposure time are used to exploit brighter regions of the image without the risk of saturation while rejecting
darker, noisier regions. The images captured with a longer exposure time are used to exploit darker regions of the image while rejecting brighter, possibly saturated regions. In this way, all parts of the image are properly exposed.

A major shortcomings of the image bracketing method is the inability to capture moving objects. Movement of the camera or of the target will cause blurring, shadowing, or even doubled images when the array of images are merged. An additional step of image registration [14] can alleviate some of the issues if it is a minor movement of the camera, at the expense of additional processing. However, dramatic changes, for example a person walking into the scene, can effectively ruin the image. Again, additional processing can minimize the effects.

Another method of generating HDR images is to utilize specialized hardware to capture the HDR image in one shot. Single-shot, HDR capture is desirable since, especially for the amateur photographer, this makes for easy image creation. Additionally, with one shot, blur risks are minimized. Current methods of implementation use multiple sensors to effectively capture multiple images at the exact same time. With these methods, image bracketing or harnessing the power of the individual sensors can be used to greatest effect since there is no chance of movements and the processing can be optimized for each sensor. As a negative, there are three cameras tied together with additional processing so the cost is at least three times that of a standard camera.

### 3.2 HDR Image Analysis in CFA Sampled Images

The analyses of a CFA sampled image that were detailed in the previous section did not assume that pixels are saturated. To extend the analysis to the case of HDR images, care needs to be taken to understand the interactions between the color channels.
Understanding that there is high probability of aliasing between the luminance and the chrominance 2 components indicates that successful image recreation may not always be possible. Overlapping Fourier support between the luminance and chrominance 1 contents are possible, but less likely. If aliasing occurs, the image is not completely lost; other methods need to be explored. As was observed in the previous section, there are details contained within the chrominance 1 and chrominance 2 components, although the high frequency content is not included. This is obvious when looking at Figure 7. The luminance has the most high frequency content, thus it has the most detail of the three component images. Additional care must be taken to properly isolate the chrominance 1 component from the luminance component, and when that is done, the process of demosaicing can occur with some additional follow-on step for the HDR image generation.

Our method of generating an HDR image with a single input image begins again by analyzing a standard RGB image. In more specific terms, the image can be represented as

\[ X_n = \begin{bmatrix} R_n \\ G_n \\ B_n \end{bmatrix} \]  

(21)

where \( n \) represents any pixel in the image. This can then be decomposed further to be

\[ X_n = X_n^{LP} + X_n^{HP} \]  

(22)

where the low-pass component of \( X_n \) is given by

\[ X_n^{LP} = \begin{bmatrix} R_n^{LP} \\ G_n^{LP} \\ B_n^{LP} \end{bmatrix} \]  

(23)
and the high-pass component of \( X_n \) is given by \( X_{n}^{HP}[1 1 1]^T \) because, keeping in mind that the high frequency content is likely to be similar to neutral in color, \( R_n^{HP} = G_n^{HP} = B_n^{HP} \). With that, the image can be represented as

\[
X_n = \begin{bmatrix} R_n^{LP} \\ G_n^{LP} \\ B_n^{LP} \end{bmatrix} + X_{n}^{HP} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{24}
\]

where \( X_{n}^{HP} \) becomes a scalar.

Referring back to equation (13), another form of the image can be presented as

\[
X_n = M Z_n \tag{25}
\]

where \( X_n \) is the sensor image and \( Z_n \) is in the camera space. Solving for \( Z_n \) and putting it in the form of (22) yields

\[
Z_n = M^{-1} X_n^{LP} + M^{-1} X_{n}^{HP} \tag{26}
\]

Therefore, combining (26) and (24) leads to

\[
Z_n = M^{-1} \begin{bmatrix} R_n^{LP} \\ G_n^{LP} \\ B_n^{LP} \end{bmatrix} + M^{-1} X_{n}^{HP} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{27}
\]

where

\[
Z_n^{LP} = M^{-1} \begin{bmatrix} R_n^{LP} \\ G_n^{LP} \\ B_n^{LP} \end{bmatrix} \tag{28}
\]
and

\[ Z_n^{HP} = X_n^{HP} M^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \] (29)

