Image Quality Modeling and Characterization of Nyquist Sampled Framing Systems with Operational Considerations for Remote Sensing

Rey Jan D. Garma

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

Recommended Citation


This Dissertation is brought to you for free and open access by the Thesis/Dissertation Collections at RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact rit scholarworks@rit.edu.
Image Quality Modeling and Characterization of Nyquist Sampled Framing Systems with Operational Considerations for Remote Sensing

by

Rey Jan D. Garma

B.S., United States Air Force Academy, 2005
M.S., University of Southern California, 2009

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Chester F. Carlson Center for Imaging Science
College of Science
Rochester Institute of Technology

July 2015

Signature of Author

Accepted by

Coordinator, Ph.D. Degree Program
Ph.D. Degree Dissertation

The Ph.D. Degree Dissertation of Rey Jan D. Garma has been examined and approved by the dissertation committee as satisfactory for the dissertation required for the Ph.D. degree in Imaging Science.

Dr. John R. Schott, Dissertation Advisor

Dr. Robert D. Fiete, Committee Member

Dr. Jie Qiao, Committee Member

Dr. Robert Parody, Outside Chair

Date
Image Quality Modeling and Characterization of Nyquist Sampled Framing Systems with Operational Considerations for Remote Sensing

by

Rey Jan D. Garma

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

Abstract

The trade between detector and optics performance is often conveyed through the Q metric, which is defined as the ratio of detector sampling frequency and optical cutoff frequency. Historically sensors have operated at $Q \approx 1$, which introduces aliasing but increases the system modulation transfer function (MTF) and signal-to-noise ratio (SNR). Though mathematically suboptimal, such designs have been operationally ideal when considering system parameters such as pointing stability and detector performance. Substantial advances in read noise and quantum efficiency of modern detectors may compensate for the negative aspects associated with balancing detector/optics performance, presenting an opportunity to revisit the potential for implementing Nyquist-sampled ($Q \approx 2$) sensors. A digital image chain simulation is developed and validated against a laboratory testbed using objective and subjective assessments. Objective assessments are accomplished by comparison of the modeled MTF and measurements from slant-edge photographs. Subjective assessments are carried out by performing a psychophysical study where subjects are asked to rate simulation and testbed imagery against a $\Delta$NIIRS scale with the aid of a marker set. Using the validated model, additional test cases are simulated to study the effects of increased detector sampling on image quality with operational considerations. First, a factorial experiment using Q-sampling, pointing stability, integration time, and detector performance is conducted to measure the main effects and interactions of each on the response variable, $\Delta$NIIRS. To assess the fidelity of current models, variants of the General Image Quality Equation (GIQE) are evaluated against subject-provided ratings and two modified GIQE versions are proposed. Finally, using the validated simulation and modified IQE, trades are conducted to ascertain the feasibility of implementing $Q \approx 2$ designs in future systems.
Disclaimer

The views expressed herein are those of the author and do not reflect the official policy of the U.S. Air Force, Department of Defense, or the U.S. Government.
Acknowledgements

We would like to thank the National Reconnaissance Office for funding this work.
# Contents

Abstract ii  
Disclaimer iii  
Acknowledgements iv  
Contents v  
List of Figures viii  
List of Tables xiv  

1 Introduction 1  
2 Objectives 4  
   2.1 Objectives ................................................. 4  
   2.2 Tasks ....................................................... 5  
   2.3 Contribution to Field ..................................... 7  
3 Background and Theory 8  
   3.1 Linear Shift-Invariant Systems ............................ 9  
   3.2 Image Quality ............................................... 10  
   3.3 NIIRS and GIQE ............................................ 14  
   3.4 Remote Sensing Systems .................................. 19  
   3.5 Framing Arrays ............................................. 29  
   3.6 Q ............................................................. 31  
   3.7 Image Chain Simulation .................................... 35  
       3.7.1 Radiometry ........................................... 35  
       3.7.2 Optics ............................................... 37  
       3.7.3 Aberrations ......................................... 39  
       3.7.4 Detector ............................................. 41  
       3.7.5 Motion ................................................. 44
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4 Temporal Coverage Trades</td>
<td>161</td>
</tr>
<tr>
<td>A Detailed Flow Chart of Digital Model</td>
<td>163</td>
</tr>
<tr>
<td>B Detailed Flow Chart of Testbed</td>
<td>165</td>
</tr>
<tr>
<td>C Objective Model Validation Results</td>
<td>167</td>
</tr>
<tr>
<td>D Standard Deviations of Collected Data</td>
<td>172</td>
</tr>
<tr>
<td>Bibliography</td>
<td>175</td>
</tr>
</tbody>
</table>
# List of Figures

3.1 Image quality assessment categories. .................................................. 11  
3.2 (A) Edge spread function (B) LSF found by taking the derivative of the ESF  
(C) MTF found by taking the Fourier Transform of the LSF ..................... 13  
3.3 Geometry of GSD look angles. .......................................................... 16  
3.4 RER measurement. ........................................................................... 17  
3.5 H measurement. ............................................................................... 18  
3.6 Basic collection geometry for space-based remote sensing platforms. ..... 20  
3.7 Collection geometry for line-scanners. Geometry is the same for whiskbroom  
scanners, with the exception of using a linear array of detectors in the along-  
track direction. .................................................................................. 21  
3.8 Zenith smear when fore or aft-looking. ............................................ 22  
3.9 Range smear when fore or aft-looking. ............................................ 23  
3.10 Azimuth smear when side-looking. ................................................... 23  
3.11 Collection geometry for pushbroom scanners. ............................... 25  
3.12 Range smear due to change in range over TDI stages (Auelmann, 1997). 25  
3.13 Azimuth-like smear due to TDI timing misalignment for TDI arrays off-center. 26  
3.14 Clocking smear for systems with TDI stages. .................................. 27  
3.15 Collection geometry for framing sensors. ........................................... 28  
3.16 Step and Stare CONOP for framing sensors - center pixel of array remains  
fixed over point of interest over the integration time, i.e., the spacecraft  
"nods" while tracking the target on the ground. ..................................... 29  
3.17 Q sampling factor mapped to PSF. A Q-factor of 0.8 results in two samples  
per central lobe of the PSF. Q=1(Q₁) is 2.44 samples per central lobe, and  
Q=2 (Q₂) is 4.88 samples. ................................................................. 32  
3.18 The effect of adjusting different Q parameters when transitioning from a Q₁  
design to Q₂. .................................................................................... 33  
3.19 Comparison of pixel size and smear for Q₁ and Q₂ designs. (A) Pixel size  
for a Q₁ design with 1/4 pixel smear (B) Relative pixel size for a Q₂ system.  
Pixel smear increases by a factor of Q to 1/2 pixel ............................... 33  
3.20 MTF for a circular aperture at λ = 550⁻⁶ mm with f-numbers f/8 and f/16. 39  
3.21 Aberrated wavefront optical path delay. ......................................... 40  
3.22 Lens MTF at f/8 and λ = 550⁻⁶ mm relative to aberration MTFs at W_{RMS} =  
0.1, 0.12 waves. ............................................................................. 41
3.23 Lens MTF at at f/8 and $\lambda = 550^{-6} mm$ relative to the modeled pixel MTF. 42
3.24 Combined Lens and Pixel MTFs relative to the modeled vertical and horizontal CTE MTFs. 43
3.25 (A) Zenith smear component (B) Range smear component (C) Azimuth smear component (D) Combined 45
3.26 Smear at isoplanatic patch. 45
3.27 Cross scan and along scan components combine to form overall smear vector. 46
3.28 (A) Smear/blur MTFs for 0.25, 0.5, 0.75, and 1.0 pixels. (B) Jitter MTFs for 0.1, 0.2, 0.3, and 0.4 pixels. 47
3.29 (A) and (B) Theoretical photon transfer curves for room-temperature and cooled CCDs. (C) Noise contributions for warm CCD operating in shot-noise limited region. (D) Noise contributions for cooled CCD operating in shot-noise limited region. (E) Noise contributions for warm CCD operating in low signal region (F) Noise contributions for cooled CCD operating in low signal region. 50
3.30 (A) Radially averaged image amplitude spectrum, with linear fit coefficient, $\alpha = 1.21$ (B) Model image using $f = 50 mm$ at $Q_1$ (C) Filtered Image (D) Zoomed unfiltered simulation image (E) Zoomed filtered simulation image. 53
3.31 Illustration of $3^3$ full factorial experiment tradespace. 57
4.1 Illustration of lab setup adjustability. 61
4.2 Gain histogram for precise measurement of gain for STF-8300M CCD. 63
4.3 Plot of average digital count vs. exposure time for STF-8300M. Linear fit generated for points between 5,000 and 50,000 DC. 65
4.4 Linearity residual plot for STF-8300M. 65
4.5 STF-8300M dark current data at $15^\circ C$ and $-15^\circ C$. The slope of each line is the dark current rate at the respective temperatures; $0.16e^{-}/pix/sec$ at $15^\circ C$ and $0.014e^{-}/pix/sec$ at. 66
4.6 STF-8300M dark current data at $15^\circ C$ and $-15^\circ C$ compared to an extrapolation of manufacturer data. 67
4.7 STF-8300 quantum efficiency curve. 68
4.8 Laboratory setup for collection QE data. 68
4.9 Measured (red) vs. specified (blue) quantum efficiency curves. 69
4.10 (A) Photo of hexagonal aperture present in 16mm and 25mm lenses. (B) Photo of octagonal aperture present in 50mm lens. (C) Photo of hexagram aperture present in 35mm lens. 70
4.11 Modeled apertures of (A) hexagon, (B) octagon, and (C) hexagram. 70
4.12 Comparison of (A) circular, (B) hexagonal, (C) octagonal, and (D) hexagram aperture OTFs. 70
4.13 Comparison of circular aperture to (A) hexagon, (B) octagon, and (C) hexagram. Actual aperture shapes shall be modeled due to the magnitude deviations from a circular aperture. 71
4.14 MTF plots for 16mm lens at f/8 and f/16. Lines shown are tangential and sagittal MTFs for diffraction limited (black), 0°, and 3° fields. 72
List of Figures

4.15 Seidel diagrams for 16mm lens at f/8 and f/16. ............................. 72
4.16 Field curvature and distortion for 16mm lens at f/8 and f/16 with a 3° field. 73
4.17 Total wavefront RMS for 16mm lens at f/8 and f/16 with a 3° field. .... 73
4.18 (A) Plot of measured vs. modeled MTF error as a function of defocus aberration. (B) Measured MTF plotted against modeled MTF with zero and 0.26λ defocus aberration .......................... 74
4.19 (A) Plot of measured vs. modeled MTF error as a function of wavefront RMS error. (B) Measured MTF plotted against modeled MTF with zero and 0.09λ wavefront RMS error .......................... 75
4.20 Sensitivity of GSD to target distance and look angle. ..................... 76
4.21 High resolution scan of B-52 transparency with edges selected throughout image to determine MTF. ................................. 77
4.22 MTF of test target. .................................................................... 78
4.23 QA-61 test chart by Applied Image Inc. ....................................... 79
4.24 MTF of test target. .................................................................... 80
4.25 Photon transfer curve of EPSON 1640XL scanner. ...................... 81
4.26 Normalized photon transfer curves of EPSON 1640XL, EPSON 845 scanners and STF-8300M camera. ........................................ 81
4.27 Noise measurements at high resolution versus noise measurements after blurring and downsampling for f = 50mm at f/8. .................. 82
4.28 Lumitile drawings provided by Evo-Lite (A) Side View (B) Front View ... 83
4.29 Measured spectrum of the actual fluorescent source and a representative LED source (peak normalized). .............................. 84
4.30 Transmission of Edmund Optics red filter. ................................. 84
4.31 LED and fluorescent weighting functions. .................................. 85
4.32 Generalized image simulation process. A detailed illustration of the digital model’s workflow is seen in Appendix A. ......................... 88
4.33 Flowchart showing the source of each MTF used in objective validation of image chain simulation. .............................................. 88
4.34 Slant edge method for generating super-resolution data. ................ 90
4.35 Subjective image quality assessment process. ......................... 91
4.36 Two image tiles selected from the transparency (A) Nadir image of B-52 (B) Nadir image of F-102 (top) and F-86 (bottom) ............... 93
4.37 Estimated ΔNIIRS range for varying focal length and f-number. .... 94
4.38 Δ NIIRS predictions using the relationship defined by Equation 4.23. 95
4.39 (A) Q1, low smear, 1× integration time B) Q1, low smear, 4× integration time 98
4.40 (A) Q1, low smear, 1× integration time, low detector performance B) Q1, low smear, 1× integration time, high detector performance .... 100
4.41 Set up for DNIIRS rating hardware. ......................................... 103
4.42 Diffuse surround lighting for DNIIRS rating stations. ................... 103
4.43 MATLAB GUI for performing DNIIRS ratings. .......................... 104
4.44 Calibration curves for translating desired DNIIRS increments to actual DNIIRS increments as measured by each GIQE. .............. 105
5.1 Modeled apertures of (A) 50mm lens at f/8 (octagon aperture) and (B) 35mm lens at f/16 (hexagram aperture). The 50mm f/8 exhibited the lowest RMSE between the measured testbed MTF and the measured simulation MTF (0.004), whereas the 35mm f/16 lens exhibited the worst at 0.02. 109

5.2 Modeled apertures with respect to the circle used to define aperture diameter when calculating f-number. (A) Octagon (B) Hexagon (C) Hexagram. Red and blue circles are inscribed and circumscribed circles, respectively. 109

5.3 Mean simulation and testbed ∆NIIRS ratings for each configuration with 95% confidence intervals. Note the discrepancy between testbed and simulation ratings for the f=35mm, f/16 configuration. Future studies should adjust the defined radius if a hexagram aperture is used. 111

5.4 Simulation vs. Testbed ∆NIIRS ratings for varying focal length and f-number. The coefficients with 95% confidence bounds are: Slope = 1.01 (0.98, 1.04), Intercept = -0.009 (-0.03, 0.009). $R^2 = 0.95$. 112

5.5 Registration differences between (A) marker and model (B) marker and unregistered testbed. 113

5.6 Registration differences between (A) marker and registered testbed (whole pixel translation only) (B) Subpixel translation and rotation. 113

5.7 Mean simulation and testbed ∆NIIRS ratings for each SNR level with 95% confidence intervals. 115

5.8 (A) RER of simulated imagery (B) G of simulated imagery. 115

5.9 Simulation vs. Testbed ∆NIIRS ratings for varying SNR. The coefficients with 95% confidence bounds are: Slope = 0.99 (0.97, 1.01), Intercept = 0.00 (-0.01, 0.01). $R^2 = 0.97$. Non-weighed values are as follows: Slope = 1.00 (0.93, 1.07), Intercept = -0.01 (-0.06, 0.03). $R^2 = 0.70$. 116

5.10 Mean simulation and testbed ∆NIIRS ratings for each DOE treatment with 95% confidence intervals. Gray bars indicate treatment groupings, where t is relative integration time (every six data points), LS = Low Smear (0.4 Q1 pixels), HS = High Smear (0.45 Q1 pixels), and Low, Medium and High detector performance are within each grouping of LS and HS. 118

5.11 Main effects plots for Q, Integration Time, Pointing Smear, and Detector Performance. 119

5.12 Interaction plots for Q, Integration Time, Pointing Smear, and Detector Performance. 122

5.13 (A) Transition from $Q_1$ to $Q_2$ at 1× Integration Time (B) Transition from $Q_1$ to $Q_2$ at 4× Integration Time at low (red) medium (green) and high (blue) detector performance. 123

5.14 ∆NIIRS for low, medium, and high detector performance as a function of Q at low smear superimposed on data graphically obtained from Fiete and Tantalo (1999). 125

5.15 Plot of GIQE predictions vs subjective ratings for simulated images. 127

5.16 Plot of GIQE predictions vs subjective ratings for simulated images with end points removed. 128
5.17 Plot of GIQE5 predictions vs subjective ratings for simulated images - Q\textsubscript{2} sampling at configurations with f/16. ................................. 130
5.18 Plot of GIQE predictions vs subjective ratings for SNR degraded imagery at (A) Q\textsubscript{1} and (B) Q\textsubscript{2}. .................................................. 131
5.19 Plot of GIQE predictions vs subjective ratings for SNR degraded imagery at Q\textsubscript{1} (left column) and Q\textsubscript{2} (right column) when (A) (B) H is removed from GIQE 3 and 4 (C) (D) G and H terms are removed from GIQE 3 and 4 ........................ 132
5.20 Plot of GIQE predictions vs subjective ratings for SNR degraded imagery when G/SNR coefficient = 1.5 for GIQE4 at (A) Q\textsubscript{1} and (B) Q\textsubscript{2}. ............... 133
5.21 Mean subject ratings and 95% CIs relative to GIQE 3, 4, and 5 predictions. 134
5.22 Mean subject ratings and 95% CIs relative to GIQE 3, 4, and 5 predictions when RER of only a vertical edge is used to calculate ∆NIIRS. ............. 135
5.23 Calculated RER when utilizing the geometric mean and the average. ..... 135
5.24 Regression of subject ratings against GIQE predictions for factorial experiments. .......................................................... 136
5.25 Plot of GIQE3 and smear-corrected GIQE3 predictions and regression. .. 137
5.26 Residuals for factorial experiments data plotted as a function of Q. ........ 137
5.27 ∆NIIRS plot of Q\textsubscript{1} and Q\textsubscript{2} SNR degraded images for SNR = 33 and SNR = 5.4 ......................................................... 138
5.28 Comparison of predictions for GIQEs 3 and 3.1 using factorial experiments data. ................................................................. 139
5.29 Comparison of predictions for GIQEs 3 and 3.2 using factorial experiments data. ................................................................. 140
5.30 Comparison of predictions for GIQEs 3, 3.1, and 3.2 using validation data where focal length and f-number are varied. .......................... 140
5.31 Comparison of predictions for GIQEs 3, 3.1, and 3.2 using validation data where SNR is degraded for Q\textsubscript{1} (A and B) and Q\textsubscript{2} (C and D) images. ........ 141
5.32 Temporal coverage is determined by comparing the maximum solar zenith angle achieved when ∆NIIRS reached -0.2 relative to the baseline system. ........................... 143
5.33 Signal variation is modeled using a cosine projection and estimated atmospheric transmission. Atmospheric optical density at noon is calculated using an estimated VIS transmittance of 0.6 vertically through 100km of atmosphere 144
5.34 Q\textsubscript{2} designs operating at 70, 80, 90, and 100 percent Q\textsubscript{1} diameter. 80, 90, and 100% designs experience a 129, 47, and 15 min loss in coverage respectively. This amounts to 22.7, 8, and 2.6 percent coverage loss. ............. 145
5.35 Temporal coverage of Q\textsubscript{1} design at low detector and pointing performance compared to A) Q\textsubscript{2} designs with high pointing stability B) Q\textsubscript{2} designs with high detector performance and C) Q\textsubscript{2} designs with high detector performance and pointing stability. ........................................ 146
5.36 ∆NIIRS curve comparisons between the baseline configuration, Q\textsubscript{2} at 80% diameter with low detector and pointing performance, Q\textsubscript{2} at 80% diameter with low detector and high pointing performance, Q\textsubscript{2} at 80% diameter with high detector and low pointing performance, and Q\textsubscript{2} at 80% diameter with high detector and high pointing performance. ........................ 148
5.37 ΔNIIRS curve comparison between the baseline configuration, $Q_1$ with high detector performance, and $Q_2$ at 100% $Q_1$ diameter and high detector performance. .................................................. 149

5.38 ΔNIIRS curve comparison between the baseline configuration, $Q_1$, $Q_2$ at 80%, 90%, and 100% $Q_1$ diameter with high detector and pointing performance. Blue dots represent a cutoff of -0.2 ΔNIIRS with respect to the baseline $Q_1$ system, and red dots represent a cutoff of -0.2 ΔNIIRS with respect to the maximum performance of each system. ................................. 150

5.39 Plot of rating differences from the mean as a function of rating order. ... 151

5.40 Plot of standard deviation a function of rating order. ............................... 152

5.41 Two image tiles selected from the transparency (A) Nadir image of B-52 (B) Nadir image of F-102 (top) and F-86 (bottom) ................................. 153

5.42 Plot of average ΔNIIRS between B-52 and jet images. P-value = 0.0259. ... 154

5.43 Plot of average ΔNIIRS between left (L) and right (R) rating stations. P-value = 0.43. ................................. 155

A.1 Inputs and Outputs of Digital Model ................................. 164

B.1 Testbed flow chart. ................................. 166

C.1 MTF and RMSE plots for $f= 16$mm and $FN = 16$ ................................. 167
C.2 MTF and RMSE plots for $f= 16$mm and $FN = 8$ ................................. 168
C.3 MTF and RMSE plots for $f= 25$mm and $FN = 16$ ................................. 168
C.4 MTF and RMSE plots for $f= 25$mm and $FN = 8$ ................................. 169
C.5 MTF and RMSE plots for $f= 35$mm and $FN = 16$ ................................. 169
C.6 MTF and RMSE plots for $f= 35$mm and $FN = 8$ ................................. 170
C.7 MTF and RMSE plots for $f= 50$mm and $FN = 16$ ................................. 170
C.8 MTF and RMSE plots for $f= 50$mm and $FN = 8$ ................................. 171
List of Tables

3.1 GIQE versions 3.0 and 4.0 coefficients. ........................................... 15
3.2 Range of values used in GIQE4 regression study. ................................. 19
3.3 Range and Timing Smear for Ikonos II and Quickbird 2 ......................... 26
3.4 Systems trades when Q increases in relative proportions. ....................... 34
3.5 Treatment table for $3^3$ factorial experiment .................................... 57

4.1 Estimated noise characteristics for $T_{CCD} = -15^\circ C$ ......................... 62
4.2 STF-8300M Dark Current Specifications ............................................ 67
4.3 Measured vs specified densities for QA-61 test target. ............................ 79
4.4 Two-term gaussian coefficients. ......................................................... 80
4.5 Estimated scanner noise contributions (in electrons) for each configuration. 82
4.6 Camera configurations that shall undergo objective IQA model validation . 90
4.7 GIQE-based predictions of $\Delta NIIRS$ values for each configuration. $\Delta NIIRS$ calculated with respect to $f = 25\text{mm}$ operating at $Q_1$. .................................................. 94
4.8 $\Delta NIIRS$ values associated with $NE\Delta\rho$ using Equation 4.23. ............ 95
4.9 Treatments for design of experiments. .................................................. 97
4.10 Additional treatments collected for design of experiments. ....................... 97
4.11 QE and read noise levels for low and high signal to noise performance. .... 99
4.12 Treatment table for mixed 3-level and 2-level four factor experiment. Ex-
perimental factors are Q-sampling, smear level (pointing stability), signal-
to-noise performance level, and relative integration time. ......................... 101
4.13 Treatment table for additional data collected at $Q = 1.7$. ....................... 102

5.1 Summary of defocus error and RMSE for modeled MTFs. ........................ 110
5.2 ANOVA of initial DOE data. Source is significant at the 5% level when p-value< 0.05. ................................................................. 119
5.3 Parameters used for estimating smear .................................................. 120
5.4 ANOVA of DOE data combined from first and second collection. Source is significant at the 5% level when p-value< 0.05. ................................. 123
5.5 Quadratic fit of combined design of experiments data to test for curvature. Term is significant at the 5% level when p-value< 0.05. ......................... 124
5.6 Linear fit of combined design of experiments data after eliminating quadratic terms for Q and Detector Performance. Term is significant at the 5% level when p-value< 0.05. ............................................. 124
<table>
<thead>
<tr>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.7</td>
</tr>
<tr>
<td>5.8</td>
</tr>
<tr>
<td>5.9</td>
</tr>
<tr>
<td>5.10</td>
</tr>
<tr>
<td>5.11</td>
</tr>
<tr>
<td>5.12</td>
</tr>
<tr>
<td>5.13</td>
</tr>
<tr>
<td>D.1</td>
</tr>
<tr>
<td>D.2</td>
</tr>
<tr>
<td>D.3</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The remote sensing community is reliant on continued efforts to produce imaging systems with ever-increasing performance capabilities at lower cost. Unfortunately, developments are often hampered by real world complications highlighted especially in space-based remote sensing where mission requirements, schedule constraints, and budgetary limitations are compounded by physical restrictions such as size, weight, and access to the spacecraft. These conditions impose system trades throughout the acquisition lifecycle, ultimately resulting in a configuration that ideally balances requirements with cost and schedule. This effort specifically addresses the tradeoff between detector sampling and optical resolution, a specification often summarized by the Q-metric.

Q is defined as the ratio between detector sampling frequency and optical cutoff frequency. Presently remote sensing platforms are intentionally designed to exhibit detector-limited performance (Boucher et al., 2010) in exchange for increases in overall system robustness. Though this trade comes at the cost of larger optics and aliasing induced by undersampling, a shift toward a theoretically ideal Q may reduce image quality due to reductions in SNR, field of view, and increases in pointing sensitivity (Fiete, 1999). Modern low noise, large format detectors show promise in counteracting the shortfalls associated with an ideally sampled system, and may have the potential to significantly reduce the cost of space borne remote sensing platforms by extracting improved performance from smaller optics. The research effort reported here focuses on establishing a validated image chain simulation through comparison of model outputs to laboratory results. Furthermore, the validated model is used to explore and characterize the effects of fundamental design parameters on image quality, with consideration of the operational trades inherent to a Q=2 design.
Because our goal is to model the potential performance of systems that leverage current and projected advances in detector performance, a trusted model of system performance is required. Model validation is driven by an end-to-end image chain simulation configured to mirror the laboratory environment, the main components of which include a charge coupled device (CCD) camera, fixed focal length refractive lenses, a light source, and target. The first phase of validation experiments focuses on objective measures such as MTF and resolution and seeks to establish a baseline confidence level in our modeling techniques. While these metrics bring to light an imaging system’s performance boundaries, they fail in predicting the interpretability of images produced due to the complexities of the human visual system (HVS). Thus, a cornerstone of model validation centers on a psychophysical study intimately tied to the National Image Interpretability Rating Scales (NIIRS) (Irvine, 1997) and the General Image Quality Equation (GIQE) (Leachtenauer et al., 1997).

NIIRS is a subjective, task-based image quality metric that quantifies the interpretability of an image using a single numerical value. To bridge the gap between NIIRS ratings and physical collection parameters, the GIQE was developed, enabling system designers to predict NIIRS performance and conduct trade studies prior to building and testing a new sensor system (Leachtenauer et al., 1997). Though the GIQE is a widely used model, an inherent limitation stems from the fact that it was derived through a regression of empirical data. Thus, low confidence may be attributed to predictions for systems whose specifications extend beyond the envelope of the original data set. This condition holds true for the proposed effort as the GIQE was developed using systems whose Q sampling roughly equaled one. Because the HVS is integral to NIIRS and the GIQE, a psychophysical study must be the primary vehicle in which interpretability of imagery is validated.

Several studies within the Remote Sensing community have employed psychophysical experiments to evaluate image quality produced by simulations. Fiete and Tantalo (1999) evaluated effects of increased along-scan sampling on image quality through analyst-assigned ∆NIIRS ratings and concluded that a ∆NIIRS increase of 0.35 was achieved when moving from a Q=1 to a Q=2 design. Similarly, Thurman and Fienup (2008, 2010) conducted psychophysical studies to analyze the GIQE and the effects of aberrations using a digital Snellen eye chart. These studies implemented a ∆NIIRS evaluation by having test subjects match model outputs to a reference set of images whose image quality was varied by adjusting the Ground Sample Distance (GSD). Although knowledge obtained from such research is valuable, the central limitation is that the simulations lack laboratory or operational validation, calling into question the legitimacy of model outputs. Using similar techniques,
the proposed study shall validate the image chain simulation by comparison of images that have been pushed through the digital model and optical testbed respectively. Participants in the experiment are asked to assess the interpretability of the images through comparison with a common marker set, serving as a link between testbed and model imagery. Provided that the model is correctly simulating the test bed, the results of the data should fall on a unit-slope line passing through the origin.

With a validated model, the study characterizes the effects of system parameters on image quality for $Q=2$ designs. Fiete (1999) used simulations and GIQE predictions to illustrate the complex interactions of smear, jitter, SNR, etc. on image quality when considering designs with ideal sampling. Results of the study indicated that increasing sampling above $Q=1$ required $Q^2$ longer integration times to maintain SNR, $Q^3$ increases in pointing accuracy to counteract smear, and $Q^2$ increases in pixel count to maintain field of view (FOV). Similarly, Cochrane et al. (2013) explored system trades for $Q=2$ designs utilizing a physical testbed capable of introducing smear, atmospheric haze, and noise. Both studies conducted a qualitative comparison of output images with obvious changes to image quality based on simulation/testbed configuration. Cochrane et al. (2013) assert that image quality cannot degrade when increasing $Q$ provided that aperture and integration time remain fixed relative to a $Q=1$ design. When examining results, images at high SNR with constant aperture and integration time increase in quality with increased $Q$ - at low SNR differences in image quality were negligible. Scenarios that varied integration time to compensate for SNR losses associated with increased $Q$ appeared to decrease in quality for both low and high SNR cases. Though both efforts provide a visual feel for changes to image quality, they fail to determine the actual interpretability differences between each image and did not use statistical measures to convey the significance of each parameter. The proposed effort attempts to address these shortfalls by exploring the tradespace using $\Delta$NIIRS ratings in conjunction with Design of Experiments (DOE) approaches to collecting and analyzing data. Using these methods, the study seeks to provide a quantitative and statistically-based comparison of the input parameters’ effect on image quality for ideally sampled imaging systems. This is designed to give us a model to allow consideration of performance opportunities associated with large format, low noise detectors with improved pointing stability.
Chapter 2

Objectives

The overarching objective of this research effort is to advance the current state of knowledge pertaining to earth observing, high Q remote sensing systems. This shall be accomplished through laboratory validation of a high Q model with a monolithic aperture observing a panchromatic light source followed by application of the validated model to more stressing cases. As alluded to in the previous chapter, model validation is a complex process, thus it is helpful to divide this endeavor into smaller “sub-objectives,” each with specific tasks. Section 2.1 outlines the main objectives associated with this effort and Section 2.2 is an itemization of tasks that must be accomplished in order to fulfill each objective. Lastly, Section 2.3 summarizes the study’s contribution to the field of Remote Sensing.

2.1 Objectives

1. Develop a mastery of the principles and phenomenology related to the imaging chain and image quality with application to remote sensing.

2. Define and create a set of virtual and real world images to appropriately assess objective and interpretability-based image quality metrics

3. Develop a test and evaluation master plan for objective and subjective image quality assessment experiments

4. Construct a laboratory testbed capable of capturing imagery under high Q, low SNR conditions
5. Design an end-to-end image chain simulation capability mirroring the laboratory environment

6. Evaluate the image quality of simulation and testbed outputs utilizing objective and interpretability-based measures

7. Utilize validated model to explore performance boundaries of monolithic aperture imaging systems operating under high Q, low SNR conditions

2.2 Tasks

1. Develop a mastery of the principles and phenomenology related to the imaging chain and image quality with application to remote sensing. Applicable topics include:

   - Application of linear systems theory to determine and analyze the system-level optical transfer function of a monolithic aperture system imaging under incoherent conditions.
   
   - Effect of various noise sources including read noise, dark noise, photon noise, etc., on image quality.
   
   - Typical image reconstruction algorithms used to compensate for performance losses and their effects on image quality.
   
   - Objective and interpretability-based image quality metrics and measurement techniques used in real-world systems.

2. Define and create a set of virtual and real world images to appropriately assess objective and interpretability-based image quality metrics

   - Obtain slant edge and resolution (hyperbolic, tri-bar, etc.) images to measure objective performance metrics.
   
   - Define requirements for scene content and tasks so that interpretability performance may be properly assessed using untrained analysts.
   
   - Using the aforementioned requirements, obtain a high resolution image and ensure that actual image sampling far exceeds detector sampling.
• Devise a method for displaying the high resolution image to the testbed and model.
• Create a marker-set from the high resolution image.

3. Develop a test and evaluation master plan for objective and subjective image quality assessment experiments
   • Utilize design of experiments best practices to craft a test plan that minimizes uncertainty of data collection from testbed.
   • Define a psychophysical test environment and methodology that consistently determines the interpretability of simulation and testbed images. Factors to consider include room lighting, display calibration, seating position, image interpolation, and others.
   • Create a schedule for data collection

4. Construct a laboratory testbed capable of capturing imagery at high Q
   • Obtain hardware and software required for the laboratory. Main components include a lab computer, CCD camera, fixed focal length lenses, filters, and light source.

