

**AN ANALYSIS OF MULTIMODAL SENSOR FUSION FOR TARGET
DETECTION IN AN URBAN ENVIRONMENT**

By

Brian M. Flusche

B.S. United States Air Force Academy, 2002

M.S. Air Force Institute of Technology, 2006

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

Rochester Institute of Technology

January 21, 2011

Signature of the Author _____

Accepted by _____
Coordinator, Ph.D. Degree Program Date

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE

ROCHESTER INSTITUTE OF TECHNOLOGY

ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Brian M. Flusche
has been examined and approved by the
dissertation committee as satisfactory for the
dissertation thesis required for the
Ph.D. degree in Imaging Science

Dr. John R. Schott, Dissertation Advisor

Dr. Carl N. Salvaggio

Dr. Michael G. Gartley

Dr. William F. Basener

Date

DISSERTATION RELEASE PERMISSION
CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY

Title of Dissertation:

**AN ANALYSIS OF MULTIMODAL SENSOR FUSION FOR TARGET
DETECTION IN AN URBAN ENVIRONMENT**

I, Brian M. Flusche, hereby grant permission to Wallace Memorial Library of R.I.T. to reproduce my dissertation in whole or in part. Any reproduction will not be for commercial use or profit.

Signature_____

Date

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

Abstract

This work makes a compelling case for simulation as an attractive tool in designing cutting-edge remote sensing systems to generate the sheer volume of data required for a reasonable trade study. The generalized approach presented here allows multimodal system designers to tailor target and sensor parameters for their particular scenarios of interest via synthetic image generation tools, ensuring that resources are best allocated while sensors are still in the design phase. Additionally, sensor operators can use the customizable process showcased here to optimize image collection parameters for existing sensors.

In the remote sensing community, polarimetric capabilities are often seen as a tool without a widely accepted mission. This study proposes incorporating a polarimetric and spectral sensor in a multimodal architecture to improve target detection performance in an urban environment. Two novel multimodal fusion algorithms are proposed—one for the pixel level, and another for the decision level. A synthetic urban scene is rendered for 355 unique combinations of illumination condition and sensor viewing geometry with the Digital Imaging and Remote Sensing Image Generation (DIRSIG) model, and then validated to ensure the presence of enough background clutter. The utility of polarimetric information is shown to vary with the sun-target-sensor geometry, and the decision fusion algorithm is shown to generally outperform the pixel fusion algorithm. The results essentially suggest that polarimetric information may be leveraged to restore the capabilities of a spectral sensor if forced to image under less than ideal circumstances.

Acknowledgments

I would like to extend a heartfelt thanks to the people who helped me through this project. My advisor, Dr. John Schott, helped me to lay a solid foundation for the experiments by teaching me to look for weaknesses from the perspective of a potential reviewer. Dr. Michael Gartley took the time to help dissect problems with me whenever hit a wall, and I always came away from our meetings with a slew of good ideas to investigate. Dr. William Basener provided me the MATLAB code for his anomaly detection algorithm, and teamed with Dr. David Messinger to ensure I understood the theory behind the calculations. Together, these people helped me to quickly build on past successes rather than attempting to duplicate previous effort.

This project required an enormous amount of computer simulation and analysis, and would never have been completed if I was forced to develop it on my own, on a single computer. Scott Brown and Dave Pogorzala provided an immense amount of help with the DIRSIG model. Scott helped me understand the intricacies in the simulation, enabling me to work as a user and keeping me from getting bogged down by software development issues of the model, while Dave helped me build the synthetic scene that I later analyzed. Dr. Carl Salvaggio ensured I had access to RIT's Research Computing resources, and Paul Mezzanini worked tirelessly to ensure the machines were always available, immediately tackling and solving any problems that arose. These individuals provided me a phenomenal capability and helped me reap an enormous time savings.

Whenever there's an actual field experiment, it takes a team of people to make it happen smoothly. Thanks to Dr. Michael Gartley and Jason Faulring for operating the sensors, thanks to Dr. Michael Presnar for registering the WASP-Lite data, and thanks to Lingfei Meng for measuring truth data and helping with crowd control.

Next I owe a big thanks to Cindy Schultz for her administrative support. Cindy repeatedly went out of her way to help me resolve any issues that arose with regard to RIT policies and procedures. Thanks to the Air Force team and all those who came out for our afternoon PT sessions—their friendship helped me to stay upbeat.

Finally, I owe a special thanks to my wife Amy—her understanding, love and encouragement never wavered and made this all possible.

Brian M. Flusche

Table of Contents

	Page
Abstract.....	iv
Acknowledgments.....	v
Table of Contents.....	vii
List of Figures.....	xi
1 Introduction.....	1
2 Project Overview.....	6
2.1 Research Questions.....	6
2.2 Objectives.....	7
3 Background.....	11
3.1 Section Overview.....	11
3.2 Existing Systems.....	11
3.3 Data Fusion Levels.....	16
3.4 Decision Level Fusion Techniques.....	18
3.5 Potential Sensing Modalities.....	21
3.6 Polarimetric Remote Sensing.....	25
3.7 Previous Multimodal Data Fusion Demonstrations.....	27
3.8 Existing Simulation Software.....	34
3.9 Atmospheric Compensation.....	46
3.10 Detection Algorithms.....	49
3.11 Section Summary.....	53
4 Methodology.....	54
4.1 Section Overview.....	54

4.2	Defining the Scenario	54
4.3	Defining the Sensor	55
4.4	Designing a Model Scene	58
4.5	Varying Relevant Parameters	61
4.6	Simulating Registration Errors	65
4.7	Compensating Atmospheric Effects.....	66
4.8	Fusing Sensor Outputs.....	67
4.9	Quantifying the Impact.....	72
4.10	Independently Optimizing CEM & TAD Geometry.....	76
4.11	Registering the Images	77
4.12	Establishing the Maximum Attainable Benefit.....	78
4.13	Section Summary	79
5	Preliminary Studies	80
5.1	Section Overview	80
5.2	Background and Covariance Matrices	81
5.3	Sensing Modality Correlation	87
5.4	SCR Calculation.....	90
5.5	ROC Curve <i>FAR</i> Threshold	95
5.6	SPI Algorithm: TAD vs. RX.....	103
5.7	Contrast vs. DOLP	109
5.8	Establishing Sensitivity to Geometry.....	118
5.9	Section Summary	125

6	Results	127
6.1	Section Overview	127
6.2	Co-located Fusion with Perfect Registration	128
6.3	Co-located Fusion with Single Pixel Registration Error	133
6.4	Multiple Sensor Platforms	138
6.5	Phenomenology Investigation.....	142
6.6	Missed Dark Targets: Sampling or SNR?	148
6.7	Section Summary	157
7	Simulation Supporting Multimodal Sensor System Design Requirements.....	160
7.1	Section Overview	160
7.2	Method.....	161
7.3	Results	163
7.4	Section Summary	170
8	Field Verification	172
8.1	Section Overview	172
8.2	Scenario Design	172
8.3	Image Acquisition	175
8.4	Scene Clutter Assessment.....	178
8.5	Fusion Algorithm Evaluation.....	180
8.6	Section Summary	189
9	Multi-View Polarimetric Fusion.....	190
9.1	Section Overview	190
9.2	Multiple PI Image Theory	191

9.3	Multiple PI Assessment Method	193
9.4	Results	200
9.5	Section Summary	207
10	Conclusion.....	210
	References.....	214

List of Figures

Figure	Page
Figure 1. Potential levels of data fusion and their relation to any image processing steps. Image courtesy of Pohl <i>et al.</i> , 1998.	16
Figure 2. Data processing scheme for hyperspectral polarimetric imaging system used by Cavanaugh <i>et al.</i> (2006).	33
Figure 3. Synthetic polarimetric imaging scene generated by Sadjadi and Chun (2004) using the IRMA software package. The scene is displayed as an intensity value for the S_0 (top), S_1 (middle), and S_2 (bottom) Stokes vectors.	37
Figure 4. Synthetic CameoSim scenes used by Harvey <i>et al.</i> (2008) for a laser imaging application.	39
Figure 5. (Left) Example of a house generated with Rhinoceros. (Right) Screen capture of the Tree Professional software package illustrating a rendered Norway maple. Images courtesy of Ientilucci and Brown (2003).	41
Figure 6. (Left) Actual image of a Humvee under camouflage netting. (Right) Synthetic image of the same Humvee under camouflage netting, demonstrating accurate reproduction of a highly detailed model and shadowing phenomenology. Images courtesy of Barcomb <i>et al.</i> (2004).	42
Figure 7. (Left) Actual polarimetric images of vehicles under a cold, clear night sky. (Right) Synthetic image generated by DIRSIG based on the actual scene. Images courtesy of Gartley <i>et al.</i> (2007).	44

Figure 8. The TAD algorithm operating on data notionally distributed in two dimensions. The closest pairs of points are deemed background and linked by dotted lines, with different colors implying different classes. Pixels linked to their nearest neighbor by longer solid lines are likely anomalous. Image courtesy of Basener and Messinger (2009).51

Figure 9. The TAD algorithm produces level sets of arbitrary topology (solid lines), allowing detection of anomalous members (circled) both inside and outside of the background’s convex hull. Image courtesy of Basener and Messinger (2009).52

Figure 10. (Left) DIRSIG rendering of Megascene 1, Tile 1. (Right) Google Earth snapshot of the region of Rochester, NY nominally represented by Megascene 1, Tile 1.58

Figure 11. Subaru wagon model provided by AFOSR for use in DIRSIG. Notice the level of detail—a side mirror, front grille and even door handles have been incorporated.59

Figure 12. The solar zenith, sensor zenith and sensor relative azimuth angles were all varied for this trade study while the hyperspectral (blue rectangle), multispectral (red rectangle) and polarimetric (green rectangle) sensor GSD values were fixed.61

Figure 13. For each time of day, the sensor azimuth angles of interest were $\pm 20^\circ$ from the line representing the specular reflection coming from the center of the scene. The red arcs represent circular flight paths with varying sensor zenith angles.63

Figure 14. Registration error introduced between high spatial resolution polarimetric pixels, labeled P1 – P36, and low spatial resolution hyperspectral super pixel labeled H1..... 65

Figure 15. In the pixel fusion algorithm, DOP (black) will be conceptually stacked onto existing spectral bands (shown as RGB above) to create the fused data set..... 68

Figure 16. When the area under a ROC curve was integrated to a user-defined threshold (vertical black dashed line), an interpolated detection rate (red circle) was determined by assuming the ROC curve is piecewise linear (red dashed line) between the two nearest measured values (blue x)..... 74

Figure 17. The ideal multispectral (MS), hyperspectral (HS) and polarimetric (PI) sensing geometries were established for each time of day. 76

Figure 18. The larger GSD hyperspectral (black) and smaller GSD multispectral (red) CEM ROC curves were created by analyzing the nadir scenario at 0800 with a global covariance matrix. The background mean vector was constructed using the entire scene (solid lines), a 3x3 sliding window with the pixel of interest masked out (dashed lines) or a 5x5 sliding window with the pixel of interest masked out. 83

Figure 19. The larger GSD hyperspectral (black) and smaller GSD multispectral (red) CEM ROC curves were created by analyzing the off-nadir scenario at 0800 with a global covariance matrix. The background mean vector was constructed using the entire scene (solid lines), a 3x3 sliding window with the pixel of interest masked out (dashed lines) or a 5x5 sliding window with the pixel of interest masked out. 84

Figure 20. The larger GSD hyperspectral (black) and smaller GSD multispectral (red) SPI ROC curves were created by analyzing the nadir scenario at 0800 with a global covariance matrix. The background mean vector was constructed using the entire scene (solid lines), a 3x3 sliding window with the pixel of interest masked out (dashed lines) or a 5x5 sliding window with the pixel of interest masked out. 85

Figure 21. The larger GSD hyperspectral (black) and smaller GSD multispectral (red) ROC curves were created by analyzing a sensor zenith angle of 50° and a sensor azimuth angle of 175° at 0800 with a global covariance matrix and local background mean vector drawn from a 5x5 neighborhood with a 3x3 neighborhood around the pixel of interest masked out. The solid lines were obtained using the CEM algorithm from Sec. 3.10, while the dotted lines were obtained using the SPI algorithm from Sec. 4.8. 86

Figure 22. For an image acquired at 0800, with a sensor zenith angle of 40° and a sensor azimuth angle of 175° from the sun, the TAD score for each pixel is plotted as a function of the pixel's multispectral CEM score. Very little correlation exists, since $r = 0.1023$ 88

Figure 23. For an image acquired at 0800, with a sensor zenith angle of 40° and a sensor azimuth angle of 175° from the sun, ROC curves generated with either the multispectral CEM or TAD scores for each pixel are plotted, confirming very little correlation exists between the two metrics. 89