Reviewing the content of (29) and knowing that \( M^{-1} \) is the color conversion matrix from the sensor image to the camera space, then when performing the matrix multiplication \( M^{-1}[1 \ 1 \ 1]^T \), the values generated can be regarded as the color brightness factor for each of the red, green, and blue color components, respectively. In specific terms, this leads to

\[ Z^{HP} = X^{HP} [\vec{N}_R + \vec{N}_G + \vec{N}_B] \] (30)

where

\[ \vec{N}_R = M^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \] (31)

\[ \vec{N}_G = M^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, and \] (32)

\[ \vec{N}_B = M^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \] (33)
Converting from vector notation of $\vec{N}_R$, $\vec{N}_G$, and $\vec{N}_B$, to matrix notation using $K_R$, $K_G$, and $K_B$, respectively, is presented as

$$Z^{HP} = X^{HP} \begin{bmatrix} K_R \\ K_G \\ K_B \end{bmatrix}$$  \hspace{1cm} (34)$$

The white balance process that occurs in the camera pipeline is denoted by a diagonal matrix, $W$, such that the values $W_R$, $W_G$, and $W_B$ represent the red, green, and blue white balance matrix elements, respectively. Additionally, the matrix, $P$, is the conversion matrix that is in (10). Building off of (10) and substituting in $Z_n$ produces

$$\begin{bmatrix} L \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} W_R \\ W_G \\ W_B \end{bmatrix} = \begin{bmatrix} Z^{LP} + X^{HP} \begin{bmatrix} K_R \\ K_G \\ K_B \end{bmatrix} \end{bmatrix}. \hspace{1cm} (35)$$

Focusing only on the high frequency components leads to

$$\begin{bmatrix} L^{HP} \\ C_1^{HP} \\ C_2^{HP} \end{bmatrix} = X^{HP} PW \begin{bmatrix} K_R \\ K_G \\ K_B \end{bmatrix} = X^{HP} P \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = X^{HP} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \hspace{1cm} (36)$$

which, by the very properties of $P$, forces all of the high frequency content into the luminance component, thus reinforcing the findings in the previous section. This obviates the optimal value for $K_R$, $K_G$, and $K_B$ to be

$$K_R = W_R^{-1} \hspace{1cm} (37)$$

$$K_G = W_G^{-1} \hspace{1cm} (38)$$
\[ K_B = W_B^{-1} \]  

such that, finally,

\[
\begin{bmatrix}
W_R \\
W_G \\
W_B
\end{bmatrix}
\begin{bmatrix}
K_R \\
K_G \\
K_B
\end{bmatrix} =
\begin{bmatrix}
1 \\
1 \\
1
\end{bmatrix}
\]  

\[ (40) \]

### 3.3 Theoretical Approach to HDR Imaging

For the experimental purposes that follow, the raw images were generated by down-sampling standard existing RGB images to verify that the down-sampling process provides an adequate starting point, and upon completing the interpolative demosaicing process, the output image favorably compares to the original input RGB image.

Based on the above analysis, we propose a unique method of HDR image generation that utilizes only one shot to capture and generate an HDR image. With a minor modification to the firmware internal to a typical digital camera, this method could be easily be implemented using current digital camera hardware. As can be seen in Figure 1, one of the first steps performed in the existing digital camera pipeline is to perform a white balance which amplifies the red and blue channel components to match the intensity of the green channel components. The underlying combination of the camera spectral responsivity, the transmittances of the different color filters, and the scene content can lead to different signal levels between the color channels. In performing the white balance, if there are any regions in the green channel that are already saturated, then the red and blue channels will be amplified to the level of equal saturation. Additionally, since red and blue channels have an overall lower magnitude than the green color channel, this step of white balance also introduces amplified noise, thus increasing the total noise content of these color channels. If this white balance step were skipped and
only the green color channel were saturated, most information that is saturated could be recovered by observing the red and blue color channels. Conversely, with respect to the Bayer CFA, there are twice as many green samples as there are red or blue samples, thus, the red or blue channels cannot completely recreate information that is missing from the green channel.

The method we propose for single-shot HDR image generation corrects saturated image content by borrowing across the image’s spectral content to fill in the lost data.

