Empirical Satellite Characterization for Realistic Imagery Generation

Nilay Vijay Mokashi
nm6794@rit.edu

Follow this and additional works at: https://scholarworks.rit.edu/theses

Recommended Citation

This Thesis is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
EMPIRICAL SATELLITE CHARACTERIZATION FOR REALISTIC IMAGERY GENERATION

by

Nilay Vijay Mokashi

B.E., Savitribai Phule Pune University, India, 2016

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Chester F. Carlson Center for Imaging Science Rochester Institute of Technology

June 08, 2018

Signature of the Author

Accepted by

Coordinator, M.S. Degree Program
The M.S. Degree Thesis of Nilay Vijay Mokashi has been examined and approved by the thesis committee as satisfactory for the thesis requirement for the M.S. degree in Imaging Science.

Dr. John P. Kerekes, Thesis Advisor

Dr. Michael Gartley

Dr. Zoran Ninkov

Date ____________________________
Empirical Satellite Characterization for Realistic Imagery Generation

by

Nilay Vijay Mokashi

Submitted to the
Chester F. Carlson Center for Imaging Science
in partial fulfillment of the requirements
for the Master of Science Degree
at the Rochester Institute of Technology

Abstract

There is an increasing interest in the use of machine learning deep networks to automatically analyze satellite imagery. However, there are limited annotated satellite imagery datasets available for training these networks. Synthetic image generation offers a solution to this need, but only if the simulated images have comparable characteristics to the real data. This work deals with analysis of commercial satellite imagery to characterize their imaging systems for the purpose of increasing the realism of the synthetic imagery generated by RIT’s Digital Imaging and Remote Sensing Image Generation (DIRSIG) model.

The analysis was applied to satellite imagery from Planet Labs and Digital Globe. Local spatial correlation was leveraged for noise estimation and the EMVA1288 standard was used for noise modeling. Real world calibration targets across the world were used together with the slanted edge method based on the ISO 12233 standard for estimation of the sensor optical systems’ point spread function (PSF). The estimated camera models were then used to generate synthetic imagery using DIRSIG. The PSF was applied within DIRSIG using its in-built functionality while noise was added in post processing. Analysis similar to real imagery was performed on the simulated scenes to verify the application of the model on synthetic scenes. Future work is recommended to further characterize the various imagery products produced by the satellite companies to better represent artifacts present in these processed images.
Acknowledgments

I find myself extremely fortunate and grateful to have been a part of the graduate program at the Center for Imaging Science at Rochester Institute of Technology. While camera technology always fascinated me, it was at the CIS that I learnt all the intricate details of this fascinating world of imaging. The program was a mesmerizing journey for me, providing me with the skillset and tools to study and analyze the plethora of information that digital images bring to us. Everyone at the CIS felt like a family member helping me grow as a scientist and achieve my career goals.

Many people directly or indirectly helped me in my research. I was fortunate to have Dr. John Kerekes as my thesis advisor. His continuous support and inspiration were key factors helping me throughout this work. Apart from technical knowledge, I believe he played a vital role in shaping me as a professional entering the industry. I would like to thank Kitware Inc. for funding this research. I don’t have enough words to thank DIRS lab and especially Jared van Cor, DIRS staff member, who helped me in this work by tirelessly running endless number of iterations in DIRSIG. My thesis committee members, Dr. Zoran Ninkov and Dr. Michael Gartley, were kind to give me invaluable tips and suggestions to better this work. I would also like to thank Dr. Jan van Aardt for helping me during the tough times. While the above mentioned people played a crucial role in my thesis work, I believe the entire CIS family has contributed to the completion of my master’s degree.

Last but not the least, I would like to thank my family, my parents and brother, who have been a constant source of motivation for me. Without their unwavering support, I would not have landed in the U.S. to pursue my dreams.
# Contents

List of Figures ................................................................. vii  
List of Acronyms .............................................................. ix  

1 Introduction and Objectives ............................................. 1  

1.1 Introduction ............................................................... 2  
1.2 Objectives ............................................................... 4  
1.3 Thesis Overview ......................................................... 4  

2 Background and Related Theory ........................................ 5  

2.1 Satellite Imagery Overview ........................................... 6  
2.2 Camera Systems .......................................................... 7  
  2.2.1 Optical Sub-System .................................................. 7  
  2.2.2 Electronic Sub-System ............................................... 9  
2.3 Description of Imageries Used ....................................... 11  
  2.3.1 Digital Globe .......................................................... 11  
  2.3.2 Planet Labs ............................................................ 14  
2.4 DIRSIG .......................................................... 15  
2.5 EMVA 1288 Standard ..................................................... 16  
2.6 Slant Edge Method for MTF ............................................. 20  

3 Approach ............................................................................. 25  

3.1 Method for Noise Estimation ........................................... 26  
3.2 Method for PSF Estimation ............................................. 29  
3.3 Data Used ................................................................. 32  
  3.3.1 Digital Globe .......................................................... 33  
  3.3.2 Planet Labs ............................................................ 34  
3.4 Targets Used ............................................................. 34  
3.5 Process for Generating Simulated Scene ............................. 36  
3.6 Verification of Models ................................................... 38  

4 Results and Discussion .................................................... 40  

4.1 Noise Model Estimation ................................................ 41  
  4.1.1 WorldView-2 Real Data ............................................. 41  
  4.1.2 WorldView-2 Simulated Data .................................... 46  
  4.1.3 PlanetScope Real Data ............................................. 47  
  4.1.4 PlanetScope Simulated Data ..................................... 52  
  4.1.5 Summary of Noise Model Results ............................... 53  
4.2 Point Spread Function (PSF) Estimation ............................. 54  
  4.2.1 PlanetScope Real Data ............................................. 59  


5 Conclusion and Future Work 69
5.1 Conclusion 69
5.2 Future Work 71
# List of Figures

2.1 A Simple Cassegrain Telescope. ................................. 7
2.2 Optical System of GaoFen4 Satellite. .......................... 8
2.3 Spectral Response of WorldView-2. ............................. 12
2.4 Simplified DIRSIG Flowchart. .................................. 16
2.5 Camera Physical Model. .......................................... 17
2.6 Noise Sources In A Camera. ..................................... 17
2.7 MTF Explanation. .................................................. 20
2.8 Relationships between ESF, LSF and MTF. ................. 21
2.9 Generation of Finely Sampled LSF. .............................. 23
2.10 ISO 12233 Method for SFR Calculation. ...................... 24

3.1 Digital Globe Radiometric Calibration. ......................... 26
3.2 Projecting Pixels Onto An Edge. ................................. 30
3.3 Undersampled Edge Responses Measured Across Tilted Edge. .................................................. 31
3.4 Sample Panchromatic Image Taken By WorldView-2. .... 33
3.5 Sample Color Image Taken By PlanetScope. .................. 34
3.6 Calibration Target at CalVal, Shadnagar, India. .......... 35
3.7 Calibration Target at Boutou, China. .......................... 36
3.8 TRONA Scene. .................................................... 39
3.9 Synthetic PSF Target Generation. ............................... 39