Figure 24. Signal to clutter ratio (SCR) as a function of sensor zenith angle with the sensor azimuth angle fixed at 175° from the sun, the solar zenith angle fixed at 55° ,

sensor SNR = 200 and GSD = 0.5 m. The hyperspectral case (black) hovered just below SCR = 30, while the multispectral case (red) decayed from SCR = 20 near nadir to SCR = 5 at the most oblique angle. Note that the DOLP case (dashed blue line) varied from about SCR = 2 to SCR = 5 near the sun's specular reflection, while the TAD case (dotted blue line) peaked at an SCR near 20 at a zenith angle of 50°. 91

Figure 25. RGB image of 256 band hyperspectral COMPASS image collected over the region of Rochester NY similar to that represented by Megascene 1. A zoomed view shows an isolated reddish vehicle (boxed) used to identify a target pixel in the image. (Right) RGB rendering of DIRSIG's Megascene 1. Several red target vehicles can be observed along with six different models of decoy vehicles.....92

Figure 26. With the solar zenith angle fixed at 55°, the sensor zenith angle at 50° and a relative sensor azimuth angle 175° from the sun, the TAD algorithm was used to identify the most anomalous 0.5% of pixels in the urban scene. The anomalous pixels were identified as either target or background material and plotted as a histogram. The scale for number of pixels has been truncated to enable the target pixel histogram bars to be observed.94

Figure 27. The area under the CEM ROC curve (SNR = 200) is shown as a function of the user-controlled FAR threshold for the multispectral (solid) and hyperspectral (striped) scenarios at both the 0600 (blue) and 0800 (red) times of day. Note the error bars, indicating the displayed area under the ROC curve is representative of a typical value after random noise has been applied.97

Figure 28. The area under the CEM ROC curve (SNR = 100) is shown as a function of the user-controlled *FAR* threshold for the multispectral (solid) and hyperspectral (striped) scenarios at both the 0600 (blue) and 0800 (red) times of day. Note the error bars, indicating the displayed area under the ROC curve is representative of a typical value after random noise has been applied.98

Figure 29. The area under the CEM ROC curve (SNR = 500) is shown as a function of the user-controlled *FAR* threshold for the multispectral (solid) and hyperspectral (striped) scenarios at both the 0600 (blue) and 0800 (red) times of day. Note the error bars, indicating the displayed area under the ROC curve is representative of a typical value after random noise has been applied.99

Figure 30. The area under the CEM ROC curve (SNR = 15) is shown as a function of the user-controlled *FAR* threshold for the multispectral (solid) and hyperspectral (striped) scenarios at both the 0600 (blue) and 0800 (red) times of day. Note how the multispectral area hovers near zero for all *FAR* values, making a meaningful comparison to this data point impractical. Further, the error bars indicate a high degree of variability in the area under the ROC curve after random noise has been applied. 100

Figure 31. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 200. 101

Figure 32. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 100. 101

Figure 33. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 75. 102

Figure 34. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 60. 102

Figure 35. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 50. 103

Figure 36. Although using TAD on all three Stokes bands (S_0 , S_1 and S_2) adds value with the SPI algorithm, even better performance is obtained using only two Stokes bands (S_1 and S_2)..... 105

Figure 37. The TAD scores for each pixel are plotted as derived from either three Stokes bands (S_0 , S_1 and S_2) or only two Stokes bands (S_1 and S_2). The two-band case produces better results when exploited by the SPI algorithm because including the S_0 band flags bright false alarms (boxed) as polarimetrically anomalous pixels. 106

Figure 38. ROC curves were calculated for the hyperspectral SPI curves derived from applying TAD to the S_1 and S_2 bands (dotted black line), applying RX to the S_1 and S_2 bands (dotted red line), and applying RX to the S_0 , S_1 and S_2 bands (dashed red line). The hyperspectral CEM curve (black solid line) is shown for the same viewing geometry, indicating the value added by incorporating polarimetric information. .. 107

Figure 39. The TAD and RX scores used in Figure 38 are shown for each pixel in the scene. TAD and RX generally rank anomalous pixels similarly, but TAD seems to weight more anomalous pixels with a higher score than RX. 108

Figure 40. For a given image, ROC curves were calculated using the hyperspectral CEM score (solid black line), the SPI algorithm fusing the CEM and TAD scores (dotted black line), the SPI algorithm fusing the CEM and RX scores (dashed black line) and the SPI algorithm fusing the CEM and square-rooted RX scores (dashed red line). Re-mapping the RX scores by taking the square root of each pixel’s score appears to give performance similar to using the TAD score..... 109

Figure 41. An RGB image of the synthetic scene was observed with a solar zenith angle of 34° , a sensor zenith angle of 30° and a sensor azimuth angle of 170° from the sun. 110

Figure 42. The S_0 information from the scene was used to test the effect of incorporating high resolution contrast information into the fusion algorithms. 111

Figure 43. Each pixel in the scene was evaluated both by applying RX to just the S_0 band and then by applying RX to all the Stokes bands. 112

Figure 44. Each pixel in the scene was evaluated both by applying RX to just the S_0 band and then by applying RX to both the S_1 and S_2 bands. 113

Figure 45. The polarimetric information from the scene was incorporated as DOLP values in an additional band for the SPOT algorithm. 114

Figure 46. SPOT scores are shown for the multispectral case when using the S_0 information (left) or DOLP (right)..... 115

Figure 47. SPOT scores are shown for the hyperspectral case when using the S_0 information (left) or DOLP (right)..... 115

Figure 48. The multispectral (red) and hyperspectral (black) SPOT ROC curves are shown when derived from the S_0 information (dots) or DOLP (dashes). The CEM scores (solid lines) are included for comparison. 116

Figure 49. The TAD scores calculated by examining the S_0 , S_1 and S_2 bands are shown as an intensity image. Note how many vehicles show up as anomalous bright spots. 117

Figure 50. The RX scores derived from the S_0 band are displayed as an intensity image. Note that the vehicles are less visible in this image. 117

Figure 51. The multispectral (red) and hyperspectral (black) SPI ROC curves are shown when incorporating RX scores derived from the S_0 information (dots) or TAD scores derived from the S_0 , S_1 and S_2 bands (dashes). The CEM scores (solid lines) are included for comparison..... 118

Figure 52. RGB images of sensor-reaching radiance calculated by DIRSIG for the cases described in Table 5. (Left) Image 1, captured near the sun’s specular reflection with a sensor zenith angle of 50° and relative azimuth angle of 180° . (Center) Image 2 is at the same zenith angle as Image 1, but only 90° relative azimuth from the sun. (Right) Image 3 is azimuthally in the specular reflection (180°), but at a sensor zenith angle of only 10° 119

Figure 53. Hyperspectral ROC curves were generated from Image 1 (with the sensor near the sun’s specular reflection) using the CEM (solid line), SPI (dotted line) and SPOT (dashed line) algorithms..... 120

Figure 54. Hyperspectral ROC curves were generated from Image 2 (with the sensor at a 90° azimuth angle from the sun) using the CEM (solid line), SPI (dotted line) and SPOT (dashed line) algorithms..... 121

Figure 55. Hyperspectral ROC curves were generated from Image 3 (with the sensor at a 180° azimuth angle from the sun, but only a 10° zenith angle) using the CEM (solid line), SPI (dotted line) and SPOT (dashed line) algorithms..... 123

Figure 56. Hyperspectral ROC curves were generated from the series of images described in Table 5 using the CEM (solid lines), SPI (dotted line) and SPOT (dashed line) algorithms. Image 1 (black) was the scenario where polarimetric information increased performance the most, and it was also the scenario where the spectral sensor performed the poorest..... 124

Figure 57. Several RGB DIRSIG images are displayed to illustrate the multiple sensor registration process for a solar zenith angle of 20°. (Left) The polarimetric data was drawn from an image captured at a sensor zenith angle of 30° and a sensor azimuth angle of 185°. (Center) The hyperspectral data was drawn from an image captured at a sensor zenith angle of 30° and a sensor azimuth angle of 195°. (Right) The hybrid image contains registered polarimetric and hyperspectral data from multiple sensors. 139

Figure 58. RGB DIRSIG image rendered with a solar zenith angle of 55°, a sensor zenith angle of 40°, and a sensor azimuth angle of 180°. 142

Figure 59. The easiest targets in the scene for the CEM algorithm to find were either well-lit target pixels or targets in a homogenous background. Note that these images

are all shown at the lower polarimetric GSD (0.5 m) for later comparison to the SPI and TAD algorithms. 143

Figure 60. The hardest targets in the scene for the CEM algorithm to find were shadowed targets, targets producing a specular reflection from the sun, and targets located in a heavily cluttered environment. Note that these images are all shown at the lower polarimetric GSD (0.5 m) for later comparison to the SPI and TAD algorithms. ... 144

Figure 61. The easiest targets in the scene for the TAD algorithm to find were well-lit edges of target vehicles. Note that these RGB images are all shown at the polarimetric GSD (0.5 m)..... 145

Figure 62. Shadowed and obscured targets vehicles were among the least anomalous target pixels in the scene. Note that these RGB images are all shown at the polarimetric GSD (0.5 m)..... 145

Figure 63. The SPI algorithm had the most impact on well-lit edges of target vehicles. Note that these RGB images are all shown at the polarimetric GSD (0.5 m). 147

Figure 64. The SPI algorithm had the least impact on shadowed targets or bright targets with a low CEM score. Note that these RGB images are all shown at the polarimetric GSD (0.5 m). 148

Figure 65. An ROI of Figure 57 contains several shadowed vehicles, one of which is a target..... 149

Figure 66. The HS CEM ROC curves were derived from scenes rendered with both the standard fidelity sampling (dashed line) and the higher fidelity spatial and pBRDF sampling (solid line). 150

Figure 67. HS SPI ROC curves were derived from scenes rendered with both the standard fidelity sampling (dashed line) and the higher fidelity spatial and pBRDF sampling (solid line). 151

Figure 68. HS SPOT ROC curves were derived from scenes rendered with both the standard fidelity sampling (dashed line) and the higher fidelity spatial and pBRDF sampling (solid line). 151

Figure 69. (Left) The DOLP metric derived from scenes rendered with the standard fidelity sampling (2x2 oversamples) for this trade study. (Right) The DOLP metric derived from scenes rendered at the higher fidelity sampling (5x5 oversamples, more extensive pBRDF) for this trade study..... 152

Figure 70. (Left) The TAD score derived from scenes rendered with the standard fidelity sampling (2x2 oversamples) for this trade study. (Right) The TAD score derived from scenes rendered at the higher fidelity sampling (5x5 oversamples, more extensive pBRDF) for this trade study..... 153

Figure 71. (Left) The HS CEM score derived from scenes rendered with the standard fidelity sampling (2x2 oversamples) for this trade study. (Right) The HS CEM score derived from scenes rendered at the higher fidelity sampling (5x5 oversamples, more extensive pBRDF) for this trade study..... 154

Figure 72. (Left) DOLP metric for ROI in Figure 64 at SNR = 200. (Right) DOLP metric for the same ROI at SNR = 10,000. In the lower SNR case, the added noise masked any signal from the vehicles. Even in the higher SNR case, the vehicles were barely distinguishable from the background. 155

Figure 73. (Left) S_1 metric for ROI in Figure 64 at SNR = 200. (Right) S_1 metric for the same ROI at SNR = 10,000. In the lower SNR case, the added noise washed out any structure in the signal, while in the higher SNR case, the vehicles and glass windows in the building were distinguishable from the background. 156

Figure 74. (Left) S_2 metric for ROI in Figure 64 at SNR = 200. (Right) S_2 metric for the same ROI at SNR = 10,000. In the lower SNR case, the added noise washed out any structure in the signal, while in the higher SNR case, the vehicles were barely distinguishable from the background..... 156

Figure 75. The percent increase in area under the ROC curve when using the SPI fusion algorithm as opposed to the CEM spectral algorithm for a given off-nadir viewing geometry is shown as a function of sensor zenith angle for solar zenith angles of 20° (green), 34° (red), 55° (blue) and 66° (black). The spectral SNR was set to 200 (top left), 100 (top right), 80 (bottom left) or 50 (bottom right)..... 164

Figure 76. The percent increase in area under the ROC curve when using the SPI algorithm off-nadir as opposed to the CEM algorithm at nadir is shown as a function of sensor zenith angle for solar zenith angles of 20° (green), 34° (red), 55° (blue) and 66° (black). The spectral SNR was set to 200 (top left), 100 (top right), 80 (bottom left) or 50 (bottom right). 165

Figure 77. The area under the ROC curves generated by the off-nadir CEM (spectral data only—solid red line), SPI (spectral and polarimetric data—dotted red line) and TAD (polarimetric data only—dashed black line) algorithms was calculated for a range of spectral GSD values with the polarimetric GSD held constant at 0.5 m. 166

Figure 78. The SCR for the red and green target vehicles were calculated as a function of sensor zenith angle for a fixed sensor azimuth angle of 190° and four solar zenith angles: 66° (top left), 55° (top right), 34° (bottom left) and 20° (bottom right). The SCR was derived from either the spectral reflectance values (solid lines) or polarimetric TAD scores (dashed lines)..... 167