Referring back to (34), the optimal form of the high pass color content can then be stated to be

\[ X^{HP} = \frac{Z_{R}^{HP}}{K_{R}} + \frac{Z_{G}^{HP}}{K_{G}} + \frac{Z_{B}^{HP}}{K_{B}} \]  \hspace{1cm} (41)

where \( K_{R}, K_{G}, \) and \( K_{B} \) are the respective color filter transmittances and act as a color balancing coefficient. Again, (41) makes the assumption that there is no saturation in the color channels. However, if the green color channel is saturated
as is presented in Figure 9a, meaning the direct reconstruction information is not available, then (36) becomes

\[
P \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ 0 \end{bmatrix}
\]

(42)

which indicates that the chrominance 2 portion remains unaffected, but the luminance and chrominance 1 portion are altered. We propose that when this occurs, the green channel portion is discarded and the high pass color content then becomes

\[
\hat{Z}_{HP} = \frac{Z_{HP}^{R}}{K_{R}} + \frac{Z_{HP}^{B}}{K_{B}}
\]

(43)

where \( \hat{Z}_{HP} \) represents the estimated high pass color content of \( Z \) and the estimated low pass color content for the green channel becomes

\[
\hat{Z}_{LP}^{G} = \frac{Z_{LP}^{R}}{K_{R}} + \frac{Z_{LP}^{B}}{K_{B}}
\]

(44)

Performing this modification to the combinations of color channels has the practical effect of relaxing the constraint on the chrominance 1 component, as shown in Figure 9b, thus improving the reconstruction. Therefore, recombining the portions of the image and eliminating green saturation yields

\[
\hat{Z}_{HP} = \frac{Z_{HP}^{R}}{K_{R}} + \left[ \frac{Z_{HP}^{R}}{K_{R}} + \frac{Z_{HP}^{B}}{K_{B}} \right] + \frac{Z_{HP}^{B}}{K_{B}}
\]

(45)

which can then be applied to (27) to generate an HDR image that also has eliminated saturation that has occurred in the green color channel.

Since the white balance process is built into all digital cameras and cannot be disabled, the first step to proving the feasibility of single shot HDR image generation is to capture and create a non-white-balanced representative image. Since
the red and blue color samples, even in the raw image format, have been balanced to match the intensity of the green samples, an unmodified version of each color channel must be generated artificially. To simulate this, the image must be processed to appear as if it were captured at different camera f-stops, thus appearing darker or brighter depending on the number of camera f-stops selected. For this, the HDRShop image processing tool [20] is used to adjust the appearance of images for faster or slower relative exposure, and then the images are saved for later use. Despite being a useful tool, it does inject a nonlinear color adjustment, which does appear as a green tint to Figure 10c. Since the blue pixels have the lowest signal level or lowest digital count with respect to the combined observed scene, transmittance of the color filter, and the spectral responsivity of the focal plane array in the camera, the RGB image that will provide the blue layer is simulated to be arbitrarily at camera f-stop -1, or one f-stop faster shutter speed than the base image. Thus, that image will appear darker than the base input image. Red pixels have the next higher signal level or next higher digital count, so the RGB image that will provide the red layer is simulated to be arbitrarily at camera f-stop 0, or equal to the base image. Green pixels typically have the highest signal level or highest digital count, thus the RGB image that will provide the green color channel is simulated to have been arbitrarily captured at camera f-stop 2, or two steps slower shutter speed than the base image. This third image is the brightest of the three. Taking the respective red, green, or blue layer from these three manufactured RGB
images thus produces a composite image that has a very dark blue channel layer, a medium intensity red channel layer, and a very bright and possibly saturated green channel layer. The three representative images are shown in Figure 8, signifying the red, green, and blue content of the final manufactured image.

Using the non-amplified representative image, the representative image is down-sampled and demosaicked as done in the previous section. Once the chrominance 1, chrominance 2 and luminance components are generated, the image is reconstructed as previously discussed. However, in this case, the reconstructed green layer is masked such that at the highest intensity level, when saturation most likely occurs, the mask is zero. Similarly, when the reconstructed green layer is at the lowest intensity level, when saturation is least likely to occur, the mask is one. Between those two end points, the mask is a linear transition from maximum to minimum based on the intensity of the reconstructed green layer. To generate the final green layer for the HDR composite image, the mask is multiplied by the reconstructed green layer and the inverse mask is multiplied by an average of the
reconstructed red and reconstructed blue layers, or

\[ Z_{G}^{HP} = (\text{greenmask} \times Z_{G}^{HP}) + \left( (1 - \text{greenmask}) \times \frac{(Z_{R}^{HP} + Z_{B}^{HP})}{2} \right). \]  