4.1 Scene Used For WorldView-2 Noise Analysis. ............... 41
4.2 Noise vs Signal For WorldView-2. ............................. 42
4.3 Noise² vs Signal For WorldView-2. ............................ 42
4.4 SNR vs Signal for WorldView-2. ............................... 43
4.5 Histogram For WorldView-2. .................................... 43
4.6 Cumulative Histogram For WorldView-2. ..................... 44
4.7 Noise² vs Signal Best-Fit Line for WorldView-2. .......... 45
4.8 DIRSIG Simulated Scene on WorldView-2 Platform. .... 46
4.9 Zoomed View of WorldView-2 Simulated Scene. .......... 46
4.10 Scene Used for Analysis of PlanetScope Imagery. ........ 47
4.11 Noise vs Signal for PlanetScope. ............................. 48
4.12 Noise² vs Signal for PlanetScope. ........................... 48
4.13 SNR vs Signal for PlanetScope. ............................... 49
4.14 Histogram for PlanetScope. ................................... 49
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.15</td>
<td>Cumulative Histogram for PlanetScope.</td>
<td>50</td>
</tr>
<tr>
<td>4.16</td>
<td>Noise$^2$ vs Signal Best-Fit Line for PlanetScope.</td>
<td>51</td>
</tr>
<tr>
<td>4.17</td>
<td>DIRSIG Simulated Scene on PlanetScope Platform.</td>
<td>52</td>
</tr>
<tr>
<td>4.18</td>
<td>Zoomed View of PlanetScope Simulated Scene.</td>
<td>52</td>
</tr>
<tr>
<td>4.19</td>
<td>Sampling of PSF.</td>
<td>54</td>
</tr>
<tr>
<td>4.20</td>
<td>Diffraction Limited PSF Cut-Section.</td>
<td>55</td>
</tr>
<tr>
<td>4.21</td>
<td>Target Used for PSF Estimation.</td>
<td>56</td>
</tr>
<tr>
<td>4.22</td>
<td>Edge Profiles for Multiple Rows.</td>
<td>57</td>
</tr>
<tr>
<td>4.23</td>
<td>Super-Resolved Edge Spread Function.</td>
<td>58</td>
</tr>
<tr>
<td>4.24</td>
<td>PSF for PlanetScope.</td>
<td>59</td>
</tr>
<tr>
<td>4.25</td>
<td>Top View of PlanetScope PSF.</td>
<td>59</td>
</tr>
<tr>
<td>4.26</td>
<td>Single row undersampled ESFs for WorldView-2.</td>
<td>60</td>
</tr>
<tr>
<td>4.27</td>
<td>WorldView-2 Super-Resolved ESF.</td>
<td>61</td>
</tr>
<tr>
<td>4.28</td>
<td>WorldView-2 Sub-Pixel Sampled LSF.</td>
<td>61</td>
</tr>
<tr>
<td>4.29</td>
<td>PSF for WorldView-2.</td>
<td>62</td>
</tr>
<tr>
<td>4.30</td>
<td>Top View of WorldView-2.</td>
<td>62</td>
</tr>
<tr>
<td>4.31</td>
<td>PSF Estimated for WorldView-2 Simulated Image.</td>
<td>63</td>
</tr>
<tr>
<td>4.32</td>
<td>PSF Estimated for PlanetScope Simulated Image.</td>
<td>63</td>
</tr>
<tr>
<td>4.33</td>
<td>Comparison Between PlanetScope and WorldView-2 PSF.</td>
<td>64</td>
</tr>
<tr>
<td>4.34</td>
<td>Airy Disk Pattern.</td>
<td>65</td>
</tr>
<tr>
<td>4.35</td>
<td>Comparison of Empirical and Diffraction Limited PSF for WorldView-2.</td>
<td>66</td>
</tr>
<tr>
<td>4.36</td>
<td>Comparison of Empirical and Diffraction Limited PSF for PlanetScope.</td>
<td>67</td>
</tr>
</tbody>
</table>
### List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC</td>
<td>Analog-Digital Converter</td>
</tr>
<tr>
<td>CCD</td>
<td>Charge Coupled Device</td>
</tr>
<tr>
<td>DIRS</td>
<td>Digital Imaging and Remote Sensing</td>
</tr>
<tr>
<td>DIRSIG</td>
<td>Digital Imaging and Remote Sensing Image Generation</td>
</tr>
<tr>
<td>EMVA</td>
<td>European Machine Vision Association</td>
</tr>
<tr>
<td>ESF</td>
<td>Edge Spread Function</td>
</tr>
<tr>
<td>FOV</td>
<td>Field of View</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full Width at Half Maximum</td>
</tr>
<tr>
<td>GSD</td>
<td>Ground Sampling Distance</td>
</tr>
<tr>
<td>HWHM</td>
<td>Half Width at Half Maximum</td>
</tr>
<tr>
<td>IFOV</td>
<td>Instantaneous Field of View</td>
</tr>
<tr>
<td>ISO</td>
<td>International Standards</td>
</tr>
<tr>
<td>ISS</td>
<td>International Space Station</td>
</tr>
<tr>
<td>LSF</td>
<td>Line Spread Function</td>
</tr>
<tr>
<td>MS</td>
<td>Multispectral</td>
</tr>
<tr>
<td>MTF</td>
<td>Modulation Transfer Function</td>
</tr>
<tr>
<td>NIR</td>
<td>Near Infrared</td>
</tr>
<tr>
<td>Pan</td>
<td>Panchromatic</td>
</tr>
<tr>
<td>PSF</td>
<td>Point Spread Function</td>
</tr>
<tr>
<td>SIG</td>
<td>Synthetic Imagery Generation</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SSO</td>
<td>Sun Synchronous Orbit</td>
</tr>
<tr>
<td>TDI</td>
<td>Time Delay Integration</td>
</tr>
<tr>
<td>TOA</td>
<td>Top-of-Atmosphere</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction and Objectives
1.1 Introduction

Imaging how our planet Earth looks like when viewed from outside has always fascinated humans. Cartography probably being one of its major manifestations; albeit being restricted to imagination and inferences without having ever seen the planet from above. It all changed when Wright brothers flew the first aircraft in early 20th century. As the World Wars began aircrafts became very powerful assets because not only could you fight using them but could also get a bird’s eye view of the enemy’s territory. Inadvertently, aerial photography gained importance. However, to infer accurate data from the photographs required equally accurate calibration of the cameras.

The early uses of aerial photogrammetry can be found in mapping. One of the first references to camera calibration is that of Deville in Canada (Field, 1946) in 1910. There was a rapid increase in the aerial photographic methods employed by countries for reconnaissance and mapping beginning with the Second World War. It continued into the Cold War era with the aircrafts getting replaced by satellites for this function.

Today, we have a vast network of satellites floating above the Earth equipped with state-of-the-art cameras that help us monitor the world closely. Applications vary from reconnaissance to generating navigation maps for public to studying the changes in environment. However, considering the criticality of these applications, we need the data to have sufficient accuracy. Thus, the calibration and characterization of cameras aboard satellites becomes a significant step in extracting data from them. This work therefore deals with development of in-situ methods for modeling the imaging systems used on commercial satellites.
The subsequent part of this work deals with using generated models for generating realistic synthetic imagery. Synthetic imagery generation (SIG) is the process of generating computerized simulated images based on physical models. Building effective SIG models requires in-depth knowledge and use of the whole image chain.

SIG is used by sensor designers to pre-validate their sensor designs by using synthetic imagery to evaluate the sensor parameters with respect to the desired application. System operators use SIG to simulate the sensors of their interest to decide the time and order of the imaging system operation. Another interesting application of SIG can be found in training purposes. SIG can be used to generate background scenes to train pilots on aircraft simulators.

Another prime application of SIG is its use in algorithm development. This, in fact, forms the motivation for our work. With the introduction of machine learning and deep learning algorithms, extracting meaningful information from visual data has seen great advances recently. There are various algorithms that can be used for object detection, object recognition, change detection etc. These artificial algorithms basically generate a model that maps the raw data to meaning information. However, the algorithms need to be trained first in order to have high accuracy in their working. Training requires data similar to that would be used for testing along with the ground truths. Typically, the volume of data required for efficiently training such models is very high. But only so much data is available from commercial satellites which would include a variety of geographies, observing conditions, objects in the scene etc. SIG can be used to generate images of scenes with these variations which can later be used for training the models. To have good accuracy, the synthetic images should mimic the real satellite data as much as possible; not just
visually but also in their analytical characteristics. Having model of the camera systems used on satellites thus becomes crucial.

1.2 Objectives

The primary objective of this work is to analyze commercial satellite imagery to model the cameras used for capturing them and use it to increase the realism of the synthetically generated simulated scenes. Specific tasks include:

- Modeling noise performance of imaging systems used on commercial satellites.
- Estimating the point spread function of imaging systems used on commercial satellites.
- Using the camera model to synthesize simulated imagery using DIRSIG.
- Verification of model by analyzing simulated imagery.

1.3 Thesis Overview

Chapter 2 contains background theory necessary for understanding the significance of the work along with past work done in the related areas. Chapter 3 discusses the approach used practically for measuring the PSF and noise functions along with the description of data used. Chapter 4 is a compilation of the results obtained and discusses their implications observations made through them. Chapter 5 concludes the work and gives an insight into some ideas for future that could follow this work.
Chapter 2

Background and previous work
2.1 Satellite imagery

Companies launch their satellites equipped with imaging systems into various orbits. Perhaps the most important parameter for any aerial imaging system is its resolution or better called as Ground Sampling Distance (GSD) in the context. GSD determines the size of real world object represented by a single pixel in the satellite image. Few important parameters affecting the GSD are: focal length of the telescope, image sensor pixel pitch and the orbital altitude.

Longer focal length leads to smaller field of view and thus better angular resolution, i.e. smaller GSD. Image sensor pixel pitch determines the size of the pixel and as one would expect, smaller the pixel, better (smaller) the GSD. However, it should be noted that smaller pixel also leads to lesser light gathering in any given amount of time and thus lower SNRs. Finally, orbital altitude has an obvious effect on GSD with higher altitude corresponding to higher GSDs (assuming rest all factors are the same). The spectral filters can also alter GSD as sometimes it is necessary to perform binning in order to gather measureable signal levels. This work mainly deals with Panchromatic imagery.