Figure 79. The percent increase in area under the ROC curve for the red target (left) and the green target (right) when using the SPI algorithm off-nadir as opposed to the CEM algorithm off-nadir is shown as a function of sensor zenith angle for solar zenith angles of 20° (green), 34° (red), 55° (blue) and 66° (black). The spectral and polarimetric SNR values were set to 200. 169

Figure 80. The ROI for the multispectral / polarimetric data fusion experiment is outlined in red. This image shows the conditions for the morning data collect, with a solar zenith angle of 26° , a sensor zenith angle of 48° , and a relative azimuth angle between the sensor and the sun of 160° 173

Figure 81. (Left) The morning data collect represented favorable polarimetric sensing conditions. All vehicles in the scene were sunlit. (Right) The afternoon data collect represented poor polarimetric sensing conditions. Additionally, many of the target vehicles were shadowed. 174

Figure 82. The vehicles staged throughout the scene were 1:32 die cast metal model cars. 174

Figure 83. (Left) Registered RGB WASP-Lite bands captured during the morning scenario. (Right) Registered RGB WASP-Lite bands captured during the afternoon scenario. 176

Figure 84. (Left) Processed S_0 image derived from the morning scenario. (Right) Processed S_0 image derived from the afternoon scenario. 177

Figure 85. (Left) RGB image of multispectral data registered to the polarimetric imagery derived from the morning scenario. (Right) RGB image of multispectral data registered to the polarimetric imagery derived from the afternoon scenario.... 177

Figure 86. (Left) The green band of the cropped fused morning data subset is shown as an intensity image after converting to reflectance values. (Right) The green band of the fused afternoon data subset is shown as an intensity image after converting to reflectance values..... 178

Figure 87. The MS CEM score maps are shown as intensity images, with brighter pixels being ranked as more target-like. (Left) Morning (Right) Afternoon 181

Figure 88. The DOLP values for each pixel are shown as intensity images, with brighter pixels being ranked as more polarizing. (Left) Morning (Right) Afternoon 182

Figure 89. The TAD scores for each pixel are shown as intensity images, with brighter pixels being ranked as more anomalous. (Left) Morning (Right) Afternoon 183

Figure 90. The SPI decision fusion algorithm scores for each pixel are shown as intensity images, with brighter pixels being ranked as more target-like. Red boxes indicate regions containing the vehicles in the scene. (Left) Morning (Right) Afternoon ... 184

Figure 91. The SPOT pixel fusion algorithm scores for each pixel are shown as intensity images, with brighter pixels being ranked as more target-like. Red boxes indicate regions containing the vehicles in the scene. (Left) Morning (Right) Afternoon ... 185

Figure 92. The CEM, SPI and SPOT ROC curves were derived from the morning scenario. 187

Figure 93. The CEM, SPI and SPOT ROC curves were derived from the afternoon scenario. 188

Figure 94. One sensor platform could potentially capture multiple polarimetric (PI) images of a scene in addition to a multispectral (MS) image. Although the images are obtained from different sensor zenith angles $SZ1$, $SZ2$ and $SZ3$, the aircraft altitude is held constant. 191

Figure 95. The S_1 and S_2 values for each pixel in the scene, as derived from sensor zenith (SZ) angles of 10, 20, 30, 40, 50, 60 and 70 degrees, were combined into a 14 band stack. The TAD algorithm was applied to the stacked data cube to identify anomalous pixels, and the $T_{10...70}$ vector was constructed by collecting the TAD scores for each target pixel (T) in the image, while background pixels (B) were ignored. 196

Figure 96. The S_1 and S_2 values for each pixel in the scene, as derived from one particular sensor zenith (SZ) angle, i , of interest were combined into a 2 band stack. The TAD algorithm was applied to the stacked data cube to identify anomalous pixels, and the T_i vector was constructed by collecting the TAD scores for each target pixel (T) in the image, while background pixels (B) were ignored. 196

Figure 97. The areas under the multispectral ROC curves for the CEM, SPI, SPI_{Full}, SPI_{Top2} and SPI_{Top1} algorithms were calculated for both the red and green target vehicles. Acquiring multiple polarimetric images generally improved target detection performance, but most of the impact could be attained by logically choosing the single best polarimetric image. 202

Figure 98. The areas under the hyperspectral ROC curves for the CEM, SPI, SPI_{Full}, SPI_{Top2} and SPI_{Top1} algorithms were calculated for both the red and green target vehicles. Acquiring multiple polarimetric images generally improved target detection performance for the red target, while incorporating polarimetric information generally decreased performance for the green target. In all cases, choosing the best two polarimetric images produced essentially the same impact as using the complete polarimetric image set. 204

Figure 99. The areas under the multispectral ROC curves were calculated for the SPI_{Top2} fusion algorithm, with a varying degree of registration error in the polarimetric images, and compared to the purely spectral CEM algorithm performance for both the red and green target. Incorporating registration error produced some decrease in fusion performance, but still outperformed simply relying on spectral information. 206

Figure 100. The areas under the hyperspectral ROC curves were calculated for the SPI_{Top2} fusion algorithm, with a varying degree of registration error in the polarimetric images, and compared to the purely spectral CEM algorithm performance for both the red and green target. Incorporating registration error

produced some decrease in fusion performance, but still outperformed simply relying
on spectral information. 207

1 Introduction

In August 2008, Dr. Pete Rustan, then director of the Ground Enterprise Directorate of the National Reconnaissance Office, identified high value target location and tracking as one of the fundamental problems facing the intelligence community [Rustan 2008]. Researchers with the Air Force Research Laboratory have stressed the importance of detecting targets hidden in natural and man-made clutter, and described efforts to exploit hyperspectral imaging capabilities for automated target recognition [Eismann 2006]. However, the corresponding false alarm rates are often higher than desired, especially for small targets in urban areas with a large amount of man-made clutter.

Faced with the possibility of acquiring a new sensor system, the Defense Acquisition Guidebook directs government acquisition personnel to consider cost as an independent variable [DAG 2010]. Individual system capability requirements are defined via a range spanning from thresholds, which must be met, to objectives, which represent the ideal capability. The emphasis on controlling costs often means that the best possible system is not actually built—rather, the government trades extra capability in one area to meet requirements in others in order to stay within the budget.

In general, trade studies are used to assess the impact of varying capabilities within the requirement range. As one way to potentially achieve a reduction in spectral imaging false alarm rates, it is essential to explore how combining spectral imaging capability with a second sensing modality will affect target detection capability in an urban environment. The resulting multimodal sensor system might become effective over a wider range of imaging conditions, more effective under ideal conditions, or some

combination of both. In addition, conditions may be identified where no increase in performance is observed; potentially preventing an expensive lesson learned only after the final system has been deployed.

Several existing commercial space-based systems incorporate multispectral imaging capability, so the potential exists to task a sensor in that class for an urban target detection scenario. However, the coarse spectral resolution of these bands limits the ability to make fine distinctions between targets and background with reasonably similar spectra, contributing to the false alarm rate. Adding a second sensor modality to a multispectral sensor may significantly improve target detection performance by weeding out some false alarms. Although performance may improve, the final multimodal system may still not meet the detection accuracy requirements.

Alternatively, a hyperspectral sensor with a dramatically larger number of spectral bands, as opposed to the four to ten in a common multispectral sensor, could capture more spectral data for use in a target detection algorithm. As a result, the initial number of false alarms is likely to be reduced, while no doubt driving a corresponding increase in cost. When a second imaging modality is included in the system, the potential again exists for an increase in performance. Further, because of the increased spectral resolution, it is expected that the hyperspectral system's target detection performance will be superior to the multispectral system under similar viewing conditions. Once the potential performance increase is known, the user can decide whether the increase in cost is justified.

The outcomes of this study provide a series of contributions to the field of remote sensing. Two novel fusion algorithms were developed by leveraging the strengths of

existing tools to fuse information at either the pixel or decision level. First, a new pixel-level fusion algorithm was demonstrated to incorporate multiple sensing modalities for target detection applications. The pixel fusion algorithm combines polarimetric information with multispectral or hyperspectral measurements. Second, a generalized decision-level fusion algorithm was showcased. The decision fusion algorithm combines the outputs from different sensing modalities after a different target detection algorithm has been applied to each data set. This dissertation breaks new ground by merging relevant theory in the fields of spectral and polarimetric remote sensing for target detection applications.

The foundation of this work was defining a process to quantify the effect of incorporating both spectral and polarimetric information into an automated target detection scenario via simulation with the Digital Imaging and Remote Sensing Image Generation (DIRSIG) model [Ientilucci and Brown 2003]. Specifically, the effect of first fusing data from a polarimetric sensor with data from a multispectral sensor, then fusing the same data from a polarimetric sensor with data from a hyperspectral sensor were quantified. A scenario with an urban environment was examined, and the available trade space was mapped as a function of the viewing geometry, illumination conditions and signal-to-noise ratios (SNR) of the chosen sensors.

Once the initial results demonstrated that incorporating additional polarimetric information may enable suitable performance with a less capable spectral sensor, a series of trade studies was carried out to assess how varying the spectral SNR, spectral ground sample distance (GSD), or target spectrum affected the impact of spectral and polarimetric data fusion via the spectral / polarimetric integration (SPI) algorithm for a

notional multimodal sensor. A field experiment was designed to further exploit the DIRSIG simulation results and apply the fusion algorithms to real data by constructing a small-scale scene of model cars and bits of urban clutter.

Finally, emerging research at RIT suggested yet another avenue to explore with the spectral/polarimetric fusion concept. Therefore, a process was defined to model a particular set of image acquisition scenarios, to determine which polarimetric image (from a set of many) will produce the most impact on target detection performance, to quantify the impact of incorporating polarimetric information from multiple viewing geometries and to evaluate the performance degradation introduced by a reasonable degree of registration error.

In summary, perhaps the most important contribution to the remote sensing community from this dissertation was the demonstration of a generalized approach to performing trade space evaluations via synthetic image generation tools. Simulation is an attractive alternative for cutting-edge systems because of the sheer volume of data required for a reasonable trade study, especially when multimodal sensors are considered. However, the available trade space is enormous, and compromises must be made in terms of what points are sampled because of finite computational resources. A generalized trade space evaluation tool therefore helps to quickly bound the problem by sampling the most extreme cases, and areas in between can then be sampled in ever finer resolution to hone in on any dramatic changes in performance. System designers can thus tailor the target and sensor parameters to their particular scenario of interest and determine how to best allocate their resources. Further, users of existing joint multispectral/hyperspectral

and polarimetric sensors are presented with a method to determine the optimum tasking conditions for their hybrid system.

2 Project Overview

2.1 Research Questions

The foundation of this research was an investigation into the effect of fusing existing multispectral or hyperspectral sensor data with polarimetric information from a second sensor in an urban target detection application. This project assumed a multispectral or hyperspectral sensor with fixed capabilities had been designed and combined a polarimetric sensor with each of the above mentioned sensors to assess the benefits from a multimodal system. A performance improvement was defined as an enhanced probability of detection at a fixed probability of a false alarm when the receiver operating characteristic (ROC) curve for the multimodal system was compared to the ROC curve for only the given multispectral or hyperspectral system under the same conditions. For the initial part of this study, two potential scenarios were examined.

In the first case, the polarimetric sensor was co-located with the multispectral or hyperspectral sensor, such that the viewing geometries for both sensors were identical. I tested whether fusing data from the two sensors improved automated target detection performance for a range of viewing geometries, illumination conditions and reasonable sensor SNR values.

In the second case, the polarimetric sensor was on a different platform from the multispectral or hyperspectral sensor, with each sensor optimally positioned for a particular set of illumination conditions. With this scenario, I established the maximum improvement in performance that could be expected from the given multimodal system across a range of reasonable SNR values for each instrument.

2.2 Objectives

There were a number of steps required to complete the study at the foundation of this project, and they are outlined below.

1. **Define a relevant scenario.** The comments in Sec. 1 indicate a useful analysis of sensor performance consists of attempting to detect a specific type of vehicle target in an urban environment under varying illumination and viewing conditions, with the targets located in varying degrees of concealment and sensors having a range of SNR values. Because of the vast number of combinations in this trade space, and lack of available hardware or actual simultaneous data, simulation presents a practical method to explore the options.
2. **Select an effective combination of sensor modalities.** Because some of the bands in a hyperspectral system are highly correlated, a point can be reached where little information is added by considering more data [Prasad and Bruce 2008]. Along these same lines, Petrakos et al (2001) found that maximizing the increase in accuracy of a combined classifier depends on fusing classifiers that often do not agree—ideally, they would be uncorrelated with each other, yet correlated with the target. Therefore, a study was carried out to demonstrate that classifiers exploiting the chosen sensing modalities were relatively uncorrelated for the given target detection scenario.
3. **Design a model scene.** To plausibly explore the trade space, a synthetic image generation model was required to render a radiometrically correct urban scene.