5. Design an end-to-end image chain simulation capability mirroring the laboratory environment
   • Characterize each component of the system under test to determine key parameters in the simulation effort
   • Using these data, develop equivalent scene radiance, MTF, and noise models of the testbed under chosen configurations.

6. Evaluate the image quality of simulation and testbed outputs utilizing objective and interpretability measures.
   • Run reference tests under modest Q and SNR to ensure that results match known expected performance
   • Measure objective quality metrics under high Q, low SNR conditions using methods outlined in ISO12233
Chapter 2. **Objectives**

- Compare objective image quality metrics of simulation and testbed
- Conduct a psychophysical experiment using untrained analysts to assess interpretability of model and testbed outputs
- Perform a ΔNIIRS analysis on psychophysical test data. If results indicate that model does not accurately predict performance of testbed, determine sources of error which may contribute to discrepancies

7. **Utilize validated model to explore performance boundaries of monolithic aperture imaging systems operating at high Q**

- Given performance increases of current imaging systems, investigate predicted image quality while accounting for trades experienced while operating at high Q, e.g., decreased signal, decreased field of view, increased pointing/jitter sensitivity

### 2.3 Contribution to Field

The contribution of this work to the field of remote sensing are two-fold:

1. Firstly, the proposed work is unique in that elements of a laboratory testbed and psychophysical study are combined to validate the interpretability performance of a model within well understood performance boundaries (Q=1, high SNR), and at the edges of current operating limits (Q=2, low SNR). This draws contrast to current work by grounding model outputs to a known hardware configuration in a controlled environment. The validation process goes beyond standard imaging system objective assessments (e.g. modulation transfer function (MTF) and resolution) and even qualitative assessments by using ΔNIIRS studies to provide quantitative evidence of model interpretability performance.

2. Secondly, a rigorously validated model provides the mechanism to evaluate interpretability performance of systems operating beyond the limits of well characterized conditions and specifications. Given that the current GIQE is inherently limited to the data set used in its regression, it is no surprise that performance predictions have been shown to breakdown for system specifications that extend past these boundaries. Using the validated model to characterize interpretability performance of configurations at these limits may provide insight into how the GIQE may be modified to accurately predict performance at Q=2 and low SNR.
Chapter 3

Background and Theory

This chapter details the background and theoretical foundation necessary to develop an approach that accomplishes the objectives outlined in Chapter 2. It begins with a detailed look at linear shift-invariant systems in order to lay the fundamental principles central to the modeling and validation effort. This is followed by a discussion on image quality and sets bounds on the examination by establishing the definition of quality that pertains to the research. Because image quality assessments (IQA) are central to evaluating the performance of the testbed and model, Section 3.2 covers various methods for assessing image quality and presents a case for which methods will be used in the analysis. Both objective and subjective assessments are proposed, the latter of which will be evaluated according to the NIIRS scale, described in Section 3.3. The section also explains the various terms in the GIQE, an empirically derived model that is used to predict the NIIRS performance of an imaging system. Because image quality is also affected by sensor type and collection geometry, sensor types as well as the distortions and degradations introduced by different collection scenarios are covered. This effort concentrates specifically on the Q sampling factor, which is derived in Section 3.6. The chapter then breaks down how the laboratory environment was modeled, highlighting specific assumptions used to develop the mathematical representation of real-world objects. Lastly, the chapter describes the filtering process used to reconstruct testbed and simulation imagery prior to conducting image quality assessments.
3.1 Linear Shift-Invariant Systems

In its most basic form, a system is a process in which a set of inputs are mapped to a set of outputs. For imaging systems, inputs and outputs may be real or complex-valued, and are functions of a two-dimensional independent space variable, i.e., coordinates on an image plane. Given a system $S\{}$ operating on an input function $f(x, y)$ with an output $g(x, y)$, the expression may be written as

$$S\{f(x, y)\} = g(x, y)$$

(3.1)

Using Equation 3.1, Goodman (2005) presents the following development for cases when linearity and space-invariance are imposed. For a set of functions, $f_1(x, y)$ and $f_2(x, y)$, a system $S$ is said to be linear if it abides by the following superposition property:

$$S\{\alpha_1 f_1(x, y) + \alpha_2 f_2(x, y)\} = \alpha_1 S\{f_1(x, y)\} + \alpha_2 S\{f_2(x, y)\}$$

(3.2)

where $\alpha_1$ and $\alpha_2$ are complex-valued constants. Using the sifting property of the dirac-delta function, $f(x, y)$ may be expressed as a decomposition

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta)\delta(x - \xi, y - \eta) \, d\xi \, d\eta$$

(3.3)

Therefore, a system operating on $f(x, y)$ can be written as

$$g(x, y) = S\{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta)\delta(x - \xi, y - \eta) \, d\xi \, d\eta\}$$

(3.4)

It is seen from Equation 3.4 that $f(x_1, y_1)$ may be treated as a weighting function to $\delta(x_1 - x_0, y_1 - y_0)$, and using the linearity property, allows the output to be expressed as

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta)S\{\delta(x - \xi, y - \eta)\} \, d\xi \, d\eta$$

(3.5)

where $S\{\delta(x_1 - x_0, y_1 - y_0)\}$ is said to the the system’s impulse response (IPR), or from an imaging perspective, the Point-Spread Function (PSF). Re-writing the IPR as
the Superposition integral is formed

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) h(x, y; \xi, \eta) \, d\xi \, d\eta$$

Furthermore, a system is said to be space-invariant if the operation of $S\{}$ is unaffected by the position of the input. In other words, a system $S$ is said to be space-invariant if the following condition is met:

$$S\{f(x - x_0, y - y_0)\} = g(x - x_0, y - y_0)$$

Although real-world imaging systems do not meet this requirement due to perturbations caused by the environment and imperfections in the components, a system may be sufficiently invariant over a small portion of the field called an isoplanatic patch. Given this assumption, Equation 3.5 becomes

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi, \eta) h(x - \xi, y - \eta) \, d\xi \, d\eta = f(x, y) * h(x, y)$$

where $*$ is the convolution operator. Equation 3.9 is the mathematical foundation upon which the modeling and validation effort will be built.

### 3.2 Image Quality

The definition of image quality varies wildly depending on the circumstances in which it is evaluated. In order to properly frame the study, ”quality” is ultimately viewed from the context of task-based image assessment by human observers, where the ability to perform more complex interpretation tasks is associated with higher quality. Based on that definition, it is possible to determine image quality measures applicable to the validation effort. Numerous works classify image quality assessments (IQA) into two broad categories: subjective and objective (Thung and Raveendran, 2009, Eskicioglu and Fisher, 1993, Boberg, 1993).
The former involves using humans to evaluate the quality of the images and may either be expert or non-expert participants (Eskicioglu and Fisher, 1993). Non-expert participants are considered to be "average" viewers, whereas experts are considered to be trained individuals (e.g. image analysts) that are qualified to provide a more refined assessment of image quality. Objective assessments use various methods to compute the quality of an image based on the number and types of data provided for analysis.

![Image Quality Assessment Categories](image.png)

**Figure 3.1:** Image quality assessment categories.

From Figure 3.1, objective assessments are classified into full-reference, reduced-reference, and no-reference measurement techniques. Using full-reference methods image quality is calculated by comparing a distorted image to a reference image which is assumed to contain truth data. Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) are commonly used full-reference metrics, however these have been shown to be inadequate at determining image quality, especially in cases where the image must be evaluated visually (Thung and Raveendran, 2009, Eskicioglu and Fisher, 1993, Janssen, 2001, Jain, 1981). Reduced-reference methods seek to accomplish the same comparison, but with only partial information about the reference image. Unfortunately in the case of aerial or space-based imagery, reference images are often unavailable, therefore no-reference, or blind IQA are essential. No-reference IQA employ algorithms to quantify image quality using data present only in the distorted image and are generally categorized into one of the following trends: 1) distortion-specific approaches 2) training based approaches 3) natural scene statistics (NSS) approaches (Saad et al., 2012). Because the image-chain simulation will distort the
test image by applying blur and noise, the first of these approaches will be used to determine the Modulation Transfer Function (MTF) and resolution from distorted imagery.

The MTF may be derived from the PSF, line spread function (LSF), or the edge spread function (ESF) (Boreman, 2001). Presenting a point source to the imaging system allows the PSF to be captured, from which the two-dimensional MTF is derived by taking the Fourier transform of the image. This is illustrated by substituting a dirac-delta function into Equation 3.9

$$g(x, y) = \delta(x, y) \ast h(x, y) = h(x, y)$$

(3.10)

$$|\mathcal{F}\{PSF(x, y)\}| = MTF(\xi, \eta)$$

(3.11)

By displaying a line to the camera, a one dimensional MTF perpendicular to the line may be found.

$$g(x, y) = \delta(x)1(y) \ast h(x, y) = LSF(x)$$

(3.12)

$$|\mathcal{F}\{LSF(x)\}| = MTF(\xi)$$

(3.13)

Difficulties associated with both of these methods stem from finite detector sampling and ensuring that the spatial extent of the point and line sources are representative of dirac delta and line delta functions. Presenting an knife-edge to the camera circumvents the latter and a method that addresses the former is discussed in the following chapter. The mathematical representation for a knife edge is a step function, $STEP(x)$, the one-dimensional response from which is the ESF.

$$g(x, y) = STEP(x)1(y) \ast h(x, y) = ESF(x)$$

(3.14)

From the ESF, the LSF is found by
\[ LSF(x) = \frac{d}{dx} (ESF(x)) \] \hspace{1cm} (3.15)

Once the LSF is derived from the ESF, Equation 3.13 may be used to determine the MTF. Figure 3.2 summarizes this process.

![Diagram of Edge Spread Function, Line Spread Function, and Modulation Transfer Function](image)

**Figure 3.2:** (A) Edge spread function (B) LSF found by taking the derivative of the ESF (C) MTF found by taking the Fourier Transform of the LSF

Although an objective assessment will provide some indication of the consistency between laboratory and simulation outputs, the definition of quality this study is concerned with ultimately requires subjective assessment by human observers. The two general approaches to subjective IQA are 1) single stimulus, or absolute evaluations, where users assess the quality of an image by assigning it to a category in a given rating scale and 2) double stimulus, or comparative evaluations, where users are given a set of images and are asked to rate the image quality of each with respect to one another (Thung and Raveendran, 2009, Eskicioglu and Fisher, 1993). These IQA are carried out as psychophysical experiments and have the potential to further understanding of the mechanisms behind the perception of quality, or in this case interpretability. In addition, careful consideration must be taken when conducting such experiments as the results are highly dependent on a number of factors to include the expertise of the user, display type (e.g. software vs. hardcopy), recognition task, etc. This effort is constrained to using NIIRS, the subjective rating scale predominantly used by the remote sensing community. A caveat to using this scale is that trained analysts are required to assign absolute NIIRS values to imagery. As this study does not employ trained analysts, a comparative image scaling assessment, or (∆NIIRS), is used to determine relative differences in interpretability.
3.3 NIIRS and GIQE

NIIRS is a subjective numerical scale that is an interpretability-based assessment of aerial imagery (Irvine, 1997). NIIRS contains ten levels ranging from 0 to 9, each associated with numerous interpretability tasks that increase in difficulty with a higher numerical rating, e.g., "Detect Large Buildings" is associated with NIIRS 2 and "Identify the spare tire on a medium sized truck" is associated with NIIRS 6. Based on these examples, it is seen that each interpretability task, or criteria, consist of three elements, namely the recognition level, an object, and a qualifier. Using these criteria, trained analysts assign a NIIRS rating to an image by determining whether the recognition task associated with the object could be accomplished if the object were present in the image. The NIIRS scale was designed so that a unity change on the scale roughly corresponds to a ground sample distance (GSD) factor of two. Although NIIRS was constructed on contain 10 discrete levels, a ∆NIIRS assessment on imagery is often made, and is independent of the absolute NIIRS rating assigned to an image (Fiete, 1999). A ∆NIIRS of 0.1 or below is difficult to detect, whereas changes above 0.2 are easily perceptible (Fiete and Tantalo, 1999).

The usefulness of NIIRS stems from an ability to convey the quality of an image using a single metric, drawing contrast to statistical or more complex MTF-based measures that may be cryptic and at times misleading in nature. Though this metric has proven successful in communicating image quality requirements between project stakeholders, neither the levels nor the criteria define performance specifications associated with the collection system. Decomposing a NIIRS value into quantitative system specifications is a necessity when conducting trade studies on the overall system design; thus a link between the subjective NIIRS assessment and an imaging system’s design parameters was formed through the GIQE. The GIQE was the product of a regression that derived an empirical relationship between system performance specifications and analyst-assigned NIIRS values. Currently, three versions of the GIQE (3, 4, and 5) are available to the public, with GIQEs 3 and 4 being the most widely used. These versions utilize inputs measured after image processing and take on the form

\[
NIIRS = c_0 - c_1 \log_{10}(GSD_{GM}) + c_2 \log_{10}(RER_{GM}) - c_3 H - c_4 \frac{G}{SNR} \tag{3.16}
\]

where \(GSD_{GM}\) is the geometric-mean GSD in inches, \(RER_{GM}\) is the geometric-mean of the normalized Relative Edge Response, \(H\) and \(G\) are the height overshoot and noise gain
caused by the sharpening, respectively, $c_0 - c_4$ are regression coefficients (Table 3.1), and SNR is the Signal-to-Noise Ratio (Leachtenauer et al., 1997).

**Table 3.1: GIQE versions 3.0 and 4.0 coefficients.**

<table>
<thead>
<tr>
<th></th>
<th>$c_0$</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIQE 3.0</td>
<td>11.81</td>
<td>-3.32</td>
<td>3.32</td>
<td>-1</td>
<td>-1.48</td>
</tr>
<tr>
<td>GIQE 4.0 (RER &gt; 0.9)</td>
<td>10.251</td>
<td>-3.32</td>
<td>1.559</td>
<td>-0.334</td>
<td>-0.656</td>
</tr>
<tr>
<td>GIQE 4.0 (RER &lt; 0.9)</td>
<td>10.251</td>
<td>-3.16</td>
<td>2.817</td>
<td>-0.334</td>
<td>-0.656</td>
</tr>
</tbody>
</table>

For GIQE 4.0, $GSD_{GM}$ is the geometric mean of the projected pixel pitch footprint on the ground plane measured in the X and Y directions, whereas in GIQE 3.0 it is measured perpendicular to the line of sight. The following development focuses on its definition with respect to GIQE 4.0.

$$GSD_{GM} = (GSD_X GSD_Y \sin(\alpha))^{1/2}$$ (3.17)

where $\alpha$ is the angle between the along-scan and cross-scan directions. Given the geometry in Figure 3.3, $GSD_X$ and $GSD_Y$ are given as

$$GSD_X = \frac{p_x}{f} H \sqrt{\cos^2 \gamma + \frac{\sin^2 \gamma}{\cos^2 \psi}}$$ (3.18)

$$GSD_Y = \frac{p_y}{f} H \sqrt{\sin^2 \gamma + \frac{\cos^2 \gamma}{\cos^2 \psi}}$$ (3.19)
where $f$ is the focal length, $p_x$ and $p_y$ are the detector pitch in the x and y directions, $H$ is the sensor height, $\psi$ is the zenith angle, and $\gamma$ is the angle between the x-axis and the ground plane. Given $p_x = p_y$ and $\gamma = 0$, i.e., when $GSD_x$ and $GSD_y$ are aligned with the cross and along track directions respectively, $GSD_{GM}$ reduces to

$$GSD_{GM} = \frac{p}{f} H \frac{1}{\sqrt{\cos \psi}}$$

The RER term accounts for the system’s MTF, which include the optics, detector, atmospheric and environmental effects, motion, and the sharpening kernel. It is a slope measurement of the systems response to a normalized edge 0.5 pixels before and after the occurrence of the edge (Fig. 3.4 ).
Mathematically, RER may be computed as

\[ RER = ER(0.5) - ER(-0.5) \]  \hspace{1cm} (3.21)

where \( ER \) is the systems response to an edge, seen in Driggers et al. (1997) as

\[ ER_x(x) = \frac{1}{2} + \frac{1}{\pi} \int_{0}^{\infty} \frac{MTF_x(\xi)}{\xi} \sin(2\pi x \xi) d\xi \]  \hspace{1cm} (3.22)

Edge overshoot is represented by the \( H \) term which acts as a penalty for over-sharpening during the image restoration process. It is measured between 1 and 3 pixels away from the edge in 0.25 pixel increments. For a monotonically increasing edge response, \( H \) is the value of the edge response function 1.25 pixels from the occurrence of the edge, otherwise, \( H \) is calculated as the maximum value of the edge response function (Fig. 3.5).
The noise gain, $G$, is computed by taking the root of the sum of the squares of the sharpening kernel values:

$$G = \sqrt{\sum_{m=1}^{M} \sum_{n=1}^{N} (kernel_{m,n})^2}$$

(3.23)

where

$$\sum_{m=1}^{M} \sum_{n=1}^{N} (kernel_{m,n}) = 1$$

(3.24)

Lastly, the SNR is calculated as the signal differential between a 7% and 15% lambertian reflectance target divided by the noise

$$SNR = \frac{Signal_{15\%} - Signal_{7\%}}{Noise}$$

(3.25)

It is important to note that the SNR for GIQE regression data was model-generated, therefore knowledge of the optics, detector, noise characteristics, as well as the illumination and atmospheric conditions was required to determine the SNR for each image. Because the
GIQE was developed from a regression, its accuracy is bound by the data used in the study, the range of which is seen in Table 3.2.

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Mean</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSD (in)</td>
<td>3</td>
<td>20.6</td>
<td>80</td>
</tr>
<tr>
<td>RER</td>
<td>0.2</td>
<td>0.92</td>
<td>1.3</td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td>10.66</td>
<td>19</td>
</tr>
<tr>
<td>SNR</td>
<td>2</td>
<td>52.3</td>
<td>130</td>
</tr>
<tr>
<td>G/SNR</td>
<td>0.01</td>
<td>-</td>
<td>1.8</td>
</tr>
<tr>
<td>H</td>
<td>0.9</td>
<td>1.31</td>
<td>1.9</td>
</tr>
</tbody>
</table>

GIQE 3.0 was originally released in 1994 to the unmanned aerial vehicle development community having a measured standard error of 0.4 NIIRS and GIQE 4.0 followed in 1997. A total of 359 images were used to generate the GIQE 4.0 coefficients, half of which were reserved for testing and validation (Leachtenauer et al., 1997). To set a baseline for developing GIQE 4.0 coefficients the original GIQE construct was regressed against analyst-provided NIIRS ratings. Initial results found that the coefficients for $RER/GSD$ and $G/SNR$ were much lower than those found in GIQE 3.0, prompting analyses that led to RER being dual-sloped in GIQE 4.0 (Table 3.1). Based on this change, the updated coefficients used in GIQE 4.0 resulted in $R^2 = 0.986$ and a standard error of 0.282 for the development half of the images. Using the validation half of the image set, GIQE 4.0 produced a standard error of 0.307 and $R^2 = 0.934$. The GIQEs serve as a starting point for predictive modeling in this work, and latter chapters explore the capabilities and shortfalls of each.

### 3.4 Remote Sensing Systems

Thus far the discussion has centered around the quality of the output from remote sensing instruments, i.e., subjective image quality in terms of NIIRS. This section details various remote sensing systems, each with varying effects on image quality based on their collection geometry. Generally speaking, remote sensing payloads gather strips of data as the platform (e.g. satellite, aircraft, ground vehicle) moves in the along-track direction. Limiting the scope of this discussion to space-based imaging, the basic collection geometry is illustrated in the figure below.
Chapter 3. *Background and Theory*

Depending on how the sensor collects data, remote sensing systems can further be divided into three general categories: cross-track scanning (line and whiskbroom scanners), along-track scanning (pushbroom scanners), and framing imagers (Schott, 2007). Cross-track scanning sensors incorporate mirrors within the optical system to sweep the detector’s instantaneous field of view (IFOV) in the cross-track direction. If only one detector element is used, these systems are referred to as line scanners, which generate images by sampling the detector output and tiling together individual surface resolution elements in the cross-track and along-track directions (Fig. 3.7). Because only one detector element is used, this design allows for fairly simple optics as only the central portion of the optical field of view is seen by the detector. However, the number of samples in the cross-track direction, coupled with the along-track size of the detector’s ground instantaneous field of view (GIFOV) severely limit the dwell time of the detector over each resolution element. Similarly, whiskbroom scanners employ a linear array of detectors, each scanning a line in the cross-track direction. Dwell time is increased due to the elongated FOV in the along-track direction, however this improvement places tighter requirements on optical quality and may require the use of a scan line corrector to eliminate gaps along the ground track caused by satellite motion. Examples of scanning systems are the Landsat instruments such as the Multispectral Scanner (MSS) and Thematic Mapper (TM) and the Enhanced Thematic Mapper Plus (ETM+).
Imaging Optics
Scanning Mirror
Detector

**Figure 3.7:** Collection geometry for line-scanners. Geometry is the same for whiskbroom scanners, with the exception of using a linear array of detectors in the along-track direction.

Line and whiskbroom scanners are subject to various geometric distortions which vary in severity depending on the stabilization capability of the platform and scanning precision of the payload. Specific perturbations include pointing (roll, pitch, yaw) errors, velocity over altitude (V/H) errors, and tangent distortions which are corrected using the global positioning system (GPS), inertial navigation systems (INS), and geometric resampling. In addition to these distortions, relative motion between the detector and ground object induce image smear. For line and whiskbroom scanners, Earth rotation is generally not a factor because it is much slower relative to the scan rate - although it should be taken into account when registering the imagery. An obvious form of smear arises from interactions between pointing stability and integration time, however proper sizing of these specifications does not guarantee low and consistent levels of smear over the entire image. Forms of smear inherent to collection geometry, platform motion, and the number of detector elements are unavoidable and vary the amount of smear based on a pixels location in the image - these include zenith smear, azimuth smear, range smear (Auelmann, 1997). The two limiting cases for these smear contributions are when the line of site (LOS) points in the orbit plane when fore or aft-looking, and when the LOS is pointed to the side of the platform.

When fore or aft-looking, zenith smear results in an expansion or compression of the detector array ground projection in the along-track direction due to a change in zenith angle contributed by the satellite velocity vector component perpendicular to the target-to-satellite vector (Fig 3.8). This type of smear increases linearly with the number of detector elements.
and depends on the direction of the linear array - it is maximal when the array is parallel to, and along the ground track.

\[ \text{Smear}_{\text{zenith}} = -\frac{N t_{\text{int}} v_{\text{sat}} \tan(\psi) \cos(\phi - \theta)}{2r_{\text{slant}}} \] (3.26)

where \( N \) is the number of elements in the detector array, \( t_{\text{int}} \) is the integration time, \( v_{\text{sat}} \) is the magnitude of the satellite velocity vector, \( r_{\text{slant}} \) is the slant range between the target and the sensor, \( \psi \) is the zenith angle, \( \phi \) is the angle between \( v_{\text{sat}} \) and the satellite local horizontal (zero for circular orbits), and \( \theta \) is the offset pointing angle from the nadir point.

In addition to changes in the zenith angle, a velocity contribution along the LOS vector causes the detector ground projection to change scale during the integration time based on a change in range. If the center pixel in the array is tuned to experience the least amount of smear, the amount of smear increases proportionally with the number of pixels away from the center of the array and manifests in the along-track direction for whiskbroom sensors. This type of smear, generally referred to as range or zoom smear (Fig. 3.9), is calculated as (Auelmann, 1997)

\[ \text{Smear}_{\text{range}} = \frac{N t_{\text{int}} v_{\text{sat}} \sin(\phi - \theta)}{2r_{\text{slant}}} \] (3.27)
When scanning a target directly to the side of the platform, a velocity contribution, $v_x$, perpendicular to the LOS vector causes a rotation of the ground projection about the local vertical. Such a condition induces azimuth smear, given by (Auelmann, 1997)

$$\text{Smear}_{\text{azimuth}} = \frac{N_{f\text{int}}}{2r_{\text{slant}}} v_{\text{sat}} \sin (\psi) \cos (\phi)$$

(3.28)

For the side-viewing case, zenith and range smear manifest as
\[
\text{Smear}_{\text{zenith-side}} = \frac{N t_{\text{int}} v_{\text{sat}}}{2r_{\text{slant}}} \tan (\psi) \sin (\phi) \sin (\theta) \tag{3.29}
\]

\[
\text{Smear}_{\text{range-side}} = \frac{N t_{\text{int}} v_{\text{sat}}}{2r_{\text{slant}}} \sin (\phi) \cos (\theta) \tag{3.30}
\]

In both cases, \( \phi = 0^\circ \) for a circular orbit, resulting in zero zenith and range smear when the satellite is side viewing.

Pushbroom sensors eliminate the need for scanning mechanisms by using platform motion to scan a linear array of detectors oriented perpendicular to the along-track direction. This configuration allows for higher geometric fidelity compared to line and whiskbroom scanners because the sampling interval in the cross-track direction remains fixed. The reduction in mechanical complexity aides in increasing the system’s lifespan while improving radiometric sensitivity through increases in dwell time. Further increases in dwell time may be achieved through back-scanning or "nodding" the imaging platform, however changes in range and the LOS angle limit the along-track distance that can be imaged in a single pass. Alternatively, "stacking" linear arrays allows the sensor to image large areas while effectively increasing integration time through time delay integration (TDI) given by

\[
t_{\text{int}} = \frac{N_{\text{TDI}}}{\text{line rate}} \tag{3.31}
\]

where \( N_{\text{TDI}} \) is the number of TDI stages. A similar configuration also allows for the acquisition of multispectral data by using various spectral filters for each row in the array. The gains in mechanical simplicity and sensor sensitivity are balanced by losses in FOV and increased sensor complexity due to the difficulties of manufacturing large linear arrays. Examples of pushbroom systems include the Satellite Pour l’Observation de la Terre (SPOT) High Resolution Visible (HRV) sensor and the Ikonos Optical Sensor Assembly (OSA).
Similar to line and whiskbroom scanners, pushbroom sensors are susceptible to range and azimuth smear, however the effect of zenith smear is minimal due to the limited number of pixels in the along-track direction. If the spacecraft is fore or aft-looking, the difference in range experience by each TDI stage (Fig. 3.12) contributes smear in the cross-track direction.

Range smear for pushbroom sensors with this type of collection geometry is calculated using Equation 3.27, where $N$ is now the number of pixels in the cross-track direction, and $t_{\text{int}}$ may be derived from Equation 3.31. Table 3.3 lists cross-track range smear taken from (Auelmann, 2012) for Ikonos II and Quickbird 2.
Table 3.3: Range and Timing Smear for Ikonos II and Quickbird 2

<table>
<thead>
<tr>
<th></th>
<th>Ikonos II</th>
<th>Quickbird 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude (km)</td>
<td>682</td>
<td>451</td>
</tr>
<tr>
<td>Pixels</td>
<td>13,500</td>
<td>28,000</td>
</tr>
<tr>
<td>TDI stages</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Zenith Angle, (\psi) (deg)</td>
<td>52</td>
<td>32</td>
</tr>
<tr>
<td>Offset Angle, (\theta) (deg)</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>Range Smear (pix)</td>
<td>0.14</td>
<td>0.27</td>
</tr>
<tr>
<td>Timing Smear (pix)</td>
<td>0.32</td>
<td>0.37</td>
</tr>
</tbody>
</table>

For side-viewing geometries, the azimuth smear experienced by pushbroom sensors whose integration time is increased by back-scanning is calculated using Equation 3.28. Similarly, the use of TDI stages in this collection scenario contributes azimuth-like smear due to timing misalignment between the center pixel and pixels at the edge (Fig 3.13).

![Figure 3.13: Azimuth-like smear due to TDI timing misalignment for TDI arrays off-center.](image_url)

In this case, line rate timing is optimized for the pixel located at the center of the array and its subsequent TDI stages, therefore the handoff of a ground resolution element between TDI stages is mismatched for pixels at the edges of the array. Timing smear is calculated as (Auelmann, 1997)

\[ \text{Smear}_{\text{Timing}} = \frac{N}{2} N_{\text{TDI}} \text{IFOV} \sin \psi \]  

(3.32)
where IFOV is the sensor’s instantaneous field of view. It can also be shown that

\[ N_{TDI \ IFOV} = \frac{v_{sat} t_{int}}{r_{slant}} \]  

(3.33)

thus it is seen that timing smear for pushbroom scanners using TDI is equivalent to azimuth smear. Timing smear levels for Ikonos II and Quickbird 2 with side offset angles equal to the forward offset angles used when calculating range smear are seen in Table 3.3. In addition to timing misalignment, the clocking structure of each TDI stage inherently contributes an amount of smear inversely proportional to the number of clocking stages present in the TDI stages. Because these sensors are based on CCD technology, the charge generated by one TDI pixel are moved to the subsequent stage by a clock cycle that adjusts the voltage at each gate. Smear occurs because the clock cycle moves charge in discrete steps, whereas the IFOV sweeps over the ground track continuously. Remote sensing applications generally employ two or four-phase detectors, each experiencing 1/2 and 1/4 pixel smear, respectively, in the along-track direction (Smith et al., 1999). Lastly, an inherent penalty for both types of scanning sensors is the along-track or cross-track smear experienced due to the continuous motion of the IFOV as it is scanned over the integration time. This type of smear, termed scan smear, may be mitigated by properly sizing \( t_{int} \) and line rate, so that smear is minimized while producing desirable signal levels.

\[ \text{Figure 3.14: Clocking smear for systems with TDI stages.} \]
Framing arrays are extensively used on aerial platforms, however vast improvements in detector performance are continually making these sensors more attractive for remote sensing from space. In a framing sensor, the optics project an image of the ground onto a two dimensional array of detectors. Because sampling in the cross and along track directions are fixed and entire frames are captured at once, the main advantage of this detector is geometric fidelity, however high quality optics must be used in order to mitigate aberrations at the edges of the FOV. In addition, large format 2-D arrays can be costly to manufacture, though continued advances in detector technology allow the implementation of these sensors to be more commonplace. Spacecraft which have employed these sensors include the Hubble Space Telescope (HST) Wide-Field Planetary Camera (WFPC), Urthecast High Resolution Camera (HRC), and the Skybox sensor.

If the LOS vector remains unchanged, these sensors offer no signal-to-noise advantage over pushbroom scanners because the integration time must be sized according to the time it takes for the sensor to shift one GIFOV - equivalent to the line rate of a pushbroom. Furthermore, using a framing array in such a manner does not eliminate the presence of scan smear because it is essentially being operated as a pushbroom sensor. This concept of operation (CONOP) is problematic especially for high resolution imaging because of the satellite’s ground velocity and decreases in signal attributed to a smaller GIFOV. For satellites in low earth orbit (LEO), the ground velocity is roughly $7m/ms$, thus for a system designed to achieve a $0.5m$ GSD, integration time must be constrained to less than $1/14ms$ to limit scan smear in the along-track direction. In addition, a smaller GSD limits the amount of signal each pixel collects due to a smaller ground projection - a concept discussed further in Section 3.6. The implementation of a step and stare CONOP (Fig. 3.16) for
framing arrays is beneficial in that dwell time is improved by adjusting the LOS so that the aim-point is fixed over the integration time for the center of the array.

This maneuver also introduces orbital smear (range, azimuth, zenith) for pixels at the edge of the array caused by range and aspect changes, however framing arrays provide an advantage in that no scan smear (which is often the largest source of smear) is experienced.

3.5 Framing Arrays

Two types of framing arrays that have been implemented in remote sensing from space are the CCD and the complementary metal oxide semiconductor (CMOS). Because of their relative maturity, CCDs top CMOS sensors as the preeminent visible and ultraviolet wavelength image sensor used in scientific Earth observing missions. The emergence of CCDs came about because of a commercial and scientific need for a solid state imaging detector to replace tube-type detectors (Waltham, 2013, Janesick, 2001). At the same time the CCD was being developed, NASA required such a detector for the Hubble Space Telescope (HST), then called the Large Space Telescope (LST). Similarly, the Jet Propulsion Laboratory required a detector with greater sensitivity, stability, and reliability than the vidicon tubes used in past missions such as Mariner, Viking, and Voyager. Since their inception, some of the main improvement areas for CCDs include charge transfer efficiency (CTE), read noise, and quantum efficiency (QE), all of which dramatically improved the signal-to-noise performance of modern framing arrays (Janesick, 2001).
CTE is crucial to maintaining the signal accuracy and plays a larger role with the growing size of focal plane arrays. Early CCDs manufactured by Bell Laboratories exhibited CTEs on the order of 99.5%, experiencing a 40% loss in signal after only 100 transfers. The poor CTE performance of early CCDs was a result of traps at the Si-SiO$_2$ interface of surface channel CCDs. This was mitigated through the development of buried channel CCDs, which added an n-type layer between the substrate and oxide layers. As a result, the potential well was pushed deeper into the substrate away from the interface traps. Modern CCDs produce CTEs better than 99.9999% (Waltham, 2013), allowing a pixel at the corner furthest from the amplifier in a 4k × 4k CCD to experience < 1% signal loss (8k transfers).