Another important aspect of aerial imagery is the way images are captured. The two main variants are whisk-broom and push-broom scanners. In the whisk-broom scanner, a mirror constantly keeps scanning the area in a direction perpendicular to that of the motion of satellite. The light reflected off the mirror is collect by the image sensor, which captures Instantaneous Field of View (IFOV) image. On the other hand, in push-broom scanners an optical system is used so that light is incident upon a linear image sensor. Thus, the image is constructed one row at a time as the satellite goes on sweeping the area. While push-
broom sensors give imagery without any pixel distortions, whisk-brooms tend to have a wider swath width and thus cover larger areas.

2.2 Camera systems

Camera systems consist of two main components: optics and electronics. Optics deals with the image formation of a scene. In other words, the optics gathering and mapping photons from object-space in the real world to image-space on the sensor. While the electronics deals with capturing those photons in image-space, converting them into electrons and storing or reading them out to the system desired.

2.2.1 Optical sub-system

The optics of a camera system consists of mirrors and lenses that lead to formation of image of an object on the sensor plane. Typically, the optical systems are of Cassegrain telescope type consisting of two mirrors: a parabolic primary and a hyperbolic secondary. This helps in reducing the size of the telescope while having a long focal length. Image sensor is attached at the focal plane of the secondary mirror. Figure 2.1 shows an example of a simple Cassegrain telescope.

![Figure 2.1: A simple Cassegrain telescope. Credits: Wikipedia](image-url)
The real systems are much more complex though. Figure 2.2 below shows an example of an optical system used on the satellite GaoFen 4 built by China.

![Optical system of GaoFen 4 satellite](image)

**Figure 2.2: Optical system of GaoFen 4 satellite**

When a point source emits light, the optics of the camera focuses the light onto the image-plane. Because of the fundamental nature of light, it cannot be focused on an infinitesimally small point in the image-plane. The light waves produce a diffraction pattern at the focal point. The pattern consists of a central bright disc followed by concentric rings of bright and dark patches, when viewed from the image plane. Thus, a point object gets imaged as an extended object. The output of an imaging system given a point source as in input is known as its impulse response or Point Spread Function (PSF). Any object can be thought of as a collection of points. As the imaging systems used are of linear nature, the final image can be thought of as superposition of impulse responses of multiple points in the object-plane. Mathematically the obtained image can be represented as convolution of the scene and the PSF.

\[
g(x, y) = f(x, y) \ast h(x, y)
\]

Eq. (1)
where \( g(x,y) \) is the image, \( f(x,y) \) is the object while \( h(x,y) \) is the impulse response or PSF. As the PSF is essentially the impulse response it can be given by:

\[
h(x_0, y_0) = \delta(x - x_0, y - y_0) \ast h(x, y)
\]

Eq. (2)

Where \( \delta \) is the Dirac-Delta function which is defined to be zero everywhere except at \((x_0,y_0)\) where it has unit volume.

### 2.2.2 Electronic sub-system

The electronics of a digital imaging system consists majorly of an image sensors along with some other circuitry. The image sensor uses photoelectric effect to convert the incident photons into electrons and transports them along the circuitry to be read out as a measurable current. Therefore, a direct relationship exists between the digital readout of the sensor and the number of photons that were incident upon it. However, due to presence of electronics and their underlying physical mechanisms the signal gets corrupted with noise by several means before getting read. The mechanism of light generation also introduces uncertainties in the number of photons emitted per unit time by a source. The noise sources can thus be differentiated by two different categories:

1. **Signal dependent noise (N_{SD})**: Noise introduced due to mechanism of light generation or in other words, uncertainty in arrival of photons. It is also called as Shot Noise and it follows Poisson statistics. It follows that the variance equals the mean of the observed values. That is, shot noise is equal to the square root of the signal or, \( \sigma^2 = \mu \)

2. **Signal independent noise (N_{I})**: Noise introduced due to electronic circuitry. It can be further broken down into 3 categories: read noise, dark current noise and quantization noise.
Read noise is an inherent type of noise found in CCDs. It is invariant to incident light, temperature or exposure time. It is a fixed amount of noise that gets added to every image captured by the camera. The readout noise arises due to imperfections in the electronics involved. It is introduced at the stage where the signal is converted from charges ($e^-$) to voltage at the output. Readout noise is zero mean noise and follows Gaussian distribution.

As light hits the detector, electrons are formed within the silicon lattice, captured in potential wells, and read out as a signal subsequently. But, electrons can also form without any incident light due to the inherent kinetics, thanks to the thermal energy, of the electrons within the silicon. This is known as Dark signal. These electrons too get captured in the potential wells of the CCD, and counted as part of the signal. As the formation of dark current relies on the thermal energy of the electrons, its rate can vary with the sensor temperature, which in turn varies mainly due to the time for which it is exposed to light.

The quantization noise arises due to rounding off errors in the signal values at the Analog-Digital conversion (ADC). Consider a 11-bit image sensor with a full well capacity of 100,000 electrons. There are $2^{11} = 2048$ digital levels available to represent the brightness levels perceived by 100,000 photo-electrons. Assuming uniform gain over entire range of signals, about 50 electrons represent a single digital count. Thus, given a digital count $x$, it could any number of electrons ranging between $50(x-1)$ and $50x$. This error introduced due to quantization of signal levels manifests in form of quantization noise. It should be noted that the quantization could be considered to be uniform over the entire dynamic range as it is not a function of the intensity of the incident radiation.
2.3 Description of imageries used

This work is mainly based on imagery products released by two major commercial satellite companies: Digital Globe and Planet Labs. From Digital Globe’s constellation, WorldView-2 satellite was used while PlanetScope was used from Planet Labs’ constellation.

2.3.1 Digital Globe

Table 1: Digital Globe satellites overview

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>WorldView-2</th>
<th>WorldView-4</th>
<th>GeoEye-1</th>
<th>QuickBird</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude &amp; Orbit</td>
<td>770 km Sun</td>
<td>617 km Sun</td>
<td>681 km Sun</td>
<td>450 km Sun</td>
</tr>
<tr>
<td></td>
<td>Synchronous Orbit</td>
<td>Synchronous Orbit</td>
<td>Synchronous Orbit</td>
<td>Synchronous Orbit</td>
</tr>
<tr>
<td>Average revisit time</td>
<td>1.1 days</td>
<td>&lt;1 day</td>
<td>8.3 days @ 10° off nadir</td>
<td>2.4 days</td>
</tr>
<tr>
<td></td>
<td>3.7 days @ 20° off nadir or less</td>
<td>4.5 days @ 20° off nadir or less</td>
<td>2.1 days @ 35° off nadir</td>
<td>5.9 days @ 20° off nadir or less</td>
</tr>
<tr>
<td>Spectral Bands</td>
<td>Pan + 8 MS</td>
<td>Pan + 4 MS</td>
<td>Pan + 4 MS</td>
<td>Pan + 4 MS</td>
</tr>
<tr>
<td>Swath Width</td>
<td>16.4 km</td>
<td>13.1 km</td>
<td>15.3 km</td>
<td>16.8 km</td>
</tr>
<tr>
<td>GSD (At Nadir)</td>
<td>Pan: 0.46 m MS: 1.85 m</td>
<td>Pan: 0.31 m MS: 1.24 m</td>
<td>Pan: 0.41 m MS: 1.65 m</td>
<td>Pan: 0.61 m MS: 2.44 m</td>
</tr>
<tr>
<td>Bit Depth</td>
<td>11-bit</td>
<td>11-bit</td>
<td>11-bit</td>
<td>11-bit</td>
</tr>
</tbody>
</table>
Table 1 gives a summary of important characteristics of some of the major satellites in Digital Globe’s satellite constellation. Data from WorldView 2 was used for this work. Among the various imagery products available, the ‘Map-ready Ortho’ product was used. It contains data in a panchromatic band and 8 multispectral bands. All the bands of the focal plane use a special technique called as Time Delay Integration (TDI) to effectively increase the SNR by gaining higher exposure times.