The synthetic scene must be geometrically plausible, with the material characteristics in the scene characterized to a high degree of spatial, spectral, and polarimetric fidelity. The model must then be capable of reasonably propagating radiation from any sources through the atmosphere and correctly portraying any interactions with objects in the scene. Finally, the model must be able to accurately simulate the effect of a variety of sensor characteristics on the image.

4. **Render multiple versions of the scene, varying a range of relevant parameters.** Schott (2007) explained that the sensor reaching radiance depends, among other things, on the illumination source's zenith and azimuth angles, the sensor's zenith and azimuth angles, and the shape factor of the target. In addition to effects from the overall sensor reaching radiance, Schott (2009) noted the ability to distinguish the target from the background is also highly sensor dependent, depending in part on SNR for most systems and the degree of polarization (DOP) for any polarimetric sensor. Therefore, a range of potentially useful values for each of these variables was identified, and the simulation executed for each case to isolate the effect of changing one variable at a time.
5. **Compensate the spectral data for atmospheric effects.** The simulated radiance reaching the sensor is a function of both the scene characteristics and any modeled atmospheric effects. In an actual target detection application based on information from spectral libraries, it is reasonable to expect a user to perform some type of atmospheric compensation, but the focus of this project was not to determine which atmospheric compensation technique is best. Therefore, considering the vast number of different simulated scenes to be analyzed, a

reasonable atmospheric compensation technique was applied with an emphasis on simplicity and speed.

6. **Simulate some degree of registration error between the two modalities.**

Although the synthetic imagery for a given version of the scene consisted of perfectly registered bands for both sensing modalities, exact registration between the two modalities is unlikely in a practical application. Therefore, a small amount of offset was induced to account for residual errors generated from applying a registration routine to actual data.

7. **Fuse co-located sensor outputs in a meaningful manner.** After establishing the

baseline sensor performance, the data captured from the two sensors was combined via a fusion algorithm to enhance the probability of target detection. The type of data to be combined was identified, and then a methodology was constructed to analyze the result and exploit the unique contributions of each sensor.

8. **Quantify any impact on target detection.** When the fusion product had been assembled, improvement was defined in terms of an increased probability of target detection over the data for a fixed probability of a false alarm. In essence, the ROC curve for the fusion product was compared to the ROC curve of the original sensor for each of the cases in Step 5, and any areas of performance improvement or detriment were noted. A crucial part of this analysis was having enough pixels to quantify low enough false alarm rates for a practical application, so a balance was struck between the extent of the modeled scene and the required computational effort.

9. **Identify viewing geometries expected to maximize target detection improvement for dispersed sensors.** Since the different sensor modalities will likely have different measures of merit for any target detection algorithm employed, the ideal viewing geometries for each sensor are unlikely to be identical. The data produced using each individual modality was investigated, and compared to relevant theory, to identify each sensor's optimum viewing geometry.
10. **Register the images generated by each sensor.** Before data fusion can occur, the data were registered into a common space. A process for aligning common scene features in images from each sensor was determined and executed for the limited number of viewing geometries identified in Step 9.
11. **Establish the maximum attainable target detection improvement with the decision level fusion algorithm for dispersed sensors.** The fusion algorithm applied in Step 7 should produce the best attainable probability of target detection at a fixed false alarm rate with data obtained from each sensor when positioned in its optimum viewing geometry. Therefore, the performance of the fused sensor systems was compared with the original sensor's performance, in its optimum viewing geometry, to again quantify via ROC curves the maximum attainable improvement in performance.

3 Background

3.1 Section Overview

This section describes previously published research and demonstrated operational capabilities to examine the effect of multimodal data fusion in a target detection scenario. Before developing an experimental method, the current state of the art is assessed by examining existing sensor systems, understanding the different data fusion levels, and reviewing established decision level fusion techniques. Further, a brief overview of existing sensing modalities is conducted, with an emphasis on the emerging field of polarimetric remote sensing. After a recap of previously published multimodal sensor fusion efforts, the capabilities of several synthetic image generation models are examined. Finally, several existing atmospheric compensation methods and target detection algorithms are assessed for potential use in this project.

3.2 Existing Systems

Several commercial space-based operational sensors are currently capable of high spatial resolution panchromatic imaging, with some also capturing a modest number of spectral channels with a significantly lower ground sample distance (GSD). In particular, the satellites operated by DigitalGlobe and GeoEye are among the most recognized and illustrate the current commercial space-based state of the art.

The IKONOS satellite, launched in 1999 and currently operated by GeoEye, was among the first commercial satellites to permit public access to high spatial resolution imagery. In addition to providing a 0.82 m panchromatic GSD at nadir, the sensor also has four multispectral bands—red, green, blue, and near-IR—with a GSD of 4 m

[GeoEye 2009]. Additionally, the sensor can capture images up to 60° off nadir, potentially imaging the same target from several different points of view.

Further pushing the envelope, DigitalGlobe's QuickBird satellite, launched in 2001, is capable of 60 cm panchromatic GSD and 2.4 m multispectral GSD at nadir [DigitalGlobe 2009]. Although QuickBird's multispectral capability is similar to IKONOS's, with red, green, blue, and near-IR channels, QuickBird is different because it can alter its pointing direction in both the across-track and along-track direction. However, the satellite's maximum slew in the across-track direction is only $\pm 30^\circ$ off nadir. To meet the growing public demand for high spatial resolution imagery, and reduce the image collection burden on QuickBird, DigitalGlobe launched the WorldView-1 satellite in 2007. Although the sensor only has a panchromatic band, it is capable of a 50 cm GSD at nadir with the ability to slew $\pm 45^\circ$ off nadir, with higher angles being selectively available [Digital Globe 2009].

The next generation of commercial satellites sought to combine a high spatial resolution, multispectral imaging capability and significant spacecraft agility. The GeoEye-1 satellite was launched in 2008, boasting a panchromatic GSD of 0.41 m at nadir and the traditional red, green, blue and near-IR multispectral bands with a GSD of 1.65 m at nadir. Most notably, the satellite has the incredible ability to rotate or swivel forward, backward, or side-to-side—essentially allowing imaging in any direction, at different times of day [GeoEye 2009]. DigitalGlobe's WorldView-2 satellite, launched in late 2009, is described as having similar spatial resolutions with a panchromatic GSD of 0.46 m and a multispectral GSD of 1.84 m at nadir [DigitalGlobe 2009]. However, the sensor has a unique combination of spectral bands: the typical red, green, blue, and near-

IR bands are included, along with additional red edge, yellow, coastal, and near-IR2 bands. The four additional multispectral bands are planned to provide enhanced information for studying regions near water. As with WorldView-1, the satellite can slew $\pm 45^\circ$ off nadir with higher angles selectively available. Table 1 summarizes the key capabilities of operational commercial high spatial resolution imaging satellites.

Table 1. Summary of key capabilities for operational high spatial resolution commercial imaging satellites, as advertised by Digital Globe and GeoEye.

Satellite	GSD (Panchromatic / Multispectral)	Number of Spectral Channels	Max Slew Angle (Across-Track)	Along-Track Pointing Capability?
IKONOS	0.82 m / 4 m	4	60°	No
QuickBird	0.60 m / 2.4 m	4	30°	Yes
WorldView-1	0.50 m / ---	1	45°	No
GeoEye-1	0.41 m / 1.65 m	4	Any angle	Yes
WorldView-2	0.46 / 1.84 m	8	45°	No

In contrast to sensors with high spatial resolution and a few spectral bands, other operational systems, dubbed hyperspectral sensors, have demonstrated the ability to acquire spectral information over a significantly larger number of bands. When the incoming signal is divided into such fine spectral bands, a trade must occur between spatial, spectral and radiometric resolution [Schott 2007]. To compensate for the decreased number of signal photons in each narrow spectral band, the GSD must be increased or a lower instrument signal-to-noise ratio (SNR) must be accepted. Two of the most widely known hyperspectral sensors are NASA's Airborne Visible Infrared Imaging Spectrometer (AVIRIS) and the Hyperion instrument onboard the EO-1 satellite.

AVIRIS's 224 spectral bands span the region from 0.4 μm to 2.5 μm in 10 nm increments, achieving a spectral resolution much higher than that of the previously described satellites. Since the instrument can be flown on different aircraft, the GSD depends on the chosen altitude. Kruse *et al.* (2003) assessed AVIRIS's capabilities to quantify the instrument's SNR, and calculated an SNR around 150 for most channels for a 50% reflector. Encouraged by results obtained from AVIRIS, NASA launched the EO-1 satellite in 2000, using the Hyperion instrument to demonstrate high spectral resolution imaging from space. After the one year technology demo was complete, the U.S. Geological Survey (USGS) assumed operational control of the satellite. The Hyperion instrument's capabilities are quite similar to AVIRIS's, resolving 220 spectral bands and spanning the region from 0.4 μm to 2.5 μm in 10 nm increments with a 30 m GSD [USGS 2009]. Additionally, the spacecraft can point $\pm 20^\circ$ off nadir to acquire imagery. Boeing's Hyperion validation report highlighted that the sensor SNR for a 30% reflector depended on the spectral region of interest: 140-190 in the VNIR, 96 near the center of the SWIR (~ 1225 nm), but only 38 near the long edge of the SWIR (~ 2125 nm) [Pearlman 2003].

In addition to space-based systems driving toward high spatial or spectral resolution, the US military has emphasized flexibility in remote sensing applications, effectively employing unmanned aerial vehicles (UAVs) like the Global Hawk and Predator. The Global Hawk is designed to loiter at an altitude of 65,000 ft and is equipped with electro-optical (EO), infrared (IR) and synthetic aperture radar (SAR) sensors [Leachtenauer and Driggers 2001]. The EO sensor response spans the 0.4-0.8 μm range and the IR sensor is sensitive across the 3.6-5.0 μm range. Further, the EO sensor

is specified to provide a visible National Image Interpretability Rating Scale (NIIRS) value of 6.5 at a 45° elevation angle (28 km slant range) and the IR sensor is specified to provide an infrared NIIRS of 5.5 at a 45° elevation angle (28-km slant range). Schott (2007) calculated that a 10-inch GSD roughly corresponded to a NIIRS of 6.9, so the visible GSD specification of the Global Hawk could be about 0.3 m for 45° off nadir, and should improve as the viewing angle approached nadir.

According to the official US Air Force factsheet, the Predator system is a remotely piloted vehicle with an operational ceiling of 7620 m equipped with a variable-aperture TV camera and a variable-aperture IR camera. Specifically, the EO camera has a 10:1 zoom, with the focal length varying from 16-160 mm and the FOV varying from 2.3 by 1.7 degrees to 23 by 17 degrees [Leachtenauer and Driggers 2001]. The GSD in the visible range therefore depends on the zoom setting and aircraft altitude, providing a significant degree of flexibility.

The above examples of existing systems demonstrate what levels of spatial, spectral and signal resolution are currently attainable. Ideally, a new system would seek to maximize all three, but practically, designers must make trades within the three spaces to meet mission requirements. This dissertation attempts to identify any regions within the trade space where an increase in capability might provide a dramatically improved target detection performance in an urban environment. One way to improve target detection capability might be to combine data from different sensors, a process known as data fusion.

3.3 Data Fusion Levels

When approaching the issue of fusing data from multiple sources, the first task is to determine the level of fusion that will occur. Essentially, one must determine the point where information from the first source will first interact with information from the second source. The three commonly accepted levels of data fusion in remote sensing applications are the pixel level, feature level and decision level [Pohl *et al.* 1998]. Figure 1 shows how the three fusion levels are related to each other.

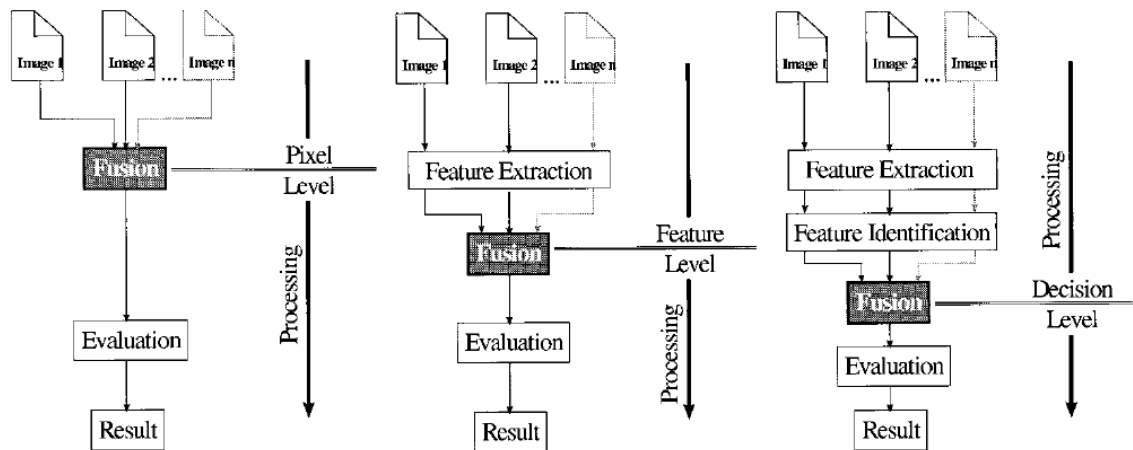


Figure 1. Potential levels of data fusion and their relation to any image processing steps. Image courtesy of Pohl *et al.*, 1998.