Read noise is often considered the limiting factor in CCD noise performance and ranged between 30e$^-$ and 100e$^-$ rms for early devices. Read noise was greatly reduced by a cancellation technique called correlated double sampling (CDS), which acquires the CCD output before ($S_1$) and after ($S_2$) the signal has been transferred to the output node. By subtracting $S_1$ and $S_2$, CDS eliminates a main component of read noise, reset noise, because it remains unchanged between the two samples. As a result, read noise becomes a function of CCD readout rate, where faster readout rates produce higher read noise. The best of modern detectors achieve read noise levels of $\sim 6e^-$ when read at 1MHz and $\sim 2e^-$ at 100kHz (Waltham, 2013).

Lastly, implementation of thinned, back-illuminated CCDs allowed vast improvements in QE. Early front-side illuminated CCDs required photons to pass through semi-transparent electrodes (i.e. the gate structures) before reaching the substrate. These structures, combined with the wavelength-dependent absorption depths of photons restricted peak QE performance to roughly 45%, with significant losses in the blue region. Modern back-illuminated CCDs implement back-surface doping and anti-reflective (AR) coatings, producing a peak QE > 90% and significant improvements in sensitivity to shorter wavelengths. In summary, current and foreseeable detectors are likely to incorporate all of these improvements resulting in large arrays with high QE and very low noise. Such performance gains directly address earlier shortfalls associated with a Q-optimized design, the specifics of which are discussed in the following section.
3.6 Q

Q is defined as the ratio between the detector sampling frequency and the optical cutoff frequency (Fiete, 1999). For this study, the detector is assumed to be a 2-D array with square pixels and a fill factor of one, allowing the detector pitch to equal the width of each pixel. Using these assumptions, the detector sampling frequency is defined to be

\[
2\xi_N = \frac{1}{p} \quad (3.34)
\]

where \( p \) is the detector pixel pitch. Assuming incoherent lighting conditions, diffraction-limited optics with a circular aperture, and a given f-number \( FN \), the optical cutoff frequency, \( \rho_0 \), is

\[
\rho_0 = \frac{1}{\lambda FN} \quad (3.35)
\]

which results in the following equation for \( Q \).

\[
Q = \frac{\text{Detector Sampling Frequency}}{\text{Incoherent Optical Cutoff Frequency}} = \frac{1}{\frac{1}{\lambda FN}} = \frac{\lambda FN}{p} \quad (3.36)
\]

Spatially, \( Q \) may be interpreted as how finely the detector samples the optical point spread function (Fig. 3.17). The number of samples per optical PSF may be derived by using Equation 3.36 and the traditional definition for the width of a diffraction-limited incoherent PSF for a circular aperture, which is defined to be the diameter of the first zero in the optical PSF, or \( 2.44\lambda FN \). Solving Equation 3.36 for \( p \) and dividing the optical PSF width by that value yields

\[
\text{Samples}_{PSF} = \frac{2.44\lambda FN}{\frac{\lambda FN}{Q}} = 2.44Q \quad (3.37)
\]

Using \( Q=2 \), the ideal sampling for an incoherent diffraction-limited circular aperture is 4.88 pixels over the optical PSF, which is in contrast to the standard metric of 2 samples over the PSF.
In the frequency domain, modifying each Q term results in changes to the nyquist frequency, cutoff frequency, shape of the MTF, or a combination of the three. Simply put, a system with a Q-factor of 2 is nyquist sampled, whereas systems with \( Q < 2 \) introduce aliasing for frequencies above the nyquist frequency of the detector. In order to understand the effect of adjusting various parameters of Q in the frequency domain, it is helpful to express the performance of the detector and optics from the perspective of the object. Using the object plane as the reference coordinate system, the optical cutoff frequency in object space is

\[
\rho_{0_{obj}} = \frac{D}{\lambda f H} = \frac{D}{\lambda H}
\]  

(3.38)

where \( D \) is the optical diameter and \( H \) is the distance from the imaging system to the object plane. Similarly, the sampling frequency in object space is

\[
2\xi_{N_{obj}} = \frac{1}{p H} = \frac{1}{GSD}
\]  

(3.39)

Figure 3.18 illustrates the effect of adjusting different parameters in Q.
Case A occurs when either the diameter of the optics is reduced or a longer wavelength is used, i.e., $Q$ is increased while keeping GSD constant. Conversely, case B is achieved by decreasing $H$, increasing the focal length, or decreasing the pixel pitch, each having the net effect of increasing $Q$ and decreasing GSD. The primary focus of this effort is on the latter case, attempting to boost image quality by increasing $Q$. This tradespace is thoroughly explored by Fiete (1999), which demonstrates that adjusting the $Q$-factor leads to a number of secondary effects, such as changes to signal, pointing sensitivity, and field of view, that may decrease image quality. A reduction in signal occurs because the projected area onto the ground plane for a $Q_2$ design is one quarter the size of a pixel for $Q_1$ (Fig. 3.19).

From Figure 3.19 it can be seen that signal decreases approximately as

\[ \text{Signal} \propto \frac{1}{Q} \]
\begin{equation}
\text{Signal} \propto \frac{1}{Q^2}
\end{equation}

(3.40)

Such losses may be acceptable when considering improved CTE, noise, and QE in modern sensors, which aide in boosting the system’s overall signal-to-noise performance. It is also seen that the reduction in pixel size increases the smear experienced by a \( Q_2 \) system as

\begin{equation}
\text{Smear} \propto Q
\end{equation}

(3.41)

when expressed in terms of pixels. It should be noted that if smear is expressed in terms of ground distance, smear does not increase if no compensation for the loss in signal takes place. The decrease in signal for a \( Q_2 \) design may be counterbalanced by increasing integration time or through TDI, however smear may be increased to upwards of \( Q^3 \) if the original amount of signal is to be recovered. Lastly, if pixel density is kept constant, the field of view for a \( Q_2 \) design decreases as

\begin{equation}
\text{FOV} \propto \frac{1}{Q}
\end{equation}

(3.42)

However, if pixel density is increased by increasing the number of array elements, the FOV will remain the same, but the amount of data collected per frame increases by \( Q^2 \). The following table summarizes the trades associated with increasing the Q-factor of an imaging system.

**Table 3.4: Systems trades when Q increases in relative proportions.**

<table>
<thead>
<tr>
<th></th>
<th>( Q=1 )</th>
<th>( Q=2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>1×</td>
<td>( \frac{1}{4} \times ) (Alleviated by improved sensor performance or increased ( t_{int} ))</td>
</tr>
<tr>
<td>Smear (Pixel Fraction)</td>
<td>1×</td>
<td>2× (Up to 8× with signal compensation)</td>
</tr>
<tr>
<td>FOV</td>
<td>1×</td>
<td>( \frac{1}{4} \times ) (1× if pixel density increases)</td>
</tr>
<tr>
<td>Data</td>
<td>1×</td>
<td>4× (1× if pixel density constant)</td>
</tr>
</tbody>
</table>

It is seen that balancing detector and optics performance can lead to negative consequences (reduced signal, increased smear, reduced FOV), however partially compensating for these impacts (e.g. increased detector signal-to-noise performance), introduces the potential to significantly decrease GSD and/or optical diameter which may result in considerable cost savings and performance increases.
3.7 Image Chain Simulation

As alluded to in Section 3.1, the laboratory environment is treated as a Linear, Shift-Invariant (LSI) system, therefore the effect of each component on the object, assumed to be lambertian, is modeled as a convolution of each respective impulse response. Equation 3.9 then becomes

\[ g(x, y) = f(x, y) \ast h_1(x, y) \ast h_2(x, y) \ast \ldots \ast h_n(x, y) \]  

(3.43)

where \( h_n(x, y) \) corresponds to the impulse response of each modeled component, e.g., aperture, aberrations, detector, etc. Taking the fourier transform of Equation 3.43, and assuming that each component operates independently on an incoherent irradiance image, the total system transfer function may be found by a point-by-point multiplication of the individual subsystem transfer functions (Boreman, 2001).

\[ G(\xi, \eta) = F(\xi, \eta)MTF_{\text{system}}(\xi, \eta) \]  

(3.44)

\[ MTF_{\text{system}}(\xi, \eta) = MTF_1(\xi, \eta)MTF_2(\xi, \eta)\ldots MTF_n(\xi, \eta) \]  

(3.45)

The following discussion details the development of each component transfer function.

3.7.1 Radiometry

In real-world imaging platforms the total sensor reaching radiance is a combination of many sources that must be accounted for to accurately perform quantitative radiometric analysis. The contribution of each source is quantified through a governing equation which feeds into the first link of the image chain. In order to simulate complex real-world scenes, synthetic image generation (SIG) models and radiative transfer code are used to produce geometrically and radiometrically accurate inputs. The controls imposed by the laboratory environment in addition to key assumptions such as a lambertian object eliminate the need for such models and allow for an elementary evaluation of the radiometric processes involved. Because the laboratory uses a light box to uniformly illuminate a transparency, the source,
with respect to the object, may be sufficiently modeled as an infinite plane source. With this assumption in place, broad-band source irradiance between any two spatial coordinates on the transparency is equal

\[ E(x_1, y_1; \lambda) = E(x_2, y_2; \lambda) = E_{\text{source}}(\lambda) \quad (3.46) \]

Assuming a lambertian object (over the narrow acceptance angle of the sensor) with transmittance function \( \tau_{\text{object}}(x, y, \lambda) \), the radiance may be expressed as

\[ L_{\text{object}}(x, y, \lambda) = \frac{M(x, y, \lambda)}{\pi} = \frac{\tau_{\text{object}}(x, y, \lambda) E_{\text{source}}(\lambda)}{\pi} \quad (3.47) \]

where \( M(x, y, \lambda) \) is the exitance. From the f-number and lens transmission, \( \tau_{\text{lens}}(\lambda) \), the irradiance onto the camera sensor is calculated from the G-number, given by

\[ G# = \frac{1 + 4(FN)^2}{\tau_{\text{lens}}(\lambda)\pi} \quad (3.48) \]

where the relation is

\[ E_{\text{sensor}}(x, y, \lambda) = \frac{L_{\text{object}}(x, y, \lambda)}{G#} = \frac{E_{\text{source}}(\lambda)\tau_{\text{lens}}(\lambda)\pi}{\pi[1 + 4(FN)^2]} \tau_{\text{object}}(x, y, \lambda) \quad (3.49) \]

Taking into account exposure time and the transmission of a filter, the number of photons collected by each pixel is then found using

\[ S_{\text{photons}}(x, y) = \frac{A_{\text{pix}} t_{\text{exp}}}{[1 + 4(FN)^2]hc} \int_0^\infty E_{\text{source}}(\lambda)\tau_{\text{lens}}(\lambda)\tau_{\text{filter}}(\lambda)\tau_{\text{object}}(x, y, \lambda)\lambda d\lambda \quad (3.50) \]

where \( A_{\text{pix}} \) is the area of the CDD pixel, \( h \) is Planck’s constant, and \( c \) in the speed of light. From the QE, the signal in electrons is
Chapter 3. Background and Theory

The relationship between \( S_{\text{electrons}} \) and digital counts is experimentally determined, but may be roughly estimated using the full-well depth and the number of bits used by the analog to digital conversion circuit

\[
g_{\text{CCD}} = \frac{N_{\text{Full Well}}}{2^n} \tag{3.52}
\]

where \( N_{\text{Full Well}} \) is the full-well depth of the CCD in electrons, \( n \) is the number of bits in analog to digital units (ADU), and \( g_{\text{CCD}} \) is the gain of the CCD in \( e^-/\text{ADU} \), also referred to as the quantum step equivalency (QSE) (Fiete, 2010). Using the CCD gain, the signal in digital counts is found by

\[
S_{\text{DC}}(x, y) = \frac{A_{\text{pix}} t_{\text{exp}} g_{\text{CCD}}}{1 + 4(FN)^2} \hbar c \int_0^\infty Q \left( \frac{\lambda}{\text{QE}(\lambda)} \right) E_{\text{source}}(\lambda) \tau_{\text{lens}}(\lambda) \tau_{\text{filter}}(\lambda) \tau_{\text{object}}(x, y, \lambda) \lambda d\lambda \tag{3.53}
\]

From Equation 3.53 it is seen that all of the terms, with the exception of \( \tau_{\text{object}}(x, y) \), are constants and therefore only scale the result of the convolution operation.

3.7.2 Optics

Under incoherent lighting conditions the optics are linear with intensity rather than field, therefore Equation 3.9 becomes

\[
g_i(x, y) = f_i(x, y) * |h(x, y)|^2 \tag{3.54}
\]

where \( g_i(x, y) \) is the intensity of the image, \( f_i(x, y) \) is the intensity of the object, and \( h(x, y) \) is the coherent impulse response. Taking the Fourier Transform of Equation 3.54 yields
\[ G_i(\xi, \eta) = F_i(\xi, \eta)F\{ |h(x, y)|^2 \} \quad (3.55) \]

Applying the Wiener-Khinchin theorem to \( F\{ |h(x, y)|^2 \} \), Equation 3.55 becomes

\[ G_i(\xi, \eta) = F_i(\xi, \eta)(H(\xi, \eta) \star H(\xi, \eta)) = F_i(\xi, \eta)H(\xi, \eta) \quad (3.56) \]

where \( H(\xi, \eta) \) is the coherent transfer function or Amplitude Transfer Function (ATF), \( \star \) is the autocorrelation, and \( H(\xi, \eta) \) is the incoherent transfer function, often called the Optical Transfer Function (OTF). For symmetrical pupils, the ATF is expressed as a function of the pupil function in frequency coordinates defined as

\[ H(\xi, \eta) = P(\lambda z_i, \lambda \xi, \lambda \eta) \quad (3.57) \]

where \( \lambda \) is the wavelength and \( z_i \) is the distance from the exit pupil to the image plane. Using a diffraction-limited circular aperture of diameter \( 2w \) and assuming \( z_i \) is the focal length, \( f \), the ATF takes the form

\[ H(\xi, \eta) = CIRC\left( \frac{\rho}{w/\lambda f} \right) \quad (3.58) \]

where \( \rho = \sqrt{\xi^2 + \eta^2} \). Performing an autocorrelation of the ATF leads to the OTF

\[ \mathcal{H}(\xi, \eta) = \begin{cases} \frac{2}{\pi} [\arccos(\frac{\rho}{2\rho_0}) - \frac{\rho}{2\rho_0} \sqrt{1 - \frac{\rho^2}{2\rho_0^2}}], & \text{if } \rho \leq 2\rho_0 \\ 0, & \text{otherwise} \end{cases} \quad (3.59) \]

and taking the absolute value of the OTF yields the \( MTF_{\text{optics}} \)

\[ MTF_{\text{optics}} = |\mathcal{H}(\xi, \eta)| \quad (3.60) \]

Figure 3.20 shows the MTF for a 25\text{mm} lens at \( \lambda = 550^{-6}\text{mm} \) with a circular aperture of \( f/8 \) and \( f/16 \).
3.7.3 Aberrations

The diffraction-limited pupil described in the previous section assumes converging, spherical waves, which is generally not the case when taking into account real-world effects such as atmospheric turbulence, manufacturing defects, and inherent properties of the optics used. Aberrations may be modeled as an optical path delay (OPD), which is essentially the difference between the spherical and aberrated wavefronts (Fig. 3.21).

\[
W_{\text{OPD}}(x, y) = W_{\text{Ab}}(x, y) - W_{\text{Sphere}}(x, y)
\] (3.61)
Chapter 3. *Background and Theory*

The OPD manifests as a phase error imposed on the ATF, therefore Equation 3.57 becomes

\[
H(\xi, \eta) = CIRC\left(\frac{\rho}{w/\lambda f}\right) \exp(ikW_{OPD}(x,y)) \quad (3.62)
\]

\(W_{OPD}(x, y)\) is generally modeled as a Siedel or Zernike polynomial series, the former of which is preferred by most optical designers because the terms are directly relatable to lens aberrations, i.e., defocus, spherical, coma, etc., and the position in the image (Voelz, 2011). Such aberrations are space variant, violating a key assumption in developing the mathematical construct for the system. In order to constrain the model to a convolution-based approach, Fiete (2010) highlights a model for aberrations developed by Shannon (1985) that uses the root-mean-square (RMS) wavefront error, \(W_{RMS}\), to calculate the optical quality factor (OQF) transfer function

\[
MTF_{OQF} = \begin{cases} 
1 - (W_{RMS}/0.18)^2[1 - 4(\rho/\rho_0 - \frac{1}{2})^2], & \text{if } \rho \leq \rho_0 \\
0, & \text{otherwise} 
\end{cases} \quad (3.63)
\]

Figure 3.22 shows the aberration MTF relative to the \(f/8\) lens MTF in Figure 3.20.
3.7.4 Detector

The sensor being modeled for this effort is a full-frame CCD assumed to have a fill factor of one and square pixels. Each pixel of the CCD is a metal-oxide-semiconductor (MOS) capacitor that is able to collect and transfer photoelectrons by manipulating the voltages at the capacitor gates. The signal detected by the CCD is essentially a sum of all the photoelectrons generated in each pixel, therefore the effect of each may be described as a low-pass filter (LPF) that averages the detected signal over pixel area. This may be modeled as a 2-D RECT function given by (Boreman, 2001, Fiete, 2010).

\[ h_{pixel}(x, y) = RECT\left(\frac{x}{dx}, \frac{y}{dy}\right) \]  

(3.64)

where \(dx\) and \(dy\) are the dimensions of the pixel. Taking its Fourier Transform, the transfer function of the pixel is

\[ MTF_{pixel}(x, y) = \frac{\sin(\pi dx \xi)}{\pi dx \xi} \frac{\sin(\pi dy \eta)}{\pi dy \eta} = SINC(dx \xi, dy \eta) \]  

(3.65)

Given a specified pixel pitch of 5.4\(^{-3}\)mm for the STF-8300 camera, Figure 3.23 shows the modeled pixel MTF relative to the lens MTF at \(f = 25\)mm and \(f/8\).
Figure 3.23: Lens MTF at at \( f/8 \) and \( \lambda = 550^{-6} \text{mm} \) relative to the modeled pixel MTF.

The efficiency that photoelectrons are moved from pixel to pixel is called the charge transfer efficiency (CTE) which contributes to crosstalk, a phenomenon where the signal from one pixel induces a false signal in an adjacent or nearby pixel. Because of CTE, electrons generated in one pixel may be left behind during the transfer process, contributing erroneously to the signal of the adjacent pixel. Fiete (2010) models this as (Dereniak and Crowe, 1984)

\[
MTF_{CTE} = \exp \left( -N_{transfers} (1 - CTE) \left[ 1 - \cos \left( \frac{\pi \xi}{\xi_{nyquist}} \right) \right] \right)
\]  

(3.66)

where \( N_{transfers} \) is the number of transfers to the output amplifier, \( CTE \) is the CTE of the serial register, and \( \xi_{nyquist} \) is the nyquist frequency of the detector. The modeled camera has an array size of \( 3326 \times 2504 \) pixels and a serial register CTE of 0.999995 producing the CTE transfer function in Figure 3.24.
Another contributor to crosstalk is diffusion, where electrons generated in one pixel are collected by its neighbor. Such spurious signals are generated in part by the absorption depth of the incoming photon which is wavelength dependent. In the visible region, the absorption depth of light in silicone increases with wavelength, therefore photo electrons generated by red light will be generated deeper into the substrate than blue. This increased distance also increases the likelihood that an electron generated in one pixel will be captured by an adjacent pixel as it travels toward the gate, and is modeled by (Seib, 1974)

$$MTF_{DIFF} = \frac{1 - \exp[\alpha_{abs}(\lambda)L_D]}{1+\alpha_{abs}L(\xi)}$$

(3.67)

where $\alpha_{abs}(\lambda)$ is the photon absorption coefficient, $L_D$ is the depletion depth, $L_{Diff}$ is the diffusion length, and $L(\xi)$ is

$$L(\xi) = \frac{L_{Diff}}{\sqrt{1 + (2\pi L_{Diff})^2}}$$

(3.68)

A consequence of using CCD’s is sampling, which is assumed to occur at the center of each pixel and separated by a distance $p$, defined to be the detector pitch. This process is modeled as a COMB function given by
producing a sampled signal in the spatial domain. Finally, taking the Fourier Transform of Equation 3.69 produces a periodic spectrum.

\[ G_{\text{sampled}}(\xi, \eta) = [F(\xi, \eta)H_{\text{system}}(\xi, \eta)] \ast \text{COMB}(p_x\xi, p_y\eta) \]  

(3.70)

Using such a detector to translate and relay signals introduces additional sensitivities that must be considered especially when operating the imaging system in a real-world environment.

### 3.7.5 Motion

The integration time inherent to all imaging systems gives rise to motion-induced effects that appear as smearing or blurring in the final image (Boreman, 2001, Fiete, 2010). Section 3.4 covered various imaging platforms and geometries, each introducing varying types and amounts of smear to the image. The spatially-variant nature of smear across the focal plane (Fig. 3.25) violates a fundamental modeling assumption; however constraining the simulation to a sufficiently small patch on the image plane allows this assumption to hold (Fig. 3.26).
Smith et al. (1999) breaks down such degradations into cross-scan and along-scan components which contribute to an overall smear vector, with magnitude $\Delta d$. 

Figure 3.25: (A) Zenith smear component (B) Range smear component (C) Azimuth smear component (D) Combined

Figure 3.26: Smear at isoplanatic patch.
Chapter 3. Background and Theory

\[ \Delta d = \sqrt{\text{Smear}_{A/S}^2 + \text{Smear}_{C/S}^2} \text{ [pixels]} \] (3.71)

Assuming that the along-scan direction is parallel to the x-axis and that no cross-scan component is present, the blur kernel is modeled as

\[ h_{\text{blur}}(x, y) = \text{RECT}(\frac{x}{\Delta d}) \] (3.72)

Taking the Fourier Transform produces a SINC function in the \( \xi \)-axis extruded along the \( \eta \)-axis.

\[ H_{\text{blur}}(\xi, \eta) = SINC(\Delta d \xi) \] (3.73)

If a cross-scan component were added, the RECT would be rotated about some angle, \( \theta_{\text{blur}} \), calculated as

\[ \theta_{\text{blur}} = \arctan\left(\frac{\text{Smear}_{C/S}}{\text{Smear}_{A/S}}\right) \] (3.74)

and consequently by the Fourier rotation theorem, the SINC is rotated by the same angle (Fiete, 2010). Similarly, vibrations characterized to be high-frequency (many oscillations over \( t_{\text{int}} \)), and occurring in random directions brings about jitter, which can be modeled as a Gaussian function given by

\[ h_{\text{jitter}}(\rho) = \frac{1}{\sigma_{\text{jitter}} \sqrt{2\pi}} \exp\left(\frac{-\rho^2}{2\sigma_{\text{jitter}}^2}\right) \] (3.75)
where $\sigma_{\text{jitter}}$ is the standard deviation of the random motion, expressed as pixels in the image plane. Taking the Fourier Transform, the transfer function for jitter is then (Fiete, 2010)

$$H_{\text{jitter}}(r) = e^{-2(\pi \sigma_{\text{jitter}} \rho)^2}$$ (3.76)

Figure 3.28 shows the transfer functions for blur (0.25-1.0 pixels) and jitter (0.1-0.4 pixels).

In addition to motion perturbations, noise sources play a role in degrading the interpretability of images.

### 3.7.6 Noise Sources

Attributes of the environment and system components add random variations to the measured signal called noise, which is quantified by the standard deviation of its statistical distribution. Two key assumptions made by the modeling effort are that the noise sources are independent and additive in nature. The latter assumption adds a noise term to Equation 3.9 resulting in Equation 3.77, whereas the former allows noise sources to be summed in quadrature, leading to an expression of the total noise in Equation 3.78.

$$g(x, y) = f(x, y) * h(x, y) + n(x, y)$$

### Figure 3.28

(A) Smear/blur MTFs for 0.25, 0.5, 0.75, and 1.0 pixels. (B) Jitter MTFs for 0.1, 0.2, 0.3, and 0.4 pixels.
Because photon flux varies with time, an inherent uncertainty exists based on the amount of signal reaching the sensor. Such noise is appropriately named photon arrival, or shot noise, and follows a Poisson distribution, thus the standard deviation is calculated by taking the square root of the incoming signal in photons. The signal actually produced by the detector varies by wavelength as a function of it’s QE, therefore shot noise in electrons is expressed as

$$\sigma_{\text{shot}} = \sqrt{S_{\text{photons}}QE}$$

(3.79)

where $S_{\text{photons}}$ is the signal arriving at the sensor in photons, and $QE$ is the quantum efficiency. Shot noise represents the theoretical limit for noise performance, thus the best-case for signal-to-noise is

$$\frac{S}{N} = \frac{S}{\sqrt{S}} = \sqrt{S}$$

(3.80)

At the sensor there is added uncertainty due to random contributions from the electronics, namely dark current, read noise, and quantization noise. Dark current is the result of silicone imperfections, the majority of which occur near the substrate to oxide interface (Janesick, 2001). Such imperfections provide stepping-stones, i.e., intermediate states, that allow thermally generated electrons to escape into the conduction band which are indistinguishable from photoelectrons. Dark current is generally quantified as a rate in $\text{e}^- \text{s}^{-1}$ and is therefore dependent on the integration time. Equation 3.81 shows the calculation for dark current, where $DC$ is the dark current in $\text{e}^- \text{s}^{-1}$ and $t_{\text{int}}$ is the integration time in $\text{s}$.

$$\sigma_{\text{dark}} = DCt_{\text{int}}$$

(3.81)

The most effective way to remove dark current is to cool the CCD, which precludes electrons from reaching an intermediate state by removing thermal energy. Similarly, read noise, $\sigma_{\text{read}}$, is an artifact of imperfections in the read-out electronics which introduce uncertainty when converting the number of collected photoelectrons to a voltage used by the camera.
electronics. It is typically the largest contributor of noise in the CCD electronics and is therefore the limiting factor to a CCD’s performance, especially in low light conditions where shot noise is not the dominant factor. During the readout process the signal also undergoes digitization, producing additional uncertainty by restricting continuous signals to integer values. This type of noise is termed quantization noise, $\sigma_{\text{quantization}}$, and may be estimated by

$$\sigma_{\text{quantization}} = \frac{g_{\text{CCD}}}{\sqrt{12}}$$

(3.82)

Combining each of the aforementioned noise contributions yields

$$\sigma_{\text{total}} = \sqrt{\sigma_{\text{shot}} + \sigma_{\text{dark}} + \sigma_{\text{read}} + \sigma_{\text{quantization}}}$$

(3.83)
Figure 3.29: (A) and (B) Theoretical photon transfer curves for room-temperature and cooled CCDs. (C) Noise contributions for warm CCD operating in shot-noise limited region. (D) Noise contributions for cooled CCD operating in shot-noise limited region. (E) Noise contributions for warm CCD operating in low signal region (F) Noise contributions for cooled CCD operating in low signal region.

To provide a sense of relative noise quantities at different signal levels and different operating temperatures, theoretical photon transfer curves derived from manufacturer specified noise levels are shown (Fig. 3.29). Low and high signal levels were selected for room-temperature ($25^\circ C$) and max-cooling ($-15^\circ C$) cases and the contribution from each noise source is plotted. The low signal (non-shot-noise limited) level is defined as the point where the...
photon transfer curve (blue line) deviates from the shot noise limited case (green line). It is seen, especially at cold temperatures, that the main driving factors for noise performance are shot and read noise. As expected, read noise becomes a significant contributor at low signal levels.

### 3.8 Image Reconstruction

After degradation images generally undergo sharpening via a MTF compensation (MTFC) kernel. Such compensation is captured by the G and H terms in the GIQE, each acting as a penalty for over-sharpening an image. Because of this, system designers generally tune post-processing techniques to produce the highest possible NIIRS rating for their system. An example of such a parameter is seen in the Wiener Filter when conducting image restoration in the frequency domain. The Wiener filter seeks to find an image, \( \hat{f} \), that minimizes error

\[
e^2 = E\{(f - \hat{f})^2\} \tag{3.84}
\]

where \( E \) is the expected value operator and \( f \) is the reference image (Gonzalez et al., 2003). In the frequency domain the Wiener filter is implemented as

\[
W(\xi, \eta) = \frac{\mathcal{H}^*(\xi, \eta)}{|\mathcal{H}(\xi, \eta)|^2 + c \frac{\Phi_n(\xi, \eta)}{\Phi_o(\xi, \eta)}} \tag{3.85}
\]

where \( \mathcal{H}(\xi, \eta) \) is the system transfer function, \( \mathcal{H}^*(\xi, \eta) \) is its complex conjugate, \( \Phi_n(\xi, \eta) \) is the noise power spectrum, \( \Phi_o(\xi, \eta) \) is the object’s power spectrum, and \( c \) is the tuning parameter, usually a constant. With increasing \( c \), the filter is more effective at removing noise with the compromise of increased blur, thus the tradespace for optimizing the tuning parameter is seen as a balancing act between RER, H, and G. This tuning ability was brought about with the advent of soft copy image viewing - a concept outside the scope of current and previous versions of the GIQE as they assumed hard-copy images, essentially fixing RER, H, and G to constant values. Though tuning the MTFC to optimize image quality allows for customized processing, it is problematic in that inconsistencies are introduced when comparing the image quality between competing designs, i.e., system A may use different post processing algorithms than system B. This notion is directly addressed in the development of GIQE 5.0, where the parameters used are un-enhanced values, and the
gain and edge overshoot terms are eliminated (Griffith, 2012). This effort uses an implementation of the Wiener filter found in Cochrane et al. (2013), where $c = 1$ for all images. Using $c = 1$ allows for unbiased restoration of images and ensures that the minimum error solution (in an RMSE sense) is reached. Assuming circular symmetry, Equation 3.85 may also take the form

$$W(\rho) = \frac{1}{\text{OTF}(\rho)} \frac{1}{1 + \frac{c}{\text{SNR}(\rho)^2}}$$

(3.86)

where $\rho$ is the radial spatial frequency and $\text{SNR}(\rho)$ for an $N \times N$ pixel image is defined as

$$\text{SNR}(\rho) = \frac{P/N}{\sqrt{P/N^2 + \sigma_{\text{read}}^2 + \sigma_{\text{dark}}^2}} \frac{|F(\rho)|}{|F(0)|} \text{MTF}_{\text{system}}(\rho)$$

(3.87)

where $P$ is the total number of signal photoelectrons in the image and $F(\rho)$ is the object’s Fourier Transform. Based on the findings in Field (1987), Cochrane et al. (2013) models the object amplitude spectrum as

$$F(\rho) = \frac{1}{\rho^\alpha}$$

(3.88)

where $\alpha$ is typically 1 for many scenes. The coefficient $\alpha$ may be solved by substituting Equation 3.88 into Equation 3.44 and taking the logarithm.

$$\log_{10} [G(\rho)] = -\alpha \log_{10} [\rho] + \log_{10} [\text{MTF}_{\text{system}}(\rho)]$$

(3.89)

$\alpha$ is determined by performing a linear regression on the radially averaged image amplitude spectrum minus the modeled MTF for spatial frequencies where the signal is substantially larger than the noise. An example of this process is seen in Figure 3.30.
Figure 3.30: (A) Radially averaged image amplitude spectrum, with linear fit coefficient, $\alpha = 1.21$ (B) Model image using $f = 50\text{mm}$ at $Q_1$ (C) Filtered Image (D) Zoomed unfiltered simulation image (E) Zoomed filtered simulation image
In summary, the unfiltered output image (either from the camera or the simulation) is used to estimate the object and noise power spectrums. These estimates are then used to form the Wiener filter, which is applied to the original unfiltered image.

### 3.9 Design of Experiments

Due to resource limitations (personnel, time, equipment, etc.), it is important to plan and conduct experiments in such a manner that maximizes efficiency while ensuring the types and amount of data collected may be statistically analyzed so that the conclusions drawn are both valid and objective. The practices behind Design of Experiments (DOE) ensure that both of these goals are met and provides a structured approach to addressing the objectives at hand. Montgomery (2008) outlines seven steps to using a statistical approach in designing and analyzing an experiment:

1. Recognition of and statement of the problem.
2. Choice of factors and levels.
3. Selection of a response variable.
5. Performing the experiment.
6. Data analysis.
7. Conclusions and recommendations.