(46)

This eliminates saturation that occurs in the green color channel while filling in saturated information with a scaled average of the red and blue color channels. Figure 10 shows a comparison between the original base image, a highly saturated image, and the lightly color-corrected HDR image as generated using this algorithm. A final additional step of color and gamma correction must be performed to insure that the output image has a realistic color content that is not skewed towards red or blue since the original tri-stimulus input intensities are essentially ignored. The amount of color correction required will be dependent upon the chosen CFA/sensor pair and can easily be tuned to the end-user’s preference.
4 MULTISPECTRAL IMAGING

With standard digital imagery, there are typically three colors – red, green, and blue – that form three individual layers of the resultant image. Those three colors are optimized to capture and display images in a similar manner to human vision. However when greater detail is needed, additional layers can be added focusing on different bands of wavelengths in the visible spectrum. Multispectral imaging [17] is the name given to images that contain more than the standard three colors, and hyperspectral imaging refers to images that contain so many layers sampled at discrete wavelengths that the effective wavelengths captured tends toward an analog spectrum, not merely a handful of discrete samples or wide bandwidths of the visible spectrum.

4.1 Spectral Imaging Background

For a standard RGB image, the image is generated using a Charge Coupled Device (CCD) that has a tri-stimulus Bayer CFA overlaid to provide wavelength selectivity. This spacial separation allows for the demosaicing of the image into the three color channels. Finally, color balance and gamma correction are performed, as seen in Figure 1. However, for hyperspectral imaging, there is a need for many more image layers to capture as much information as possible, with each layer representing an individual, or very narrow, wavelength of visible light. There is no equivalent CFA designed to quickly and efficiently generate a 31 layer, or larger, multispectral image. A CFA design is not a mere extension of the standard tri-stimulus model, and global analysis of multispectral imagery is needed to understand the spatial and spectral correlations.
4.2 Spectral Image Analysis

Similar to the approaches presented for standard RGB and HDR images, the information content of a multispectral image can be analyzed using frequency domain techniques. Consider a three dimensional discrete Fourier transform applied to an LxMxN image, where L and M are the length and width of the image and N is the number of discrete sampled spectral slices, as shown in Figure 11. Of particular note is that the spatial bandwidth is high near spectral DC while the the spatial bandwidth is low for spectral high-pass. Conversely, the spectral bandwidth is high near spatial DC while the spectral bandwidth is low for spatial high-pass.

Let \( x(\lambda, \vec{n}) \) be the three dimensional spatial-spectral signal and \( s(\lambda, \vec{n}) \) represents a sampling function in both spatial and spectral domains, where \( \lambda \) is the spectral component and \( \vec{n} \) is the spatial location vector. With that, the observed sensor data, \( y(\vec{n}) \) can be written as

\[
y(\vec{n}) = \int x(\lambda, \vec{n}) \, s(\lambda, \vec{n}) \, d\lambda.
\]  \hspace{1cm} (47)
If \( m(\lambda, \vec{n}) = x(\lambda, \vec{n}) s(\lambda, \vec{n}) \), then from the convolution theorem it can be stated that

\[
M(\omega_\lambda, \omega_\vec{n}) = X(\omega_\lambda, \omega_\vec{n}) * S(\omega_\lambda, \omega_\vec{n})
\]  

(48)

which is the Fourier transform of \( m(\lambda, \vec{n}) \), and \( \omega_\lambda \) is the spectral frequency term and \( \omega_\vec{n} \) is the spatial frequency vector. From the inverse Fourier transform, it can be stated that

\[
m(\lambda, \vec{n}) = \iint M(\omega_\lambda, \omega_\vec{n}) e^{j(\omega_\lambda \lambda + \omega_\vec{n}^T \vec{n})} d\omega_\lambda d\omega_\vec{n}.
\]  

(49)

Substituting equation (49) into (47) yields

\[
y(\vec{n}) = \iiint M(\omega_\lambda, \omega_\vec{n}) e^{j(\omega_\lambda \lambda + \omega_\vec{n}^T \vec{n})} d\omega_\lambda d\omega_\vec{n} d\lambda.
\]  

(50)

Rearranging the terms produces

\[
y(\vec{n}) = \iint e^{j\omega_\lambda \lambda} d\lambda M(\omega_\lambda, \omega_\vec{n}) e^{j\omega_\vec{n}^T \vec{n}} d\omega_\lambda d\omega_\vec{n}.
\]  