Pixel level data fusion describes combining information from each source at the lowest possible processing level, in effect merging different measurements of the same physical parameter. One example of this type of data fusion is the spatial resolution enhancement of multispectral data using higher spatial resolution panchromatic information [Price 1987]. Given a lower GSD multispectral super pixel, the signal strength of the smaller GSD panchromatic pixels can be used to infer what the spectral characteristics of the smaller pixels might have been. The resulting product is a multispectral image with a GSD equivalent to the original higher spatial resolution

panchromatic image. Although the two images must be very well geometrically registered to avoid artifacts and misinterpretation, pixel level fusion preserves the most original data [Pohl *et al.* 1998].

Feature level data fusion implies some processing has already taken place to identify objects consisting of multiple pixels. Pohl *et al.* (1998) described a segmentation application where individual buildings were extracted from an urban scene by exploiting extent, shape and neighborhood, citing a dissertation written in French [Mangolini 1994]. Further, the corresponding structures used in feature level fusion tend to relax the geometric accuracy required, but some information is lost during the feature extraction process.

Decision level fusion consists of individually processing images, then combining the results in some way to reinforce common interpretations and resolve discrepancies. The final product thus provides a better understanding of the scene. Petrakos *et al.* (2001) cautioned that the correlation between classifiers used in decision level fusion can limit the potential accuracy increase. In brief, if two classifiers always agree or always disagree, no new information is gained by combining their outputs. Table 2 demonstrates how to determine the maximum potential increase in effectiveness obtained by fusing outputs from two different classifiers.

Table 2. Table of dichotomous outcome for two classifiers, based on Petrakos *et al.* (2001)

<u>Classifier Comparison</u>	Classifier 2 – Correct classification	Classifier 2 – Incorrect classification
Classifier 1 – Correct classification	Region 1	Region 2
Classifier 1 – Incorrect classification	Region 3	Region 4

Region 1 represents the pixels that will never be misclassified in a combination. If the number of pixels in Region 1 is divided by the total number of pixels, the result is the lower limit of the classification accuracy of any combined classification scheme. Region 4 represents the pixels that will never be classified correctly by a given combination, and dividing the number of pixels in Region 4 by the number of total pixels provides an upper bound on the fused classification accuracy. Regions 2 and 3 are therefore a relative measure of classifier effectiveness, and illustrate the potential performance increase attainable with a data fusion algorithm.

This section has divided data fusion efforts into three levels, based on the amount of processing that occurs before fusion. At the pixel level, raw measurements of the same parameter are combined in some fashion, making the result highly dependent on accurate registration. Fusing data at the feature level involves identifying objects through their shape, extent, or neighborhood, slightly reducing the dependence on accurate registration but preserving less of the original information. Finally, fusing data at the decision level consists of processing sensor outputs independently, and often combining metrics like a classification score rather than measurements of actual data. Since data from ill-correlated sensors can be combined via decision level fusion, a great opportunity exists to exploit decision level fusion in multimodal target detection scenarios. Therefore, specific decision level fusion techniques will be investigated.

3.4 Decision Level Fusion Techniques

Decision level fusion techniques are based either on mathematical manipulations of some score metric from different sensors, or some form of voting based on the opinion

of all sensors involved. The most common mathematical manipulations are a linear opinion pool (LOP), logarithmic opinion pool (LOGP) and a pseudo-inverse solution. In contrast, the much simpler voting schemes are based either on some form of majority rule or using the result with the highest confidence level.

The linear opinion pool (LOP) has the form shown in Eq. 1, where the group probability of belonging to class w_j is given by $C_j(X)$ if n data sources are used

$$C_j(X) = \sum_{i=1}^n \lambda_i p(w_j | x_i) \quad (1)$$

where $X = [x_1, \dots, x_n]$ is an input data vector, where each x_i is a source-specific pattern which is multidimensional if the data source is multidimensional, $p(w_j | x_i)$ is a source specific posterior probability and λ_i ($i = 1, \dots, n$) are source-specific weights which control the relative influence of the data sources [Petrakos *et al.* 2001]. One weakness of the LOP is that it shows dictatorship when Bayes' theorem is applied and it is not externally Bayesian [Benediktsson and Swain 1992]. The second technique, the logarithmic opinion pool (LOGP), has the form shown in Eq. 2, where the group probability of belonging to class w_j is given by $L_j(X)$ if n data sources are used [Petrakos *et al.* 2001]:

$$L_j(X) = \prod_{i=1}^n p(w_j | x_i)^{\lambda_i} \quad (2)$$

The LOGP differs from the LOP in that it is unimodal and less dispersed, while also treating the data sources independently [Benediktsson *et al.* 1999]. Additionally, a zero vote from any source is an automatic veto. In either case, the weights should be high when the data sources are expected to contribute to higher accuracy, and low when the

additional data provides little added value. Finally, a pseudo-inverse solution can be obtained by solving the problem with a linear regression model, where the inputs are the spatial and spectral class-conditional probabilities and the outputs are the observed data values [Benediktsson *et al.* 2004].

Decision level fusion techniques based on voting have been demonstrated in hyperspectral classification applications. Perhaps the simplest method is to apply the classification result from the sensor with the maximum degree of confidence in its classification result [Benediktsson *et al.* 2004]. A more sophisticated approach, qualified majority voting (QMV), is designed to better exploit classifiers of disparate expertise. The QMV fusion algorithm allows each classifier to influence the decision, but varies the impact each classifier has on the final decision by modifying the vote by some weighting factor [Cheriyadat *et al.* (2003)]. The final decision is then the outcome with the most votes. In practice, the weighting factor is based on the degree of confidence the experimenter has in each classifier.

Although several different decision level fusion schemes have been described, the ideal fusion method is highly application and data dependent [Pohl *et al.* 1998]. In particular, the issues of differing classifier expertise and potential veto power are especially important when considering potential fusion techniques. In a multimodal sensor target detection application, if the sensors measure ill-correlated physical phenomena under the same illumination conditions, it is likely that more confidence will be placed in the results from one sensor than from another, and the LOP, LOGP and QMV methods account for this. However, in a LOGP, if one sensor is sure that a given pixel is not a target, the pixel is ignored regardless of the other sensors' opinions. The

veto effect could be good or bad, but the issue of how to address veto power must be considered in any decision level fusion scheme. In general, the way veto power is handled depends on the sensing modalities used in a particular scenario.

3.5 Potential Sensing Modalities

Several different sensing modalities could be exploited for the urban target detection problem, and each has potential strengths and weaknesses that must be considered. The modalities potentially applicable for use in this work are synthetic aperture radar (SAR), panchromatic imaging, multispectral imaging, hyperspectral imaging, polarimetric imaging, and light detection and ranging (LIDAR).

SAR data has been previously analyzed via simulation with automatic target detection algorithms [Douglas *et al.* 2004]. A significant amount of data is required to accurately map the trade space associated with incorporating a second imaging modality. This study therefore requires either an existing cache of SAR and spectral urban data, obtained simultaneously, for a variety of viewing geometries and illumination conditions. Alternatively, the data could be generated through rigorous simulation of a notional scene as observed by both modalities under the same variety of viewing conditions. However, since no such cache of simultaneous SAR and spectral urban data is freely available for analysis, and since RIT doesn't currently have access to a rigorous simulation program capable of integrating SAR with other optical modalities, SAR must be ruled out for consideration in this study.

Panchromatic imaging is often an attractive sensing modality, as the existing satellites described in Sec. 3.2 show. Since photons are integrated over a wide spectral

range, the instrument SNR can be quite high. Alternatively, the increase in signal photons captured at the focal plane means the detector size, and therefore the GSD, can be much smaller while still maintaining the SNR of a similar multispectral system. Although a small GSD panchromatic image allows an analyst to more easily identify features, panchromatic imagery is of limited use in automated target detection scenarios. The nature of panchromatic imagery means only one image is available for processing, and target detection would probably occur by using a spatial matched filter. In the urban target detection scenario addressed by this project, the use of a spatial matched filter becomes problematic for several reasons. First, because of the variety of different viewing geometries, matched filters must be generated for all possible orientations of the target vehicle. Next, the background is expected to be quite cluttered, meaning the target will likely be partially obscured by trees, buildings, or other vehicles, reducing the matched filter's effectiveness. Finally, if the sensor is positioned at a fixed altitude but is agile enough to look significantly off nadir, the range to the target and therefore scale on the ground can change significantly. The variety of scaled replicas of the matched filter required for every possible vehicle orientation would dramatically increase the computational requirements for this scenario, making it an unrealistic option.

Multispectral imaging, referring to data captured with up to ten spectral bands, is especially attractive for this project because of the relatively widespread availability of these sensors. Multispectral systems range in complexity from the advanced satellites outlined in Sec. 3.2 to airborne systems like the WASP-Lite (2009), flown in a Cessna by RIT's Laboratory for Imaging Algorithms and Systems. In addition, incorporating a few bands of spectral information means images can be analyzed by either spatial or spectral

techniques, or a combination of both, for a target detection application. Although splitting the incoming signal spectrally impacts the instrument SNR, a reasonable GSD can still be maintained. One can easily envision a multispectral system mounted on a UAV orbiting an urban environment searching for targets, making this a critical modality to consider when mapping the available trade space.

Hyperspectral systems, with tens or hundreds of spectral bands, acquire a massive amount of data. Schott (2007) summarizes many effective existing hyperspectral target detection algorithms, but those algorithms which increase performance within an imaging modality tend to require more processing time or human involvement, or both. Additionally, new articles frequently appear in the literature detailing better hyperspectral target detection and classification algorithms [Fauvel *et al.* 2008] [Huang and Zhang 2008] [Prasad and Bruce 2008]. Since it seems that no consensus has occurred within the community with regards to a universally “best” algorithm, the challenge is to find a reasonably effective and efficient target detection algorithm to model the baseline system performance for this scenario.

Polarimetric remote sensing in the visible and thermal infrared is a relatively new and largely undeveloped field [Schott 2009]. Anomaly detection algorithms have achieved some degree of success with polarimetric data in separating man-made materials from the natural background [Cavanaugh *et al.* 2006]. For the urban target detection scenario, a significant portion of the background is man-made material, so the anomaly detection capability may be reduced. Further, polarimetric sensing is highly sensitive to illumination and viewing geometries [Devaraj *et al.* 2007]. However, since polarimetric anomaly detection algorithms exploit different phenomenon than spectral

algorithms, a polarimetric remote sensor could potentially complement a multispectral or hyperspectral sensor for urban target detection applications.

As an alternative to passive systems that are affected significantly by solar illumination conditions, active LIDAR systems have also been used in seaborne target detection applications [Van den Heuvel *et al.* 2008]. Although range information can be extracted from reflective scene elements, identifying specific targets in a cluttered urban environment seems to be a computationally intensive task subject to some of the same limitations as panchromatic imaging. If done correctly, LIDAR data could be used to aid registration of the polarimetric data with the spectral data [Lach *et al.* 2009]. Further, LIDAR data has been fused with aerial imagery to extract individual building footprints, with a GSD of about 1 m, but several hundred pulses were needed per pixel [Zabuawala *et al.* 2009]. However, it may be a challenge with LIDAR to obtain this resolution over a notional 1 km² urban environment, in a timely fashion, without being apparent to observers below.

In summary, SAR, panchromatic EO, multispectral, hyperspectral, polarimetric and LIDAR systems could all potentially be combined in some fashion as part of a decision level data fusion experiment. Although most of these modalities have been extensively employed in the remote sensing field, polarimetric sensing is an emerging capability. Before investigating previous fusion demonstrations, a high-level overview of polarimetric remote sensing will be conducted.

3.6 Polarimetric Remote Sensing

A full explanation of the field of polarimetric remote sensing would require an entire text, so this section is intended to provide a basic overview on how to manipulate measured polarimetric information for later exploitation. Polarimetric information is derived from a combination of the signal obtained by viewing the scene through various linearly polarized filters, and typically expressed as a vector of Stokes parameters based on irradiance [Schott 2009]. However, the direction of the polarization in the sensor reaching radiance can be expressed through the unnormalized Stokes parameters, \hat{S} , as shown in Eqs. 3 - 6 below:

$$\hat{S}_0 = L_H + L_V \quad (3)$$

$$\hat{S}_1 = L_H - L_V \quad (4)$$

$$\hat{S}_2 = L_{+45} - L_{-45} \quad (5)$$

$$\hat{S}_3 = L_R - L_L \quad (6)$$

where \hat{S}_0 is the total incident radiance, \hat{S}_1 is related to horizontal polarization, \hat{S}_2 is related to the polarization at 45° , and \hat{S}_3 is related to the circular polarization of the incident radiance. The \hat{S}_1 , \hat{S}_2 and \hat{S}_3 values can be either positive or negative, depending on which type of polarization dominates. In Eqs. 3-6, L represents the sensor reaching radiance while the subscript denotes the use of a polarizing filter oriented to transmit only radiance polarized horizontally (H), vertically (V), diagonally ($+45$ or -45) or circularly (R or L).