In the process of accomplishing each of these, it is important to keep in mind the three basic principles of experimental design: replication, randomization, and blocking. Replication is the repetition of experiments and measurements, a process that aides in quantifying and analyzing experimental error that results from measurement uncertainty and variation. A key assumption of using statistical methods to analyze data is that the observations or errors must be independently distributed random variables. Randomization allows the experiment to abide by this requirement through a random selection of experimental order. Lastly blocking allows for greater precision by performing experiments and analysis on separate portions that are homogenous.
NIST/SEMATECH (2013) highlights four engineering problem areas to which DOE may be applied.

1. Comparative
2. Screening/Characterizing
3. Modeling
4. Optimizing

The first portion of this effort focuses on modeling, as the intent is to construct a simulation that is able to predict the output of the testbed with maximal accuracy. The relationship between the testbed and model image ratings is expected to be linear, thus the model takes the form

\[ \hat{y}_n = \hat{\beta}x_n + \hat{\alpha} + \epsilon_n \] (3.90)

where \( x_n \) and \( y_n \) are data coordinates, \( \hat{\beta} \) is the slope estimator, \( \hat{\alpha} \) is the intercept estimator, and \( \epsilon_n \) is the error. Each of the \( n \) points is encompassed by a margin of error in the \( x \) and \( y \) axes that is minimized by taking multiple ratings at each of the data points. The margin of error is calculated using

\[ ME = z \pm \frac{\sigma_n}{\sqrt{M}} \] (3.91)

where \( ME \) is the margin of error, \( z \) is the z-score of the confidence interval, \( \sigma_n \) is the standard deviation of the ratings at the \( n^{th} \) data point, and \( M \) is the number of ratings per data point \( n \). With \( N \) data points the variance of the slope estimator is then

\[ V[\hat{\beta}] = \frac{\hat{\sigma}^2}{\sum_{n=1}^{N} (x_n - \bar{x})^2} \] (3.92)

where \( x_n \) is the \( x \)-coordinate of each point, \( \bar{x} \) is the mean \( x \)-coordinate, and \( \hat{\sigma}^2 \) the squared standard error of the regression, defined as
The variance of the intercept estimator is

\[
\sigma^2 = \frac{\sum_{n=1}^{N} (y_n - \hat{y}_n)^2}{N - 2} \tag{3.93}
\]

The variance of the intercept estimator is

\[
V[\hat{\alpha}] = \left( \frac{1}{N} + \frac{\bar{x}^2}{\sum_{n=1}^{N} (x_n - \bar{x})} \right) \tag{3.94}
\]

Lastly, the coefficient of determination may be used as an indicator of how well the testbed data matches the model, and is calculated with

\[
r^2 = \frac{\sum_{n=1}^{N} (\hat{y}_n - \bar{y})}{\sum_{n=1}^{N} (y_n - \bar{y})} \tag{3.95}
\]

Provided that the model outputs are able to predict testbed performance with high accuracy, the following segment of this effort seeks to explore how various parameters, namely integration time, pointing stability, detector performance, and Q-sampling affect image quality. This type of study falls in under the screening/characterizing problem area, and may be approached through a factorial experimental design, which examines how each factor influences the response variable (\(\Delta NIIRS\)) for all level combinations. For example, if three factors at three levels are to be investigated, the resulting 3\(^3\) factorial design requires 27 separate treatments (factor-level combinations) with \(M\) replicants (Equation 3.91) for each treatment and is represented pictorially below.
Figure 3.31: Illustration of $3^3$ full factorial experiment tradespace.

In tabular form, these treatments may be displayed as follows:

**Table 3.5: Treatment table for $3^3$ factorial experiment**

<table>
<thead>
<tr>
<th>Factor A</th>
<th>Factor B</th>
<th>Factor B</th>
<th>Factor B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (1)</td>
<td>Med (2)</td>
<td>High (3)</td>
<td>Low (1)</td>
</tr>
<tr>
<td>A₁B₁C₁</td>
<td>A₂B₂C₁</td>
<td>A₂B₂C₁</td>
<td>A₃B₃C₁</td>
</tr>
<tr>
<td>A₁B₁C₁</td>
<td>A₂B₂C₁</td>
<td>A₂B₂C₁</td>
<td>A₃B₃C₁</td>
</tr>
<tr>
<td>A₁B₁C₁</td>
<td>A₂B₂C₁</td>
<td>A₂B₂C₁</td>
<td>A₃B₃C₁</td>
</tr>
<tr>
<td>Med (2)</td>
<td>Low (1)</td>
<td>Med (2)</td>
<td>High (3)</td>
</tr>
<tr>
<td>A₁B₁C₁</td>
<td>A₂B₂C₁</td>
<td>A₂B₂C₁</td>
<td>A₃B₃C₁</td>
</tr>
<tr>
<td>A₁B₁C₁</td>
<td>A₂B₂C₁</td>
<td>A₂B₂C₁</td>
<td>A₃B₃C₁</td>
</tr>
<tr>
<td>A₁B₁C₁</td>
<td>A₂B₂C₁</td>
<td>A₂B₂C₁</td>
<td>A₃B₃C₁</td>
</tr>
<tr>
<td>High (3)</td>
<td>Low (1)</td>
<td>Med (2)</td>
<td>High (3)</td>
</tr>
<tr>
<td>A₁B₁C₁</td>
<td>A₂B₂C₁</td>
<td>A₂B₂C₁</td>
<td>A₃B₃C₁</td>
</tr>
<tr>
<td>A₁B₁C₁</td>
<td>A₂B₂C₁</td>
<td>A₂B₂C₁</td>
<td>A₃B₃C₁</td>
</tr>
<tr>
<td>A₁B₁C₁</td>
<td>A₂B₂C₁</td>
<td>A₂B₂C₁</td>
<td>A₃B₃C₁</td>
</tr>
</tbody>
</table>

An Analysis of Variance (ANOVA) is conducted from the data gathered to perform statistical analysis using the following hypotheses:

\[
H_0 : A_1 = A_2 = A_3 = 0 \\
H_1 : \text{at least one } A_i \neq 0
\]  

(3.96)

for each factor (three total)

\[
H_0 : A_iB_j = 0 \\
H_1 : \text{at least one } A_iB_j \neq 0
\]  

(3.97)

for each two-interaction effect (three total - AB, AC, and CB) and
for the three-interaction effect. Assuming normally distributed errors and observations, the
sums of their squares are distributed as chi-square with degrees of freedom $DOF_{Error}$ and
$DOF_{Effect}$, respectively. These may be evaluated using upper-tail, one-tail, F-tests, with the
test statistic for each effect type (factors, two-way interactions, and three-way interactions)
calculated as

$$
F_0 = \frac{SS_{Effect}/(DOF_{Effect})}{SS_{Error}/(DOF_{Error})} = \frac{MS_{Effect}}{MS_{Error}}
$$

where $SS_x$ is the sum of squares, and $MS_x$ is the mean squares. The null hypothesis would
be rejected if

$$
F_0 > F_{\alpha,DOF_{Effect},DOF_{Error}}
$$

where $\alpha$ is the significance level.

Graphically the results may be examined using normal probability plots, box plots, and
histograms. The normal probability plot is used to assess whether or not the data is
normally distributed, the results of which are used to refine regression models. Box plots
for each factor illustrate which factors contribute significantly to the response variable, and
histograms reveal structure in the data by depicting the number of distributions, shape,
and the existence of any outliers. Because ratings for each configuration occur in random
order, a run order plot may be used to determine if there were any temporal effects (fatigue,
learning, etc.) on the ratings. It should be noted that the design covered here was intended
to be a theoretical exercise, and does not represent the design used - specifics of the actual
design are covered in the approach.
3.10 Summary

This chapter was a general overview of the concepts necessary for establishing a viable approach to accomplishing the stated objectives. The imaging chain is assumed to be an LSI system, allowing the effect of each link to be modeled as a series of convolutions with the input object. Though this assumption does not hold true for real world effects such as atmospheric distortions, certain optical aberrations, and motion-induced disturbances, limiting the analysis to isoplanatic patches mitigates the space variant nature of large scenes. Numerous methods, generally categorized into objective (MTF, resolution) and subjective (NIIRS) measures, may be used to assess image quality of the output. As part of the chain, artifacts such as geometric distortion, blur, and smear are introduced, each varying with sensor type and collection geometry. Though relatively new to scientific Earth observing missions, CCD framing arrays show promise in minimizing these effects through increased geometric fidelity and the elimination of certain smear effects inherent to cross-track and along-track scanning sensors. In addition, vast improvements in CCD performance such as CTE, read noise, and QE have greatly improved the signal-to-noise performance of these sensors. These factors present an opportunity to investigate the possibility of considering Nyquist-sampled designs, a concept that has previously been unattractive due to increasingly strict performance requirements when operating at Q=2. Though the GIQE has been successful at predicting the image quality of remote sensing designs, the fact that the model is based on a regression of Q=1 data limits its ability to accurately extrapolate performance of Q=2 systems. Thus, using well established DOE practices, this effort seeks to produce a validated image chain simulation for ideally sampled framing sensors. Using the same techniques, the validated simulation shall be used to explore and characterize tradespace variables central to operating systems at Q=2.
Chapter 4

Methods and Approach

This chapter explains the details behind implementing ideas covered in Chapter 3 to conduct laboratory experiments with the overarching goal of validating image chain simulation outputs for systems with $Q=2$. Section 4.1 describes the notional laboratory configuration, detailing sensor and optics capabilities as well as the accuracy required for translation and rotation stages. As with all real-world systems, the specifications stated by the manufacturer may not match those of the actual system under test, thus this section also covers tests used to characterize the CCD. In addition, given camera specifications, requirements for real-world (film transparency) and digital (scanned film transparency) target quality and illumination spectra are developed. Because such experiments often uncover irregularities across the sensor plane, Section 4.2 describes a method for calibrating the camera output to account for these discrepancies. Section 4.3 describes model implementation and Section 4.4 describes the initial phase of model validation through objective assessments. Finally, Section 4.5 explains the method for subjective IQA and provides estimates on margins for error as well as time required to collect $\Delta$NIIRS data.

4.1 Laboratory Description

The laboratory set up consists of a CCD, lens, light source, and target mounted to an optics table (Fig. 4.1). Camera translation along the $x$, $y$, and $z$ is achieved by moving the assembly along horizontal ($x,z$) and vertical ($y$) rails. The camera’s position is adjusted so that the optical axis is in-line with the center of the the light box and different points of
interest (POI) are photographed by translating the target along x and y so that the POI lies on the optical axis. Given the current configuration of the mounts, rotation about all of the camera’s axes is fixed - however x-z plane leveling is checked using a precision bubble level.

![Diagram of lab setup adjustability](image)

**Figure 4.1:** Illustration of lab setup adjustability.

The light source for the transmissive target is a light box, which ensures uniform incoherent lighting from light emitting diodes (LEDs) through a diffusive window. A detailed illustration of the laboratory workflow is seen in Appendix B.

### 4.1.1 Sensor

Because this effort specifically addresses the modeling performance of systems with a $Q$-factor of 2, lab hardware must be appropriately sized so that the system under test is capable of reaching that level of sampling. The imaging system being tested (detector, lenses, filters) is configured around the pixel size of the sensor due to the selection of lenses and filters available. The camera used for this study is a Santa Barbara Instrument Group (SBIG) STF-8300M, a high-quality commercial off the shelf (COTS) product. At the core of the camera assembly is the Truesense KAF-8300, an 8.9 megapixel image sensor with $5.4\mu m \times 5.4\mu m$ pixels. Because signal drops by a factor of $Q^2$ (Fiete, 2010), a sensor with low noise is preferred. The manufacturer specified value for read noise and dark noise are and $9.3e^-$, and $0.02e^-/pix/s$ (at $-15^\circ C$) respectively. Given a dark current doubling temperature of $5.8^\circ C$, dark current is suppressed through a single-stage thermoelectric cooler capable of reaching temperatures $-40^\circ C$ below ambient with a temperature regulation accuracy of $\pm 0.1^\circ C$. Assuming an ambient temperature of $25^\circ C$, Table 4.1 shows the noise performance of the CCD over a range of integration times.
Table 4.1: Estimated noise characteristics for $T_{\text{CCD}} = -15^\circ$C

<table>
<thead>
<tr>
<th>Integration Time (s)</th>
<th>Read Noise (e-)</th>
<th>Dark Noise (e-)</th>
<th>Quantization Noise (e-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>9.3</td>
<td>0.002</td>
<td>0.107</td>
</tr>
<tr>
<td>0.5</td>
<td>9.3</td>
<td>0.01</td>
<td>0.107</td>
</tr>
<tr>
<td>1</td>
<td>9.3</td>
<td>0.02</td>
<td>0.107</td>
</tr>
<tr>
<td>5</td>
<td>9.3</td>
<td>0.1</td>
<td>0.107</td>
</tr>
</tbody>
</table>

Laboratory tests found that an integration time of 0.5 seconds provided adequate signal without detector saturation, thus read noise was determined to be the dominant noise contributor. When imaging slant edges to measure system MTF, noise was further suppressed by averaging multiple images as blur from high frequency jitter was not found to be an issue. In order to mitigate any unwanted effects caused by the shutter resulting from changes to integration time, signal between configurations was held constant by adjusting the intensity of the light source. On the other hand, in cases where higher noise was required additive white gaussian noise (AWGN) was digitally supplemented to both testbed and simulation images in order to precisely control SNR.

Though the camera came from a reputable manufacturer and detailed CCD specifications are available, the specific unit was not fully characterized, thus performance testing is necessary prior to conducting experiments. Due to the loss in signal experienced at higher values of $Q$, it was critical to characterize sensor noise sources, namely read noise and dark current. Other aspects of the sensor that were tested include CCD gain, which provides a conversion factor between photoelectrons collected and digital counts output by the camera. This specification is intimately tied to CCD linearity, the level of consistency exhibited by the sensor over a range of input values - a critical measure especially in cases where absolute signal level must be known. The final quantity that was measured is quantum efficiency ($\text{QE}$), a wavelength dependent term that serves as an indication of how well the detector converts incident photons to photoelectrons. Determining the CTE was also considered, however access limitations in addition to a risk of damaging the CCD precluded this investigation from taking place.

4.1.1.1 Gain, Read Noise, and Linearity

Gain and read noise is calculated by measuring the mean and standard deviations of pixel values in two pairs of images with significantly different signal levels, i.e., a pair of bias
and flat-field images. Bias images are taken using zero exposure time, while flat fields are taken by uniformly illuminating the sensor using an OL Series 455 integrating sphere. Illumination levels are adjusted using a micrometer-controlled variable aperture, and must be high enough such that the sensor is operating in the shot-noise limited region, but lower than the full well depth (saturation) of the sensor - typically $1/3 - 2/3$ of full well. Flat field images attained a mean digital count (DC) value of $\sim 40,000$ - well into the shot-limited regime but also far below saturation (65,535 DC). Given these frames, a rough gain estimate is calculated using

$$g_{CCD} = \frac{(\bar{F}_1 + \bar{F}_2) - (\bar{B}_1 + \bar{B}_2)}{\sigma^2(\bar{F}_1 - \bar{F}_2) - \sigma^2(\bar{B}_1 - \bar{B}_2)}$$  \hspace{1cm} (4.1)$$

Where $\bar{F}_n$ and $\bar{B}_n$ are the flat and bias frame averages. The denominator contains the standard deviations of the differenced flat and bias frames. Janesick (2001) describes a method of using a multitude of randomly selected subframes within the image to produce a gain histogram enabling accurate measurement of gain (Fig. 4.2). Randomly selecting 15,000 tiles with a size of $100 \times 100$ pixels resulted in a gain measurement of 0.373, a near-exact match to the specification of 0.37.

![Gain histogram for precise measurement of gain for STF-8300M CCD.](image)

Given the number of pixels used, uncertainty in gain may be calculated from
Chapter 4. *Methods and Approach*

\[
\sigma_{g_{CCD}}^2 = \left[ \frac{\partial g_{CCD}}{\partial S_{DC}} \right]^2 \sigma_S^2 + \left[ \frac{\partial g_{CCD}}{\partial S_N} \right]^2 \sigma_N^2
\]  

(4.2)

which simplifies to Equation 4.3 when signal is in the shot noise limited level.

\[
\sigma_{g_{CCD}}^2 = \left[ \frac{2}{N_{pix}} \right]^{1/2} g_{CCD}
\]  

(4.3)

Variance in gain was calculated to be \(5.3 \times 10^{-3} e^-\). Given CCD gain, the read noise in electrons was calculated to be \(9.94e^-\) (Equation 4.4), near the specification of \(9.3e^-\).

\[
\sigma_{\text{read}} = g_{CCD} \sigma (B_1 - B_2) \sqrt{2}
\]  

(4.4)

CCD linearity is the degree to which the output signal is proportional to the photons received by the detector. This specification may be represented as a non-linearity measurement, i.e., two what extent the data deviates from a line of best fit between a range of exposure levels. For the STF-8300M camera, the manufacturer specification is within 1% of linear between \(\sim 5,000\) and \(\sim 50,000\) DC. This measurement is intimately tied to the previous section as it measures gain stability as a function of signal. Linearity is measured by systematically varying the signal captured from uniform illumination. From these data nonlinearity may be measured to within 1% by calculating linearity residuals using Equation 4.5 (Janesick, 2001)

\[
LR = 100 \left( 1 - \frac{S_{DC,m}/t_{exp,m}}{S_{DC}/t_{exp}} \right)
\]  

(4.5)

where \(LR\) is the Linearity Residual, \(S_{DC,m}\) is the mid-scale signal level in digital counts, \(t_{exp,m}\) is the exposure time associated with the mid-scale signal level, and \(S_{DC}\) is signal in digital counts for each exposure time, \(t_{exp}\). Plotting the residuals as a function of average signal serves to illustrate the linearity of the detector over the camera’s dynamic range (Fig. 4.4).
Figure 4.3: Plot of average digital count vs. exposure time for STF-8300M. Linear fit generated for points between 5,000 and 50,000 DC.

Figure 4.4: Linearity residual plot for STF-8300M.

From the data it is seen that the CCD is well within ±1% linear between 5,000 and 50,000 DC and remains within the ±1% threshold until 62,000 DC.
4.1.1.2 Dark Current

Dark current at a specific temperature is calculated by taking dark frames (no light exposure) with varying exposure times and plotting the signal in electrons vs. exposure time.Calculating the slope of the line gives the dark current, seen in the figure below.

![Figure 4.5: STF-8300M dark current data at 15°C and −15°C. The slope of each line is the dark current rate at the respective temperatures; 0.16e⁻/pix/sec at 15°C and 0.014e⁻/pix/sec at −15°C.](image)

Dark current at 15°C and −15°C was measured and compared to manufacturer specifications by extrapolating dark current performance (Table 4.2) using Equation 4.6

\[ \sigma_{dark} = \sigma_{dark_0} 2^{(T-T_0)/Tdoubling} \]  

(4.6)

where \( \sigma_{dark} \) is the dark current, \( \sigma_{dark_0} \) is the baseline dark current at a given baseline temperature, \( T_0 \) is the baseline temperature, \( T_{doubling} \) is the dark current doubling temperature, and \( T \) is the specified temperature.
Table 4.2: STF-8300M Dark Current Specifications

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Dark Current (at 60°C)</td>
<td>200 e−/pix/sec</td>
</tr>
<tr>
<td>Baseline Temperature</td>
<td>60°C</td>
</tr>
<tr>
<td>Dark Current Doubling Temperature</td>
<td>5.8°C</td>
</tr>
</tbody>
</table>

Figure 4.6 indicates that the camera exhibits superior dark current performance compared to the quoted baseline for the given temperature range - a quality which is not uncommon according to the Santa Barbara Instrument Group.

![Graph showing dark current data comparison](image)

**Figure 4.6:** STF-8300M dark current data at 15°C and −15°C compared to an extrapolation of manufacturer data.

### 4.1.1.3 Quantum Efficiency

QE, also referred to as the spectral response, is the ratio of photoelectrons generated by photons incident on the image sensor. This measurement is paramount as it characterizes the behavior of the sensor at each wavelength. Due to its wavelength dependency (Fig. 4.7), QE is generally mapped over a range of the spectrum.
Chapter 4. *Methods and Approach*

A three-port integrating sphere is used to distribute monochromatic light to the CCD camera and NIST calibrated diode with a known detector area (Fig. 4.8).

![Diagram of laboratory setup](image)

**Figure 4.8:** Laboratory setup for collection QE data.

A median value is calculated from a uniformly illuminated portion of the CCD, and using the exposure time and gain, the signal in \( e^-/s/cm^2 \) is given by

\[
S_{e^-/s/cm^2} = \frac{S_{DCGCCD}}{Apix_t_{exp}} \tag{4.7}
\]

The signal in \( pho/s/cm^2 \) is calculated from the NIST diode using
\[
S_{\text{pho/s/cm}^2} = \frac{S_W \lambda}{hc} \quad (4.8)
\]

where \(S_W\) is the signal from NIST diode in \(\text{Watts/cm}^2\). QE is then found with

\[
QE = \frac{S_{e^-/s/cm^2}}{S_{\text{pho/s/cm}^2}} \quad (4.9)
\]

Because a red filter shall be used to reach the desired Q sampling, QE measurements were taken over 550nm-700nm. It is seen in Figure 4.9 that the measured QE closely matches the specification, deviating no more than 2.9% over the desired range and having a root-mean squared error of 1.47%.

![Figure 4.9: Measured (red) vs. specified (blue) quantum efficiency curves.](image)

4.1.2 Optics

The lenses used include Edmund Optics C-mount Fixed Focal length lenses ranging from 16\(mm\) to 50\(mm\) (Fig. 4.10). It should be noted that the cutoff frequency used for Q is derived from that of a circular aperture, whereas the apertures in the available lenses vary in shape due to their aperture-blade configuration. The modeled apertures are seen in Figure 4.11.
Despite this difference, the overall shape of the OTF is highly similar to that of a circle, as shown in Figure 4.12.
Figure 4.13: Comparison of circular aperture to (A) hexagon, (B) octagon, and (C) hexagram. Actual aperture shapes shall be modeled due to the magnitude deviations from a circular aperture.

Figure 4.13 shows vertical and horizontal profiles of the OTFs which illustrate that cutoff frequencies for each shape are either identical or very close to that of a circular aperture. Based on these similarities, it is assumed that the cutoff frequency definition used for Q is adequate as it stands; however differences in magnitude, especially for the hexagram case, indicate that it may be prudent to implement individual shapes for each iris in the optical model. This was done for all data presented here. Another fundamental assumption is that the optics are diffraction limited, a level of quality that is difficult to achieve under certain conditions. Zemax black box files provided by Edmund Optics were used to determine the extent to which each lens approached the diffraction limit at f/8 and f/16 apertures. The 16mm lens exhibited the worst performance and its MTFs are seen in Figure 4.14.
It is seen that the on-axis and $3^\circ$ field sagittal MTF’s approach the diffraction limit, however the off-axis tangential MTFs are degraded indicating astigmatism. This is confirmed by the Seidel diagrams seen in Figure 4.15. The lens also exhibited very low levels of distortion (Fig. 4.16) which were undetectable when viewing a distortion grid. Lastly, total wavefront error is seen in Figure 4.17.
Because astigmatism has a substantial effect on tangential MTF at the edge of the field of view, image tiles are limited to a 1.7° field. In this area, distortion and wavefront RMS are constrained to roughly 0.01% and $2 \times 10^{-2}$, respectively - levels assumed to be negligible.

Due to the manual focus ring, a larger contributor to image quality degradation is defocus aberration, the quantity of which is estimated from an image of the slant edge target. Fiete (2010) highlights a model of defocus aberration in Holst (1995) given as

$$MTF_{\text{Defocus}} = \frac{2J_1 \left[ 8\pi W_{pp} \rho_n (1 - \rho_n) \right]}{8\pi W_{pp} \rho_n (1 - \rho_n)} \quad (4.10)$$

where $W_{pp}$ is the peak-to-peak wavefront error in waves and $\rho_n$ is defined in terms of frequency, $\rho$, and coherent cutoff frequency, $\rho_0$. 

**Figure 4.16:** Field curvature and distortion for 16mm lens at f/8 and f/16 with a 3° field.

**Figure 4.17:** Total wavefront RMS for 16mm lens at f/8 and f/16 with a 3° field.
\[ \rho_n = \frac{\rho}{2\rho_0} \]  

(4.11)

Because configuration changes involve the installation and focusing of each lens and filter, a slant edge image is taken prior to photographing film transparencies to estimate the amount of defocus aberration and ensure that the modeled MTF matches the state of the testbed at the time images are captured. From the slant edge, the testbed MTF is derived and compared to the modeled MTF. The amount of defocus aberration is then adjusted until the error between modeled and measured MTFs is minimized. This process is shown in Figure 4.18 for the 25mm lens.

![Figure 4.18: (A) Plot of measured vs. modeled MTF error as a function of defocus aberration. (B) Measured MTF plotted against modeled MTF with zero and 0.26\(\lambda\) defocus aberration](image)

Results for modeling aberrations using Equation 4.10 are favored over using Equation 3.63, seen in Figure 4.19. The overall RMSE using defocus aberration is lower and the shape of the modeled MTF more closely matches the measured MTF.
Figure 4.19: (A) Plot of measured vs. modeled MTF error as a function of wavefront RMS error. (B) Measured MTF plotted against modeled MTF with zero and 0.09 wavefront RMS error.

4.1.3 Mount Accuracy

In addition to ensuring the camera meets the aforementioned requirements, rotation and translation stages must be properly specified in order to have knowledge of key parameters to within a defined percent error. The following development conducts a high level sensitivity analysis in order to determine the positional and pointing requirements for the lab set up. The positional and pointing accuracy of lab hardware may significantly affect GSD, a key parameter in image quality assessments. Assuming absolute knowledge of pixel pitch and focal length (i.e. zero percent error for these parameters), a sensitivity analysis of GSD may be performed.
Figure 4.20: Sensitivity of GSD to target distance and look angle.

From Figure 4.20 it is seen that an accuracy of $\pm 4^\circ$ and $\pm 2\,mm$ for look angle and target distance respectively result in a GSD error of less than 0.25%. The expected margin of error for distance and look angle are $\pm 1\,mm$ and $\pm 2^\circ$, respectively, leading to a maximum GSD error of 0.08%. Translation and rotation stages meeting these specification are easily obtainable and relatively inexpensive, therefore the purchase of such hardware posed minimal risk to schedule and budget constraints.

4.1.4 Target Specifications

Transmissive targets are used in these experiments to increase dynamic range, help mitigate surface reflections, simplify the lighting configuration, and increase the efficiency of providing enough signal to the camera. Natural scenes may be generated by one of two methods: 1) If an original film positive exists, a high resolution scan of the print will be performed to generate a digital copy for model ingestion. 2) If the scene is originally in digital form, a high resolution transparency or film positive must be created. A number of transmissive targets have been provided by Exelis Inc., therefore this section shall concentrate on the former.

4.1.4.1 Film Transparency Characterization

To ensure consistent results and minimize erroneous contributions to the system transfer function it is critical to select proper viewing geometry and object characteristics. Effects of
aberrations at the edges of the sensor array are mitigated by constraining the maximum field to be $1.7^\circ$ (Section 4.1.2). In addition to aberrations, the MTF of the target transparency must be of sufficient quality, i.e., $\text{MTF} \approx 1$ over the frequency range of the camera for the chosen viewing distance ($2m$). Typical film resolution is on the order of $50-100 \frac{[\text{lp}]}{[\text{mm}]}$, thus the limiting factor for transparency resolution is the optical quality of the intermediate platform. Transparency MTF was determined by taking a high resolution scan (Fig. 4.21) of a B-52 film positive provided by Exelis Inc., finding edges in the image, and determining the image MTF using the slant-edge method.

![Image of B-52 transparency with edges selected](image-url)

**Figure 4.21:** High resolution scan of B-52 transparency with edges selected throughout image to determine MTF.

As the MTF of the high-resolution image undoubtedly contains contributions from the scanner, the transparency MTF was determined by dividing the image MTF by the scanner MTF.
Chapter 4. *Methods and Approach*

It is seen that the transparency MTF is nearly 1 at the Nyquist frequencies of both viewing distances. Significant overshoot is also observed, a phenomenon commonly seen in film prints known as *Mackie Lines* (Spencer, 1973). Though this effect would normally be filtered out for synthetic image generation (SIG) of real world scenes, this artifact is left in place because the simulation is meant to represent the laboratory set up.

### 4.1.4.2 Scanner Characterization for Digital Natural Scenes

Because the chosen method for generating natural scenes requires transposing an image to digital format, a QA-61 test target (Fig 4.23) with slant edge targets and known density values is used to characterize scanner performance. The scanner used for this translation is the EPSON Expression 1640XL, typically used for archiving film transparencies. Scanner MTF is measured using *sfmt3* (Burns, 2009) and the slant edge targets on the QA-61 target. The four MTFs measured from each of the slant edges is averaged, and assuming a circularly symmetric MTF, the scanner’s MTF is formed (Fig. 4.22).
Using the density targets, the scanner’s Opto-Electronic Conversion Function (OECF) is measured to form a curve fit that translates raw digital counts to transmission. First, the QA-61 test target densities are measured using an X-Rite 500 Series Spectrodensitometer to verify the specified density values (Table 4.3).

Table 4.3: Measured vs specified densities for QA-61 test target.

<table>
<thead>
<tr>
<th>Target #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spec</td>
<td>0.600</td>
<td>0.700</td>
<td>0.800</td>
<td>0.900</td>
<td>1.000</td>
<td>1.200</td>
<td>1.400</td>
<td>1.600</td>
<td>1.500</td>
<td>1.300</td>
</tr>
<tr>
<td>Measured</td>
<td>0.609</td>
<td>0.678</td>
<td>0.826</td>
<td>0.894</td>
<td>0.997</td>
<td>1.235</td>
<td>1.435</td>
<td>1.636</td>
<td>1.540</td>
<td>1.309</td>
</tr>
<tr>
<td>Target #</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>Spec</td>
<td>1.100</td>
<td>1.000</td>
<td>0.900</td>
<td>0.800</td>
<td>0.700</td>
<td>0.500</td>
<td>0.300</td>
<td>0.100</td>
<td>0.200</td>
<td>0.400</td>
</tr>
<tr>
<td>Measured</td>
<td>1.101</td>
<td>0.994</td>
<td>0.931</td>
<td>0.800</td>
<td>0.690</td>
<td>0.470</td>
<td>0.268</td>
<td>0.083</td>
<td>0.169</td>
<td>0.394</td>
</tr>
</tbody>
</table>

Measured densities, $D$, are then converted to transmittance, $\tau$, using Equation 4.12 and plotted against digital count to generate a curve fit.

$$\tau = \log_{10} \left( \frac{1}{D} \right)$$  \hspace{1cm} (4.12)

A two-term gaussian of the form

$$f(x) = a_1 \exp\left(-\left(\frac{x - b_1}{c_1}\right)^2\right) + a_2 \exp\left(-\left(\frac{x - b_2}{c_2}\right)^2\right)$$  \hspace{1cm} (4.13)
is used to model the transmittance data which resulted in $R^2 = 0.9997$ using the coefficients in Table 4.4.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>0.8583</td>
</tr>
<tr>
<td>b1</td>
<td>6.97E+04</td>
</tr>
<tr>
<td>c1</td>
<td>1.78E+04</td>
</tr>
<tr>
<td>a2</td>
<td>0.3271</td>
</tr>
<tr>
<td>b2</td>
<td>4.56E+04</td>
</tr>
<tr>
<td>c2</td>
<td>1.98E+04</td>
</tr>
</tbody>
</table>

The resulting curve fit is shown in Figure 4.24

![Figure 4.24: MTF of test target.](image)

It is seen that the model conforms to the data points well and stays within an acceptable range of transmittances when extrapolating to the low and high extremes of digital count, i.e., digital count values of 0 and 65535 do not result in transmittance values of $<0$ and $>1$, respectively. From this curve pixel values in the digital image are calibrated to represent transmittance.

In addition to MTF and OECF, scanner noise is characterized by subtracting two consecutive scans of the QA-61 target and plotting of the mean signal versus standard deviation of the density targets. Unfortunately, the data exhibits an unexpected trend shown in Figure 4.25, where the standard deviation decreases with mean signal.
Chapter 4. *Methods and Approach*

To isolate the discrepancy, additional images of the QA-61 target are acquired at lower resolution (remove target structure) using a different scanner and the STF-8300M camera - the normalized photon transfer curves are shown in Figure 4.26.