(51)

It is known that

\[
\delta(\omega_\lambda - 0) = \int e^{j\omega_\lambda \lambda} d\lambda,
\]  

(52)

which lead to the following

\[
y(\vec{n}) = \iiint \delta(\omega_\lambda - 0) M(\omega_\lambda, \omega_\vec{n}) e^{j\omega_\vec{n}^T \vec{n}} d\omega_\lambda d\omega_\vec{n}
\]  

(53)

and finally

\[
y(\vec{n}) = \int M(0, \omega_\vec{n}) e^{j\omega_\vec{n}^T \vec{n}} d\omega_\vec{n}.
\]  

(54)

From (48) and by the definition of convolution, we have

\[
M(\omega_\lambda, \omega_\vec{n}) = \iiint X(\nu_\lambda, \nu_\vec{n}) S(\nu_\lambda, \nu_\vec{n}) d\nu_\lambda d\nu_\vec{n}.
\]  

(55)
Substituting equation (55) into (54) yields

\[ y(\vec{n}) = \iiint X(\nu_\lambda, \nu_n) S(0 - \nu_\lambda, \omega_n - \nu_n^* e^{j \omega_n^* T \vec{n}} d\nu_\lambda d\nu_n d\omega_n \]  \hspace{1cm} (56)

which can be extended into analysis of any filters.

One particular case of interest is an idealized narrow band measurement of the wavelength. This is the case representative of putting color filters over the lens of a panchromatic camera. As the measurements become narrower, it can be approximated as a periodic impulse function in wavelength, such that \( s(\lambda, n) = \delta(\lambda - \lambda_o) \), then for the idealized case becomes

\[ S(\omega_\lambda, \omega_n) = \sum_n \delta(\lambda - \lambda_o) e^{-j \omega_\lambda \lambda_o} e^{-j \omega_n^* T \vec{n}} d\lambda \]  \hspace{1cm} (57)

\[ S(\omega_\lambda, \omega_n) = \sum_n \delta(\lambda - \lambda_o) e^{-j \omega_\lambda \lambda_o} e^{-j \omega_n^* T \vec{n}} d\lambda \]  \hspace{1cm} (58)

\[ S(\omega_\lambda, \omega_n) = \sum_n e^{-j \omega_\lambda \lambda_o} e^{-j \omega_n^* T \vec{n}} \]  \hspace{1cm} (59)

and finally

\[ S(\omega_\lambda, \omega_n) = e^{-j \omega_\lambda \lambda_o} \delta(\omega_n^*) \]  \hspace{1cm} (60)

which contains the spatial DC term \( \delta(\omega_n^*) \). Substituting equation (60) into equation (56) yields

\[ y(\vec{n}) = \iiint X(\nu_\lambda, \nu_n) e^{-j \nu_\lambda \lambda_o} \delta(\omega_n^*) e^{j \nu_n^* T \vec{n}} d\nu_\lambda d\nu_n d\omega_n \]  \hspace{1cm} (61)

where \( \nu_\lambda \) represents points along the spectral reference axis and \( \nu_n^* \) represents vectors along the spatial axis in the Fourier domain. Further manipulation leads to

\[ y(\vec{n}) = \iiint X(\nu_\lambda, \nu_n) e^{-j \nu_\lambda \lambda_o} e^{j \nu_n^* T \vec{n}} d\nu_\lambda d\nu_n \]  \hspace{1cm} (62)
Comparing equation (62) to (49), it becomes obvious that this is an inverse Fourier transform which therefore produces a series of three-dimensional spectral contours in the Fourier domain that are repeated periodically along the reference plane, with likely overlap, which would ultimately cause aliasing. Figure 12 presents a representation of this situation. The shaded area of the contour is the spectral content of the image, similar to that shown in Figure ??, and with DC being the center slice. These spectral contours are only repeated along the spectral axis, $\omega_\lambda$, because the spatial component was set to zero for this case of interest. Of particular note is the amount of wasted spectral area that could be used to fit additional samples of the spectral contour.