Another way to express the polarization state of the incoming radiation is through the normalized Stokes vector, \mathbf{S} , arranged as shown in Eq. 7 [Schott 2009]

$$\mathbf{S} = \begin{bmatrix} \hat{S}_0 / \hat{S}_0 \\ \hat{S}_1 / \hat{S}_0 \\ \hat{S}_2 / \hat{S}_0 \\ \hat{S}_3 / \hat{S}_0 \end{bmatrix} = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 1 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} \quad (7)$$

The degree of polarization (DOP) and the degree of linear polarization (DOLP) in the incoming radiance are calculated via Eqs. 8 and 9.

$$DOP = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \quad (8)$$

$$DOLP = \frac{\sqrt{S_1^2 + S_2^2}}{S_0} \quad (9)$$

In practice, $S_3 \approx 0$ for most passive sensing applications, so $DOP \approx DOLP$ [Schott 2009].

Because the polarimetric components of the Stokes beam obey linear superposition, the incoming energy can be represented as the sum of a completely polarized component and a randomly polarized, or completely unpolarized, component by incorporating the DOP as shown in Eq. 10 [Schott 2009]:

$$\mathbf{S}_{\text{TOT}} = \begin{bmatrix} DOP \cdot S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} + (1 - DOP) \begin{bmatrix} S_0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (10)$$

Again, for a passive sensing application, the radiance vector can be simplified by neglecting circular polarization, as shown in Eq. 11:

$$\mathbf{S}_{\text{TOT}} = \begin{bmatrix} DOLP \cdot S_0 \\ S_1 \\ S_2 \end{bmatrix} + (1 - DOLP) \begin{bmatrix} S_0 \\ 0 \\ 0 \end{bmatrix} \quad (11)$$

If a polarization sensitive sensor is incorporated, the instrument is represented as a Mueller matrix, \mathbf{M}_I , and relates the Stokes vector of the incident radiance, \mathbf{S}_I , to the Stokes vector recorded by the detector, \mathbf{S}_D , as shown in Eq. 12 [Schott 2009]:

$$\mathbf{S}_D = \mathbf{M}_I \cdot \mathbf{S}_I \quad (12)$$

The ability to detect an object's polarimetric signature is highly dependent on the sun-target-sensor viewing geometry [Schott 2009]. Specifically, the Fresnel reflection coefficients of polarization states parallel and perpendicular to the plane of incidence, of a specular surface, depend differently on incident angle. As a result, nadir view or illumination tends to induce no polarization difference, while larger angles induce a larger DOP. However, as angles get more oblique, the DOP again decreases. Further, the specular component, rather than the diffuse component, of the reflectance tends to induce polarization [Schott 2009]. These effects suggest the presence of a polarimetric sensing “sweet spot” located azimuthally near the specular component of the reflection from the target and at some middle declination angle. Now that the fundamentals of polarimetric information have been described, an in-depth investigation of previous multimodal data fusion demonstrations can be conducted.

3.7 Previous Multimodal Data Fusion Demonstrations

Several examples of multimodal fusion demonstrations can be found in the literature, exploiting different fusion levels for different applications. In imaging applications, spectral data is often fused with either spatial or temporal data for land-use

classification. However, spectral data is more commonly fused with polarimetric data for anomaly detection or target detection scenarios.

One application of data fusion was based solely on the relationship between bands of spectral data. Hyperspectral data was partitioned into contiguous subspaces, maximizing the discrimination information in each subspace while minimizing the statistical dependence between subspaces [Prasad and Bruce 2008]. Each subspace was then treated as a separate source in a multisource classification problem. Several decision level fusion algorithms were employed, classified as either hard decision fusion or soft decision fusion. In hard decision fusion, like QMV, a final decision was based on some weighted vote from each data subspace. Hard decision fusion techniques had the advantage of proving relatively insensitive to inaccurate estimates of posterior probabilities. Alternatively, a soft decision fusion technique, like LOP or LOGP used some class membership function from every classifier to make the final decision. For the hyperspectral subspace problem, soft decision fusion techniques were more likely to provide stable and accurate classification.

Another application relied solely on hyperspectral data, but used decision level fusion to leverage multi-temporal data in an application distinguishing between two aquatic vegetation species [Prasad *et al.* 2008]. Several common classification algorithms were used as baselines to determine the fusion system's effectiveness, and majority voting was employed as the decision fusion scheme in all cases. In the first baseline method, linear discriminant analysis (LDA) was conducted for each day, and then the LDA outputs from all dates were combined into one common feature space for analysis with a single classifier. In the second baseline approach, hyperspectral

signatures from all dates were combined into a single dataset, and the hyperspectral space was partitioned into multiple smaller subspaces, with LDA and classification performed on each subspace. The third baseline algorithm again involved merging hyperspectral signatures from all dates into a single dataset, projecting the data into a new space based on LDA (learned from training data), then applying a single classifier. The proposed multi-temporal decision fusion system consisted of classifying each day's hyperspectral data via Gaussian maximum likelihood (GML), then combining those classification results over multiple days via majority voting. For this two-class application, the multi-temporal fusion system outperformed the baseline algorithms that didn't fully exploit the temporal information.

In an attempt to better exploit complementary imaging modalities, airborne LIDAR data has been fused with hyperspectral data at the pixel level for a seafloor and land cover classification problem [Macon *et al.* 2008]. Since the LIDAR data and hyperspectral data were acquired under highly similar environmental conditions, the LIDAR data was first used to correct the hyperspectral data to produce a geo-referenced image for each flight line. Radiation transfer equations were then used to estimate the seafloor reflectance from the data. Finally, the results of LIDAR processing were used to establish an elevation for each pixel, permitting fine corrections of the seafloor reflectance data when known water properties were taken into account.

Another complementary set of imaging modalities used for decision level fusion was multi-frequency polarimetric SAR with a panchromatic EO sensor [Yang and Moon 2003]. For a land-cover classification application, each data source was independently classified using the maximum likelihood classification method, based on a subset of

training data, and the classification results were merged using the Dempster-Shafer theory of evidence. The Dempster-Shafer theory measures conflicting results via user defined belief and plausibility metrics, and assigns the pixel of interest to the most likely class. Specifically, the belief metric was the sum of user-defined mass functions for each class, and the plausibility metric was the maximum class-conditional probability assigned by any classifier. No general rules exist for defining the class mass functions, although this step was identified as the most crucial one in the algorithm [Yang and Moon 2003]. Using a mass function based on all the classification results for pixels in a window around the pixel of interest, the fusion algorithm was found to enhance classification accuracy when results from the EO and SAR system were compared to results from just the EO system.

One of the more sophisticated techniques to fuse multispectral and panchromatic data was an algorithm relying on principal component analysis (PCA) and the discrete wavelet transform (DWT) [Chen *et al.* 2005]. First, both images were geometrically registered and the multispectral image was spatially resampled to match the panchromatic image. Next, the first principal component (PC) was obtained from the multispectral image via PCA. Then, the panchromatic was histogram matched to the PC image, and wavelet decomposition was performed. The low frequency coefficients of the PC image were then set as the low frequency coefficients of the fused image, but the high-frequency components were chosen from either the PC image or the panchromatic image via a series of decision rules. Finally, an inverse DWT and inverse PCA were performed to generate the spatially enhanced fused image.

Attempts to fuse hyperspectral and polarimetric data have occurred at different decision levels for a wide variety of applications, essentially limited by the capabilities of the hardware available. In one example, LWIR hyperspectral imagery was combined with polarimetric information to determine the near-surface wind vector, sea-surface temperature, and surface thermal emission properties [Iannarilli *et al.* 2000]. The hyperspectral and polarimetric data were acquired simultaneously using the InfraRed Polarimetric HyperSpectral Imager (IRPHSI) developed by Aerodyne, which was designed to be mounted on a ship and consisted of a single focal plane array, conventional optics, and no moving parts. Birefringent crystals were used to influence the polarimetric information such that the polarization information modulates the spectral information when captured at the focal plane. A demodulation scheme based on the Stokes vectors was used to extract the polarimetric information, and the primary advantage of this approach was that the hyperspectral and polarimetric images were perfectly registered in both time and space. From the modulation pattern imposed on the spectral data, a pixel fusion technique was carried out to solve for the wave slope, which was then used to determine the near-surface wind vector. Further, the ability to solve for the wave slope inversely depended on the GSD, since the polarimetric effects tended to average out over larger sample sizes. This last point illustrates the importance of high-spatial resolution in any polarimetric system designed for use in anomaly or target detection.

One particular target detection application employing polarimetric sensing was a system designed to exploit spectral features in visible and near-IR polarimetric images [Duggin and Loe 2002]. In that setup, thin film polarizers were manually rotated

between image captures to obtain the polarization measurements required to define three of the four Stokes parameters. Images of a model military vehicle, painted with camouflage paint and in a shadow under a platform covered in live leaves, were taken at discrete wavelengths over the range from 400 nm to 1 μm . The intent was to determine whether polarimetric information could assist in differentiating man-made targets from the background. In an example of fusion at the pixel level, the polarimetric images from the near-infrared were analyzed with an NDVI algorithm and contrast in the resulting image provided some ability to separate the target from the background.

In an attempt to further exploit the combination of spectral and polarimetric imaging, an anomaly detection experiment was carried out with a ground-based hyperspectral polarimetric imaging test bed [Cavanaugh *et al.* 2006]. The imaging test bed consisted of SWIR and VNIR hyperspectral capabilities, three polarimetric channels with linear polarizers set at 0, 60 and 120 degrees, and one high resolution panchromatic channel. Anomaly detection was accomplished by computing the PC transform of a hyperspectral image, classifying each pixel via a K-means classifier, and highlighting any pixels below some threshold probability of belonging to any of the classes. Separately, the polarimetric images were combined to calculate the DOLP of each pixel in the scene. Since the scale of the hyperspectral imagery exactly matched the polarimetric image, the images were combined by simply overlaying them such that the resulting pixel value was a multiplication of the hyperspectral anomaly score and the DOLP. Figure 2 summarizes the data fusion procedure laid out by Cavanaugh *et al.* (2006). Their preliminary results showed some ability to detect small man-made targets amidst a natural background, although a number of false alarms were also flagged.

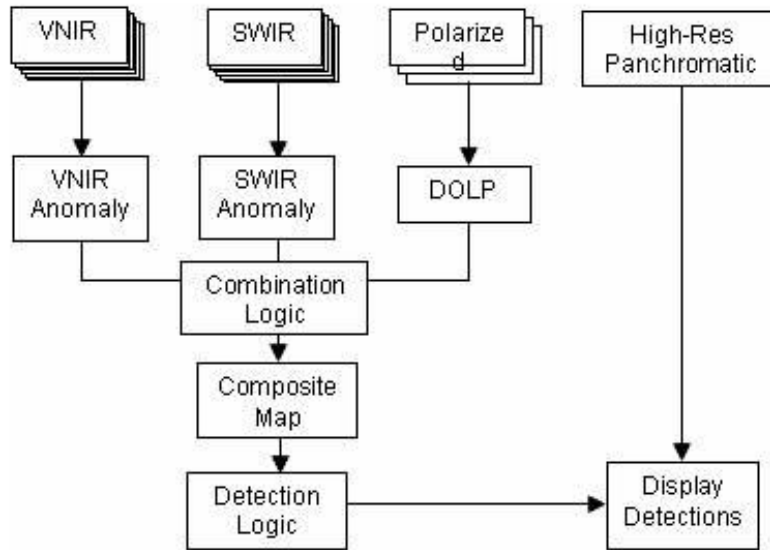


Figure 2. Data processing scheme for hyperspectral polarimetric imaging system used by Cavanaugh *et al.* (2006).

Another system with both hyperspectral and polarimetric imaging capabilities was demonstrated by Alouini *et al.* (2009). The hyperspectral imager spanned the 800 - 2100 nm range with a spectral resolution of 5 nm. The system was operated in either active or passive mode, and in the active mode nonlinear optical crystals were used to create one of five different signal wavelengths. To acquire images, the camera was operated at twice the pulse repetition rate, effectively interleaving an active image with a passive image, while rotating the polarization state between active images. For each image, the passive image was subtracted from the active image to eliminate the contribution of ambient light. Therefore, the system effectively became a multispectral sensor, acquiring polarization images at five different wavelengths. In their ground-based target detection application, one scene was analyzed with small metal plates located 20 m away from the sensor against different types of backgrounds. Further, a metric called the peak-to-sidelobe ratio (PSR) was defined, representing the conspicuousness of the peak

corresponding to the position of the target in the particular image, with a higher PSR providing easier target detection. Alouini *et al.* found that the PSR increased with wavelength for the intensity image, but that the PSR decreased with wavelength for the polarimetric image, and concluded that the intensity and polarimetric images for diffuse objects tended to behave in a complementary manner in most cases.