Both scanners show the same trend while the STF-8300M exhibits noise qualities more in-line with theoretical predictions, i.e., noise is proportional to signal level. Scanner data indicate the presence of either software or hardware noise compensation, thus a "read noise" equivalent may be difficult to ascertain due to the black-box nature of both units. A curve fit may be formed to the data, however it is unclear whether the trend continues for values below 10,000 digital counts, therefore additional data must be taken if a curve fit is used to calculate scanner-added noise. To form an initial estimate, the mean of Figure 4.25 was used to estimate scanner noise contribution. Because noise is added after blurring and downsampling the digital image, the simulation is used to form a relationship between input and output noise (Fig 4.27). A high resolution digital flat field image with additive white
gaussian noise at levels seen in Figure 4.25 is generated. The image is then blurred and down sampled using a 50\text{mm} lens operating at f/8.

![Figure 4.27](image)

It is seen that noise in the input image is reduced to roughly 5.5% of the original amount, resulting in a scanner noise contribution of 5.3 digital counts, or 1.96 electrons when multiplied by the CCD gain. The amount of noise added by the scanner may be compensated by adjusting the amount of digitally added CCD read noise (Equation 4.14).

\[
\sigma_{\text{read}}^2 - \sigma_{\text{scanner}}^2 = \sigma_{\text{read,comp}}^2
\]  

(4.14)

Using Equation 4.14, \(\sigma_{\text{read,comp}}\) is calculated to be 9.09e\(^{-}\) when \(\sigma_{\text{read}} = 9.3e\(^{-}\) and \(\sigma_{\text{scanner}} = 1.96e\(^{-}\). Because more downsampling and blurring shall occur for other configurations, it is anticipated that this level of compensation is the limiting case for the model. Estimated scanner noise contributions (in electrons) for each of the configurations is shown below.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>50\text{mm}</th>
<th>35\text{mm}</th>
<th>25\text{mm}</th>
<th>16\text{mm}</th>
</tr>
</thead>
<tbody>
<tr>
<td>f/8</td>
<td>1.96</td>
<td>1.36</td>
<td>0.98</td>
<td>0.62</td>
</tr>
<tr>
<td>f/16</td>
<td>1.19</td>
<td>0.85</td>
<td>0.60</td>
<td>0.38</td>
</tr>
</tbody>
</table>

**Table 4.5:** Estimated scanner noise contributions (in electrons) for each configuration.
4.1.5 Light Source

Lighting for the object must be incoherent, uniform, and consistent. Due to the camera’s cooling ability and low dark current, it is assumed that adequate signal levels may be reached with little noise contribution from integration time. To ensure incoherent lighting, the angular subtense of the source, with respect to the object, must be much larger than the angular subtense of object’s angular spectrum and angular diameter of the imaging system’s pupil (Goodman, 2005). Initial tests used to fine-tune laboratory processes utilized fluorescent bulbs, however LED lighting was used during final experimentation as the spectrum exhibited more favorable characteristics. Incoherent lighting for testbed photographs was achieved through the use of a light box, where the light from two strips of LED’s at the edges of frame is dispersed by a reflective grid closely positioned behind a diffuse window (Fig. 4.28).

In addition to meeting these requirements, the source spectrum in conjunction with the optics and QE must achieve $Q \approx 2$ from the available camera configurations. This is computed as a weighted average of $Q$ over the bandwidth utilized. Because of its variability, designers generally exclude the source spectrum from weighted average calculations, however due to the controlled laboratory environment it was used. This ensures that the entire
system is configured to perform at the required sampling rates for each source. Spectra for fluorescent and LED lighting is seen below.

![Normalized Spectrum](image1.png)

**Figure 4.29:** Measured spectrum of the actual fluorescent source and a representative LED source (peak normalized).

Though panchromatic systems typically span the entire visible spectrum, the limited number of available apertures (f/2.8 - f/16) requires the use of a filter to achieve desirable values of Q. Utilizing a minimum aperture of f/16 with a broadband red filter (Fig. 4.30) produces a balance between a broad spectrum and a weighted average $Q \approx 2$.

![Filter Transmission](image2.png)

**Figure 4.30:** Transmission of Edmund Optics red filter.

The weighting function (Fig. 4.31) is calculated by multiplying the normalized source spectrum, optics transmission (including red filter), and detector QE.
The weighted average $Q$ for fluorescent lighting and LED lighting are calculated to be 1.98 and 1.91, respectively, using

$$Q = \frac{\sum_{n=1}^{N} w_n Q_n}{\sum_{n=1}^{N} w_n}$$ (4.15)

where $w_n$ is the weighting function, and $Q_n$ is the system $Q$-value associated with each wavelength.

### 4.2 Image Calibration

Raw images collected from cameras often contain artifacts that are a result of inherent properties of the system (e.g., dark current, hot pixels, optical vignetting) and external contaminants (e.g., dust, cosmic rays). The camera must be calibrated to account for the variations introduced by these artifacts so that the images viewed accurately depict the light levels of the object. A calibrated image is generated by performing the following operations on the raw image

$$I_{cal} = \frac{I_{raw} - D_{master}}{F_{master}}$$ (4.16)

where $I_{cal}$ is the calibrated image (science exposure), $I_{raw}$ is the raw science exposure, $D_{master}$ is the master dark frame, and $F_{master}$ is the master flat frame. A master dark
frame accounts for the noise contribution of dark current and must be taken under the same imaging conditions (e.g., f-number, exposure time, temperature) that science exposures are taken to ensure accurate calibration. Master dark frames for each camera configuration are generated by taking multiple frames at each layout and constructing a median image from the entire set (Equation 4.17). The master dark frame is also key to creating a master flat frame.

\[
D_{\text{master}} = \text{Median} (D_1...D_n) \tag{4.17}
\]

A flat field is the response of the camera system to a uniformly illuminated source, and provides insight into variations across the image plane due to distortions introduced by the overall system. Uniform illumination is provided by the OL Series 455 integrating sphere and multiple flat fields are collected for each camera configuration. Once collected each flat is dark subtracted (i.e., subtract the master dark frame) and mean scaled. \(F_{\text{master}}\) is found by generating a median image from a set of \(n\) images.

\[
F_{\text{master}} = \text{Median} \left( \frac{(F_1 - D_{\text{master}})}{\text{Mean} (F_1 - D_{\text{master}})} \ldots \frac{(F_n - D_{\text{master}})}{\text{Mean} (F_n - D_{\text{master}})} \right) \tag{4.18}
\]

### 4.3 Model Implementation

The image chain model is implemented in the Matlab environment on an Apple iMac 3.4 GHz Core i7 with four cores and 16 GB of memory. Measured source spectra from the lab as well as empirically determined camera specifications are used as model inputs. High-resolution scans of the film transparencies are generated using an EPSON Expression 1640XL scanning at 1200ppi. Scanner performance is measured in Section 4.1.4 using the QA-61 target and found to have an MTF of 0.75 at the camera’s Nyquist frequency when configured for high resolution \((f = 50mm, z = 2m)\). Due to these losses the scanned image is filtered to compensate for scanner MTF. In addition to verifying scanner resolution, pixel values are calibrated to represent transmittance using a curve fit between digital counts and known density values on the QA-61 target. Spectral information for each pixel is generated by multiplying the measured spectrum of the source with the digital transmittance image.
The wavelength-dependent system MTF is formed using the lens aperture (hexagon, octagon, or hexagram depending on the lens), aberrations, and pixel mtf (assumed to be square). Based on the analysis in Section 4.1.2, wavefront error and distortion resulting from the optical elements are shown to be negligible. Because lenses must be manually focused, defocus aberration is identified as the main contributor to blurring. Using Equation 4.10 to model defocus produced more favorable results than Equation 3.63 when comparing measured and modeled MTFs. Due to hardware limitations in the lab, motion-induced blurring or smearing are digitally introduced, modeled as a space-invariant smear over the entire image. The final blurred image is generated by integrating the spectrally varying images across the bandwidth of the weighting function. This image is then down-sampled to match the pixel pitch of the CCD and pixel values are scaled to represent units of electrons.

Fixed pattern noise and image banding may be effectively removed using the calibration techniques outlined in Section 4.2, thus the remaining noise contributions are assumed to be additive white gaussian. Baseline noise sources include shot, read, and quantization noise, though additional noise are added to reach desired SNR levels when evaluating SNR-degraded imagery. Noise is generated as a separate image using a random number generator with poisson distribution for shot noise and zero-mean additive white gaussian (AWGN) for other sources. The noise image is added to the scene after it has been blurred and downsampled.

Once noise is added, pixel values are scaled to represent digital counts, and converted to 16-bit integers representing the output of the CCD. Using the method found in Cochrane et al. (2013) to estimate the object power spectrum, Wiener reconstruction is performed using a tuning factor equal to one for all images. Objective and subjective image quality assessments are then performed on the resulting image in order to validate model performance.
4.4 Model Validation through Objective IQA

The Initial stage of model validation is conducted through a three-way comparison between the known theoretical MTF imposed by the simulation, and the measured MTFs of the testbed and simulation output. Comparison of the model and testbed MTF outputs to a known theoretical function ensures that the simulation is a true representation of the testbed. The MTF measurement technique is similar to those discussed in the previous chapter.
Chapter 3 demonstrated that it is possible to derive the system transfer function from the PSF, LSF, and ESF. In order to ensure adequate signal while eliminating the need to account for the finite size of the source, objective IQA will derive the system transfer function from the ESF. Unfortunately, detector-induced factors such as sampling and noise impose limits on the accuracy of this measurement especially at high frequencies (Boreman, 2001, Reichenbach et al., 1991). Because of the derivative operation used in finding the MTF, high frequency components in areas with low SNR are amplified. In order to combat this effect on the data, multiple frames are calibrated and averaged reducing noise by

$$N_n = \frac{N}{\sqrt{n}} \quad (4.19)$$

Though this method introduces blurring for signals varying in space or time, given the stability of the camera and target it is assumed in this case that the image is sufficiently still to mitigate this effect. In addition to noise, finite sampling on the detector plane may limit the application of the transfer function derivation methods. For a Q-factor of 1, transfer function estimates are limited due to detector undersampling, which results in aliasing. Because the initial phase of model validation includes an objective assessment at Q=1, a method for increasing sampling while preserving the characteristics of the test bed must be implemented. Reichenbach et al. (1991) describes a modification to the ESF or knife edge technique that accounts for detector undersampling by imaging a slant edge. This method artificially increases sampling across the edge by registering scan lines that are inherently offset due to the slant (Fig. 4.34) and is equivalent to the cosine projection of each pixel onto the edge.
This method reiterates the importance of the steps performed in Section 4.2 as it assumes identical responses from each pixel over the edge. Once registered, edge data is averaged and resampled to further reduce noise artifacts before calculating the MTF. The software used for this process is sfrmat3 written and provided by Peter D. Burns. Camera configurations used for objective validation are listed in the table below.

**Table 4.6: Camera configurations that shall undergo objective IQA model validation**

<table>
<thead>
<tr>
<th>Q</th>
<th>Focal Length (mm)</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>35</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>16</td>
</tr>
</tbody>
</table>

Though objective assessments can provide accurate measures of model performance, it
is often difficult to translate MTF in terms of HVS interpretability. Similarly, there is low confidence in GIQE predictions for systems operating at $Q_2$ as the images used to perform the regression were at $Q_1$. Because of this, a second layer of validation seeks to add confidence in image interpretability accuracy through a $\Delta$NIIRS study.

4.5 Model Validation through Subjective IQA

Subjective assessment takes place by instructing observers to match output imagery from the test-bed and model to reference images (i.e. marker set) that span a quality scale (Fig 4.35) common to the Remote Sensing community. $\Delta$NIIRS ratings of test bed images are then plotted against ratings of model outputs under an identical configuration, which serves to illustrate how well the image quality of model outputs conforms to images taken in the lab.

![Figure 4.35: Subjective image quality assessment process.](image)

A set of reference images, usually referred to as *marker sets*, is typically generated by using the relationship between NIIRS and GSD defined by the GIQE, which follows that every halving or doubling of GSD results in a $\Delta$NIIRS of +1 or -1, respectively. Given that
$\Delta \text{NIIRS}$ of 0.2 results in a perceptible change to image quality for average viewers, the scale of reference images shall employ a resolution 0.1. $\Delta \text{NIIRS}$ in terms of GSD varies as (Thurman and Fienup, 2008)

$$\Delta \text{NIIRS}_{\text{GSD}} = -3.32 \log_{10} \left( \frac{GSD}{GSD_0} \right)$$

(4.20)

where $GSD_0$ and $GSD$ are the sample distances of the baseline and comparative images, respectively. Images are rated on a calibrated monitor that displays the test bed and reference images through a user-controlled flicker function, and allowing the observer to zoom in and out of each image. To determine a $\Delta \text{NIIRS}$ rating for each input image, observers are asked to select the reference image that best represents the input image in terms of interpretability (as it relates to the NIIRS scale). The marker set and images rated images were presented in random order.

For each data point, an adequate number of ratings must be collected such that the remainder of samples after an analysis of variance (ANOVA) results in a reasonable margin of error about the mean value. Assuming a normal distribution curve, the margin of error about each point is computed using Equation 3.91. For this effort a confidence level of 95% is chosen resulting in a z-score of 1.96. Additionally, Leachtenauer (2004) indicates that typical standard deviations for $\Delta \text{NIIRS}$ ratings ranges between 0.2-0.3 while the desired ME is generally limited to $< \pm 0.1$. This results in a sample size of

$$n = 1.96 \frac{0.3}{0.1} \approx 35$$

(4.21)

An independent sample is defined as each subject-to-image combination, e.g., one subject rating three different images degraded using a specific camera configuration results in three independent samples for that configuration. Two image tiles from the transparency are selected, thus 18 subjects are required to meet the calculated sample size. The two scenes are shown below.
Selecting a baseline camera configuration from the focal lengths provided, an estimate of ∆NIIRS values may be calculated for each point on the plot. The baseline configuration is defined to be $f = 25\text{mm}$ and $Q_1$, and will serve as the reference point on the ∆NIIRS plot, i.e., coordinates of $(0, 0)$. Assuming that only GSD changes between each focal length, ∆NIIRS for each focal length at $Q_1$ with respect to the baseline may be calculated using Equation 4.20. The ∆NIIRS spacing between each focal length under a $Q_2$ design is the same, however these points are shifted lower with respect to the $Q_1$ points due to a decrease in RER. It should be noted that for the actual imagery collected additional changes to RER, G, and H are expected due to filtering, however this exercise is only meant to provide a rough estimate of the ∆NIIRS range. SNR between $Q_1$ and $Q_2$ is kept constant by adjusting illumination levels, thus the shift in ∆NIIRS between $Q_1$ and $Q_2$ points is isolated to a change in RER resulting from decreased optical diameter between f/8.0 and f/16, i.e., decreased optical cutoff frequency. This change may be calculated using Equation 4.22, and results in a ∆NIIRS range of approximately -1 through 1 (Fig. 4.37).

$$\Delta NIIRS_{RER} = 3.32 \log_{10} \left( \frac{RER}{RER_0} \right)$$ (4.22)
Figure 4.37: Estimated $\Delta$NIIRS range for varying focal length and f-number.

Table 4.7: GIQE-based predictions of $\Delta$NIIRS values for each configuration. $\Delta$NIIRS calculated with respect to $f = 25\text{mm}$ operating at $Q_1$.

<table>
<thead>
<tr>
<th>Q</th>
<th>Focal Length (mm)</th>
<th>Distance (m)</th>
<th>$\Delta$NIIRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>16</td>
<td>2</td>
<td>-0.9</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>2</td>
<td>-0.6</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>2</td>
<td>-0.2</td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>2</td>
<td>0.3</td>
</tr>
<tr>
<td>1</td>
<td>35</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>2</td>
<td>0.7</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>2</td>
<td>1.0</td>
</tr>
</tbody>
</table>

With 18 subjects rating two images for each configuration and each source (testbed and simulation), this plot requires 576 ratings to generate, and will validate model outputs for $Q_1$ and $Q_2$ camera configurations at high SNR. Because SNR is a key component to this study, the interpretability of model outputs is also validated for SNR-degraded imagery.

Fiete and Tantalo (2001) show that the GIQE $\Delta$NIIRS for variations in SNR are non-linear, and statistically different from analyst ratings for changes greater than -0.2. The study alternatively formed a relationship using $NE\Delta\rho$ for a 15% reflective target, producing the following linear regression
$\Delta NIIRS_{SNR} = -(0.17 \pm 0.02) NE\Delta \rho (%)$ \hspace{1cm} (4.23)

$NE\Delta \rho$ ranged between 0.06 and 11.1, and an $NE\Delta \rho$ increase of 0.59 resulted in a -0.1 change in $\Delta NIIRS$. Though the authors warn that the regression applies only to the specific conditions used in the experiment, it is used as an estimate for determining the required noise levels for this effort. Fiete and Tantalo (2001) found that it was difficult for analysts to rate images with $NE\Delta \rho > 5\%$, therefore that value is used as the limiting case for low SNR. A series of evenly spaced points are used at $\Delta NIIRS = 0, 0.2, 0.4, 0.6, 0.8$. Defining a reference point of $NE\Delta \rho$ of 0.3, the noise levels are shown in the table below with the structure seen in Figure 4.38.

<table>
<thead>
<tr>
<th>$\Delta NIIRS$</th>
<th>$NE\Delta \rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.3</td>
</tr>
<tr>
<td>-0.2</td>
<td>1.48</td>
</tr>
<tr>
<td>-0.4</td>
<td>2.66</td>
</tr>
<tr>
<td>-0.6</td>
<td>3.84</td>
</tr>
<tr>
<td>-0.8</td>
<td>5.02</td>
</tr>
</tbody>
</table>

Table 4.8: $\Delta NIIRS$ values associated with $NE\Delta \rho$ using Equation 4.23.

![Image](image.png)

Figure 4.38: $\Delta NIIRS$ predictions using the relationship defined by Equation 4.23.
These data are also generated using a marker-set degraded in GSD, and from a sample size of 36 (18 subjects times two images per configuration) will require 360 ratings. Successful completion of validation experiments enables follow on evaluations to be conducted, utilizing the validated "virtual camera" to explore system trades when designing a $Q_2$ system.

### 4.6 $Q_2$ designs: Exploration and Characterization of System Trades

Designers of remote sensing systems have generally employed $Q_1$ designs due to the shortcomings of $Q_2$ systems, which include increased smear sensitivity and decreased MTF, SNR, and FOV. Modern sensors have the potential to address these challenges due to decreased noise, increased QE, CTE, and pixel density, ultimately resulting in improved signal-to-noise performance while maintaining FOV. The goal of experimentation is to characterize the effects of these trades in terms of image interpretability scale ($\Delta$NIIRS) common to the remote sensing community. Furthermore, the methods detailed here seek to produce quantifiable results rooted in a modern DOE approach. The four factors considered include $Q$, pointing stability (expressed as smear in fractions of a $Q_1$ pixel), detector performance (a combination of QE and read noise), and integration time due to their complex interactions and fundamental ties to the image quality trade-space. $Q$ is determined using laboratory experiments detailed in Section 4.1.5. The levels of pointing stability and detector performance are based on specifications of current spacecraft and CCDs, respectively - Sections 4.6.1 and 4.6.2 detail their origins and how they translate to $Q_1$ pixel smear and SNR. Integration time is expressed in relative terms, with $1 \times$ being the amount of time required for a $Q_1$ system to achieve an SNR of 38 for a 50% transmissive pixel, and $4 \times$ being the amount of time required for a $Q_2$ system to recoup lost signal. Though integration time is not directly identified by the GIQE, an investigation of its effect on image quality is necessary due to its direct effect on SNR and smear - two parameters that are integral when considering $Q_2$ designs. The baseline system is defined to be $Q_1$, high smear (i.e., low pointing stability), medium detector performance, and low integration time. Each factor and their levels are outlined in Table 4.9.
Upon preliminary examination of the abovementioned data, a decision was made to collect additional ratings at $Q = 1.7$ to ascertain the shape of the "knee" when approaching $Q_2$ sampling - these treatments are shown below.

**Table 4.10: Additional treatments collected for design of experiments.**

<table>
<thead>
<tr>
<th>Integration Time (C)</th>
<th>Pointing (D)</th>
<th>Q Sampling (A)</th>
<th>Detector Performance (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (i) 1×</td>
<td>Excellent (i)</td>
<td>$A_i B_i C_i D_i$</td>
<td>$A_i B_i C_i D_i$</td>
</tr>
<tr>
<td></td>
<td>Good (k) 1×</td>
<td>$A_i B_i C_i D_k$</td>
<td>$A_i B_i C_i D_k$</td>
</tr>
<tr>
<td>High (k) 4×</td>
<td>Excellent (i)</td>
<td>$A_i B_i C_i D_i$</td>
<td>$A_i B_i C_i D_i$</td>
</tr>
<tr>
<td></td>
<td>Good (k) 1×</td>
<td>$A_i B_i C_i D_k$</td>
<td>$A_i B_i C_i D_k$</td>
</tr>
</tbody>
</table>

4.6.1 Smear range estimation for design of experiments

The amount of smear experienced by an image is driven by the orbital geometry, pointing stability, and integration time of the platform (among other effects). Orbital smear is calculated using Equations 3.28, 3.29, and 3.30 for a pixel located at the far corner for a $16000 \times 16000$ conceptual large format framing detector. The nadir and azimuth angles for this geometry are defined as $30^\circ$, and $30^\circ$ with respect to the spacecraft. Pointing smear is calculated using an orbit altitude of 770 km (Worldview 2), GSD of 0.5 m (Worldview 2), and pointing stabilities of $1 \times 10^{-3}$ deg/s and $1 \times 10^{-4}$ deg/s (in the range of Worldview,
SPOT, ALOS spacecraft). As smear varies with integration time, a nominal integration time for a conceptual framing array operating within these parameters is deduced by dividing the GSD by the ground track velocity and multiplying by the maximum number of TDI stages in the Worldview 2 spacecraft (64 TDI stages). This results with an equivalent integration time of roughly 5 ms, which is used for the $1\times$ integration time case. These parameters are combined to form an estimate of smear for each of the treatments. For a $Q_1$ system operating at $1\times$ integration time and a pointing stability of $1 \times 10^{-4}$ deg/s (low smear), the resulting smear is 0.4 pixels, thus a $Q_2$ design also functioning under the same parameters experiences 0.8 pixel smear. When pointing is set to $1 \times 10^{-3}$ deg/s (high smear), $Q_1$ and $Q_2$ systems experience 0.45 and 0.9 pixel smear, respectively, at $1\times$ integration time. Smear for all images is applied in the horizontal direction, examples of which are shown below.

![Images showing different smear levels](image1.png)

**Figure 4.39**: (A) $Q_1$, low smear, $1\times$ integration time B) $Q_1$, low smear, $4\times$ integration time

### 4.6.2 Relative SNR estimation for design of experiments

Depending on collection geometry, integration time for space based platforms is generally on the order of a few milliseconds and is sized to balance the amount of smear experienced with the required SNR. The lower bound for integration time, $1\times$, is defined as the time required for a $Q_1$ system operating at low detector performance to reach an SNR of 38 for a 50% transmissive target. This level of SNR is biased toward the lower limit of the values used in Cochrane et al. (2013), an area of the tradespace where confidence in GIQE predictions is lacking. The upper bound for integration time, $4\times$, is defined as the amount
of time required for a $Q_2$ system to recoup the signal losses associated with a smaller pixel projection.

As discussed in Section 3.5, CCD framing arrays have experienced vast improvements in signal to noise performance. Because gains have been made over a number of CCD parameters (QE, read noise, CTE, etc.) one must consider both decreases in noise and increases in sensitivity when exploring the trade space. For this study, QE and read noise was used to determine the detector’s performance level. In order to tie experimental results to advances seen in the real world, the lower and medium levels of signal to noise performance for the experiment were defined relative to the performance differences between existing systems. A survey of front and back illuminated sensors indicate that the low and medium levels of detector performance shown in Table 4.11 are reasonable estimates of average QE (visible) and read noise in real world systems. The values shown for a high performance detector are conceptual values that future detectors may achieve.

Table 4.11: QE and read noise levels for low and high signal to noise performance.

<table>
<thead>
<tr>
<th></th>
<th>QE (%)</th>
<th>$\sigma_{\text{read}}$ ($e^{-}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Performance</td>
<td>40</td>
<td>13</td>
</tr>
<tr>
<td>Medium Performance</td>
<td>80</td>
<td>3.2</td>
</tr>
<tr>
<td>High Performance</td>
<td>95</td>
<td>1</td>
</tr>
</tbody>
</table>

Translating these performance specifications to relative SNR values that can be applied to each image for each configuration requires calculating a signal constant in photons, $S_{\text{photons}}$, that is scaled depending on the system configuration, e.g., if the signal level for a $Q_1$ design is $S_{\text{photons}}$, then the signal experienced by a $Q_2$ design is $S_{\text{photons}} / 4$. Similarly, if the same $Q_2$ system compensates for lost signal by increasing integration time by $4 \times$, then the resulting signal would be $S_{\text{photons}} / 4^4$, or $S_{\text{photons}}$. Assuming dark and quantization noise are negligible, noise for the system may be calculated as

$$\sigma_{\text{total}} = \left( \sigma_{\text{shot}}^2 + \sigma_{\text{read}}^2 \right)^{1/2}$$

(4.24)

where $\sigma_{\text{shot}}$ is defined by Equation 3.79. Signal is calculated as

$$S_{e^-} = S_{\text{photons}} Q E \tau_{50\%}$$

(4.25)
SNR then takes the form

$$SNR = \frac{S_{\text{photons}}QE\tau_{50\%}}{(\sigma_{\text{shot}}^2 + \sigma_{\text{read}}^2)^{1/2}}$$

(4.26)

Using a $Q_1$ system operating at $1 \times t_{\text{int}}$, $QE_{\text{LP}}$ and $\sigma_{\text{read-LP}}$ with an SNR of 38, $S_{\text{pho}}$ is calculated to be 7984.1 photons. Thus, image SNR values varying with $Q$, $t_{\text{int}}$, and signal to noise performance level may be calculated as

$$SNR_{\text{image}} = \frac{7984.1\, t_{\text{int}}\, QE\tau_{50\%}}{\left(\left(\frac{7984.1\, t_{\text{int}}\, QE\tau_{50\%}}{Q^2}\right)^2 + \sigma_{\text{read}}^2\right)^{1/2}}$$

(4.27)

where the respective values for $QE$ and $\sigma_{\text{read}}$ are defined by Table 4.11, and $t_{\text{int}}$ is either 1 or 4. Figure 4.40 shows examples of the changes in SNR exhibited at each detector performance level.

![Figure 4.40: (A) $Q_1$, low smear, 1× integration time, low detector performance B) $Q_1$, low smear, 1× integration time, high detector performance](image)

The actual SNR applied to the image is checked by applying identical levels of signal and noise to a digital 50% transmittance image. The SNR of that image is determined by dividing the mean value of the transmittance image by its standard deviation.

Using each level and calculating relative SNR and image smear values, the treatment tables (Table 4.12 and Table 4.13) are generated, and run order is randomized. Pointing stability is
expressed in fractions of a $Q_1$ pixel at 1× integration time to provide a relative comparison for the reader. This value takes into account orbital and pointing smear, and should be distinguished from the actual amount to smeared applied to the image (column 8 under "Image Attributes").

Table 4.12: Treatment table for mixed 3-level and 2-level four factor experiment. Experimental factors are Q-sampling, smear level (pointing stability), signal-to-noise performance level, and relative integration time.

<table>
<thead>
<tr>
<th>Focal Length</th>
<th>F-Number</th>
<th>Q</th>
<th>Smear Level</th>
<th>QE</th>
<th>$\sigma_{\text{read}}$</th>
<th>$t_{\text{int}}$</th>
<th>Image Attributes</th>
<th>$SNR_{T=50%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.4</td>
<td>0.4</td>
<td>13</td>
<td>1×</td>
<td>0.4</td>
<td>38</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.4</td>
<td>0.8</td>
<td>3</td>
<td>1×</td>
<td>0.4</td>
<td>56</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.4</td>
<td>0.95</td>
<td>1</td>
<td>1×</td>
<td>0.4</td>
<td>62</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.4</td>
<td>0.4</td>
<td>13</td>
<td>4×</td>
<td>1.6</td>
<td>79</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.4</td>
<td>0.8</td>
<td>3</td>
<td>4×</td>
<td>1.6</td>
<td>113</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.45</td>
<td>0.95</td>
<td>1</td>
<td>1×</td>
<td>0.45</td>
<td>38</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.45</td>
<td>0.8</td>
<td>3</td>
<td>1×</td>
<td>0.45</td>
<td>56</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.45</td>
<td>0.95</td>
<td>1</td>
<td>1×</td>
<td>0.45</td>
<td>62</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.45</td>
<td>0.4</td>
<td>13</td>
<td>4×</td>
<td>1.8</td>
<td>79</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>1</td>
<td>0.45</td>
<td>0.8</td>
<td>3</td>
<td>4×</td>
<td>1.8</td>
<td>113</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.4</td>
<td>0.4</td>
<td>13</td>
<td>1×</td>
<td>0.8</td>
<td>17</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.4</td>
<td>0.8</td>
<td>3</td>
<td>1×</td>
<td>0.8</td>
<td>28</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.4</td>
<td>0.95</td>
<td>1</td>
<td>1×</td>
<td>0.8</td>
<td>31</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.4</td>
<td>0.4</td>
<td>13</td>
<td>4×</td>
<td>3.2</td>
<td>38</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.4</td>
<td>0.8</td>
<td>3</td>
<td>4×</td>
<td>3.2</td>
<td>56</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.4</td>
<td>0.95</td>
<td>1</td>
<td>4×</td>
<td>3.2</td>
<td>62</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.45</td>
<td>0.4</td>
<td>13</td>
<td>1×</td>
<td>0.9</td>
<td>17</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.45</td>
<td>0.8</td>
<td>3</td>
<td>1×</td>
<td>0.9</td>
<td>28</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.45</td>
<td>0.95</td>
<td>1</td>
<td>1×</td>
<td>0.9</td>
<td>31</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.45</td>
<td>0.4</td>
<td>13</td>
<td>4×</td>
<td>3.6</td>
<td>38</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.45</td>
<td>0.8</td>
<td>3</td>
<td>4×</td>
<td>3.6</td>
<td>56</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>2</td>
<td>0.45</td>
<td>0.95</td>
<td>1</td>
<td>4×</td>
<td>3.6</td>
<td>62</td>
</tr>
</tbody>
</table>
Table 4.13: Treatment table for additional data collected at $Q = 1.7$.

<table>
<thead>
<tr>
<th>Focal Length</th>
<th>F-Number</th>
<th>Q</th>
<th>Image Smear</th>
<th>QE</th>
<th>$\sigma_{read}$</th>
<th>$t_{int}$</th>
<th>Image Smear</th>
<th>$SNR_{\tau=50%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>42.5</td>
<td>13.6</td>
<td>1.7</td>
<td>0.4</td>
<td>0.4</td>
<td>13</td>
<td>1×</td>
<td>0.68</td>
<td>21</td>
</tr>
<tr>
<td>42.5</td>
<td>13.6</td>
<td>1.7</td>
<td>0.4</td>
<td>0.8</td>
<td>3</td>
<td>1×</td>
<td>0.68</td>
<td>33</td>
</tr>
<tr>
<td>42.5</td>
<td>13.6</td>
<td>1.7</td>
<td>0.4</td>
<td>0.95</td>
<td>1</td>
<td>1×</td>
<td>0.68</td>
<td>36</td>
</tr>
<tr>
<td>42.5</td>
<td>13.6</td>
<td>1.7</td>
<td>0.4</td>
<td>0.4</td>
<td>13</td>
<td>4×</td>
<td>2.72</td>
<td>45</td>
</tr>
<tr>
<td>42.5</td>
<td>13.6</td>
<td>1.7</td>
<td>0.4</td>
<td>0.8</td>
<td>3</td>
<td>4×</td>
<td>2.72</td>
<td>66</td>
</tr>
<tr>
<td>42.5</td>
<td>13.6</td>
<td>1.7</td>
<td>0.4</td>
<td>0.95</td>
<td>1</td>
<td>4×</td>
<td>2.72</td>
<td>72</td>
</tr>
</tbody>
</table>

4.7 Delta NIIRS Rating Environment

As the rating scale used for the experiment is sensitive to how images are displayed, steps must be taken in both hardware and software to minimize the effects of erroneous factors that may affect the perceived image quality. Such factors include, but are not limited to, monitor display quality, light incident upon the screen, screen output calibration, and interpolation. The following sections provide details on the steps taken to minimize these effects.