Figure 2.3: Spectral response of WorldView 2. Courtesy: Digital Globe.
Table 2: Spectral band specifications for WorldView 2

<table>
<thead>
<tr>
<th>Band Name</th>
<th>Center Wavelength (nm)</th>
<th>Lower Band Edge (nm)</th>
<th>Upper Band Edge (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panchromatic</td>
<td>627</td>
<td>447</td>
<td>808</td>
</tr>
<tr>
<td>Costal Blue</td>
<td>427</td>
<td>396</td>
<td>458</td>
</tr>
<tr>
<td>Blue</td>
<td>478</td>
<td>442</td>
<td>515</td>
</tr>
<tr>
<td>Green</td>
<td>546</td>
<td>506</td>
<td>586</td>
</tr>
<tr>
<td>Yellow</td>
<td>608</td>
<td>584</td>
<td>632</td>
</tr>
<tr>
<td>Red</td>
<td>659</td>
<td>624</td>
<td>694</td>
</tr>
<tr>
<td>Red Edge</td>
<td>724</td>
<td>699</td>
<td>749</td>
</tr>
<tr>
<td>NIR 1</td>
<td>833</td>
<td>765</td>
<td>901</td>
</tr>
<tr>
<td>NIR 2</td>
<td>949</td>
<td>856</td>
<td>1043</td>
</tr>
</tbody>
</table>

The Digital Globe satellites capture raw data in terms of radiance incident on the sensor, which is later processed according to the product requirements. The Ortho (Map ready) imagery involves radiometric corrections to convert incident radiance into digital counts, sensor corrections to subtract dark current and hot pixels and ortho-rectification with a fine digital terrain model. The imagery product provides data in 16 bit format, out of which 11 significant bits contain the radiometric data. But this product is also disturbed as 8-bit imagery in some cases. To achieve this, the highest intensity in the 11-bit data is simply set to 255 and the rest values are mapped linearly accordingly.
### 2.3.2 Planet Labs

Table 3: Planet Labs satellites overview

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Planet Scope</th>
<th>Planet Scope</th>
<th>Rapid Eye</th>
<th>SkySat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude &amp; Orbit</td>
<td>400 km</td>
<td>475 km</td>
<td>630 km</td>
<td>450 km</td>
</tr>
<tr>
<td></td>
<td>International Space Station Orbit (~ 60 satellites)</td>
<td>Synchronous Orbit (~60 satellites)</td>
<td>Synchronous Orbit (5 satellites)</td>
<td>Polar inclined, circular orbit (2 satellites)</td>
</tr>
<tr>
<td>Average revisit time</td>
<td>Variable</td>
<td>Daily</td>
<td>Daily (off nadir) 5.5 days (at nadir)</td>
<td>N/A</td>
</tr>
<tr>
<td>Spectral Bands</td>
<td>B/G/R/NIR (455 nm – 860 nm)</td>
<td>B/G/R/NIR (455 nm – 860 nm)</td>
<td>B/G/R/NIR (440 nm – 850 nm)</td>
<td>B/G/R/NIR (450 nm – 900 nm)</td>
</tr>
<tr>
<td>Swath Width</td>
<td>20km x 12km</td>
<td>24.6km x 16.4km</td>
<td>77 km</td>
<td>2 km x 1.1 km</td>
</tr>
<tr>
<td>GSD</td>
<td>3.0 m</td>
<td>3.7 m</td>
<td>6.5 m</td>
<td>1.0 m – Multispectral 0.86 m – Panchromatic</td>
</tr>
</tbody>
</table>

Table 2 gives a summary of important characteristics of some of the major satellites in Planet Labs’ satellite constellation.
2.4 DIRSIG

DIRSIG is a synthetic image generation (SIG) tool developed at the Digital Imaging and Remote Sensing (DIRS) lab at the Rochester Institute of Technology. It is basically a radiative transfer engine which implements ray tracing from the fundamentals using image science principles to simulate how an imaging system would capture a particular scene. The user can describe the scene in terms of object parameters such as geometry, material properties etc., imaging system in terms of sensor characteristics and atmosphere in terms of location, climate, cloud cover, time of year etc. Using the atmosphere and scene specifications, the program calculates the radiance reaching the imager. The image is then calculated using the sensor characteristics.
2.4 EMVA 1288 standard for noise measurement

EMVA1288 is an industry standard for camera characterization defined by the European Machine Vision Association (EMVA). It provides standardized tests/methods to characterize various performance parameters of monochrome as well as color digital cameras with linear photo-response characteristics.
A number of photons ... 
... hitting the **pixel area** during exposure time

... creating a number of **electrons** ...

... forming a **charge** which is converted by a **capacitor** to a **voltage** ...

... being **amplified** ...

... and **digitized** ...

... resulting in the **digital gray value**.

Figure 2.2: Camera physical model. Credits: EMVA1288

---

![Diagram of noise sources in a camera](image)

- **dark noise**
- **quantization noise**
- **quantum efficiency**
- **system gain**
- **number of photons**
- **number of electrons**
- **digital grey value**

Figure 2.3: Noise sources in a camera. Credits: EMVA1288

\( \eta \): Quantum efficiency of the detector

\( A \): Pixel area
\( E \): Incident radiation

\( \mu_p \): Mean number of incident photons

\( \mu_e \): Mean number of generated electrons

\( t_{\text{exp}} \): Image exposure time

\( K \): System gain

\( \mu_y \): Mean signal

\( \mu_{y,\text{dark}} \): Mean output in absence of any incident light (dark signal)

\( \text{DN} \): Digital signal step size

\( \sigma_q^2 \): Total detector noise including the dark noise and the readout noise

\( \sigma_e^2 \): Shot noise

\( \sigma_q^2 \): Quantization noise.

Following the above models, we write the following equations

\[
\eta = \frac{\mu_e}{\mu_p}
\]

\[
\mu_p = \frac{AE t_{\text{exp}}}{hc/\lambda}
\]

\[
\mu_y = \mu_{y,\text{dark}} + K\eta \mu_p
\]

\[
\mu_y = \mu_{y,\text{dark}} + K\eta \frac{AE t_{\text{exp}}}{hc/\lambda}
\]

Following the Poisson statistics, the variance due to shot noise is given by

\[
\sigma_e^2 = \mu_e
\]

and the quantization noise is given by

\[
\sigma_q^2 = \frac{1}{12} DN^2
\]
As noises add in quadrature, variances add up linearly. Thus, the total variance, $\sigma_y^2$, in the measure output is given by

$$\sigma_y^2 = K^2 (\sigma_d^2 + \sigma_e^2) + \sigma_q^2$$

$$\sigma_y^2 = K^2 \sigma_d^2 + \sigma_q^2 + K(\mu_y - \mu_{y,dark})$$

This equation gives us a linear relation between the variance in the output ($\sigma_y^2$) and the mean signal level $\mu_y$. 
2.5 Slant edge method for MTF

While the noise model gives temporal response of the imaging system, the modulation transfer function (MTF) gives the spatial response. In simple words, MTF is basically a measure of how faithfully the spatial frequencies in the object-plane are reproduced in the image-plane. For this reason, it also often called as spatial frequency response (SFR). The MTF(f) at a spatial frequency f is given by

\[ MTF (f) = \frac{C(f)}{C(0)} \times 100\% \]

Where C(f) is the contrast function at frequency f. It is multiplied by 100 to normalize the MTF to 100% at low spatial frequencies. MTF is spatial response of the system with respect to the low frequencies. Thus, an extended MTF response correspond to better representation of high spatial frequencies meaning finer details or sharper images.

Figure 2.4: MTF explanation. Credits: Norman Koren, Imatest
The modulation transfer function can be calculated by performing Fourier Transform on the impulse response or the line spread function too.

\[ LSF(x) \rightarrow MTF(f) \]

The line spread function, in turn, can be calculated by differentiating the edge response or the edge spread function (ESF).

\[ LSF = \frac{\partial (ESF)}{\partial x} \]

Figure 2.5: Relationships between ESF, LSF and MTF. Credits: Crespi et al

2.6 Past Work

A typical laboratory based method for estimating noise has been using large homogenous areas to calculate the mean signal levels and standard deviations as noise. This can become hard to employ when it comes to satellite imagery though. It is hard to determine areas in satellite imagery that are truly homogenous, i.e. areas that would have differences in measured radiance solely because of the noise variations. It is difficult to find objects in natural scenes that are truly uniformly illuminated. Most of the times it is a combination of noise, atmospheric conditions and texture of the object itself.

Curran et al. introduced a new method based on Geostatistical assumptions for estimating SNR of remote sensing data. They use nugget variance as an estimate of the random noise in AVIRIS satellite data. Atkinson et al. take this method a step ahead by
postulating noise to be dependent on land cover type. Their modified geospatial method calculates SNR based on wavelength as well as land cover types.

Zhang et al. employ noise estimation methods based on homogeneity too. Their algorithm is to find means and standard deviations over homogenous as well as inhomogeneous areas. The noise estimation is done by considering some percentage of the total histogram, with different thresholds for both types.

Crespi et al. use homogenous as well as inhomogeneous areas for calculating the noise levels. They move 3x3 windows all over the image and estimate signal and noise levels in over them as samples. Later the values are divided into classes and the smallest 5% values in the cumulative histogram are considered to have arisen due to noise. For MTF calculations they look at edges, calculate the ESF, LSF and then MTF using Fourier Transform.

Among the many methods proposed for on-orbit MTF evaluation of satellite cameras, they can be divided into three broad categories. First one deals with comparing satellite images with high resolution images whose MTF is already known. The second way is to capture images of scenes in which object sizes are well known beforehand. Images of point or line objects are then analyzed for their size to determine the MTF. While the third category uses artificial as well as man-made sharp edge targets for finding the edge spreads and subsequently line spread and MTF.