4.3 Theoretical Approach to Spectral Imaging

As an alternative to the single wavelength image acquisition, consider a spectral filter array

$$s(\lambda, \vec{n}) = e^{j\theta_\lambda \lambda + j\theta_n T \vec{n}}.$$  \hspace{1cm} (63)
In this case, the sampling function is a three-dimensional spatial sinusoid. Taking a discrete time Fourier transform of \( s(\lambda, \vec{n}) \) then yields

\[
S(\omega_\lambda, \vec{\omega}_n) = \sum_n \int s(\lambda, \vec{n}) e^{-j\omega_\lambda \lambda} e^{-j\vec{\omega}_n \cdot \vec{n}} d\lambda
\]  

(64)

where \( \vec{n} \) again represents the spatial location vector and \( \vec{\omega}_\lambda \) is the spatial frequency vector. Continuing with the substitution leads to

\[
S(\omega_\lambda, \vec{\omega}_n) = \sum_n \int e^{j\theta_\lambda \lambda + j\theta_n \cdot \vec{n}} e^{-j\omega_\lambda \lambda} e^{-j\vec{\omega}_n \cdot \vec{n}} d\lambda
\]  

(65)

and finally

\[
S(\omega_\lambda, \vec{\omega}_n) = \delta(\theta_\lambda - \omega_\lambda, \theta_n - \vec{\omega}_n)
\]  

(66)

where \( \omega_\lambda \) and \( \vec{\omega}_n \) index Fourier coefficients. Substituting (66) into (56) yields

\[
y(\vec{n}) = \iiint X(\nu_\lambda, \nu_n) \delta(\theta_\lambda - 0 + \nu_\lambda, \theta_n - \vec{\omega}_n + \nu_n) e^{j\vec{\omega}_n \cdot \vec{n}} d\nu_\lambda d\nu_n d\vec{\omega}_n
\]  

(67)

\[
y(\vec{\omega}_n) = x(-\theta_\lambda, \vec{\omega}_n - \theta_n).
\]  

(68)

In other words, sampling with \( s(\lambda, \vec{n}) = e^{j\theta_\lambda \lambda + j\theta_n \cdot \vec{n}} \) has the effect of measuring \( X \) at a particular spectral frequency \( \theta_\lambda \), where \( \theta_\lambda \) is a Fourier spectral slice. Additionally, \( X \) is spatially modulated by spatial frequency \( \theta_n \). An example of this can be seen in Figure 13a, where complex conjugate point pairs are injected into an LxMxN image and processed through an inverse 3D Fourier transform. In the spectral domain though, this has the effect of additional spectral samples that fall into unoccupied spectral area, as seen in Figure 13c.

In the experiments performed for this thesis, images that contained 31 layers were analyzed, one layer for each narrow-band wavelength from 420 to 720 nm, with 10nm separation between each sample. This is a far greater number of sampled wavelengths than RGB and the contiguous wavelengths over the spectral
pass band are sampled at sufficient spectral resolution, thus these images can be regarded as hyperspectral images. Analysis and processing of multispectral images can follow in the same manner, however there is a physical limitation to the number of sampled narrow band wavelengths possible for a realizable hyperspectral imaging solution.

The visual interpretation of multispectral data is not possible directly, and as such automated techniques are usually employed to extract information from multispectral imagery [21]. For this work, the interest is in analysis of multispectral imaging through the use of color filter arrays and, as a first step, a three color RGB conversion is performed. Thus the analysis of the experimental multispectral images necessitates the development a 31-to-3 multispectral-to-RGB conversion process so that the image can be displayed on a standard monitor via the inner product with a color matching function. Keeping in mind that each of the 31 individual layers of the multispectral image represent a discrete wavelength of visible light, the color matching function essentially provides the proper amount of scaling for each layer to properly map the discrete wavelengths to the standard tri-stimulus model. Figure 14a shows the RGB projection of the multispectral image.

The multispectral image analysis is similar to the analysis performed on standard RGB images and HDR images, and that is to verify that the spectral content
of the image matches the theorized conical shape shown in Figure 11. Keeping in mind that a Fourier transform of the original multispectral image has 31 spectral layers, a method of analysis includes either projecting the original image to two dimensions and observing the Fourier transform of that 2D image, or the Fourier transform can be collapsed to two dimensions. If the multispectral image were converted to RGB and then down-sampled to a CFA image, information would be lost in the compression process. A Fourier transform is then performed on this 2D image, and generates an image that can be analyzed. The output of this experiment is shown in Figure 14b, and it can be clearly observed that around the two center axes that there is a conical shape moving along the axis.