To recap, spectral data (of various spectral resolutions) has been fused with additional spectral data, obtained either in different regions of the spectrum or at different times, to enhance classification performance. Additionally, LIDAR, SAR, and polarimetric systems have been used in conjunction with spectral data for land cover classification, anomaly detection and target detection applications. The past successes achieved with spectral and polarimetric data suggest that fusing those two data types may increase target detection performance. Further, the fusion demonstrations indicate that information can be gained by examining both how anomalous and how target-like a pixel appears. However, the key requirement for multimodal fusion applications is simultaneous data acquired by both modalities. Because of the difficulty involved in acquiring simultaneous data for two state of the art systems, under a variety of imaging conditions, simulation becomes an attractive option. Therefore, the capabilities of existing software packages must be assessed.

3.8 Existing Simulation Software

Several software simulation packages exist, with varying degrees of use and acceptance in the remote sensing community. Results from MODTRAN, IRMA, CameoSim, IRHSS, and DIRSIG have been shown in remote sensing literature and

conferences for the field. These tools often began as one specialized application, with increasing capabilities provided via extra modules developed over time.

3.8.1 MODTRAN

The MODTRAN model [Berk *et al.* 1989] is probably the most widely used and readily available of any of the atmospheric radiation propagation models [Schott 2007]. MODTRAN assumes the atmosphere is divided into a stack of homogenous layers, with the temperatures and concentrations of those layers determined either from user inputs or standard atmospheric profiles. After the user defines a sensor location and view angle, the radiance reaching the sensor from a given point is determined by incorporating the cumulative transmission effect of each of the atmospheric layers, with a spectral resolution as small as 2 cm^{-1} . Additionally, the MODTRAN model allows a sensor to be placed on the ground, looking up to space. Integrating the observed radiance from several angles effectively computes the downwelled radiance under the given weather conditions. A limited release version, MODTRAN 4P, is capable of modeling the atmospheric effects on radiation with different polarization states, a crucial requirement for synthetic polarimetric modeling efforts [Devaraj *et al.* 2007]. However, MODTRAN 4P is limited to modeling single scattering rather than accounting for multiple bounces.

3.8.2 IRMA

The US Air Force's Infrared Modeling and Analysis (IRMA) software package has the capability to generate imagery simulating sensors from the visible range, IR, millimeter wave radar, and SAR, and claims over 130 users [Savage *et al.* 2008]. The software is divided into a passive channel, a radar channel, and a LADAR channel, each viewing a common scene made of models constructed from triangular facets. Material

IDs are associated with each facet to describe its electromagnetic properties, referencing additional reflectance, thermal properties, and texture files.

Although significant effort seems to have been invested in the radar channel, interest in IRMA for this project is limited to the passive channel, which Savage *et al.* (2008) described as further divided into three separate programs: ENVIRO, PASSIVE/PPASSIVE, and SSW. ENVIRO is used to compute the heat signatures of objects and one-dimensional heat transfer between objects, calculating facet temperatures for use by the passive image generators. PASSIVE models an unpolarized signature, while PPASSIVE models generalized elliptical polarization signatures, rendering images via Z-buffer accounting for emitted radiation, diffuse and specular reflections from the sun, the sky, and the earth, atmospheric path radiance, path transmittance, sensor spectral effects, and sensor spatial effects. SSW handles sensor effects modeling such as system responsivity and digitization. Surface reflectance is modeled with diffuse and specular components, and the passive channel includes multiple bounce effects.

In one example, IRMA was used to model a polarimetric IR imaging system in a small target detection scenario [Sadjadi and Chun 2004]. The scene of interest consisted of four aircraft hangars connected by runways, with a target vehicle (either a T-72 tank or M-35 truck) parked on the grass nearby. Several simplifying assumptions were made in the modeling process. First, all surfaces were set to the same temperature, 24° C. Second, only two surface materials existed: grass, which emitted unpolarized light, and glossy paint, which emitted polarized light as described by the Fresnel equations. Finally, no sun was present in the scene. Figure 3 shows the synthetic scene as observed by the three Stokes vectors S_0 , S_1 and S_2 as described in Sec. 3.6.

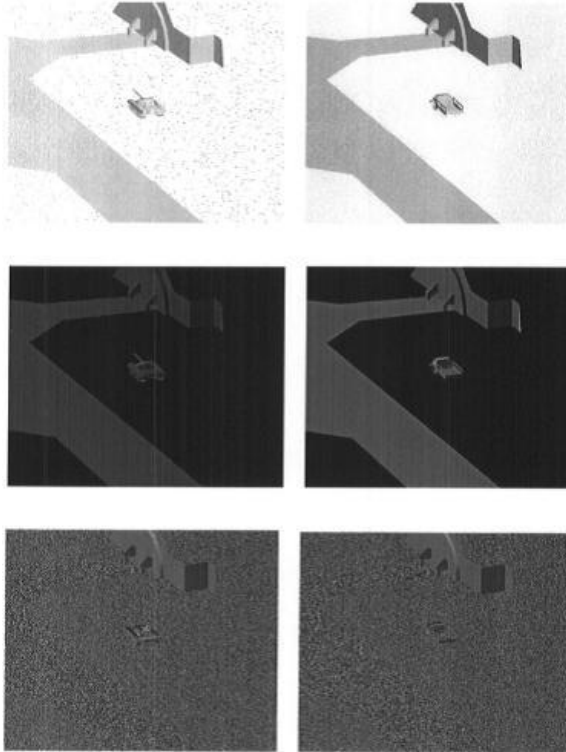


Figure 3. Synthetic polarimetric imaging scene generated by Sadjadi and Chun (2004) using the IRMA software package. The scene is displayed as an intensity value for the S_0 (top), S_1 (middle), and S_2 (bottom) Stokes vectors.

Given these constraints, Sadjadi and Chun demonstrated the ability to separate the target from the background using statistical methods in a simple scene at GSDs with as few as four pixels on target. However, IRMA's polarimetric fidelity currently seems hindered by significant simplifying assumptions, so it will be avoided for this dissertation.

3.8.3 CameoSim

The CameoSim software package is currently maintained by Lockheed Martin's UK division. Ranging from the ultraviolet to the infrared, CameoSim is based on first principles physics and radiometric interactions, while using ray tracing techniques to generate a radiance map as well as temperature images for synthetic scenes. In general, a CameoSim scene is constructed of 3-D geometric models, which are generally composed of facets [Mitchell *et al.* 2007]. Each facet then has a material associated with it, and

each material is described in a database containing a bidirectional reflectance distribution function (BRDF) along with optical and thermophysical properties. A sensor system is defined, along with known weather parameters and a time of day. An atmospheric database incorporates direct solar or lunar fluxes, sky shine fluxes, path radiance and transmission, and the local thermal environment. The synthetic image is rendered by incorporating the cumulative effects of the sensor and atmosphere in observing the scene.

Some work has been done to validate CameoSim's hyperspectral modeling capability and to determine the fidelity of reflection modeling. Although the synthetic hyperspectral data compared reasonably well to real data, areas for improvement were identified [Briottet *et al.* 2006]. Specifically, difficulties arose in capturing the correct tree density and effects of the 3-D nature of grass in the modeled scene. Problems were also noted with facets that failed to accurately represent reflections, or glints, from curved areas like rotor blades in the 3-5 μm band [Mitchell *et al.* 2007]. One approach to solving the problem was to simply use the law of reflection with an enormous number of small facets, but a severe increase in computational load resulted. In contrast, a method known as vertex averaging was demonstrated to be more accurate with far fewer facets.

More recently, Harvey *et al.* (2008) used CameoSim to analyze a laser imaging application and their synthetic scenes are shown in Figure 4 below. In their urban scenario, the sensor flew toward and then over the target, such that the degree of obscuration remained constant regardless of range or look-down angle. In the scene with the target under foliage, rendering individual leaves took a significant amount of computing time, limiting the number of target types that could be used in a given scenario.



Figure 4. Synthetic CameoSim scenes used by Harvey *et al.* (2008) for a laser imaging application. Although CameoSim seems to be reasonably rigorous, the common intense computational requirements make CameoSim an unattractive option for simulating a complicated urban environment under the vast number of illumination conditions for a serious trade study.

3.8.4 IRHSS

Kwan *et al.* (2008) described the Infrared Hyperspectral Scene Simulation (IRHSS) software suite sponsored by the US Air Force Research Laboratory and the Army Armament Research, Development and Engineering Center. In addition to models for atmospheric propagation and sensor processing, the software integrates the Multi-service Electro-optical Signature (MUSES) model to compute scene temperatures and hyperspectral radiances for the thermal IR bands (3-14 μm). As with the other modeling packages, material surface properties are linked to facets in 3-D geometry models. The software is relatively new and results lack widespread publication in remote sensing literature. IRHSS's limitation to IR spectral data severely constrains its potential use in a data fusion experiment.

3.8.5 DIRSIG

The Digital Imaging and Remote Sensing Laboratory at Rochester Institute of Technology developed the DIRSIG software package as a first-principles based ray tracing model to create sample data to test image system designs, evaluate target detection algorithms and train image analysts. Rays are cast out from each pixel on the focal plane to determine what that pixel sees, and what the sources of radiance are for the area observed. Each pixel is subsampled, with the spectral results combined via a linear mixing model [Ientilucci and Brown 2003]. The model is capable of producing hyperspectral imagery from the visible through the thermal infrared (0.4-20 μm), and can correctly model interactions for different polarizations of incoming light. The MODTRAN model is incorporated into DIRSIG to accurately account for atmospheric effects.

DIRSIG has been used to model both high resolution spectral and spatial data. Since DIRSIG is a ray tracing model, geometric detail is essential to accurately describe radiometric mechanisms. Striking a balance between the modeled field-of-view and the required computation time means the amount of detail required is inversely proportional to the expected GSD. A wide area synthetic scene dubbed Megascene 1 was constructed in DIRSIG to mimic an area on the northeast side of Rochester, NY for use in hyperspectral image analysis [Ientilucci and Brown 2003]. The area of interest has a combination of urban and suburban residential, industrial and forested areas, so detailed models of ten geometrically unique house types, along with specific commercial and government buildings were created in the Rhinoceros CAD package [Rhinoceros software 2010]. In addition, six different species of trees were modeled with the Tree

Professional software package [Tree Professional software 2010], and three geometric variants of each species were created. Figure 5 shows examples of the models produced by Ientilucci and Brown.

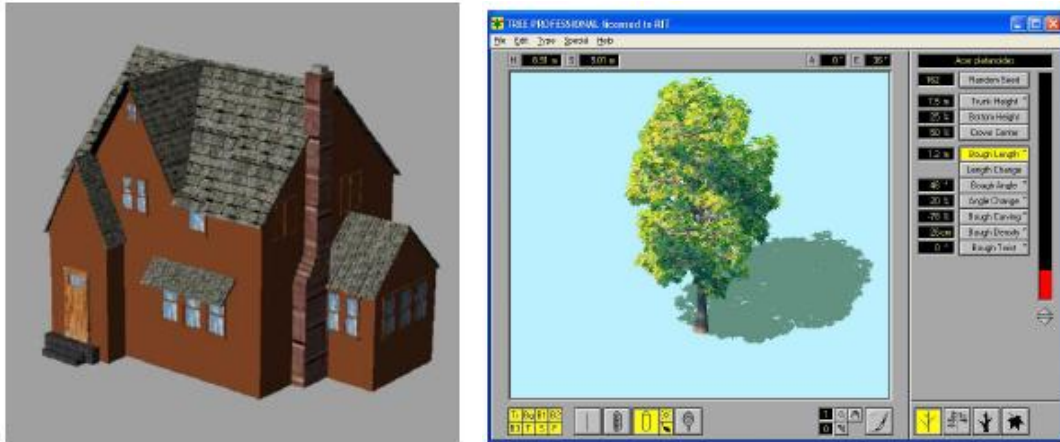


Figure 5. (Left) Example of a house generated with Rhinoceros. (Right) Screen capture of the Tree Professional software package illustrating a rendered Norway maple. Images courtesy of Ientilucci and Brown (2003).

Once the houses and trees were designed, the models were replicated and altered through scale, rotation and translation for use throughout the scene. Further, the materials attributed to each facet could be changed each time the model was reused in the scene. As a result, a set of 50 geometrically and spectrally unique houses was produced, complemented by a set of 20 geometrically and spectrally unique trees. Additionally, primary terrain materials like asphalt and grass were characterized spectrally and placed in the scene to match the results of a GML classification performed on actual data of the site. Finally, the spectral angle mapper (SAM) and spectral matched filter (SMF) algorithms were applied in a target detection scenario, and the trends of the ROC curves that resulted were consistent with results obtained from actual data.