4.7.1 Delta NIIRS Rating Environment Hardware

The monitors used for rating images are Apple iMac 27” with a resolution of 109 DPI, an 8-bit grayscale depth, and a 1000:1 contrast ratio. These specifications meet the minimum requirements outlined in Leachtenauer (2004) for screen displays. Monitors are calibrated using an X-Rite i1 Pro2 under the lighting conditions experienced while rating images. To minimize incident light onto the screen, overhead lights are turned off during rating sessions - ambient lighting is provided using lamps angled away from the monitoring stations. Walls in the rating lab are painted black and light colored equipment are covered with black cloth. Furthermore, flat-black display hoods draped with black cloth are used to isolate displays from each other (Fig. 4.41). Leachtenauer (2004) also recommends display surround lighting to decrease the disparity between display brightness and its surrounding environment - such lighting is achieved using night lights mounted behind each display. The texture of the aluminum at the rear of the display diffuses the incident light providing the desired effect seen in Figure 4.42.
4.7.2 Delta NIIRS rating Environment Software

Images are rated by subjects using a MATLAB GUI (Fig. 4.43). Each subject is assigned an ID number which is recorded at each rating session. The task is to match the image being rated to an image in a marker set which spans -1.5 to 1.5 DNIIRS relative to the baseline configuration, degraded in increments of 0.1 DNIIRS. Flickering between the image and marker set is manually controlled by pressing the "Switch" button, and marker set quality is increase/decreased by pressing the "Increase Quality" and "Decrease Quality" buttons respectively. Once a suitable marker image is found, the "Record Rating" button is pressed, which simultaneously records the subject’s rating and advances to the next image. If a
subject erroneously presses the "Record Rating" button, the "Undo" button may be used to return to the desired image. Images within a set are displayed in random order and the initial quality of the marker image is selected at random for each image that is rated to prevent any bias toward the lower or upper end of the marker set.

Generating the marker set is at the core of experimentation as it defines the increments and range of the image quality scale used. Increments are established by GSD degradation of a high resolution marker image and the range of $-1.5$ to $1.5$ DNIIRS was chosen to encompass the range seen in Section 4.5. Because marker images are generated by downsampling a high resolution scan some level of interpolation is required. Though constant MTF interpolators exist, the coefficients for such an interpolator are not available. As developing such an interpolator is outside the scope of this work, bicubic interpolation is used when downsampling marker images. The MTF imposed by bicubic interpolation changes the edge response of the image which skews the desired incremental change in DNIIRS - having an average and maximum effect of $0.02$ and $0.05 \Delta$NIIRS, respectively, over the range of the marker set. Attempts to undo the bicubic interpolation MTF via filtering led to unfavorable edge overshoot and the added complexity of variations in noise gain for each of the marker images. To account for H and RER changes in marker images caused by interpolation, calibration curves for each version of the GIQE (Fig. 4.44) are generated to represent the "true" DNIIRS increment at each marker image. Calibration curves are generated by
resampling a slant edge over the DNIIRS range of the marker set and measuring the RER and H at each level. Each GIQE is then used to compute actual DNIIRS over the range of marker images.

![Calibration curves for translating desired DNIIRS increments to actual DNIIRS increments as measured by each GIQE.](image)

**Figure 4.44:** Calibration curves for translating desired DNIIRS increments to actual DNIIRS increments as measured by each GIQE.

When regressed using a linear model, GIQE 3 and GIQE 5 have slopes and intercepts of virtually 1 and 0, with $R^2 > 0.999$. GIQE 4 has an intercept of 0 and a slope of 0.96 due to GIQE 4’s dual slope GSD dependency on RER.

### 4.8 Expected Outcomes

Based on previous studies, there are a number of foreseeable outcomes from this effort. Fiete and Tantalo (1999) found that increasing the along scan sampling rate from $Q_1$ to $Q_2$ resulted in a 0.35 NIIRS increase in image quality, thus a similar result is expected when analyzing the data for the effects of Q sampling. Similarly, Cochrane et al. (2013) found that increasing Q while keeping integration time constant (i.e. signal reduced by a factor of $Q^2$ and smear increased by a factor of Q) produced qualitatively noticeable improvements at high SNR, and little to no improvement at low SNR. These results bolster the findings of Fiete and Tantalo (1999) and indicate that there may be a level of interaction between Q and SNR. Such outcomes can also be predicted based on the GIQE, however confidence
is lacking for scenarios where $Q$ approaches 2 and the SNR is low. Using prior versions of the GIQE as a starting point, it is clear that $Q$ and signal-to-noise performance will produce significant main effects because of the established relationship between GSD, SNR, and NIIRS. It is unclear to what extent integration time and smear will have on $\Delta$NIIRS because of their interaction with both SNR and $Q$, however at some point increased smear associated with higher integration time is expected to overcome the gains associated with increased SNR. Based on mathematical relationships, two way interactions between $Q$ and signal-to-noise performance, $Q$ and smear, and integration time and smear are expected. Similarly three way interactions are expected between each of the factors, most notably $Q$-smear-integration time, and $Q$-integration time-signal to noise performance.
Chapter 5

Results

Given the methods introduced in Chapter 4, this chapter details the results obtained during objective and subjective model validation as well as the design of experiments trade space study. An in depth analysis of GIQEs 3, 4, and 5 is also conducted, evaluating the performance of each by comparing predicted ∆NIIRS to subject ratings. Given analysis outcomes, modifications that increase prediction accuracy at high Q, high smear, and low SNR are examined. Objective validation is accomplished by minimizing RMSE between modeled MTF and measured MTF, then ensuring that the MTF measured from blurred digital edges matches the modeled MTF. This validation step is only meant to provide an initial indication of model fidelity, with more emphasis placed on subsequent subjective experiments. Because trade-space experiments utilize subjective ratings based on a ∆NIIRS scale, subjective validation requires that perceived changes to image quality for simulated images match those of equivalent testbed photographs. Section 5.2 presents the findings from validation experiments that compare ∆NIIRS ratings of laboratory photographs to simulated imagery when varying focal length, f-number, and SNR. Upon conclusion of subjective validation, factorial experiments that would normally be difficult to configure and execute using laboratory equipment may be accomplished with the greater control and ease using the validated simulation. Section 5.3 discusses the findings of factorial experiments, indicating modest gains in image quality when transitioning from $Q_1$ to $Q_2$ with constant aperture diameter and decreasing GSD - encouraging results as this study focuses on images with low SNR. Performing a regression on ∆NIIRS ratings between $Q_1$, $Q_{1.7}$ and $Q_2$ found that a linear model is most suitable for the data, thus a point of diminishing returns when Q increases is not determined. Section 5.4 compares GIQE predictions to subject ratings
collected during validation and factorial experiments, and finds that GIQEs 3 and 4 perform similarly with GIQE5 having higher overall RMSEs. Two modified versions of GIQE3 are evaluated, both of which augment the equation by increasing ΔNIIRS losses at higher Q, and show promise in improving rating accuracy for high Q, high smear, and low SNR cases. Post-hoc analysis of subject ratings show slight levels of acclimation within factorial experiments, and suggest that subjects are, to an extent, able to retain rating skills between experiments. Performing an ANOVA on image type and rating station found that these factors have minimal effect on results as the differences between means are either below the perceivable threshold or have little statistical significance.

5.1 Objective Model Validation

The purpose of objective model validation is to provide an initial measure of the virtual camera’s performance. As discussed in Chapter 4, this is accomplished by conducting a three-way comparison between the modeled MTF, measured simulation MTF, and measured testbed MTF. Based on models provided by the manufacturer wavefront RMS error is found to be negligible at each aperture (f/8 and f/16), and defocus is identified as the largest aberration given the manual nature in which the lenses must be focused. Measured MTFs are computed as an average between vertical and horizontal edges, and the optimal rotation angle and defocus error are determined by minimizing RMS error between modeled and measured MTFs. Given the rotation angle and defocus error, vertical and horizontal digital edges are blurred using the optimized modeled MTF, and the average MTF measured from the blurred digital edges is compared to the modeled MTF to ensure a match. It has previously been shown that the apertures have irregular shapes (hexagon, octagon, and hexagram), thus each aperture is modeled as such, however it could not be determined whether the aperture shape is inscribed or circumscribed by a circle which defines the aperture diameter. Physical restrictions and lack of fidelity in the models provided by the manufacturer precluded such measurements, therefore the geometric mean of the inscribed and circumscribed circle is used.
Figure 5.1: Modeled apertures of (A) 50mm lens at f/8 (octagon aperture) and (B) 35mm lens at f/16 (hexagram aperture). The 50mm f/8 exhibited the lowest RMSE between the measured testbed MTF and the measured simulation MTF (0.004), whereas the 35mm f/16 lens exhibited the worst at 0.02.

The error for the 35mm f/16 lens is mainly attributed to its highly irregular shape (hexagram), resulting in a large difference (with respect to other aperture shapes) between the radii of the inscribed and circumscribed circles (Fig. 5.2). The results of objective model validation are seen in Table 5.1 and MTF plots for all of the configurations used are seen in Appendix C.

Figure 5.2: Modeled apertures with respect to the circle used to define aperture diameter when calculating f-number. (A) Octagon (B) Hexagon (C) Hexagram. Red and blue circles are inscribed and circumscribed circles, respectively.
Table 5.1: Summary of defocus error and RMSE for modeled MTFs.

<table>
<thead>
<tr>
<th>Focal Length (mm)</th>
<th>F-Number</th>
<th>Defocus Error (waves)</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>16</td>
<td>0.14</td>
<td>0.012</td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>0.28</td>
<td>0.007</td>
</tr>
<tr>
<td>25</td>
<td>16</td>
<td>0.14</td>
<td>0.009</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>0.25</td>
<td>0.010</td>
</tr>
<tr>
<td>35</td>
<td>16</td>
<td>0.13</td>
<td>0.020</td>
</tr>
<tr>
<td>35</td>
<td>8</td>
<td>0.34</td>
<td>0.010</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>0.14</td>
<td>0.008</td>
</tr>
<tr>
<td>50</td>
<td>8</td>
<td>0.31</td>
<td>0.005</td>
</tr>
</tbody>
</table>

These levels of error are of the same magnitude seen in Cochrane et al. (2013), indicating that the next phase of validation may be pursued. The following section reports the results collected for subjective validation, a step which is integral to the success of the overall effort.

5.2 Subjective Model Validation

Subjective data is collected by instructing subjects to match simulation output imagery to a reference, or marker set which is degraded in increments of 0.1 ∆NIIRS over a range of -1.5 to 1.5 ∆NIIRS. Based on an estimated standard deviation of 0.3 ∆NIIRS, the discussion in Chapter 4 suggests that it would be prudent to collect roughly 40 samples per configuration in order to meet a confidence interval specification of < 0.1 ∆NIIRS.18 volunteers participated in the experiment resulting in only 36 samples per configuration (18 subjects × 2 images each), however the standard deviations for a vast majority of the data are lower than originally estimated, allowing the confidence interval specification to be met. Subsection 5.2.1 discusses the results obtained when subjects are asked to rate images which varied focal length and f-number while Subsection 5.2.2 examines the effect of changing SNR. In both cases ratings are compared to GIQE predictions, providing an indication of where these IQEs hold and where they break down. Overall it is shown that ratings assigned to simulated imagery adhere very well to testbed photographs, however for SNR degraded imagery, specifically at $Q_1$ sampling, the confidence intervals for images at low SNR extend beyond the desired specification.
5.2.1 Varying Focal Length and F-Number

The purpose of this experiment is to determine whether perceived changes in the quality of virtual camera images match those of photographs captured in the lab when focal length and f-number are varied. This is ultimately accomplished by asking subjects to rate simulation and testbed imagery, and plotting perceived changes in quality ($\Delta$NIIRS) of the simulation against those of the testbed. Standard deviations and margins of error for testbed and model ratings are listed in Table D.1 and are graphically depicted in Figure 5.3.

![Figure 5.3](image)

**Figure 5.3:** Mean simulation and testbed $\Delta$NIIRS ratings for each configuration with 95% confidence intervals. Note the discrepancy between testbed and simulation ratings for the f=35mm, f/16 configuration. Future studies should adjust the defined radius if a hexagram aperture is used.

Standard deviations are below the originally estimated value of 0.3 $\Delta$NIIRS, resulting in confidence intervals below the specification of $\pm 0.1$ $\Delta$NIIRS despite the reduced number of subjects utilized. Observed values are individual simulation ratings from each subject, whereas known values are defined as mean testbed ratings - treated similar to input settings to the experiment. Results of the regression are seen in Figure 5.4.
The regressed coefficients are nearly ideal and 95% confidence intervals encompass both slope = 1 and intercept = 0, indicating that the perceived changes in quality of the simulation adhere well to those experienced by the testbed. A slope of 1.01 implies that imagery from the virtual camera are perceived to have marginally better quality than testbed photographs. This phenomenon may be explained by real world limitations on testbed imagery, specifically the rotational alignment of the photographed imagery to the marker set. Because subjects compare images by flickering between the marker and image, the ΔNIIRS rating GUI automatically performs image registration so that observed features are optimally aligned. Registration is limited to whole pixel translation in order to mitigate additional MTF effects that occur when translating by sub-pixel amounts or rotating. Because marker and simulation images are generated from the same high resolution scan they maintain identical rotational alignment. This is not the case for testbed imagery as equipment limitations preclude perfect rotational alignment between scanned and photographed images. This effect is seen in Figure 5.5.
Figure 5.5: Registration differences between (A) marker and model (B) marker and unregistered testbed

Figure 5.5a exhibits smoother transitions than Figures 5.5b and 5.6a, illustrating that variations between the marker and modeled images are mainly a result of pixel phasing and noise artifacts. Sharp transitions seen between the testbed and marker imagery in Figure 5.5b show signs that edges are misaligned, pointing to either rotational or scaling differences. Registration is improved in Figure 5.6a (shift testbed image 2 pixels down and to the right), however portions of the image are still misaligned. Further improvement is achieved using sub-pixel translation and rotating the testbed image by 0.35 degrees (Fig. 5.6).

Figure 5.6: Registration differences between (A) marker and registered testbed (whole pixel translation only) (B) Subpixel translation and rotation
Residual misalignments between the testbed and marker after whole-pixel translation cause slight shifts between observed features as subjects flicker between images, or "flicker shift." This is in contrast to flickering between marker and model images where perceived changes are limited to phasing and image quality rather than spatial location. Flicker shift seen in testbed imagery is a potential source of subjects’ preference (however slight) toward model images. The following section addresses simulation fidelity for perceived changes to image quality when varying SNR for $Q_1$ and $Q_2$ imagery.

### 5.2.2 Varying Signal to Noise Ratio

Signal to Noise Ratio is a key attribute when determining image quality. Because the overarching study intends to focus on the realm of high-$Q$, low-SNR, a further validation step is required as the images used for the previous experiment exhibit reasonably high-SNR. Moreover, the large standard deviations associated with rating SNR degraded imagery (Fiete and Tantalo, 2001) as well as the complex interactions between SNR, $Q$, and other image attributes (e.g. RER, H, G) add further motivation to complete this undertaking. In this experiment GSD is held constant between $Q_1$ and $Q_2$ configurations, and drawing from the results of Fiete and Tantalo (2001), a regression model (Equation 4.23) is used to estimate the amount of added noise required to achieve a perceivable change in image quality. Because GSD is held constant, the quality of $Q_2$ images is expected to suffer from reduced edge response. Standard deviations and margins of error are shown in Table D.2 and Figure 5.7.
Standard deviations in $\Delta$NIIRS ratings for $Q_1$ and $Q_2$ images are seen to increase proportionally to $\frac{1}{SNR}$ (Table D.2), a trend that is expected as subjects experience increased difficulty when rating noisy images (Fiete and Tantalo, 2001). $Q_2$ images stay within required confidence interval limits ($\pm 0.1 \Delta$NIIRS) throughout the entire range of noise levels, whereas $Q_1$ images are in excess of that range when $SNR_{\tau=8\%}$ drops below 33 ($\frac{1}{SNR} = 0.03$) for model images. Subjects note that it is more difficult to assign ratings to noisy $Q_1$ images than $Q_2$, which may explain the larger variance. To investigate this phenomenon, G and RER are examined to uncover any patterns indicative of increased difficulty when rating $Q_1$ images.

Figure 5.7: Mean simulation and testbed $\Delta$NIIRS ratings for each SNR level with 95% confidence intervals.
RER and G both trend downward with increase noise for $Q_1$ and $Q_2$ sampling, a by-product of the Wiener Filtering process. As both sampling rates exhibit similar trends in RER, and have nearly identical values and trends in G, the effect of filtering is ruled out as a possible cause. It is theorized that the improved edge response in $Q_1$ images results in decreased sensitivity to noise, allowing details to be preserved at higher noise levels. This characteristic may lead to confusion during the rating process due to the complex task of balancing the interpretability of details with noise in the image - an understandable circumstance by virtue of participants' lack of formal image interpretability training. The regression between simulation and testbed ratings is seen below.

**Figure 5.9:** Simulation vs. Testbed ∆NIIRS ratings for varying SNR. The coefficients with 95% confidence bounds are: Slope = 0.99 (0.97, 1.01), Intercept = 0.00 (−0.01, 0.01). $R^2 = 0.97$. Non-weighed values are as follows: Slope = 1.00 (0.93, 1.07), Intercept = -0.01 (−0.06, 0.03). $R^2 = 0.70$.

Because standard deviations increase with decreasing SNR, weighted linear least squares is used for the regression with weights equalling $\frac{1}{\sigma^2}$. Similar to Section 5.2.1, the coefficients as well as their confidence bounds indicate that the perceived quality changes in simulation images is comparable to the testbed for SNR degraded imagery. The lower $R^2$ is a consequence of greater variability, however such a result is expected of SNR degraded imagery.

Overall it is seen that simulated imagery, when varied by focal length, f-number, and SNR, produced equivalent changes to perceived quality when compared to the testbed. Standard
deviations remain within specification for images where focal length and f-number are varied, with mixed results for SNR degraded imagery. The high standard deviations experienced when evaluating $Q_1$, SNR-degraded imagery may be a result of numerous factors to include utilizing untrained image analysts, the complex task of relating noise degraded images to a GSD degraded marker set, and a diminished ability of noise to mask details due to higher RER in $Q_1$ images. Low standard deviations in $Q_2$, SNR-degraded images are an indication that the latter of these factors may have been the largest contributor to variance in the ratings. Though the following study focuses on low SNR images, the amount of noise experienced by images in the factorial experiment remain at levels where observed variance is within specification. With a caveat to minimum SNR, the simulation is seen as an adequate representation of the laboratory camera for configurations spanning those used during validation. Most importantly the discrepancies for low SNR data between simulation and testbed ratings are for $Q_1$ systems - the focus of the ongoing study is for higher Q designs where the agreement is very good.

5.3 Factorial Experiments

Given sufficient confidence in the modeling effort, the following phase of study explores image quality trades at high-Q, low SNR. Image quality is evaluated at various levels of Q, integration time, pointing stability (expressed as smear in fractions of a $Q_1$ pixel), and detector performance. Data collection is divided into two parts, the first containing two levels of Q ($Q_1$ and $Q_2$), integration time (1× and 4×), smear (i.e. point stability), and three levels of detector performance. Given the results of the initial collection, a follow-on set is collected at $Q_{1.7}$, low smear, 1× and 4× integration time, and low, medium, and high detector performance - all treatment combinations are shown on Tables 4.12 and 4.13. The purpose of this experiment is to determine the effect and significance of each factor on image quality, and ultimately seeks to uncover whether the performance gains of decreasing GSD (when transitioning from $Q_1$ to $Q_2$) coupled with increased detector performance offset the losses experienced due to decreased signal, MTF, and increased sensitivity to smear. The standard deviations and corresponding confidence intervals for each collection are seen in Table D.3 and plotted in Figure 5.10.
Chapter 5. Results

Figure 5.10: Mean simulation and testbed ΔNIIRS ratings for each DOE treatment with 95% confidence intervals. Gray bars indicate treatment groupings, where t is relative integration time (every six data points), LS = Low Smear (0.4 Q1 pixels), HS = High Smear (0.45 Q1 pixels), and Low, Medium and High detector performance are within each grouping of LS and HS.

Treatments 1-24 and 25-30 correspond to the first and second collections, respectively. Ratings from 18 participants are collected for the first set of data, allowing the confidence interval specification to be met for the recorded standard deviations. Though only 14 participants are identified for the second collection, the standard deviations allowed for confidence intervals to remain within desired limits. An ANOVA of the initial data set (Table 5.2) indicate that all main factors are significant (p-value < 0.05), plots of which are seen in Figure 5.11.
Table 5.2: ANOVA of initial DOE data. Source is significant at the 5% level when p-value < 0.05.

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum Sq.</th>
<th>d.f.</th>
<th>Mean Sq.</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>16.69</td>
<td>1.00</td>
<td>16.69</td>
<td>715.34</td>
<td>0.0000</td>
</tr>
<tr>
<td>Integration Time</td>
<td>5.03</td>
<td>1.00</td>
<td>5.03</td>
<td>215.38</td>
<td>0.0000</td>
</tr>
<tr>
<td>Pointing Stability</td>
<td>0.29</td>
<td>1.00</td>
<td>0.29</td>
<td>12.54</td>
<td>0.0004</td>
</tr>
<tr>
<td>SNP</td>
<td>4.75</td>
<td>2.00</td>
<td>2.38</td>
<td>101.81</td>
<td>0.0000</td>
</tr>
<tr>
<td>Q*Integration Time</td>
<td>0.83</td>
<td>1.00</td>
<td>0.83</td>
<td>35.36</td>
<td>0.0000</td>
</tr>
<tr>
<td>Q*Pointing Stability</td>
<td>0.12</td>
<td>1.00</td>
<td>0.12</td>
<td>5.26</td>
<td>0.0220</td>
</tr>
<tr>
<td>Q*SNP</td>
<td>0.56</td>
<td>2.00</td>
<td>0.28</td>
<td>12.04</td>
<td>0.0000</td>
</tr>
<tr>
<td>Integration Time*Pointing Stability</td>
<td>0.07</td>
<td>1.00</td>
<td>0.07</td>
<td>2.79</td>
<td>0.0952</td>
</tr>
<tr>
<td>Integration Time*SNP</td>
<td>1.11</td>
<td>2.00</td>
<td>0.55</td>
<td>23.73</td>
<td>0.0000</td>
</tr>
<tr>
<td>Pointing Stability*SNP</td>
<td>0.01</td>
<td>2.00</td>
<td>0.01</td>
<td>0.24</td>
<td>0.7847</td>
</tr>
<tr>
<td>Q<em>Integration Time</em>Pointing Stability</td>
<td>0.02</td>
<td>1.00</td>
<td>0.02</td>
<td>0.92</td>
<td>0.3385</td>
</tr>
<tr>
<td>Q<em>Integration Time</em>SNP</td>
<td>0.16</td>
<td>2.00</td>
<td>0.08</td>
<td>3.36</td>
<td>0.0353</td>
</tr>
<tr>
<td>Q<em>Pointing Stability</em>SNP</td>
<td>0.03</td>
<td>2.00</td>
<td>0.01</td>
<td>0.61</td>
<td>0.5417</td>
</tr>
<tr>
<td>Integration Time<em>Pointing Stability</em>SNP</td>
<td>0.03</td>
<td>2.00</td>
<td>0.02</td>
<td>0.69</td>
<td>0.5024</td>
</tr>
<tr>
<td>Q<em>Integration Time</em>Pointing Stability*SNP</td>
<td>0.01</td>
<td>2.00</td>
<td>0.01</td>
<td>0.29</td>
<td>0.7468</td>
</tr>
<tr>
<td>Error</td>
<td>19.60</td>
<td>840.00</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>49.32</td>
<td>863.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.11: Main effects plots for Q, Integration Time, Pointing Smear, and Detector Performance
The results depicted are encouraging as they show an average increase of 0.25 $\Delta$NIIRS when transitioning from $Q_1$ to $Q_2$ sampling. This is especially significant given that this effect is manifesting at low SNR and while taking into account increased pixel smear for $Q_2$ designs. Increased $\Delta$NIIRS are also associated with improved detectors, an expected outcome given that QE and read-noise performance increases are directly tied to increased SNR. $\Delta$NIIRS losses seen when increasing integration time from $1 \times$ to $4 \times$ brings to light that gains in SNR have been overcome by smear, although it should be noted that the baseline $Q_1$ pixel smear fraction is 0.4, resulting in roughly two pixels of $Q_1$ smear at $4 \times$ integration time. Such negative effects may be inhibited by constraining smear at high integration time to less than two pixels, e.g., utilizing a baseline $Q_1$ smear of 0.25 pixels, however, changing this input parameter requires re-estimating orbital smear by adjusting the location on the focal-plane where smear is measured, modifying orbit geometry, or redefining the pointing stabilities associated with each smear level. Minimal $\Delta$NIIRS losses associated with decreased pointing performance indicates that image smear for the specific scenario used in this study is dominated by orbital smear. A sign of this lies in the amount of $Q_1$ pixel smear calculated for low and high levels of pointing performance - 0.4 pixels at $1 \times 10^{-4}$ deg/s and 0.45 pixels at $1 \times 10^{-3}$ deg/s (both at $1 \times$ integration time). Recalling from Section 4.6.1, the parameters used to estimate smear are as follows:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration Time (ms)</td>
<td>5</td>
</tr>
<tr>
<td>Location (Row $\times$ Column)</td>
<td>$16000 \times 16000$</td>
</tr>
<tr>
<td>Nadir Angle (deg)</td>
<td>30</td>
</tr>
<tr>
<td>Azimuth Angle (deg)</td>
<td>30</td>
</tr>
<tr>
<td>Altitude (km)</td>
<td>770</td>
</tr>
<tr>
<td>GSD (m)</td>
<td>0.5</td>
</tr>
<tr>
<td>Low Pointing Stability (deg/s)</td>
<td>$1 \times 10^{-3}$</td>
</tr>
<tr>
<td>High Pointing Stability (deg/s)</td>
<td>$1 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

At low stability, the amount of smear added by changes to pointing is 0.15 pixels, whereas the along and cross components of orbital smear are 0.05 and 0.4 pixels respectively. For this experiment pointing smear is added to the along track component resulting in a total smear magnitude of $\sqrt{(0.05 + 0.15)^2 + 0.4^2} = 0.45$ pixels. The effect of pointing stability
may be increased in future work by adding its component to the cross direction or to the total magnitude of orbital smear, resulting in roughly 0.55 pixel smear for both cases. Because orbital smear increases proportionally to the number of pixels away from the center of the field of view, features near the center are more dominated by pointing stability than perspective changes when nodding the spacecraft to stabilize the image on the focal plane. Thus, the influence of pointing stability may also be increased by evaluating an isoplanatic patch closer to the center of the array (smear for this study is calculated at the corner of the array).

In addition to main effects, the ANOVA table (Table 5.2) shows evidence of numerous interactions occurring between each factor. The interaction plot below depicts the interaction between factors i and j at tile \((i, j)\) and \((j, i)\), where factors i and j are identified on the diagonal \((i, i)\) and \((j, j)\). For example, interactions between Q (1,1) and Integration Time (2,2) are located on tiles (1,2) and (2,1). The plot shown on (1,2) depicts the ∆NIIRS experienced when transitioning from 1× to 4× integration time for Q1 (blue line) and Q2 (dotted green line). Table 5.2 indicates a significant interaction between Q and Integration time, made obvious by the slope differences seen between Q1 and Q2 lines at plot (1,2). This interaction takes place because of the added smear that occurs as a result of increased integration time. Though the amount of smear in meters on the ground experienced by both Q1 and Q2 systems is identical, Q2 designs exhibit much lower MTFs leading to increased smear sensitivity. This outcome leads to the obvious expectation that a second two-way interaction between Q and Pointing Stability exists, though not as drastic given the minimal influence of pointing on the total smear vector explained earlier.
### Chapter 5. Results

#### Figure 5.12: Interaction plots for Q, Integration Time, Pointing Smear, and Detector Performance

The third two-way interaction is between Q and Detector Performance, plots (1, 4) and (4, 1), which suggest that \( Q_2 \) designs benefit more from detector performance increases than \( Q_1 \). Recalling the discussion of Q and SNR sensitivity in Section 5.2.2, it is postulated that the same effect is taking place - improved RER inherent to \( Q_1 \) designs allow details to be preserved in noisier images, therefore it stands to reason that increased detector performance, primarily tied to improving SNR, lead to marginal gains in perceived image quality at \( Q_1 \). The last significant two-way interaction identified by the ANOVA table is between Integration Time and Detector Performance, plots (2, 4) and (4, 2), which illustrate the negative effects of smear counteracting SNR gains of improved detectors. The ANOVA table also identifies a three way interaction between Q, Integration Time, and Detector Performance, where the common link is that each has an effect on image SNR.

Though statistically significant, \( \Delta \text{NIIRS} \) seen from changing pointing smear are marginal, thus this factor is held constant during the follow-on experiment, which sets out to determine the extent of nonlinearity associated when transitioning from \( Q_1 \) to \( Q_2 \). For this experiment pointing smear is set to its minimal value, and adjustments to integration time and detector performance are made for a sensor designed to \( Q = 1.7 \). These data are shown as the latter

<table>
<thead>
<tr>
<th>Q</th>
<th>Relative Integration Time</th>
<th>Smear in Q1 pixels</th>
<th>Detector Perf</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>Med</td>
</tr>
<tr>
<td>0.4</td>
<td></td>
<td></td>
<td>High</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Low</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Med</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>High</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Low</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Med</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>High</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>
six points in Figure 5.10. The transition from $Q_1$ to $Q_2$ for each detector performance level at low and high integration times is seen below.

![Graph A](image1)

![Graph B](image2)

**Figure 5.13:** (A) Transition from $Q_1$ to $Q_2$ at 1× Integration Time (B) Transition from $Q_1$ to $Q_2$ at 4× Integration Time at low (red) medium (green) and high (blue) detector performance.

The ANOVA table (Table 5.4) for these data indicate that all factors and two way interactions are significant - the three-way interaction is not determined to be statistically significant (p-value = 0.46).

**Table 5.4:** ANOVA of DOE data combined from first and second collection. Source is significant at the 5% level when p-value < 0.05.

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum Sq.</th>
<th>d.f.</th>
<th>Mean Sq.</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q</td>
<td>10.68</td>
<td>2.00</td>
<td>5.34</td>
<td>232.31</td>
<td>2.93E-75</td>
</tr>
<tr>
<td>Integration Time</td>
<td>2.62</td>
<td>1.00</td>
<td>2.62</td>
<td>113.79</td>
<td>2.00E-24</td>
</tr>
<tr>
<td>SNP</td>
<td>3.28</td>
<td>2.00</td>
<td>1.64</td>
<td>71.41</td>
<td>1.61E-28</td>
</tr>
<tr>
<td>Q*Integration Time</td>
<td>0.29</td>
<td>2.00</td>
<td>0.15</td>
<td>6.41</td>
<td>1.76E-03</td>
</tr>
<tr>
<td>Q*SNP</td>
<td>0.21</td>
<td>4.00</td>
<td>0.05</td>
<td>2.34</td>
<td>5.42E-02</td>
</tr>
<tr>
<td>Integration Time*SNP</td>
<td>0.85</td>
<td>2.00</td>
<td>0.42</td>
<td>18.44</td>
<td>1.70E-08</td>
</tr>
<tr>
<td>Q<em>Integration Time</em>SNP</td>
<td>0.08</td>
<td>4.00</td>
<td>0.02</td>
<td>0.86</td>
<td>4.86E-01</td>
</tr>
<tr>
<td>Error</td>
<td>13.65</td>
<td>594.00</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>31.69</td>
<td>611.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To test for curvature, the data is fit to a model with quadratic terms for detector performance, and Q-sampling, as well as all two-way interactions.
Table 5.5: Quadratic fit of combined design of experiments data to test for curvature. Term is significant at the 5% level when p-value < 0.05.