Han et al. have used DIRSIG to generate image chips for training deep learning algorithms. They have generated more than a hundred thousand synthetic image chips by varying various parameters such as atmospheric conditions, shadows, illumination changes etc. The make use of Simulation of Urban Mobility (SUMO) tool to simulate traffic.
Although it takes a lot of radiometric factors into consideration, their work does not accurately model the cameras used on satellites and thus the images lack realism.

A large part of our work for PSF estimation is based on the ISO 12233 standard for image quality analysis. Almost all of the past works in this subject employ super-resolution technique for finding the edge spread function. Fujita et al. simply find the edge spreads for different lines and collect the data points in an order to super-resolve the edge. However, this technique does not take into account the precise phase of the edge associated with a certain data point and thus can give scope for some erroneous answers.

Figure 2.9 Generation of a finely sampled LSF (c) from coarsely sampled individual LSFs for different phases (b) for the slanted line shown in (a) Credits: Fujita et al

Burns’ technique is a widely used one in digital camera image quality analysis. The edge spreads are found for rows in the image, an edge is fitted and super-resolved edge spread functions are built by projecting the data points on to the edge. The LSF is then found out using discrete Fourier transform.
Figure 2.6: ISO 12233 method for SFR calculation as explained by Peter Burns
Chapter 3

Approach
3.1 Method used for estimating noise

We established a relationship between the variance and mean signal value based on our model as:

\[ \sigma_y^2 = K^2 \sigma_d^2 + \sigma_q^2 + K(\mu_y - \mu_{y,dark}) \]

\( \mu_{y,dark} \) can be approximated by substituting radiance \( L=0 \) in the radiometric calibration equations released by the satellite companies.

For WorldView-2 the radiometric calibration equation is given as:

\[ L = Gain \times DN \times \left( \frac{\text{abscl factor}}{\text{effective bandwidth}} \right) + \text{offset} \]

Where \( L \) is radiance in \( \mu m^{-1} m^{-2} sr^{-1} \), \( DN \) is the pixel value found in imagery. The gain and offset are absolute radiometric calibration band dependent adjustment factors and are released by Digital Globe periodically. The abscl factor and effective bandwidth are TDI specific and are provided in the metadata file with every image.

<table>
<thead>
<tr>
<th>Cal Version</th>
<th>WORLDVIEW-3</th>
<th>WORLDVIEW-2</th>
<th>GEDEYE-1</th>
<th>QUICKBIRD</th>
<th>WORLDVIEW-1</th>
<th>IKONOS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2016v0.1nt*</td>
<td>2016v3.1nt*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAND</td>
<td>GAIN</td>
<td>OFFSET</td>
<td>GAIN</td>
<td>OFFSET</td>
<td>GAIN</td>
<td>OFFSET</td>
</tr>
<tr>
<td>PAN</td>
<td>0.950</td>
<td>-3.629</td>
<td>0.942</td>
<td>-2.704</td>
<td>0.970</td>
<td>-1.926</td>
</tr>
<tr>
<td>COASTAL</td>
<td>0.905</td>
<td>-8.604</td>
<td>1.151</td>
<td>-7.478</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLUE</td>
<td>0.940</td>
<td>-5.809</td>
<td>0.988</td>
<td>-5.736</td>
<td>1.053</td>
<td>-4.537</td>
</tr>
<tr>
<td>GREEN</td>
<td>0.938</td>
<td>-4.996</td>
<td>0.936</td>
<td>-3.546</td>
<td>0.994</td>
<td>-4.175</td>
</tr>
<tr>
<td>YELLOW</td>
<td>0.962</td>
<td>-3.649</td>
<td>0.949</td>
<td>-3.564</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RED</td>
<td>0.964</td>
<td>-3.021</td>
<td>0.952</td>
<td>-2.512</td>
<td>0.998</td>
<td>-3.754</td>
</tr>
<tr>
<td>REDEDGE</td>
<td>1.000</td>
<td>-4.521</td>
<td>0.974</td>
<td>-4.120</td>
<td>1.060</td>
<td>-2.954</td>
</tr>
<tr>
<td>NIR1</td>
<td>0.961</td>
<td>-5.522</td>
<td>0.961</td>
<td>-3.300</td>
<td>0.994</td>
<td>-3.870</td>
</tr>
<tr>
<td>NIR2</td>
<td>0.978</td>
<td>-2.992</td>
<td>1.002</td>
<td>-2.891</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWIR1</td>
<td>1.200</td>
<td>-5.546</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWIR2</td>
<td>1.227</td>
<td>-2.600</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWIR3</td>
<td>1.199</td>
<td>-2.309</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWIR4</td>
<td>1.196</td>
<td>-1.676</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWIR5</td>
<td>1.262</td>
<td>-0.705</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWIR6</td>
<td>1.314</td>
<td>-0.669</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWIR7</td>
<td>1.346</td>
<td>-0.512</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWIR8</td>
<td>1.376</td>
<td>-0.372</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Includes 2016 vicarious season

Figure 3.1: Digital Globe radiometric calibration coefficients Courtesy: Digital Globe

For PlanetScopes, with the ‘Ortho-analytical’ product of imagery, the pixel values themselves represent the radiance (units: \( \mu m^{-1} m^{-2} sr^{-1} \)) in digital counts
directly. They are only scaled up by 100x to eliminate any information loss during quantization. Thus, the signal bias pedestal can itself be taken as the dark signal in that case.

\[
\sigma_q^2 = \frac{1}{12} DN^2
\]

Quantization noise can be approximated to zero as the step size DN equals one in digital images.

\[
\sigma_y^2 \propto \mu_y
\]

Thus, if we plot the variance versus the mean signal value, we can fit a straight line to the data and find the model parameters.

As discussed earlier one of the most common methods for calculating the noise is to use homogenous areas in images. They are assumed to have uniform reflectance and the variations in the incoming radiance from those areas are solely attributed to the presence of noise. However, we feel that it might not hold true always for the ground sampling distances varying from sub-meter to a few meter resolutions. There is a good chance for variations to arise because of various factors such as the texture of the object, orientation of the object with respect to the light source etc. It is therefore difficult to identify truly homogenous areas in a satellite image. On the other hand, it is safer to assume local homogeneity. A better way is to use only two neighboring pixels and assume them to have high local spatial correlation as explained in Schott, 2007. For most of the objects in a satellite image such as vehicles and rooftops, we can safely assume uniformities on the order of couple of meters, which is about what is sampled by two pixels in the image, for GSD of the order of few meters or sub-meters.
Our algorithm works on just two neighboring pixels at a time. The mean signal is taken as the mean of two adjacent pixel values and the difference between them is considered as a sample of noise at that mean signal level. For two pixels i and j,

\[ DN_i = [S_i + n_i] \quad \text{and} \quad DN_j = [S_j + n_j] \]

where DN stands for Digital Number read out, \( S_i \) stands for the signal value and \( n_i \) stands for the noise value. Note that the signal value is assumed to be equal for the two pixels, thus attributing the differences in DNs solely to the noise.

\[ S = \frac{DN_i + DN_j}{2} \]

and \( \Delta n = DC_i - DC_j \)

These \( \Delta n \)'s form the samples of noise at the signal level \( S \). The variance at this mean signal level can be found as:

\[ \sigma_{\Delta n}^2 = \sigma_{DC_i}^2 + \sigma_{DC_j}^2 \]

\[ \sigma_{\Delta n}^2 = 2 \times \sigma_S^2 \]

\[ \sigma_S^2 = \frac{\sigma_{\Delta n}^2}{2} \]

The variance at the signal level \( S \) half the variance of the noise samples \( \Delta n \).

From implementation point of view, we consider small patches in the images and shift them by a single pixel to get a patch of neighboring pixels. The signal and noise samples are calculated as described above. The signal values are then sorted in ascending order and arranged in bins, each of arbitrarily chosen width of 100 DNs. The corresponding noise values are sorted accordingly. The mean of signal values and the standard deviation of noise samples for each bin are then calculated as representations of signal and noise of the whole bin, respectively. The final step is to fit a straight line to the plot of variance.
against mean signal value. It was observed from the real imagery that the data for both Digital Globe and Planet, have a peculiar characteristic in terms of noise levels. The noise was observed to be nearly constant at low noise levels and it followed the previously described linear model above a certain signal level. In the noise model, we assumed the noise to be constant in this low signal region. For signal levels above this, the linear model was followed for a large number of signal levels until the model broke down at very high signal levels. A best fit line was found for this region using least squares method, implemented by the function `polyfit` within Matlab. The final noise model thus consisted of two components:

1. A fixed noise level for low signal regions
2. A linear equation relating variance and signal for higher signal regions

### 3.2 Method used for estimating PSF

Slanted edge method as described in Section 2.4 was followed for estimating the point spread function (PSF). The LSF is essentially a sliced profile of the PSF taken at angle perpendicular to the orientation of the edge. Ideal estimation of the PSF would thus involve calculating LSF at various angles and interpolating through them to construct the PSF. However, given the limited amount of data, it is hard to get edges at various angles. It follows from the fundamentals of physical optics that for an ideal optical system, the PSF can be considered as circularly symmetric. We assume the same in our case and estimate PSF by rotating a single LSF 360°.