A full 3D Fourier transform is performed on the multispectral image, which generates a 31 wavelength spectral image. Using the same 31-to-3 multispectral to RGB color matching function, a full color spectral image is generated that is similar to Figure 3b. The RGB approximation of the full 3D Fourier transform of the original multispectral image can be seen in Figure 14c. This image, similar in effect to the images shown in Figure 3a such that it is an the spectral content is being views along an axis instead of an idealized projection, as is the case with Figure 14b. Analyzing 2D projections of a 3D object, though, can only provide limited information.
Knowing that the spectral samples can be placed at optimum locations simply by the choice of the proper $S(\omega_\lambda, \omega_n)$ opens the door to developing an multispectral CFA. Pursuing this further, we know that when $\omega_\lambda = 0$, we are observing the DC component of $X$. Furthermore, since $M = X \ast S$, we can say that $X$ is sampled one spectral Fourier slice at a time. Thus, with proper selection of $S(\omega_\lambda, \omega_n)$, we can build an idealized 2D representation of a 3D Fourier transform for multispectral images. An example of this multispectral idealized Fourier transform is shown in Figure 15. The rearranged circles and additional gray circles represent the other spectral Fourier slices that would be sampled, in this case a nine channel array, thus filling in the spectral area.

The argument can be made that we have enough information that we can choose $S(\omega_\lambda, \omega_n)$ so that we can dictate which Fourier spectral slice to observe. Knowing that we require equivalent performance to sampling at $\lambda = \Delta_\lambda k$, where $\Delta_\lambda$ is the bandwidth of the sampled wavelength and $k$ is the specific wavelength, that the set of wavelengths fall between 400 and 800 nm, and then referring to (67) and (68), using the discrete time Fourier transform, we can state that by taking the discrete Fourier transform

$$X(\omega_\lambda) = \sum_k \int x(\lambda) \delta(\lambda - \Delta_\lambda k) e^{j\omega_\lambda \lambda} d\lambda$$  \hspace{1cm} (69)
\[ X(\omega_\lambda) = \sum_k x(\Delta \lambda, k) e^{-j\omega_\lambda \Delta \lambda k} \]  
(70)

We know we have \( N \) spectral samples, thus we only need to observe the \( N \) discrete Fourier coefficients, which is presented as

\[ X(\ell) = \sum_k x(\Delta \lambda, k) e^{-j \frac{2\pi \ell k}{N}} \]  
(71)

Therefore, solving for \( \omega_\lambda \) leads to

\[ e^{-j \omega_\lambda \delta \lambda k} = e^{-j \frac{2\pi \ell k}{N}} \]  
(72)

which leads directly to

\[ \omega_\lambda = \frac{2\pi}{800 - 400nm} \ell \]  
(73)

\[ \omega_\lambda = \frac{\pi}{200} \ell \]  
(74)

This leads directly to a multispectral CFA, a direct combination of Figure 13c and Figure 15 such that each spectral Fourier slice from 1 to \( N \), where \( N \) is greater than 3 but less than 10, is sampled without aliasing any other spectral Fourier slice, and so each spectral Fourier slice in the full 3D spectral content, shown in Figure 11 is accounted for and can be reconstructed.
5 CONCLUSIONS

In the previous chapters, performing Fourier analysis on digital images was shown to provide useful information regarding the spatial-spectral content. Through analysis and simulation, it was demonstrated that there is underutilized information contained in the spectral content of image. The algorithm we propose is based on the Fourier analysis and harnesses the underutilized information in images to generate single-shot HDR images. The experimental results prove that borrowing information across spectral channels is an efficient method of eliminating saturation in the green color channel while generating HDR images. This method demonstrates moderate success and points to a simplified process of generating HDR images, possibly even opening the technology to lower-end digital cameras. The necessary embedded processing is available and the overall technology is mature enough that this becomes a refinement in implementation.

Further pursuing the Fourier analysis into multispectral and hyperspectral images, the method of designing a scalable multispectral CFA that we propose takes into account the spatial-spectral relationships contained within the images. We successfully leveraged Fourier-based spatial-color relationship that exists in standard RGB images and successfully extend the theory to include our multispectral test images. Using these methods, spectral filter arrays of varying sizes can be designed and implemented as needed. Additional research is required to optimize the design of the spectral filter array and related demosaicing strategies.
REFERENCES