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
<th>SE</th>
<th>t-Stat</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-5.79E-01</td>
<td>3.72E-01</td>
<td>-1.55E+00</td>
<td>1.20E-01</td>
</tr>
<tr>
<td>Q</td>
<td>4.98E-01</td>
<td>2.04E-01</td>
<td>2.44E+00</td>
<td>1.50E-02</td>
</tr>
<tr>
<td>Integration Time</td>
<td>1.10E-01</td>
<td>2.32E-02</td>
<td>4.75E+00</td>
<td>2.58E-06</td>
</tr>
<tr>
<td>Detector Performance</td>
<td>-1.45E-02</td>
<td>3.10E-02</td>
<td>-4.67E-01</td>
<td>6.41E-01</td>
</tr>
<tr>
<td>x1 × x2</td>
<td>-3.21E-02</td>
<td>9.48E-03</td>
<td>-3.39E+00</td>
<td>7.49E-04</td>
</tr>
<tr>
<td>x1 × x3</td>
<td>7.32E-03</td>
<td>2.43E-03</td>
<td>3.01E+00</td>
<td>2.76E-03</td>
</tr>
<tr>
<td>x2 × x3</td>
<td>-4.17E-03</td>
<td>6.98E-04</td>
<td>-5.97E+00</td>
<td>4.06E-09</td>
</tr>
<tr>
<td>x1^2</td>
<td>-9.96E-02</td>
<td>6.54E-02</td>
<td>-1.52E+00</td>
<td>1.28E-01</td>
</tr>
<tr>
<td>x3^2</td>
<td>5.63E-04</td>
<td>6.66E-04</td>
<td>8.46E-01</td>
<td>3.98E-01</td>
</tr>
</tbody>
</table>

The coefficients for the quadratic terms in Table 5.5 are negligible and their associated p-values are not significant indicating that the presence of quadratic curvature is unlikely. Eliminating the quadratic terms produces the following linear model with two-way interactions.

Table 5.6: Linear fit of combined design of experiments data after eliminating quadratic terms for Q and Detector Performance. Term is significant at the 5% level when p-value < 0.05.

<table>
<thead>
<tr>
<th>Term</th>
<th>Estimate</th>
<th>SE</th>
<th>t-Stat</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-6.59E-01</td>
<td>1.17E-01</td>
<td>-5.65E+00</td>
<td>2.44E-08</td>
</tr>
<tr>
<td>Q</td>
<td>2.04E-01</td>
<td>6.68E-02</td>
<td>3.06E+00</td>
<td>2.34E-03</td>
</tr>
<tr>
<td>Integration Time</td>
<td>1.10E-01</td>
<td>2.32E-02</td>
<td>4.74E+00</td>
<td>2.63E-06</td>
</tr>
<tr>
<td>Detector Performance</td>
<td>1.15E-02</td>
<td>4.31E-03</td>
<td>2.67E+00</td>
<td>7.78E-03</td>
</tr>
<tr>
<td>x1 × x2</td>
<td>-3.21E-02</td>
<td>9.49E-03</td>
<td>-3.39E+00</td>
<td>7.57E-04</td>
</tr>
<tr>
<td>x1 × x3</td>
<td>7.32E-03</td>
<td>2.44E-03</td>
<td>3.00E+00</td>
<td>2.78E-03</td>
</tr>
<tr>
<td>x2 × x3</td>
<td>-4.17E-03</td>
<td>6.98E-04</td>
<td>-5.96E+00</td>
<td>4.17E-09</td>
</tr>
</tbody>
</table>

It is seen that all terms are significant, reinforcing the finding that no quadratic terms, i.e. curvature, exist for the range of Q and detector performance used in the experiment. This is comparable to the findings in Fiete and Tantalo (1999), which depict a linear relationship as Q is increased from 1, 1.25, and 1.5 in the along-scan direction. The current study differs in that a linear relationship is shown for Q-values from 1, 1.7, and 1.92, however a fundamental difference (among other distinctions between the models) is that increased sampling occurs
in two-dimensions. For illustration purposes, $\Delta$NIIRS for low, medium, and high detector performance as a function of $Q$ at low smear are superimposed on data graphically obtained from Fiete and Tantalo (1999) - $\Delta$NIIRS for each detector performance level are centered to be zero at $Q=1$.

![Graph of $\Delta$NIIRS for different detector performance levels](image)

**Figure 5.14:** $\Delta$NIIRS for low, medium, and high detector performance as a function of $Q$ at low smear superimposed on data graphically obtained from Fiete and Tantalo (1999).

Given the confidence bounds, it is seen that ratings for medium and high performance detectors may fall within estimates found in Fiete and Tantalo (1999), however it should be noted that any conclusions drawn from this plot may not be reliable given the methods used to obtain the referenced data. In order to fully ascertain the location and shape of the curve as it relates to the data collected for the current study, it may be prudent to collect ratings for $Q$-sampling at tighter intervals between 1.7 and 2 and for values greater than 2.

Results from the factorial experiment indicate modest $\Delta$NIIRS increases when transitioning from $Q_1$ to $Q_2$ at low SNR, building on the qualitative findings in Cochrane et al. (2013) which concluded that transitioning from $Q_1$ to $Q_2$ designs represents no risk to image quality. Statistical evidence of $\Delta$NIIRS improvements found in this study reinforce the feasibility of $Q_2$ designs for real-world sensors, however this does not eliminate the need to conduct proper trade studies on $Q$-sampling when evaluating the requirements of a specific mission. Finding the point of diminishing returns as $Q$ transitions from 1 to 2 is valuable when determining the optimum $Q$ for a given CONOP - an attempt was made to find such a value for this work by assessing the significance of a quadratic model between three levels of
Q-sampling, however results were inconclusive in that only a linear relationship was found to be significant. In order to fully understand the location and extent of the curve, image quality at additional Q-designs between $Q_{1.7}$ and $Q_2$, as well as a point $> Q_2$ are required. Utilizing ratings of all simulated imagery, the adequacy of current models is evaluated in the following section.

5.4 GIQE Analysis

An overarching goal of this effort is to build trust and confidence in our modeling efforts, thus an examination of current models used to predict $\Delta$NIIRS is necessary. The purpose of this section is to investigate capabilities and shortfalls of the GIQE, a model widely used in the Intelligence, Surveillance, and Reconnaissance (ISR) community to predict the interpretability of an image on the NIIRS scale using sensor specifications. Three versions of the GIQE evaluated include GIQE3, 4, and 5 - the former two utilize filtered values (i.e. with MTFC sharpening) for each of the input terms while GIQE5 requires unfiltered inputs. SNR, RER, and H inputs are measured from an 8% transmissive digital flat field, and vertical and horizontal edges that undergo identical processing to simulated imagery. This analysis makes use of data collected during the trade space study, as well as simulation ratings collected as part of model validation given the consistency between testbed and simulation ratings. A concept similar to the validation effort is employed, plotting GIQE $\Delta$NIIRS predictions on the x-axis and simulation ratings on the y-axis. Results for the data collected in Section 5.2.1 are shown below.
The regression slopes indicate that the GIQE5s predict higher near the low end of the \( \Delta \text{NIIRS} \) range and lower at the high end - a finding also made evident by the end points of Figure 5.16. This is further reinforced by the 95% confidence intervals shown on Table 5.7.

Table 5.7: 95% Confidence Intervals when regressing GIQE3,4, and 5 against simulation and testbed ratings. Uncalibrated ratings utilize a constant increment of 0.1 \( \Delta \text{NIIRS} \) for each marker image. Calibrated ratings use the calibration curves detailed in Section 4.7.2, which take into account changes to the \( \Delta \text{NIIRS} \) increment due to interpolation.

<table>
<thead>
<tr>
<th></th>
<th>Calibrated</th>
<th>Uncalibrated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slope</td>
<td>Intercept</td>
</tr>
<tr>
<td>GIQE3</td>
<td>(1.05, 1.12)</td>
<td>(-0.02, 0.02)</td>
</tr>
<tr>
<td>GIQE4</td>
<td>(1.05, 1.12)</td>
<td>(-0.04, -0.01)</td>
</tr>
<tr>
<td>GIQE5</td>
<td>(1.05, 1.12)</td>
<td>(-0.10, -0.05)</td>
</tr>
<tr>
<td>GIQE3</td>
<td>(1.03, 1.09)</td>
<td>(-0.04, -0.01)</td>
</tr>
<tr>
<td>GIQE4</td>
<td>(1.03, 1.09)</td>
<td>(-0.07, -0.03)</td>
</tr>
<tr>
<td>GIQE5</td>
<td>(1.02, 1.09)</td>
<td>(-0.12, -0.08)</td>
</tr>
</tbody>
</table>

Because the regression between simulated and testbed imagery also exhibits a similar characteristic (though to a much lesser extent), 95% CIs for coefficients when regressing testbed ratings against GIQE predictions are also shown. For both cases, it is seen that 95% confidence intervals are greater than and exclude 1, leading to the conclusion that the discrepancy for which this symptom is tied to does not reside in the simulation. The effect of calibrating
ratings according to the actual $\Delta$NIIRS intervals calculated by each IQE (Section 4.7.2) is also investigated by performing the regression using uncalibrated ratings. It is seen that for GIQE3 and GIQE5 the calibration process has minimal effect, however CI’s for the GIQE4 slope (testbed and simulation) are significantly effected - an indicator of GIQE4’s dual slope with respect to GSD as a function of RER. In this specific case RER is calculated to be less than 0.9, thus the coefficient associated with GSD is -3.16 rather than -3.32 (the value used when generating the marker set). Because the discrepancy resides mainly at the endpoints and is unaffected by image source or calibration (when GSD slope matches that which the marker set is degraded), it is theorized that the lack of an end-point indication feature in the $\Delta$NIIRS rating software is the cause. Given the inexperience of participants, when rating an image at the extremes of the scale (either the lowest quality or the highest quality) there may have been a tendency to click ”Decrease Quality” or ”Increase Quality” until no perceived changes in the maker are seen, essentially overshooting the desired interpretability level. This explanation is consistent with the data in that the average rating for the lowest quality image is lower than the predicted rating, and the average rating for the highest quality image is higher than the predicted rating. Endpoint configurations are not used in the Q trade space study which is the main thrust of this work, thus the errors at these points are unlikely to have significant effects. Assuming that end points of the data are unreliable, the regression is re-accomplished with the following results.

\[
\begin{align*}
\text{GIQE DNIIRS Predictions} & \\
\begin{array}{cccc}
\text{DNIIRS Ratings} & -1.5 & -1 & -0.5 & 0 & 0.5 & 1 & 1.5 \\
\text{Slope} = 1 & & & & & & & \\
\end{array}
\end{align*}
\]

\[
\begin{align*}
\text{GIQE5 Predictions (Slope = 0.94, Int = -0.06, R^2 = 0.87)} & \\
\text{GIQE4 Predictions (Slope = 0.97, Int = -0.02, R^2 = 0.88)} & \\
\text{GIQE3 Predictions (Slope = 1.00, Int = -0.00, R^2 = 0.89)} &
\end{align*}
\]

Figure 5.16: Plot of GIQE predictions vs subjective ratings for simulated images with end points removed.
Table 5.8: 95% Confidence Intervals when regressing GIQEs 3, 4, and 5 against simulation and testbed ratings with end-points removed.

<table>
<thead>
<tr>
<th></th>
<th>Slope</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIQE3</td>
<td>(0.96, 1.05)</td>
<td>(-0.02, 0.02)</td>
</tr>
<tr>
<td>GIQE4</td>
<td>(0.93, 1.02)</td>
<td>(-0.04, 0.00)</td>
</tr>
<tr>
<td>GIQE5</td>
<td>(0.90, 0.99)</td>
<td>(-0.08, -0.03)</td>
</tr>
</tbody>
</table>

Slopes for GIQE 3 and 4 are nearly one with slope and intercept CIs that encompass unity and zero respectively, however both the slope and CI for GIQE5 indicate that predicted \( \Delta NIIRS \) may be an overestimation of subject ratings. The cause may stem from underlying assumptions made when developing this latest equation. Based on limited documentation provided, the form of GIQE5 is rooted in the relationship between unfiltered and filtered RER, varying as a function of SNR when Wiener filtering an image. From GIQE3, NIIRS predictions are made utilizing Wiener filtered RER values - from these NIIRS predictions, GIQE5 is regressed using the corresponding unfiltered RER values and takes the form

\[
NIIRS = c_0 + c_1 \log_{10}(GSD_{GM}) + c_2 \left[ 1 - e^{\left(\frac{c_3}{SNR}\right)} \right] \log_{10}(RER_{GM}) + c_4 \log_{10}(RER_{GM})^4 + \frac{c_5}{SNR} \tag{5.1}
\]

where \( c_0 = 9.57 \), \( c_1 = 3.32 \), \( c_2 = -3.32 \), \( c_3 = -1.9 \), \( c_4 = -2.0 \), and \( c_5 = -1.8 \). The terms of interest are \( 1 - e^{\left(\frac{c_3}{SNR}\right)} \), an interaction term, and \( c_4 \log_{10}(RER_{GM})^4 \), presumably a corrective term that compensates for decreased RER weighting at high SNR. The course of thought used to develop GIQE5 is problematic in that various filtering methods may adjust unfiltered RER differently from Wiener filtering. Furthermore, the form of Wiener filter used is a modified version taken from Thurman and Fienup (2010), the coefficients of which are affected by image content. It is therefore reasonable to conclude that without exact knowledge of the filter used for regression, predicted \( \Delta NIIRS \) values of GIQE5 are likely to stray from subject ratings, especially in cases where \( \Delta NIIRS \) changes due to RER and SNR (Fig. 5.17).
At high SNR the influence of RER is highly attenuated due to the interaction term resulting in minimal change between $Q_1$ and $Q_2$ images - in these cases SNR is approximately 40, thus the weighting of RER is reduced to $3.32[1 - e^{(-1.9/40)}] = 0.154$, leading to the positive bias seen for $Q_2$ configurations. Errors between $Q_1$ estimates are low because the reference point used ($\Delta$NIIRS = 0) is $Q_1$, therefore $\Delta$NIIRS between $Q_1$ configurations is dominated by GSD changes. Because of this discrepancy it is likely that the filter employed in this work differs from that used during GIQE5 development, consequently GIQE5, as it stands, may be a poor model for this specific study. Overall results in terms of RMSE are summarized in Table 5.9, which shows that GIQEs 3 and 4 perform similarly in experiments where focal length and f-number are systematically varied - GIQE5 has slightly larger error, stemming from biased $Q_2$ predictions.

<table>
<thead>
<tr>
<th>GIQE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIQE 3</td>
<td>0.176</td>
</tr>
<tr>
<td>GIQE 4</td>
<td>0.174</td>
</tr>
<tr>
<td>GIQE 5</td>
<td>0.207</td>
</tr>
</tbody>
</table>

GIQE predictions are also compared to subject ratings when evaluating SNR degraded imagery by centering $Q_1$ and $Q_2$ evaluations at the highest SNR and plotting simulation
ratings against GIQE outputs (Fig. 5.18a and 5.18b). In both $Q_1$ and $Q_2$ imagery it is seen that GIQE5 overestimates perceived losses in image quality, while GIQE4 underestimates - GIQE3 is shown to have predictions that adhere well to subject ratings. Results for GIQE3 and 4 are similar to those shown in Thurman and Fienup (2008, 2010), however lower $R^2$ values, especially at $Q_1$, are seemingly a result of using aerial photographs rather than Snellen Eye charts. The latest documentation concerning GIQE5 indicates a regression SNR range between 10 and 200, thus its poor performance may be a consequence of the low SNRs used in this experiment ranging between 33 and 1.

![Figure 5.18: Plot of GIQE predictions vs subjective ratings for SNR degraded imagery at (A) $Q_1$ and (B) $Q_2$.](image)

Eliminating the H term produces slightly improved performance for GIQE4 while decreasing the accuracy of GIQE3 (Fig. 5.19a and 5.19b) - removing both G and H terms further enhances GIQE4 predictions and results in overestimated losses for GIQE3 (Fig. 5.19c and 5.19d).
Because G decreases proportionally to SNR due to Wiener filtering, it is seen that GIQE4’s underestimation of ΔNIIRS loss may be rooted in the coefficient that scales G/SNR - in GIQE 3 and 4 they are 1 and 0.334; accordingly losses in image quality are relatively less severe in GIQE4 when SNR is reduced. When the coefficient is increased to 1.5, GIQE4’s estimations have the same regressed slope as GIQE3 with marginal increases in $R^2$. (Fig. 5.20a and 5.20b)
Chapter 5. Results

133

GIQE DNIIRS Predictions

-1.5 -1 -0.5 0 0.5
DNIIRS Ratings

-1.5
-1
-0.5
0
0.5
Slope = 1
GIQE5 Predictions (Slope = 0.7, Int = 0.046, R^2 = 0.775)
GIQE4 Predictions (Slope = 1.0, Int = 0.085, R^2 = 0.761)
GIQE3 Predictions (Slope = 1.0, Int = 0.086, R^2 = 0.754)

Figure 5.20: Plot of GIQE predictions vs subjective ratings for SNR degraded imagery when G/SNR coefficient = 1.5 for GIQE4 at (A) Q_1 and (B) Q_2.

Prediction accuracy for each IQE is summarized in the table below which lists RMSE for SNR-degraded Q_1 and Q_2 images. Results are consistent with the previous experiment, showing that GIQEs 3 and 4 perform similarly while GIQE5 has larger errors.

Table 5.10: RMS Error for GIQEs 3, 4, and 5 for experiments where SNR is degraded at Q_1 and Q_2 sampling.

<table>
<thead>
<tr>
<th></th>
<th>Q1</th>
<th>Q2</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIQE 3</td>
<td>0.406</td>
<td>0.236</td>
</tr>
<tr>
<td>GIQE 4</td>
<td>0.385</td>
<td>0.265</td>
</tr>
<tr>
<td>GIQE 5</td>
<td>0.562</td>
<td>0.521</td>
</tr>
</tbody>
</table>

GIQE predictions are also compared to results of the factorial experiments, depicting how each IQE may perform when considering operational factors such as changes in smear and SNR. Figure 5.21 plots mean subject ratings and GIQE predictions for each of the treatment combinations.
Qualitatively, it is seen that GIQE 3 and 4 perform similarly having small errors for cases with low smear (all cases for $Q_1$, and $Q_{1.7}$ and $Q_2$ at $1 \times$ integration time), however predictions are overly optimistic at $4 \times$ integration time for both $Q_{1.7}$ and $Q_2$ (cases 7-12, 19-24, and 28-30). The cause of this may be a combination of complex factors that involve the contrast sensitivity function (CSF) of the human visual system; for that reason the effects of modifying the RER input variable are examined. Utilizing the RER of only the vertical edge, i.e., the edge most affected by blur, leads to an underestimation of perceived image quality for all cases at $4 \times$ integration time (high amount of smear present) (Fig. 5.22). On the other hand average RER leads to greater overestimations made evident by Figure 5.23, which shows that the average of vertical and horizontal edge response produces higher values than the geometric mean as the disparity between vertical and horizontal RER is increased. These findings suggest that when a disproportionate amount of smear occurs along a given direction, thereby adversely affecting the one dimensional RER perpendicular to the smear vector, an additional term may be required to compensate for perceived losses to image quality.
GIQE5’s overestimation of $\Delta$NIIRS for cases where smear and SNR are high is the result of the RER-SNR interaction term which reduces the influence of RER at increased values of SNR. The effect of this term is further illustrated by the fact that GIQE5 predictions remain virtually unchanged when using only the vertical edge RER. The opposite of this effect manifests in Figure 5.21 for treatments 13, 16, and 25, (high smear, low detector performance) - low detector performance produces lower SNR; consequently the influence
of RER is increased resulting in predictions at or below GIQE 3 and 4 values. Regressing ∆NIIRS ratings against GIQE predictions is shown in Figure 5.24.

![Figure 5.24: Regression of subject ratings against GIQE predictions for factorial experiments.](image)

The results indicate that GIQEs 3 and 4 perform similarly, while GIQE5 predictions vary wildly from subject ratings. In an effort to improve GIQE3, a smear term is added to compensate for these differences, resulting in marginal improvements to regressed slope and $R^2$ (Fig. 5.25). The term takes the form

$$DNIIRS_{Smear} = -c_5(SM - SM_0)$$  \hspace{1cm} (5.2)

where $SM_0$ is the smear present in the baseline configuration and $SM$ is the amount of smear in the current image, and $c_5$ is regressed to be 0.04. An advantage of adding a separate smear term is that GIQE predictions for the validation portion are unchanged because zero smear is added to the images, though it may be seen as doubly compensating for MTF which should be addressed by changes to RER.
From Figure 5.25 it is seen that Equation 5.2 does not model subject ratings well as $\Delta$NIIRS for $Q_1$ images at $t = 4\times$ (high smear) are underestimated while images for $Q > 1$ at $t = 1\times$ are still overestimated. Examining the residuals as a function of Q-sampling shows that model accuracy decreases with increasing Q (Fig. 5.26), thus two candidate solutions that adjust $\Delta$NIIRS proportional to Q-sampling are evaluated.

The first version compensates for overestimations at higher Q by including a linear term that takes the form
\[ DNIIRS_{Qv1} = -c_5(Q - Q_0) \]  
(5.3)

where \( Q_0 \) is the sampling of the baseline system, and \( c_5 \) is 0.12. A second version adds a Q term to \( G/SNR \) that accounts for the significant interaction between Q and detector performance shown by the ANOVA tables in Section 5.3. The lowest SNR experienced by all of the treatments is 4.5, therefore it stands to reason that values below that level are irrelevant to the trade-space study. Utilizing data collected from SNR-degraded imagery, figure 5.27 performs a regression between the first three data points, \( SNR_{\tau=8\%} = 33, 5.44, \) and 3.01 - the corresponding noise equivalent change in transmittance (\( NED_{\tau} \)) values are 0.3, 1.48, and 2.66.

\[ DNIIRS_{Qv2} = -c_3 \left( \frac{(G)(Q)}{SNR} - \frac{(G_0)(Q_0)}{SNR_0} \right) \]  
(5.4)

For Equation 5.4 \( c_3 \) is calculated to be 0.82. Both equations accomplish the same effect and are unaffected by filtering as Q is a system specification rather than an image attribute. Total \( \Delta NIIRS \) for each version are calculated as
Chapter 5. Results

\[
DNIIRS_{\text{GIE}3.1} = -3.32 \log_{10} \left( \frac{GSD}{GSD_0} \right) + 3.32 \log_{10} \left( \frac{RER}{RER_0} \right) - \\
\left( \frac{G}{SNR} - \frac{G_0}{SNR_0} \right) + 1.48(H - H_0) - 0.12(Q - Q_0)
\] (5.5)

\[
DNIIRS_{\text{GIE}3.2} = -3.32 \log_{10} \left( \frac{GSD}{GSD_0} \right) + 3.32 \log_{10} \left( \frac{RER}{RER_0} \right) - \\
0.82 \left( \frac{(G)(Q)}{SNR} - \frac{(G_0)(Q_0)}{SNR_0} \right) + 1.48(H - H_0)
\] (5.6)

Note that the other coefficients for GIE3 remain the same. The regression is accomplished using the design of experiments data, with the following results.

\textbf{Figure 5.28:} Comparison of predictions for GIEEs 3 and 3.1 using factorial experiments data.
Both equations have improved regression slopes and marginally better $R^2$ values than GIQE3 with GIQE3.2 having a slight edge over GIQE3.1, GIQE3 and its smear corrected version when examining regressed slope. Performing GIQE predictions utilizing validation data shows GIQEs 3.1 and 3.2 have minimal differences from GIQE3 when GSD and RER are systematically varied.

The results are also similar for SNR degraded imagery.
When comparing RMSE for all of the data, GIQEs 3, smear correction, 3.1, 3.2, and 4 perform similarly while GIQE5 deviates from the pack on several of the experiments. GIQE3.2 has a slight edge over others, with modest improvements over GIQE3 for all experiment cases.
Table 5.11: Summary of RMSE for each GIQE for all experiments

<table>
<thead>
<tr>
<th>GIQE</th>
<th>Vary f, FN</th>
<th>$Q_1$</th>
<th>$Q_2$</th>
<th>Factorial Experiments</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIQE 3</td>
<td>0.176</td>
<td>0.406</td>
<td>0.236</td>
<td>0.174</td>
<td>0.205</td>
</tr>
<tr>
<td>GIQE 4</td>
<td>0.174</td>
<td>0.385</td>
<td>0.265</td>
<td>0.195</td>
<td>0.219</td>
</tr>
<tr>
<td>GIQE 5</td>
<td>0.207</td>
<td>0.562</td>
<td>0.521</td>
<td>0.377</td>
<td>0.383</td>
</tr>
<tr>
<td>GIQE 3 (Smear Correction)</td>
<td>0.176</td>
<td>0.406</td>
<td>0.236</td>
<td>0.174</td>
<td>0.205</td>
</tr>
<tr>
<td>GIQE 3.1</td>
<td>0.180</td>
<td>0.406</td>
<td>0.236</td>
<td>0.154</td>
<td>0.193</td>
</tr>
<tr>
<td>GIQE 3.2</td>
<td>0.172</td>
<td>0.400</td>
<td>0.233</td>
<td>0.155</td>
<td>0.191</td>
</tr>
</tbody>
</table>

An analysis of GIQE predictions on image attributes recorded from simulated images indicate the GIQEs 3 and 4 perform similarly across all experiments. When varying focal length and f-number, GIQE5 is found to exhibit a bias when predicting $\Delta$NIIRS for $Q_2$ imagery, a characteristic that stems from decreased sensitivity to RER changes at high SNR. Decreased sensitivity is rooted in an interaction term that scales RER weighting as a function of $[1 - e^{-\frac{1}{\text{SNR}}}]$, a by-product of modifying and tuning GIQE3 to provide accurate predictions of Wiener filtered images from unfiltered input values. GIQE5 also provided erroneous predictions for SNR-degraded images sampled at $Q_1$ and $Q_2$ which may be a result of the low SNRs used in the experiment. With respect to GIQE4, findings of this experiment are similar to those in Thurman and Fienup (2008, 2010) showing slight underestimations in $\Delta$NIIRS losses with increased noise. Results from both factorial experiments indicate that, for smeared imagery sampled at $> Q_1$, the perceived loss in image quality recorded by subjects is greater than predictions by the GIQEs. Two proposed modifications to GIQE3, which compensated for additional $\Delta$NIIRS losses with increased Q, produced higher accuracy across all test cases, alluding to the notion that current models must be adjusted if higher Q systems are to be evaluated.

5.5 Temporal Coverage Trade Study

The previous sections have detailed the results of experiments that validated simulation imagery, explored the effects and interactions of key system design parameters, and analyzed the performance of current models for scenarios in the high-Q low SNR domain. Using the validated simulation and GIQE 3.2, this section evaluates the feasibility of implementing
high-Q designs and projected detector and spacecraft bus performance to potentially reduce payload weight and size. The basic supposition is that gains associated with reducing GSD to increase Q, aided by increases in detector and pointing performance, may allow for reduced optical diameter while maintaining image quality - with the caveat of reduced temporal coverage (i.e. decreased performance at dawn and dusk) that stems from reduced signal. The scenario plots ∆NIIRS as a function of Sun angle to depict the relative performance of each design in terms of image quality and temporal coverage (Fig. 5.33). The cutoff for temporal coverage is defined as the point where ∆NIIRS drops to -0.2, thus if System 1 and System 2 drop to -0.2 at 80° and 65°, respectively, System 2 is deemed to have lost 30° in temporal coverage (sunrise to sunset). Assuming the sun traverses 15° per hour, this equates to 2 hours, or roughly a 19% loss in coverage.

\[ ΔNIIRS_{sys1} = -0.2 \]
\[ ΔNIIRS_{sys2} = -0.2 \]

Atmosphere
Target
Lost temporal coverage

(A)

**Figure 5.32:** Temporal coverage is determined by comparing the maximum solar zenith angle achieved when ∆NIIRS reached -0.2 relative to the baseline system.

\[ \Delta NIIRS \] is calculated by using GIQE 3.2 to process inputs which are measured from slant edges that are degraded using the virtual camera. Sun-angle is represented by adjusting \( SNR_{\tau=8\%} \) relative to a baseline (defined as 50 when sun angle = 0, i.e., noon) as illumination changes.
Signal variation is modeled using a cosine projection and estimated atmospheric transmission. Atmospheric optical density at noon is calculated using an estimated VIS transmittance of 0.6 vertically through 100km of atmosphere.

Synthetic image generation (SIG) is not used to determine SNR for varying sun angles - instead relative SNR for angles greater than zero are estimated by scaling the baseline signal, \(s_0\), that produces an \(SNR_{\tau=8\%} = 50\) at noon by a cosine projection term and atmospheric transmission estimates. An SNR of 50 is chosen as it represents the mean of the range used when developing the GIQE. From Figure 5.33 it is seen that the path from the target to the sensor is identical in either case, thus the only difference that is accounted for is the path from the source to the target. Assuming an average vertical transmission of 0.6 through 100km of atmosphere, optical density at noon, \(\delta'_\alpha\), is calculated to be

\[
\delta'_\alpha = \ln(0.6) \approx -0.5 \tag{5.7}
\]

Transmission at some angle \(\theta_n\) is then estimated to be

\[
\tau_n = \exp\left(\sec(\theta)\delta'_\alpha\right) \tag{5.8}
\]

Thus the signal at \(\theta_n\) may be represented as

\[
s_n = s_0 \cos(\theta_n) \exp(\delta'_\alpha [\sec(\theta_n) - 1]) \tag{5.9}
\]
Some modifications were made to the simulation to lessen the complexity and processing time, standardize between configurations, better represent the relative performance of detectors used for space-based remote sensing, and increase the influence of pointing smear on performance. To reduce complexity, processing time, and standardize between different configurations, optics were modeled to be circular and diffraction limited. Average quantum efficiency of the low-performance detector was raised from 40 to 65, a level that was deemed more reasonable for detectors used in high-resolution space imaging. This change also allowed for equal spacing of detector QE performance (65, 80, and 95) - read noise levels remained the same. Finally, as a result of the factorial experiments conducted in Section 5.3, pointing geometry and stability were adjusted to have 0.4 and 0.8 $Q_1$ pixel smear for low and high pointing performance, respectively. These changes are seen as reasonable and within the boundaries which the simulation was originally developed for.

The baseline system used for the first trade is defined as $Q_1$, $QE_{avg} = 65\%$, Smear $= 0.8 \times Q_1$ pixels. The performance of $Q_2$ designs with aperture diameters at 70, 80, and 90 percent of the original $Q_1$ diameter are evaluated at each detector and pointing performance level. Solar zenith angle is varied from 0 to 80° and the calculated ∆NIIRS of each design is plotted. When operating at low detector and pointing performance (Fig. 5.34) designing a $Q_2$ system at 70% of the original optics diameter does not meet the performance level of the $Q_1$ system over the entire range of solar zenith angles.

![Figure 5.34: $Q_2$ designs operating at 70, 80, 90, and 100 percent $Q_1$ diameter. 80, 90, and 100% designs experience a 129, 47, and 15 min loss in coverage respectively. This amounts to 22.7, 8, and 2.6 percent coverage loss.](image)

However, $Q_2$ designs at 80 and 90% of the original $Q_1$ diameter show promise in reducing size and weight with some losses to temporal coverage. $Q_2$ designs at 100% $Q_1$ diameter experience slightly better performance with minimal temporal losses. Though lower in performance, this plot shows the robustness of $Q_1$ designs, as $\Delta NIIRS$ is held constant across a larger swath of sun angles - i.e. slope $\approx 0$ over a larger range, whereas $Q_2$ performance begins to deteriorate almost immediately. Boosting detector performance and pointing stability has obvious benefits to $Q_2$ systems, shown in Figure 5.35 with losses and gains in temporal coverage shown in Table 5.12.

![Figure 5.35](image-url)

**Figure 5.35:** Temporal coverage of $Q_1$ design at low detector and pointing performance compared to A) $Q_2$ designs with high pointing stability B) $Q_2$ designs with high detector performance and C) $Q_2$ designs with high detector performance and pointing stability.
Table 5.12: Temporal coverage gained or lost for $Q_2$ designs with varying aperture size for low/high detector and pointing performance.