During practical implementation, we first find the edges in an image. We choose the images that have fairly homogenous areas on both the sides of the edge. For a slant
edge, the edge detected pixels are obviously not in an exact straight line. Rows were scanned to get edge profiles for each. The edge profile was differentiated to obtain preliminary line spread function for each row in image. The maxima location for the LSF gives us the location of edge pixels. A best-fit straight line was found for the edge pixels. The image was then again scanned across rows to get edge profiles. As the edge is at an oblique angle with respect to the scanning direction, pixels along rows represent different phases of the edge spread. Knowing the angle of the edge, perpendicular distance from the line was calculated for every pixel.

Figure 7: Projecting pixels onto the edge  (Credits:Mary Pagnutti, Stennis Space Cener)

where $\theta = $ edge tilt angle, $\delta = $ pixel index and $x = $ pixel’s distance from edge.
A super-sampled edge response, or Edge Spread Function (ESF), was then constructed using a large number of such under-sampled edges. The ESF was stored as a lookup table of signal values (DN’s) corresponding to their perpendicular distance from the edge (x’s). Using multiple rows across the edge, a large number of samples of the edge spread function were found which would not have been possible with a single row or with an edge orthogonal to the camera axes. Thus, the ESF was found in a super-resolved form with phases of the edge known at various sub-pixel locations. A discrete derivative of this ESF gives us the Line Spread Function (LSF).

$$LSF(i) = \frac{DN_{i+1} - DN_i}{x_{i+1} - x_i}$$

The LSF values were determined for sub-pixel sampling for up to 5 pixels on both sides of the edge. For the ease of use with DIRSIG, the LSF was up-scaled 10 times and
the x co-ordinates were rounded off to the nearest integer. A 50 super-pixel long array containing half lobe of the LSF was thus generated and any missing values were interpolated linearly. This was rotated by 360 degrees to get a full 2-D 99x99 pixel PSF. The PSF was normalized to peak value of 1. Bearing in mind the empirical nature of this work, the PSF was left in discrete form as observed instead of fitting it with a Gaussian curve.

3.3 Data used

Images from various commercial satellites were analyzed in this work. The two major companies in this field are: Digital Globe and Planet Labs. While Digital Globe has satellites with high resolution (small GSD) imagers, Planet Labs’ satellites have a wide range in resolution. Their Planet Scopes, also called as ‘doves’, are cubesats launched from the International Space Station (ISS) in its orbit and have low resolution. But their strength lies in the revisit times as they frequent over any given latitude almost daily. Planet Labs also recently acquired a satellite start-up, SkySat, which add high resolution imagery to their portfolio; although with longer revisit times. Thus, there is always a tradeoff between the desired resolution and revisit frequency and so the imagery product choice needs to be made according the need of the application.
3.3.1 Digital Globe

Figure 3.4: Sample panchromatic image taken by World View-2 satellite at 0.5 m GSD.
3.3.2 Planet Labs

Figure 3.5: Sample color image taken by PlanetScope at 3.0 m GSD

3.4 Targets used

Targets were chosen so as to give us robust results using our algorithms. For noise modeling, scenes were chosen such that they contain a wide range of radiances contained in them to get a good mapping between noise and signal at various intensities. For PSF measurement, areas with sharp edges were needed. Edges in satellite imagery are typically found at bridges or at rooftops. For the ESF to have distinct features, a good contrast on either side of the edge are also expected. We have used bridges in our initial results for PSF estimations as still water can be assumed to be fairly dark as well as homogenous.
However, on the bright side of the edge in spite of presence of visual homogeneity, there might be subtle differences in the scene itself which would result in differential radiance reaching the sensor, thus introducing errors in our measurement. We have now procured data with real calibration targets built on ground for aerial imagery. We plan to repeat the algorithms on them to get more faithful results. Figures 3.6 and 3.7 show examples of such sites.

![Figure 3.6: Calibration target at CalVal facility in Shadnagar, India. It is the only site in the world to have a colored calibration target too.](image-url)
3.5 Process for generating synthetic simulated scene

The process for generating simulated scenes based on our results consists of two major steps: incorporating PSF and adding noise.

1. The Point Spread Function (PSF) was incorporated by using functionality built-in DIRSIG. Rather than simply mathematically convolving the original image intensity matrix with the PSF kernel, DIRSIG uses a more physics based approach for generating the image with advanced ray-tracing. Modulation transfer functions of various sub-systems such as the optics, detector, platform and scan motion are multiplied together to get the final MTF, Fourier transform of which gives it the PSF. In our case however, since the analysis is performed on final imagery product itself, the estimated PSF contains effects of all the sub-systems. Given a ground
sample, the PSF is projected onto the array of pixels on the detector. The radiance is then calculated as a weighted sum over the extent of the PSF.

2. For every pixel in the image a noise factor was added to it, corresponding to the noise model estimated. A fixed amount of noise was added for low signal levels where shot noise isn’t very dominant. The fixed noise was added for signal levels (in digital counts) below 300 for the WorldView-2 images and below 8000 for the PlanetScope images. Varying noise according to the estimated linear model was added for levels between 300 to 1000 for WorldView-2 and between 8000 to 12000 for the PlanetScope. A noise corrupted pixel value is given by

\[ x'(i,j) = x(i,j) + \sigma_s(x(i,j)) \ast n(0,1) + \sigma_f \ast n(0,1) \]

where \( x' \) represents the noise corrupt image, \( x \) represents the original image, \( \sigma_s(x(i,j)) \) represents the signal dependent noise at intensity \( x(i,j) \), \( \sigma_f \) is the fixed, signal independent noise and \( n(0,1) \) is a Gaussian random number with zero mean and unit variance. Given a sufficiently large number of pixels, \( \sigma \ast n(0,1) \) generates
data samples with mean = 0 and variance = $\sigma^2$. The fixed noise, $\sigma_F$ is calculated while estimating the noise model and $\sigma_S$ is calculated as

$$\sigma_S = \sqrt{\sigma_S^2 (x(i,j) + \sigma_F^2) - \sigma_F^2}$$

$$\sigma_S = \sqrt{\sigma^2 - \sigma_F^2}$$

where $\sigma^2$ is the variance calculated as the y co-ordinate of the straight line corresponding to $x(i,j)$. An empirically derived bias level of 110 DN and 5900 DN was added to WorldView 2 and PlanetScope simulations respectively.

**3.6 Verification of models**

Analysis similar to that done on the realistic imagery was applied to simulated imagery in order to verify if the models were applied correctly. For the purpose of estimating PSF, a target was specially designed having sharp edges, homogenous patches and tilted at an angle of 30 degrees with respect to the camera axes. Simulated scenes were generated for both the platforms. For verification of noise model, a scene similar to that of the real world town Trona, CA was constructed. This allowed the scene to have a large number of pixels and also covering all dynamic range in order to derive meaningful statistics out of them.
Figure 3.8: A part of the Trona scene generated using the WorldView 2 platform.

Figure 3.9: PSF Target generated for the PlanetScope platform
Chapter 4

Results and Discussion
4.1 Noise

Noise was estimated using the adjacent pixels method outlined in section 3.1. All the plots show in this section containing error bars have been calculated using 95% confidence intervals. Initial results showed that there was slight difference in the parameters of the straight line fit when the pixels were shifted in horizontal versus in vertical direction. Thus, the final results were calculated combining data points for both the methods

4.1.1 WorldView-2 Real Data

Figure 4.1: Scene used for analysis of WorldView 2 imagery (Levels corrected for visual representation)

Figure 4.1 shows the scene that was used for analysis of WorldView 2 imagery. This 16-bit panchromatic band image was acquired by WorldView 2 over Washington, D.C. on 2016-09-02 from an altitude of 770 km. It belongs to the map ready ortho-rectified
image product and has a mean GSD of 0.488 meters collected at exposure duration of 0.0016 seconds.

Figure 4.2: Noise vs Signal for WorldView 2

Figure 4.3: \( \text{Noise}^2 \) vs Signal for WorldView 2
Figure 4.4: SNR vs Signal for WorldView-2

Figure 4.5: Histogram for WorldView 2 image under test
Figure 4.6: Cumulative histogram for WorldView 2 image

Figure 4.2 shows noise levels estimated for various signal levels on WorldView 2 as described in the approach in Section 3.1. While Figure 4.3 shows variance plotted against signal. Some erratic observations can be seen in the graph. Since the data was observed to have a bias level of 110 DN, the first two observations hold no physical significance most probably arose due to some artifacts in the image. Similarly, the noise doesn’t seem to follow the model above 1000 DN. No analysis was done to further find out the reason for this anomaly. These pixels were simply disregarded as they comprised of a very small portion of the total number of pixels as can be seen in the histogram shown in figures 4.5 and 4.6. As can be seen from the cumulative histogram, 99.51% of the total number of pixels are contained within first 1000 digital levels of the WorldView 2 data.
Figure 4.7: Noise\(^2\) vs Signal with best fit line (shown in red)

Figure 4.7 shows the final noise model estimation. The red line shows the best fit line obtained for data points having signal intensities of 350 < DN < 1000. The best-fit line equation with \(R^2 = 0.9598\) was obtained as following:

\[
\sigma_y^2 = 7.60 \mu_y - 2600
\]

For DN<350, the fixed noise was found out to be:

\[
\sigma_F^2 = 295
\]
4.1.2 WorldView-2 Simulated Data

Figure 4.8: Noise added to scene simulated on WorldView 2 platform with noise-free image on the left and noisy image on the right.

Figure 4.9: Zoomed view showing effect of adding noise with noise-free image on the left and noisy image on the right.

Figures 4.8 and 4.9 show simulations created in DIRSIG with WorldView 2 platform and the effect of adding noise to it according to our predicted noise model. The noise estimation algorithms were re-run on this image for the purpose of verification.

The best-fit line equation for $350 < \text{DN} < 1000$ with $R^2 = 0.9581$ was obtained as following:
\[ \sigma_y^2 = 7.84\mu_y - 3150 \]

For DN<350, the fixed noise was found out to be:

\[ \sigma_F^2 = 325.3 \]

Comparing with the observations from the real data, we see that the noise model equations agree for both, real and simulated imageries.

### 4.1.3 PlanetScope Real Data

![Figure 4.10: Scene used for analysis of PlanetScope imagery](image)

The image shown in Figure 4.10 was used for analysis of PlanetScope imagery. This 16-bit image was taken over Shadnagar, India from an altitude of 400 km. The data directly represents the measured top-of-atmosphere (TOA) radiance up-scaled by a factor of 100 to reduce quantization errors. The mean GSD is 3.0 m in this multispectral image of Analytic type product.
Figure 4.11: Noise vs Signal values for PlanetScope

Figure 4.12: Noise$^2$ vs Signal for PlanetScope
Figure 4.13: SNR vs Signal for PlanetScope

Figure 4.14: Histogram for PlanetScope image
Figures 4.11 and 4.12 show Noise and Noise² respectively plotted against the signal. As with WorldView 2 data, PlanetScope data also shows anomalous behavior in very high signal regions, say above 12000, as can be seen in the graphs. However, like the histogram in figure 4.14 explains, very few pixels exist for values greater than 12000. Cumulative histogram in fact shows that 99.78% of all the pixels have values below 12000.
Figure 4.16: Noise$^2$ vs Signal for PlanetScope along with the best-fit line

Figure 4.16 shows the model fitting for PlanetScope data. The red line shows the best fit line for $8500 < \text{DN} < 12000$, given by:

$$ \sigma_y^2 = 18.21 \mu_y - 147000 $$

While the fixed noise is given by:

$$ \sigma_F^2 = 2819 $$

These models were then applied to simulate scenes within DIRSIG for the PlanetScope platform. The simulated scene was near replicated from the town of Trona, CA.
4.1.4 PlanetScope Simulated Data

Figure 4.17: Simulated scene for PlanetScope

Figure 4.18: Zoomed in view showing noise added to simulated scene

Figures 4.17 and 4.18 show noise-free images on the left and noise added images on the right of the simulated scene using PlanetScope platform.

The scenes simulated for the same area as of WorldView-2, for the PlanetScope platform, were smaller in size in terms of pixel resolution due to higher GSD. As a result fewer total number pixels were available in the Planet scene. We suspect this affected the
statistics as some intensities got very less number of pixels associated with them. It was also observed that the results came to closer to the real noise model as the size of the image was increased from 400x400 to 1000x1000 (the slope of the line decreased from 102 to 42.65, closer to the real slope of 22.96). Perhaps, an even bigger scene could be used to better mimic the noise model in the simulated imagery.

It was observed that the SNR was quite high at low signal, then dropped with increasing signal values and finally became almost constant for a range of signal values that followed the linear model between variance and mean signal. The SNR performance can be seen in Figure 4.4 for WorldView-2 and in Figure 4.13 for PlanetScope.

### 4.1.5 Summary of Noise Model Results

Summary of noise model results for WorldView-2

<table>
<thead>
<tr>
<th>DN &lt; 350</th>
<th>Real Data</th>
<th>Simulated Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \sigma_F^2 = 295 )</td>
<td>( \sigma_F^2 = 325.3 )</td>
</tr>
</tbody>
</table>

\begin{align*}
350 < DN < 1000 & \quad \sigma_y^2 = 7.60 \mu_y - 2600 \\
\text{Simulated Data} & \quad \sigma_y^2 = 7.84 \mu_y - 3150
\end{align*}

Summary of noise model results for PlanetScope

<table>
<thead>
<tr>
<th>DN &lt; 85000</th>
<th>Real Data</th>
<th>Simulated Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \sigma_F^2 = 2819 )</td>
<td>( \sigma_F^2 = 4302 )</td>
</tr>
</tbody>
</table>

\begin{align*}
8500 < DN < 12000 & \quad \sigma_y^2 = 18.21 \mu_y - 147000 \\
\text{Simulated Data} & \quad \sigma_y^2 = 37.41 \mu_y - 193040
\end{align*}
4.2 Point Spread Function (PSF)

The spatial resolution was measured in form of PSF. $\frac{\lambda FN}{p}$ is the ratio of the sampling frequency to optical bandpass limit of the optical system [Fiete, 1999]. It shows the ability of the detector to finely sample the diffraction limited optical PSF. Where $\lambda$ is the center wavelength, FN is the focal ratio and $p$ is the detector pixel pitch.

![Diagram of PSF sampling](image)

Figure 4.19: Sampling of PSF. Credits: Robert Fiete

The $\frac{\lambda FN}{p}$ ratio for WorldView-2 is 0.86 while that for PlanetScope is 0.846 for a wavelength of 550 nm. Thus, both the systems sample the PSF almost as finely as each
other. In other words, it can also be inferred as, the width of the PSF for both systems should be almost equal, in terms of the pixels. The ground spot size will be different for obvious reasons. For a diffraction limited system, the PSF width is given by $2.44 \times \lambda \times FN$, where FN is the focal ratio. For light of wavelength 550 nm and f/# of 12, the theoretical PSF width from central maximum to first zero is 16.1 $\mu m$ or about 2 pixels. Thus, the FWHM width of the PSF can be expected to be around $\frac{1.02}{2.44} \times 2 = 0.836$ pixels or 6.73 $\mu m$ for both the systems in the diffraction limited region. The theoretical PSF sections are shown in Figure 4.20.

![Figure 4.20: A section through the theoretical diffraction limited PSF](image)

The Point Spread Function (PSF) was calculated for by assuming ideal optics and thus, circular symmetry. A line spread function (LSF) was measured which represents a slice through the real PSF. This LSF was rotated 360 degrees about its axis to get the PSF.
Figure 4.21: Target used for PSF estimation

Figure 4.21 shows the target used for PSF estimation of Planet imagery. It is located at Shadnagar, India (17.034426N, 78.182959E).
Figure 4.22: Edge profiles for various rows

Figure 4.22 shows edge profile plotted for 6 different rows in the image across a high contrast edge on the target. The x-axis shows pixel locations with respect to the edge in sub-pixel resolution while y-axis shows the intensity values for each of those pixels. It can be seen clearly that a number of different phases along the edge are obtained in either of the six edge profiles. Combining a number of such edge profiles gives us a super-resolved edge as shown in figure 4.23.
Figure 4.23: Super resolved edge spread function

A super-resolved edge spread function (ESF) plotted by combining 6 edges can be seen in Figure 4.23.
4.2.1 PlanetScope Real Data

Figure 4.24: PSF for PlanetScope

Figure 4.25: PlanetScope PSF as viewed from a plane parallel to the sensor plane
4.2.2 WorldView-2 Real Data

Figure 4.26: Single row undersampled ESFs for WorldView 2
Figure 4.27: WorldView-2 Super-resolved ESF