<table>
<thead>
<tr>
<th>% $Q_1$ Diameter</th>
<th>Minutes Lost (-) or Gained (+)</th>
<th>Percentage Lost (-) or Gained (+)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Pointing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance Only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-5.8</td>
<td>-1</td>
</tr>
<tr>
<td>90</td>
<td>-32.6</td>
<td>-5.7</td>
</tr>
<tr>
<td>80</td>
<td>-84.2</td>
<td>-14.8</td>
</tr>
<tr>
<td>70</td>
<td>-225.8</td>
<td>-39.7</td>
</tr>
<tr>
<td><strong>High Detector</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance Only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>51.6</td>
<td>9</td>
</tr>
<tr>
<td>90</td>
<td>33.3</td>
<td>5.9</td>
</tr>
<tr>
<td>80</td>
<td>-19.9</td>
<td>-3.5</td>
</tr>
<tr>
<td>70</td>
<td>-214.5</td>
<td>-37.7</td>
</tr>
<tr>
<td><strong>High Pointing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>and Detector</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>58.6</td>
<td>10.3</td>
</tr>
<tr>
<td>90</td>
<td>41.6</td>
<td>7.3</td>
</tr>
<tr>
<td>80</td>
<td>8</td>
<td>1.4</td>
</tr>
<tr>
<td>70</td>
<td>-76.6</td>
<td>-13.46</td>
</tr>
</tbody>
</table>

Figure 5.35 shows that the effect of boosting detector and pointing performance alleviates temporal coverage losses experienced by $Q_2$ designs, with noticeable gains for high SNR cases. For this scenario, Figure 5.36 illustrates that increases in pointing stability benefit $\Delta$NIIRS at higher SNR than detector improvements, while enhanced QE and read noise benefit $\Delta$NIIRS at larger zenith angles - an obvious outcome given that the majority of $\Delta$NIIRS result from changes to SNR (RER is also minimally effected as the Wiener filter gain is reduced proportionally to SNR). Overall, it is seen that increasing detector performance provides the largest amount of compensation for lost temporal coverage.
Chapter 5. Results

Figure 5.36: ∆NIIRS curve comparisons between the baseline configuration, $Q_2$ at 80% diameter with low detector and pointing performance, $Q_2$ at 80% diameter with low detector and high pointing performance, $Q_2$ at 80% diameter with high detector and low pointing performance, and $Q_2$ at 80% diameter with high detector and high pointing performance.

Thus far the trade-study has shown that improved detector performance and pointing stability allow for $Q_2$ designs with smaller apertures to meet or exceed the ∆NIIRS performance of a $Q_1$ design with specifications that are comparable to legacy systems. These findings allude to the notion that future systems held to the same requirements as prior missions may be designed at $Q_2$ with reduced optical diameter, ultimately reducing payload size and weight. Another perspective in which this study may be used is a scenario where the detector of a legacy system is upgraded to the current state of the art.

Though not used for Earth-observing missions, the Hubble Space Telescope (HST) serves as a prime example of increased CCD signal-to-noise performance over time. Originally launched in 1990, the HST’s main camera, the Wide Field Planetary Camera (WFPC-1), was comprised of eight $800 \times 800$ pixel CCD’s, four used for the Wide Field (WF) Camera and the other four used for the Planetary Camera. WF/PC-1 achieved read noise levels of $13e^{-}$ and a peak QE of 50%, but was shortly replace by WF/PC-2 to correct for optical defects in the primary mirror (MacKenty et al., 1992). WF/PC-2 was replaced by the Wide Field Camera 3 (WFC3) in 2009, which exhibited read noise levels of $3.2e^{-}$ and a peak QE of 63% (Dressel, 2012). This scenario is similar to that seen in Figure 5.35b, specifically for the case where $Q_2$ optics diameter and pointing performance is equal to the original $Q_1$ system. Figure 5.37 illustrates the differences between the baseline system, a $Q_1$ design with high detector performance, and a $Q_2$ design with high detector performance. For this
specific case, it is seen that $\Delta$NIIRS at high-SNR is notably improved ($\approx 0.4$ $\Delta$NIIRS), with marginal benefits to temporal coverage (9% increase). Simply upgrading the sensor and retaining the same Q-sampling, i.e. not reducing the pixel pitch, results in marginal gains to $\Delta$NIIRS. It should be noted that if spatial coverage is to be maintained, the pixel density of the $Q_2$ detector must also increase by a factor of two in each direction resulting in $4\times$ the amount of data.

![Graph](image)

**Figure 5.37:** $\Delta$NIIRS curve comparison between the baseline configuration, $Q_1$ with high detector performance, and $Q_2$ at 100% $Q_1$ diameter and high detector performance.

Assuming that both $Q_1$ and $Q_2$ systems perform at projected levels, i.e., high detector and pointing performance, Figure 5.38 serves to illustrate the potential of implementing $Q_2$ designs in future systems.
Figure 5.38: ∆NIIRS curve comparison between the baseline configuration, Q₁, Q₂ at 80%, 90%, and 100% Q₁ diameter with high detector and pointing performance. Blue dots represent a cutoff of -0.2 ∆NIIRS with respect to the baseline Q₁ system, and red dots represent a cutoff of -0.2 ∆NIIRS with respect to the maximum performance of each system.

It is seen that predicted image quality of Q₂ designs at small zenith angles is notably improved over Q₁, however as stated before, the consistency of Q₂ quality is lacking over the range of zenith angles used. One could argue that the coverage cutoff angle for these future systems should not be evaluated at -0.2 ∆NIIRS with respect to the baseline Q₁ system, therefore both the original benchmark (blue dots), and -0.2 ∆NIIRS with respect to the maximum performance of each system (red dots) are shown. When evaluating against the original standard, Q₂ designs offer improved image quality, with minimal losses in coverage. However, if -0.2 ∆NIIRS with respect to the maximum value is used, temporal coverage loss of Q₂ systems is significant.

In summary, this section demonstrated the feasibility of implementing Q₂ designs in future systems with some caveats. The overall shape of Q₁ and Q₂ plots indicate that image quality deteriorates at a higher rate for Q₂ designs as SNR is decreased - an outcome that is consistent with the findings of model validation. This demonstrates the need for Q₁ designs in missions where robustness is favored, especially for cases where SNR is reduced. On the other hand, the trades indicate that when designing a future mission to meet legacy system performance specifications, i.e., if image quality and temporal coverage requirements remain unchanged, Q₂ designs which utilize improved detector performance and/or improved pointing stability show promise in reducing payload size and weight. Furthermore, if the detector
of an existing system is able to be upgraded in subsequent blocks, it may be beneficial to implement a $Q_2$ design provided that QE and read noise are substantially improved in latter detectors. Finally, implementing $Q_2$ designs for future systems where increased pointing stability and detector performance are available mainly depends on mission requirements. System designers must take into account a multitude of trades, to include Q-sampling, when determining how to best meet user needs.

5.6 Post Hoc Analysis

Further analysis of data from varying perspectives may reveal lessons to be learned from experimental methods that should be implemented in future work. The three aspects evaluated are subject learning and fatigue, differences between rating stations, and differences between images. Because rating order is randomized to prevent biases toward a specific configuration, learning and fatigue may be tracked by plotting each subject’s deviation from the mean as a function of the order in which images are rated (Fig. 5.39). Increases in deviation magnitude may reveal fatigue, whereas decreases may indicate learning and/or acclimation to the rating procedures/environment.

![Figure 5.39: Plot of rating differences from the mean as a function of rating order.](image)

Qualitatively, deviations look flat across all experiments (with exception of SNR-degraded $Q_1$ imagery), an indication that subjects experienced minimal changes to rating ability.
as the study progressed. Recalling discussion from the previous sections, the difficulty of rating $Q_1$ SNR-degraded images is highlighted given the the abrupt change in deviations as subjects transition to factorial experiment 1 images during the second collection.

A total of three collections were made on separate dates: 1) Varying focal length, f-number, and SNR degraded images at $Q_2$ 2) SNR degraded images at $Q_1$ and factorial experiment 1, and 3) Factorial experiment 2. To determine whether subjects experienced fatigue or acclimation/learning throughout each session, standard deviations as a function of rating order are plotted in Figure 5.40 and linear models are fit to depict any trends - their slopes and 95% confidence intervals are in Table 5.13.

![Figure 5.40: Plot of standard deviation a function of rating order.](image)

**Table 5.13: Slopes and 95% confidence intervals for rating order vs standard deviation linear regression**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Slope</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vary Focal Length and FN</td>
<td>-1.13E-04</td>
<td>(-0.0014, 0.0011)</td>
</tr>
<tr>
<td>SNR-degraded $Q_2$ images</td>
<td>5.41E-04</td>
<td>(-0.0020, 0.0030)</td>
</tr>
<tr>
<td>SNR-degraded $Q_1$ images</td>
<td>-3.69E-03</td>
<td>(-0.0086, 0.0013)</td>
</tr>
<tr>
<td>Factorial Experiment 1</td>
<td>-1.29E-03</td>
<td>(-0.0018, -0.0008)</td>
</tr>
<tr>
<td>Factorial Experiment 2</td>
<td>-5.06E-03</td>
<td>(-0.0088, -0.0013)</td>
</tr>
</tbody>
</table>

The slopes for varying focal length, f-number, and degrading $Q_1$ images with SNR indicate
minimal acclimation, while the positive slope in SNR-degraded $Q_2$ images may infer fatigue; however, because 95% confidence intervals encompass zero, it is also very likely that no change in rating ability exists. Factorial experiments 1 and 2 show negative slopes and confidence intervals which may be a sign that subjects experienced minute improvements in rating ability over the course of factorial experiments. It is also seen that the slope for factorial experiment 2 is somewhat more negative than factorial experiment 1, suggesting that subjects may have retained ratings skills between the second and third collections. Such a notion is reinforced by the fact that both sets of images were degraded in similar fashion, i.e., a combination of noise and smear versus only changes to SNR, and that they were collected in succession. No conclusive evidence of fatigue is uncovered, which may be a result of utilizing a flicker method to compare images (this serves to reduce eye strain) and limiting the length of each collection to less than 1.5 hours.

Because two images are required to generate the proper number of samples per treatment, differences in ratings between each image is also evaluated. The two images (Fig. 5.41) utilized for all experiments are tiles from a larger image film transparency provided by Exelis Inc., one containing a B-52 bomber and the other containing two jet aircraft (F-102 and F-86).

![Figure 5.41: Two image tiles selected from the transparency (A) Nadir image of B-52 (B) Nadir image of F-102 (top) and F-86 (bottom)](image)

Plotting the main effects of each image on the entire data set indicates that the B-52 image
is rated, on average, 0.05 $\Delta$NIIRS higher than the image containing jets, and performing an ANOVA shows that the effect is statistically significant at the 5% level ($p$-value = 0.0259). This difference may be a result of small objects, e.g., rivets, vortex generators, engine components, present in the B-52 image that allow subjects to perceive higher levels of detail not contained in the jet image. Though a variation in mean ratings between the two images is shown to be probable, a $\Delta$NIIRS of 0.05 is below the threshold of perceivable change to image quality therefore this effect is inconsequential to the results.

![Plot of average $\Delta$NIIRS between B-52 and jet images. P-value = 0.0259.](image)

**Figure 5.42:** Plot of average $\Delta$NIIRS between B-52 and jet images. P-value = 0.0259.

In an effort to reduce data collection times, two stations were constructed to facilitate simultaneous rating sessions. Though numerous precautions were taken to ensure identical configurations, rating differences between workstations are examined to determine if any inconsistencies exist. The main effects plot shows an average $\Delta$NIIRS of 0.03 between left and right stations, however the ANOVA indicates that this difference is not significant at the 5% level ($p$-value = 0.43). These results indicate that differences in ratings between each station are not probable, an expected outcome given that the monitors are identical, collocated, and underwent the same calibration procedures.
Chapter 5. Results

155

Station Position
R L
mean
0.15
0.152
0.154
0.156
0.158
0.16
0.162
0.164
0.166

Figure 5.43: Plot of average ΔNIIRS between left (L) and right (R) rating stations.
P-value = 0.43.

In summary, these analyses indicate that unwanted variance contributions from fatigue, image type, and rating station have little to no effect on experimental results. No conclusive evidence of fatigue is found, however decreasing standard deviations within the first and second factorial experiments suggest minimal acclimation. Furthermore increasingly negative slopes between these experiments may shed light on subjects’ ability to retain rating skills. Though statistically significant at the 5% level, differences between image types are below the threshold of perceivable change, and a difference in ratings between workstations is found to be improbable.

This chapter presented the results and findings of validation activities, factorial experiments, GIQE analysis, and post-hoc evaluations. Objective validation provided an initial indication of model fidelity and was followed by a subjective validation phase which compared perceived changes in image quality between laboratory photographs and simulated imagery. Plotting ΔNIIRS ratings for photographs against model outputs for cases where focal length, aperture, and SNR were systematically varied indicated that subject-assigned ratings for the virtual camera were statistically similar to those of the laboratory. The purpose of producing a validated image chain simulation was to ease the accomplishment of factorial experiments that would normally be expensive and difficult to configure using laboratory equipment. Factorial experiments which employed the validated model revealed modest gains in image quality for $Q_2$ sampled imagery at low SNR and yielded inconclusive results for determining a point of diminishing returns as Q-sampling approached Nyquist.
An analysis of three versions of the GIQE showed similar performance between GIQEs 3 and 4, each with consistently lower RMSE that GIQE5, and utilizing GIQE3 as a baseline platform, two modified versions were assessed to show improved accuracy for factorial experiment data while retaining prediction accuracy on data obtained during subjective validation. System level trades demonstrated the feasibility of implementing $Q_2$ designs for specific scenarios, however the Q-sampling for a given system must be viewed as a trade-space variable and weighed accordingly to determine a design that best meets user requirements. Finally, other factors such as fatigue, acclimation/learning, image content between the two tiles used, and workstation location were shown to have minimal effects on results. The following chapter details lessons learned and conclusions drawn from this study, as well as potential areas for future study.
Chapter 6

Conclusions

6.1 Model Validation

Given that an integral part of this effort is to produce an image chain simulation that is an accurate representation of a physical testbed, objective validation by comparison of modeled and measured MTF proves to be a logical first step before proceeding with subjective experiments. MTF is shown to correlate well with subjective ratings (Leachtenauer et al., 1997), thus objective validation utilizing similar metrics serves as an adequate indicator of expected outcomes when collecting subjective ratings. The results of this validation step suggest that simulation outputs adhere well to those produced by the testbed, however fidelity of the modeling effort was limited by real world constraints. The largest factor was estimating aperture diameter of irregular shapes at a given f-number, a limitation that was revealed when attempting to model the 35mm lens at f/16. The aperture shape at this particular configuration was a hexagram, producing the largest range of possible radii to select from, and resulted in the largest RMS error between modeled and measured MTFs. For future studies, it is recommended that customized optics are used when constructing the testbed so that specifications can be precisely defined and known - a major limitation of this effort given that aperture shapes were defined by the available commercial of the shelf (COTS) products. Furthermore, such customization allows for aperture shapes that better represent space based sensors, e.g., central obscuration and support structures. Although error for this configuration was roughly 2× greater than others, the discrepancy had minimal effect on proceeding experiments because it was minimally used.
Results from subjective ratings indicate that observed changes in image quality for the model are equivalent to those seen in the testbed. For cases where f-number and focal length are varied at high SNR, simulation images are within $\pm 0.07 \Delta \text{NIIRS}$ over the entire range, thus there is 95% confidence that subjects would not perceive a difference in interpretability between the two images. When degrading images by decreasing SNR, model and testbed consistency are seen over the entire range for $Q_2$ images, however 95% confidence intervals of $Q_1$ images are in excess of 0.1 $\Delta \text{NIIRS}$ when $SNR_{\tau=8\%}$ is below 33 for model images. High variability in $Q_1$ SNR-degraded imagery is believed to stem from a weakening of the masking effect that noise has on details given the higher edge response typically seen in $Q_1$ sampled images. Subjects noted the difficulty of balancing the amount of noise present with the level of perceived detail when determining interpretability, as well as assigning noisy images to a GSD-degraded equivalent. The task may have been more straightforward if noisy images were matched to a noise-degraded marker set, however there is low confidence in using general models to predict $\Delta \text{NIIRS}$ as a function of SNR for this specific case given the complex interaction of MTF, SNR, and image processing on interpretability.

Using untrained analysts to rate images did not have significant negative effects on data variance with the exception of low SNR $Q_1$ images. Low variance may be attributed to a common marker set which had identical scene content to the image being rated, however marker set interpolation was a main source of concern given the MTF effects of bicubic interpolation. In addition, flickering between the image-marker pair allowed for direct comparison of details but had the downside of exaggerating rotational differences between testbed and simulation images. A constant MTF interpolator may alleviate both difficulties for future studies by minimizing MTF effects when images are down-sampled or rotated.

Though results may indicate high model fidelity, it should be noted that its usefulness is limited to the configurations of the testbed and the specific scenarios used in this study. The modeled apertures are not representative of space base systems and many framing sensors may stray from the nadir viewing geometry used. Furthermore, the assumption of an LSI imaging chain is not consistent with phenomena such as blur that may have significant effects on image quality at the edges of a scene. Because the detector model is based on the specific camera used in the lab, it will need to be modified if other detector types or modes of collection are investigated. Notable assumptions made when developing the detector model include panchromatic square pixels, short exposure times, poisson and additive white gaussian noise (AWGN), and a full-frame CCD fill factor of roughly 1. These may not be appropriate for CMOS, interline transfer CCDs, scanning sensors, or when CONOPS call
for a change to acquisition procedures, e.g., longer exposure times.

### 6.2 Factorial Experiments

After completing validation steps that ensure virtual camera images adequately represent laboratory photographs, factorial experiments utilizing the validated simulation are performed with the intent to explore image quality when pushing the boundaries of detector sampling and performance. The difficulty of producing images for these experiments is drastically reduced by using the virtual camera because environmental factors are essentially eliminated, and variables such as smear and SNR can be precisely controlled. Factorial experiments adjust four factors (Q-sampling, integration time, detector performance, and pointing accuracy) over a range of predefined values, i.e. settings, and measure $\Delta NIIRS$ assigned by experiment participants. Changes to factors relative to the baseline configuration manifest as variations in image attributes, e.g. SNR, resolution, smear, and sharpness, which are estimated at the corner of a 16000 x 16000 array using a step-and-stare CONOP and orbit altitude typically seen in modern high resolution satellites.

Results from this experiment indicate that, for the specific scenario used, gains experienced by halving GSD when transitioning from $Q_1$ to $Q_2$, i.e. $\Delta NIIRS = +1$, overshadow losses typically associated with $Q_2$ designs such as decreased SNR and MTF. Though the average change in $\Delta NIIRS$ between $Q_1$ and $Q_2$ configurations is shown to be only +0.25, these experiments define a baseline $Q_1$ $SNR_{\tau=50\%} = 38$ ($SNR_{\tau=8\%} = 13.8$), a value at the very low end of SNR typically seen in operational systems, thus scenarios with high SNR are likely to experience larger $\Delta NIIRS$ increases associated with $Q_2$ designs. It is also seen that attempting to compensate for lost signal by increasing integration time to 4 times is not beneficial due to losses in perceived image quality resulting from increased smear, however it should be noted that smear for the scenario used is calculated at the corner for the array, representing the worst case. Portions of the image closer to the center of the field of view may be less sensitive to increases in integration time (especially with improved pointing) because components of orbital smear are space-variant and increase proportionally to the number of pixels away from the center of the array. In future studies it may be valuable to investigate if image quality improves for smaller increases to integration time, e.g., 1.5 times or 3 times, at varying locations on the focal plane.
Findings of the second factorial experiment, which aimed to reveal the extent and location of curvature in \( \Delta \text{NIIRS} \) when transitioning from \( Q_1 \) to \( Q_2 \), are inconclusive as the relationship between \( \Delta \text{NIIRS} \) and Q-sampling is found to be linear over the range of \( Q_1, Q_{1.7}, \) and \( Q_2 \). Though similar to the findings of Fiete and Tantalo (1999), the linear relationship extends to higher levels of Q-sampling which is likely the result of increasing sampling in two dimensions rather than one. Interestingly, the fall off in performance gains as Q approaches 2 in Fiete’s work is similar to that seen in images with high smear (4\( \times \) integration time), arguably equivalent phenomenon given that \( \Delta \text{NIIRS} \) may be expressed either in terms of changes to GSD or RER. In other words, the delta between GSD in along scan and cross scan sampling in Fiete and Tantalo (1999) is analogous to high (vertical) and low (horizontal) RER caused by smear in this work, thus a similar falloff in quality improvement is seen when increasing two-dimensional sampling and adding a substantial amount of smear in one direction. Based on these results, it is likely that curvature for framing arrays exists between \( 1.7 < Q < 2 \), thus future studies should target this range so that the transition may be characterized.

### 6.3 GIQE Analysis

Using simulation ratings gathered from validation and factorial experiments, three IQEs specific to the ISR community (GIQEs 3, 4, and 5) are evaluated to uncover capabilities and shortfalls of the status quo. GIQEs 3 and 4 are regressed using filtered input parameters using ratings obtained from hardcopy images, while latest iteration, GIQE5, is specifically designed for soft copy evaluations utilizing unfiltered inputs and is not widely used among the community. As the images used to develop GIQEs 3 and 4 exhibited sampling rates near \( Q_1 \), predictions are expected to deviate when sampling is increased to \( Q_2 \). On the other hand, GIQE5 was developed to account for \( Q_2 \) sampling, thus its accuracy is expected to improve over GIQEs 3 and 4 for Nyquist sampled imagery. Surprisingly, an analysis of simulation ratings where focal length and f-number were varied resulted with GIQEs 3 and 4 performing similarly, with lower RMSEs than GIQE5 for both \( Q_1 \) and \( Q_2 \) images. This trend was consistent throughout the entire data-set, with GIQEs 3 and 4 producing lower RMSE for SNR-degraded imagery and factorial experiments.

GIQE5’s error seems to stem from a decreased sensitivity to changes in RER as SNR is
increased, resulting in optimistic $\Delta$NIIRS predictions for $Q_2$ configurations. This characteristic is rooted in the assumption of using a Wiener filter for processing soft copy images - in essence GIQE5 is an adaptation of GIQE3 that produces accurate NIIRS predictions for unfiltered inputs on images that have been Wiener filtered. Such an assumption is problematic in that Wiener filters may be tuned to adjust the balance between sharpening and blurring, and because coefficients for the object power spectrum estimate may change based on image content. Therefore, without knowing the exact form of the filter used, GIQE5 predictions may not be the most accurate model for this work.

On the other hand, using a specific filter to produce accurate $\Delta$NIIRS may be beneficial in that it levels the playing field between different systems. Currently, the image quality of different platforms may be tweaked to produced higher NIIRS ratings under specific conditions by fine tuning the sharpening kernel, thus it may be shown that one system performs better than another using specific post processing techniques, and vice-a-versa. With the assumption of using a specific Wiener filter in GIQE5, performance of varying systems may be directly compared as images must be sharpened in a specific fashion to ensure accurate predictions. However, it can also be argued that post processing is a critical link in the imaging chain and limiting image sharpening techniques to those used in the development of GIQE5 can mischaracterize the image quality of different systems under varying conditions.

In an effort to improve IQE performance, two variants of GIQE3 were developed to compensate for overestimations produced by the original when predicting $\Delta$NIIRS for $Q > 1$ during factorial experiments. The first variant (GIQE3.1) included a subtractor term that decreased $\Delta$NIIRS gains proportional to the difference between the baseline $Q$-sampling rate ($Q_1$) and the observed sampling rate. The second variant (GIQE3.2) scales the $G/\text{SNR}$ term by $Q$ due to the increased sensitivity of $Q_2$ designs observed when subjects evaluated SNR-degraded imagery. Both variants exhibited increased prediction accuracy for $Q > 1$ designs during the factorial experiment and maintained similar RMSEs for validation data, with the second variant having slightly lower RMSE than both GIQE3 and the 3.1.

6.4 Temporal Coverage Trades

System trades were conducted which evaluated the image quality ($\Delta$NIIRS) and temporal coverage of various designs. Temporal coverage gains and losses were defined as the zenith
angle corresponding to a -0.2 \( \Delta \text{NIIRS} \) with respect to a \( Q_1 \) system operating at low detector and pointing performance. This study utilized cosine projection and a simplified model for atmospheric transmission to adjust SNR as a function of solar zenith angle, resulting in a reduction of \( \Delta \text{NIIRS} \) with increased solar zenith angle. Simulation parameters were adjusted slightly (low QE performance = 65 and circular diffraction limited apertures), but were deemed to exist within the boundaries used for objective and subjective model validation. It must be stressed that the outcomes are confined to the specific laboratory configuration used, and may not be representative of actual earth-observing systems. Using an aperture that more resembles operational systems may be a simple change that bolsters simulation outputs in future work, however similar validation steps using an appropriate laboratory set-up should be performed.

The findings of the study bring to light the feasibility of \( Q_2 \) designs provided that detector performance and pointing are significantly improved. In cases where a future sensor must meet the performance specifications of a legacy \( Q_1 \) system, improved detector performance and pointing allow for \( Q_2 \) designs with decreased optical size and weight to be considered. If a detector is upgraded during a servicing mission or between subsequent blocks, i.e., if pointing and optical diameter remain constant over time with significant improvements to detector performance, implementing a \( Q_2 \) design may lead to notable advantages over \( Q_1 \) at high detector performance. This finding must be caveated with the fact that if pixel density is kept constant, the \( Q_2 \) design suffers from lost spatial coverage - conversely if pixel density is increased to retain field of view, data transfer requirements quadruple as the number of pixels must increase by \( Q^2 \). Lastly, when comparing \( Q_1 \) and \( Q_2 \) designs with high detector and pointing performance, \( Q_2 \) systems exhibit increased image quality for ideal situations (high SNR, low smear), whereas \( Q_1 \) demonstrates added robustness, maintaining image quality over a larger range of solar zenith angles. As this study only focused on temporal coverage, future work should consider image quality as a function of location on the focal plane. Overall, Q-sampling should be treated as any system specification, and must be balanced with numerous factors to include user needs, budget, time, and available technologies.
Appendix A

Detailed Flow Chart of Digital Model
Appendix A. Detailed Flow Chart of Digital Model

Figure A.1: Inputs and Outputs of Digital Model
Appendix B

Detailed Flow Chart of Testbed
Figure B.1: Testbed flow chart.
Appendix C

Objective Model Validation

Results

Figure C.1: MTF and RMSE plots for f = 16mm and FN = 16
Appendix C. Objective Model Validation Results

Figure C.2: MTF and RMSE plots for f= 16mm and FN = 8

Figure C.3: MTF and RMSE plots for f= 25mm and FN = 16
Appendix C. Objective Model Validation Results

Figure C.4: MTF and RMSE plots for f= 25mm and FN = 8

Figure C.5: MTF and RMSE plots for f= 35mm and FN = 16
Figure C.6: MTF and RMSE plots for \( f = 35 \text{mm} \) and \( \text{FN} = 8 \)

Figure C.7: MTF and RMSE plots for \( f = 50 \text{mm} \) and \( \text{FN} = 16 \)
Figure C.8: MTF and RMSE plots for f = 50mm and FN = 8
## Appendix D

### Standard Deviations of Collected Data

Table D.1: Standard deviations and 95% confidence intervals for testbed and simulation DNIIRS ratings when varying focal length and f-number

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Testbed</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focal Length (mm)</td>
<td>F-Number</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>16</td>
<td>f/8</td>
<td>0.15</td>
</tr>
<tr>
<td>16</td>
<td>f/16</td>
<td>0.09</td>
</tr>
<tr>
<td>25</td>
<td>f/8</td>
<td>0.12</td>
</tr>
<tr>
<td>25</td>
<td>f/16</td>
<td>0.15</td>
</tr>
<tr>
<td>35</td>
<td>f/8</td>
<td>0.13</td>
</tr>
<tr>
<td>35</td>
<td>f/16</td>
<td>0.13</td>
</tr>
<tr>
<td>50</td>
<td>f/8</td>
<td>0.15</td>
</tr>
<tr>
<td>50</td>
<td>f/16</td>
<td>0.20</td>
</tr>
</tbody>
</table>
Table D.2: Standard deviations and 95% confidence intervals for testbed and simulation DNIIRS ratings when varying $SNR_r=8\%$ at $Q=1$ and $Q=2$

<table>
<thead>
<tr>
<th>1/SNR</th>
<th>Testbed Standard Deviation</th>
<th>95% CI</th>
<th>Model Standard Deviation</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.03</td>
<td>0.20</td>
<td>0.06</td>
<td>0.23</td>
<td>0.07</td>
</tr>
<tr>
<td>0.18</td>
<td>0.29</td>
<td>0.10</td>
<td>0.37</td>
<td>0.12</td>
</tr>
<tr>
<td>0.33</td>
<td>0.44</td>
<td>0.14</td>
<td>0.37</td>
<td>0.12</td>
</tr>
<tr>
<td>0.48</td>
<td>0.46</td>
<td>0.15</td>
<td>0.42</td>
<td>0.14</td>
</tr>
<tr>
<td>0.63</td>
<td>0.51</td>
<td>0.17</td>
<td>0.54</td>
<td>0.18</td>
</tr>
<tr>
<td>0.03</td>
<td>0.12</td>
<td>0.04</td>
<td>0.18</td>
<td>0.06</td>
</tr>
<tr>
<td>0.18</td>
<td>0.19</td>
<td>0.06</td>
<td>0.24</td>
<td>0.08</td>
</tr>
<tr>
<td>0.33</td>
<td>0.24</td>
<td>0.08</td>
<td>0.24</td>
<td>0.08</td>
</tr>
<tr>
<td>0.48</td>
<td>0.24</td>
<td>0.08</td>
<td>0.28</td>
<td>0.09</td>
</tr>
<tr>
<td>0.63</td>
<td>0.27</td>
<td>0.09</td>
<td>0.26</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Table D.3: Standard deviations and 95% confidence bounds for design of experiments data.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Standard Deviation</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>2</td>
<td>0.12</td>
<td>0.04</td>
</tr>
<tr>
<td>3</td>
<td>0.17</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>0.13</td>
<td>0.04</td>
</tr>
<tr>
<td>5</td>
<td>0.17</td>
<td>0.06</td>
</tr>
<tr>
<td>6</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>7</td>
<td>0.13</td>
<td>0.04</td>
</tr>
<tr>
<td>8</td>
<td>0.10</td>
<td>0.03</td>
</tr>
<tr>
<td>9</td>
<td>0.11</td>
<td>0.04</td>
</tr>
<tr>
<td>10</td>
<td>0.12</td>
<td>0.04</td>
</tr>
<tr>
<td>11</td>
<td>0.10</td>
<td>0.03</td>
</tr>
<tr>
<td>12</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>13</td>
<td>0.21</td>
<td>0.07</td>
</tr>
<tr>
<td>14</td>
<td>0.16</td>
<td>0.05</td>
</tr>
<tr>
<td>15</td>
<td>0.18</td>
<td>0.06</td>
</tr>
<tr>
<td>16</td>
<td>0.19</td>
<td>0.06</td>
</tr>
<tr>
<td>17</td>
<td>0.22</td>
<td>0.07</td>
</tr>
<tr>
<td>18</td>
<td>0.18</td>
<td>0.06</td>
</tr>
<tr>
<td>19</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>20</td>
<td>0.14</td>
<td>0.05</td>
</tr>
<tr>
<td>21</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>22</td>
<td>0.14</td>
<td>0.05</td>
</tr>
<tr>
<td>23</td>
<td>0.13</td>
<td>0.04</td>
</tr>
<tr>
<td>24</td>
<td>0.14</td>
<td>0.05</td>
</tr>
<tr>
<td>25</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>26</td>
<td>0.17</td>
<td>0.06</td>
</tr>
<tr>
<td>27</td>
<td>0.14</td>
<td>0.05</td>
</tr>
<tr>
<td>28</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>29</td>
<td>0.14</td>
<td>0.05</td>
</tr>
<tr>
<td>30</td>
<td>0.08</td>
<td>0.03</td>
</tr>
</tbody>
</table>

First Collect

Second Collect
Bibliography


Evo-Lite. Lumitile. Product Overview.


Andy Cochrane, Kevin Schulz, Rick Kendrick, and Ray Bell. Q selection for an electro-
optical earth imaging system: theoretical and experimental results. Optics express, 21

Joseph W Goodman. Introduction to Fourier optics. Roberts and Company Publishers,
2005.

Kim-Han Thung and Paramesran Raveendran. A survey of image quality measures. In Techn-
nical Postgraduates (TECHPOS), 2009 International Conference for, pages 1–4. IEEE,
2009.

Ahmet M Eskicioglu and Paul S Fisher. A survey of quality measures for gray scale image

Anders E Boberg. Subjective vs. objective image quality. INTERNATIONAL ARCHIVES


1981.

Michele A Saad, Alan C Bovik, and Christophe Charrier. Blind image quality assessment: A
natural scene statistics approach in the dct domain. Image Processing, IEEE Transactions

Glenn D Boreman. Modulation transfer function in optical and electro-optical systems,

Ronald G Driggers, Michael Kelley, and Paul Cox. National imagery interpretation rat-
ing system and the probabilities of detection, recognition, and identification. Optical


org.

Steven L Smith, James Mooney, Theodore A Tantalo, and Robert D Fiete. Understanding
image quality losses due to smear in high-resolution remote sensing imaging systems.


Doug Griffith. General image quality equation (giqe), April 2012.