Figure 4.28: WorldView-2 sub-pixel sampled LSF
Figure 4.29: 2-D PSF for WorldView 2

Figure 4.30: 2-D PSF for WorldView 2 as seen from a plane parallel to sensor plane
4.2.3 WorldView-2 Simulated Data

Figure 4.31: PSF Estimated for WorldView 2 simulated image

4.2.4 PlanetScope Simulated Data

Figure 4.32: PSF estimated for PlanetScope simulated image
Figure 4.33: Comparison between PlanetScope and WorldView-2 PSF estimations

Figure 4.33 shows PSFs estimated for WorldView-2 and PlanetScope plotted together. The results were encouraging since both show similar width of the PSF. The diffraction limited incoherent optics PSF for a circular aperture is given by

\[
PSF_{optics}(r) = \left[ \frac{2J_1(\frac{\pi Dr}{\lambda f})}{\left(\frac{\pi Dr}{\lambda f}\right)} \right]^2
\]

Where \( r = \sqrt{x^2 + y^2} \), \( J_1(r) \) is the first-order Bessel function, \( D \) is the aperture diameter, \( f \) is the focal length and \( \lambda \) is the mean wavelength of the light under observation. The observation shown in figure 4.33 follows theory according to which the PSF width is a function of light wavelength and camera f/number only. The f/number for WorldView 2 is 12 and for PlanetScope it is 12.5.
Simulated scenes were also generated using DIRSIG scenes without any PSF applied. PSF was then applied in post processing using simple convolution. It was observed that the PSF measured for such a scene more closely resembled the actual PSF than the one obtained from DIRSIG’s built-in functionality.

![Airy Disk pattern](image)

**Figure 4.34: Airy Disk pattern. Credits: James E.H. Turner**

In a diffraction limited system, the FWHM is \((1.02/2.44)\) times the diameter of the central maximum. The central maximum diameter is given by \(2.44 \times \lambda \times FN\), which is about 16.1 \(\mu m\) for PlanetScope and WorldView-2. Thus, the theoretical FWHM width of PSF is \(1.02/2.44 \times 16.1 = 6.73 \mu m\). However, for a diffraction limited system, for proper sampling of images, the pixel size should be no greater than half the FWHM of the point source. The pixel pitch for both Digital Globe and Planet satellites is about 8 \(\mu m\) which clearly doesn’t satisfy the diffraction limited condition. As a result, higher frequencies in the scene are not faithfully recovered in the image. Moreover, our PSF estimation algorithm works on a
sharp edge, i.e. a high frequency region, which is contaminated with noise as well. Thus, the PSF widths observed are higher than the theoretical calculations. However, proper recovery of the input PSF from the simulated scenes verifies the working of the working of our algorithm.

Figure 4.35: Comparison of empirical and diffraction limited PSF for WorldView-2
Figure 4.36: Comparison of empirical and diffraction limited PSF for PlanetScope

Figures 4.35 and 4.36 show the theoretical diffraction limited PSF and the empirical PSF plotted together for WorldView-2 and PlanetScope respectively. The empirical data contains blurring due to optics, detector, atmosphere as well as the post-processing involved. The diffraction limited PSF is plotted as an ideal case showing just the optical blurring considering the aperture size, focal length and the wavelength of light observed. Clearly the empirical PSF is wider than the theoretical one since our system is not diffraction limited. It should also be noted that due all effects, the actual Ground Resolved Distance (GRD) observed in the real data is much higher than the GSD specified by the companies. This is not surprising since GSD just signifies how much ground area every pixel on the sensor plane corresponds to, considering only the distances from the optics to the sensor and ground. Thus, GRD could also be thought of as the real resolution or minimum separation distance between two points on ground that are capable of being resolved on the concerned detector. The GRD for WorldView-2 was observed to be 0.95
meters against the GSD of 0.46 meters. Similarly, the GRD for PlanetScope was observed to be about 10 meters while its GSD is 3 meters.

### 4.2.5 Summary of PSF Results

The values in Table 2 show half width at half maximum (HWHM) for the respective PSF in pixels.

<table>
<thead>
<tr>
<th></th>
<th>Real scene</th>
<th>Simulated scene using PSF within DIRSIG</th>
<th>Simulated scene using PSF convolution in post processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>WorldView-2</td>
<td>1.254</td>
<td>1.594</td>
<td>1.272</td>
</tr>
<tr>
<td>platform</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PlanetScope</td>
<td>1.187</td>
<td>1.482</td>
<td>1.248</td>
</tr>
<tr>
<td>platform</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 5

Conclusion and Future Work
5.1 Conclusion

The aim of this dissertation has been to estimate models for camera systems used on imaging satellites for generating realistic synthetic imagery. Firstly, the noise and PSF models were estimated by analyzing the real imagery from Digital Globe’s WorldView-2 and Planet Labs’ PlanetScope satellites. Secondly, synthetic imagery was generated using DIRSIG. While PSF model was incorporated using DIRSIG’s in-built functionality, noise was added externally using post processing techniques. Thirdly, the analysis was re-run on the simulated imagery and the models were verified.

We developed and used a novel technique for noise estimation which does not depend upon presence of homogenous areas in the image. Noise model relating the variance and signal was adopted from the EMVA 1288 standard. However, the model would break down in low signal regions and thus was substituted by a mean noise level in that region. Both the satellites studied, WorldView-2 and PlanetScope followed the model well for about 99.5% of the total number of pixels. The noise behavior in very high signal regions was found to be erratic and needs further investigation. A fixed noise was estimated for low signal levels: below 300 digital counts and 8000 digital counts for WorldView-2 and PlanetScope respectively. A linear relationship was established between variance and mean signal level for signal levels between 300 DN and 1200 DN for WorldView-2 and between 8000 DN and 12000 DN for PlanetScope.

PSF was estimated using the slanted edge method loosely based upon the ISO 12233 standard. Due to discrete nature of data, only a limited number of samples of edge spread can be obtained from any given direction oblique to that of the edge. A slanted edge offers samples of the edge spread function at various phases. A super-resolved ESF was
thus obtained using multiple single row ESFs. A super-resolved LSF was subsequently obtained which represented a 1-D slice through the 2-D PSF to be estimated. Assuming circular symmetry of the optics, 2-D PSF was estimated by rotating the LSF by 360 degrees. The width of the PSFs, expressed in terms of full width at half maximum (FWHM), was found as 2.50 pixels for WorldView-2 and 2.36 pixels for the PlanetScope.

However, at the end of the day, this work is empirical. A lot of is unknown about the specifications of the camera systems used aboard these satellites due to the proprietary nature of those companies. This work is, therefore, based purely on the data obtained publicly as-is from the commercial satellite companies. We suspect the released data to have gone through quite a bit of post processing even beyond the publicly released notes by companies. But since the ultimate goal of this work was to generate realistic imagery that tries to mimic the characteristics of the publicly available data itself, this work does not deal a lot with deviations from theoretical predictions at this time.

In conclusion, we feel, this work creates a good starting point for aerial synthetic imagery generation having realistic characteristics taking the utility of already well known DIRSIG, a step further. Along with other applications, this work would certainly help in creating a variety of data that could be used for the data-hungry machine learning algorithms.
5.2 Future Work

One of the important findings of these works was the deviation of the systems from the theoretical predictions. It could be attributed to the post-processing chain adopted by the companies before distributing various products. A more in-depth analysis is needed to study and characterize artifacts arising out of it to better represent them in the simulated imagery. This could lead to better modelling of noise in very low and very high signal regions where the current linear model was found to be breaking down. It was observed that a 3rd degree polynomial would fit the data in linear as well as low signal regions. However, it was not pursued further due to limitations in the scope of this work. Future work could investigate that kind of fit along with a mathematical model explaining it.

It was observed that the model verification for the PSF was better if the simulated scene was generated by convolving DIRSIG scene with PSF in post-processing rather than using the built-in functionality within DIRSIG. A deeper dive could be taken in the implementation of DIRSIG to find out the reasons for this observation.

Another important task we feel could follow up this research is to devise a method for analyzing the realistic nature of the simulated scenes. Having an objective metric to quantify the realism would help to understand the efficiency of this work. Current methods include use of human observers for Visual Turing Test (VTT) in which they are presented with real and simulated images and are asked to differentiate between the two. However, for good results, it is necessary to have the real and simulated scenes representing the same part of the world which is not easily possible for simulated scenes. Thus, a computer based approach would work far better without need of similar scenes and eliminating observational bias.
References

- "Imagery Support Data (ISD) Documentation v1.1.2" Digital Globe
- Kuester, Michele ‘Absolute Radiometric Calibration: 2016v0”. Digital Globe
- “Planet Labs Specifications: Spacecraft Operations and Ground Systems v1.0” Planet Labs June 2015