For applications based on a much narrower field-of-view, a high resolution synthetic scene dubbed MicroScene was constructed based on an existing data collection

effort [Barcomb *et al.* 2004]. High spatial (3" GSD) and spectral (0.01 μm) resolution imagery was captured using at the Rochester Institute of Technology using the Center for Imaging Science's MISI and WASP sensors at an oblique viewing angle. Three Humvees were placed in the scene under varying levels of concealment to serve as targets. After the data was collected, a spatially and spectrally accurate synthetic replica of the scene was created in DIRSIG. The synthetic imagery was validated against the images of MicroScene by analyzing both scenes with qualitative analysis, GML classification, and the RX target detection algorithm. Figure 6 demonstrates one example of the qualitative analysis performed by Barcomb *et al.*, where an actual image of a Humvee under camouflage netting was compared to the synthetic image of the same scene to highlight the phenomenology reproduced.

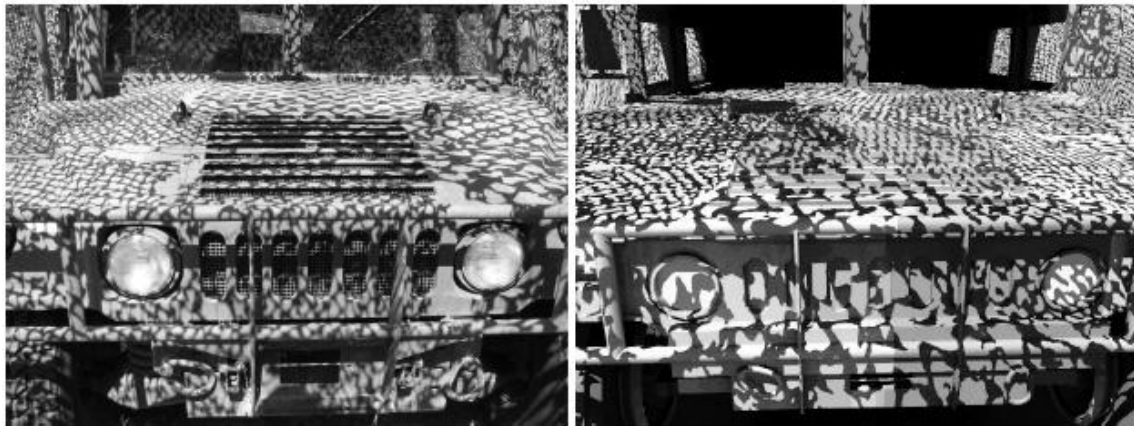


Figure 6. (Left) Actual image of a Humvee under camouflage netting. (Right) Synthetic image of the same Humvee under camouflage netting, demonstrating accurate reproduction of a highly detailed model and shadowing phenomenology. Images courtesy of Barcomb *et al.* (2004).

Although the GML classification results showed some differences depending on whether DIRSIG or truth derived training sets were used, overall the objects in the image were classified appropriately with either training set. Further, it was found that the results of the RX algorithm at the high resolution examined were dependent on the accurate

modeling of objects in the actual scene that were initially considered inconsequential, demonstrating a potential flaw in model scenes that use only a test object and a uniform background.

Building on the successful modeling of spectral effects, polarization effects were first incorporated into DIRSIG by using BRDF models based on the Torrance-Sparrow and Beard-Maxwell models to provide polarized BRDF estimations [Meyers *et al.* 2002]. By applying the effects of surface roughness and index of refraction to the different incident polarizations, the models reasonably matched measured BRDF data. However, a lack of polarimetric BRDF databases poses the single largest problem for conducting fully polarimetric radiometry simulations, while physically based BRDF models do a reasonable job of predicting the polarimetric BRDF of simple surfaces. Because materials with highly polarized properties, such as the man-made materials often sought in target detection algorithms, are often also highly specular, the full BRDF is required to correctly incorporate the directionally reflected background contributions [Gartley *et al.* 2007]. DIRSIG generally treats the polarized BRDF as a polarized specular component and an unpolarized volume component, as shown in Eq. 13:

$$f_{pBRDF} = f_{specular} + f_{volume} = f_{polarized} + f_{unpolarized} \quad (13)$$

The generalized, polarized BRDF model used for Gartley’s study was the “Shell Target” BRDF model [Shell 2005]. In this approach, the specular component was modeled using a generalized, polarized, Gaussian-distributed micro-facet approach, while the unpolarized component was modeled using a compound volume scattering and diffuse scattering term. Polarimetric measurements were taken of geometrically complex vehicles open to the night sky, and a similar scene was rendered in DIRSIG to assess the

polarimetric modeling capabilities. Figure 7 shows the actual imagery alongside the synthetic DIRSIG imagery generated by Gartley *et al.* Qualitatively, the images agree quite well in each band. However, some salt-and-pepper noise is apparent in the S_0 synthetic image of the SUV, a fault which the study's authors traced to the chosen polarized BRDF model breaking down at grazing angles rather than an issue with DIRSIG's ray tracing capabilities.

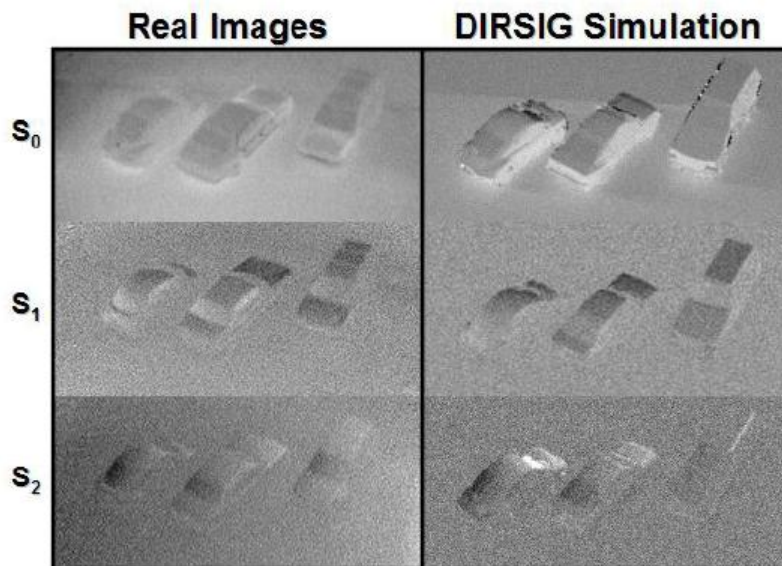


Figure 7. (Left) Actual polarimetric images of vehicles under a cold, clear night sky. (Right) Synthetic image generated by DIRSIG based on the actual scene. Images courtesy of Gartley *et al.* (2007).

Another effort to reproduce real-world polarimetric imagery in the reflective region with the DIRSIG model incorporated pBRDFs, polarized atmospheric models, and polarization-sensitive sensor models [Pogorzala *et al.* 2007]. Grass and asphalt were attributed based on the Shell BRDF model [Shell 2005], while an aluminum fire hydrant was attributed with a material based on a Priest-Germer pBRDF [Priest and Germer 2002]. The remaining objects were attributed with unpolarized spectral emissivity curves, since the authors noted the limited number of pBRDF models available for the

construction of their scene. The lack of available polarized optical properties was compensated for by a technique called bump mapping, designed to introduce pixel-to-pixel variability within material classes. In bump mapping, the surface normal of an object is deflected a specific amount based on the corresponding pixel's value in the bump map. In essence, the bump map serves to introduce small-scale variability while constraining the number of facets required for a given scenario. A qualitative comparison between the real and synthetic imagery showed that the bump mapping technique was able to recreate asphalt texture observed in real images. Further, the synthetic imagery produced a low DOP from natural materials and a high DOP from man-made materials, confirming phenomenology observed in actual data. However, the largest deficiencies in the synthetic data were associated with the lack of available pBRDF models for all materials in the scene.

As a key ingredient for scenes involving vegetation, Gartley and Basener (2009) verified the correct simulation of leaf pBRDF properties in DIRSIG. A synthetic forest attributed with leaves having pBRDFs but unpolarized trunks and branches was created, and the degree of linear polarization from the generated forest imagery was validated against publicly available POLDER measurements.

In summary, DIRSIG is a proven first principles based synthetic image generation model. The proven ability to simulate hyperspectral effects in the visible and IR regions, combined with an emerging capability to model polarimetric effects, makes DIRSIG an attractive choice for use in a multimodal sensor trade study. Further, the existence of a rigorously modeled, realistically attributed synthetic urban scene makes DIRSIG a natural choice for modeling an urban target detection scenario. Once the synthetic

imagery has been generated, the modeled atmospheric effects must be compensated to some degree for spectral target detection algorithms.

3.9 Atmospheric Compensation

As part of the analysis required for an accurate trade, it is assumed that an analyst would have performed some form of atmospheric compensation on the imagery.

However, rather than identifying the optimum atmospheric compensation technique, this study only requires a reasonable, consistent process be applied to each image. The users could then select and optimize a particular atmospheric correction technique when the parameters in their trade study have been customized based on the actual resources available. Since this study is concerned with many images in the solar reflective region of the spectrum, three primary techniques for atmospheric compensation exist. Schott (2007) showed that the radiance reaching the sensor through a bandpass in the reflective region can be approximated as shown in Eq. 14

$$L = \left[E'_s \cos \sigma \frac{r}{\pi} \tau_1 + [FE_{ds} + (1-F)E_{bs}] \frac{r_d}{\pi} \right] \tau_2 + L_{us} \quad (14)$$

where E'_s is the exoatmospheric irradiance, σ is the solar zenith angle, τ_1 is the transmission from the sun to the target, F is a shape factor describing what fraction of the hemisphere of the sky above the target is open, E_{ds} is the downwelled irradiance, E_{bs} is the reflected background irradiance onto the target, and L_{us} is the upwelled radiance reaching the sensor.

The empirical line method (ELM) is a common ground truth reflectance correction method [Schott 2007]. For clear skies at small zenith angles, Eq. 14 can be

simplified by assuming $E_s \cong E'_s \cos \sigma' \tau_1$. Further, for pixels in a relatively open space, $F \cong 1$. The main goal of ELM is to regress observed radiance values against known reflectance values in each band as shown in Eqs. 15 and 16

$$L = (E_s \pi^{-1} + L_d) \tau_2 r_d + L_u \quad (15)$$

$$L = m r_d + b \quad (16)$$

where $m = E_s \pi^{-1} \tau_2 + L_d \tau_2$ is the slope of the regression, and $b = L_u$ is the intercept. In ideal cases, carefully calibrated control panels are placed in the scene as a reference, but the users could also obtain reflectivity measurements of several Lambertian objects in the scene.

For situations without some form of ground truth, Piech and Walker (1974) described a method to take advantage of the difference in radiance levels observed at shadow edges. The radiance in a given spectral band observed just outside a shadow, L_s , depends on the shape factor, F , as shown in Eq. 17, while the radiance observed just inside the shadow, L_{sh} , cast on the same diffuse material is given by Eq. 18 [Schott 2007]:

$$L_s = [E_s \pi^{-1} \tau_2 + FL_d \tau_2] r_d + L_u \quad (17)$$

$$L_{sh} = FL_d \tau_2 r_d + L_u \quad (18)$$

Combining Eqs. 17 and 18 by substitution of r_d and rearranging yields Eqs. 19a and 19b, where $m = (E_s \pi^{-1} + FL_d)$ and $b = (1 - m)L_u$:

$$L_s = \frac{E_s \pi^{-1} + FL_d}{FL_d} L_{sh} - \frac{E_s \pi^{-1} + FL_d}{FL_d} L_u + L_u \quad (19a)$$

$$L = m r_d + b \quad (19b)$$

If L_u is assumed constant, and objects of similar shape are selected, then F is also approximately constant. Therefore, m and b are approximately constant, and can be solved for by a linear regression of the shadow radiance versus the sun radiance, producing L_u as shown in Eq. 20 [Schott 2007]:

$$L_u = \frac{b}{1 - m} \quad (20)$$

Piech *et al.* (1978) then explained how the upwelled radiance could be combined with a statistical estimate of the mean observed radiance, for a class of objects whose mean reflectance is well known, to produce the total radiance incident on the sensor. They typically used concrete, with the resulting equation form shown in Eq. 21 [Schott 2007]:

$$m = [E_s \pi^{-1} \tau_2 + L_d \tau_2] = \frac{L_{avg} - L_u}{r_{davg}} \quad (21)$$

Where L_{avg} is the mean radiance observed for many samples of the standard material chosen and r_{davg} is the mean reflectance based on known measurements. Once the slope term is known, the reflectivity of each pixel in the scene can be estimated.

Several atmospheric correction methods based on radiative transfer models exist [Schott 2007]. In a typical case, as many known parameters from the data collect as possible are entered into MODTRAN, where the transmittance, upwelled radiance, downwelled radiance and solar radiance are calculated and combined to produce the total radiance reaching the sensor for a nominal reflectivity target. From that information, in conjunction with the observed radiance, an estimate of the reflectivity of every pixel in the scene can be calculated.

In closing, the shadow method, ELM and MODTRAN-based methods are all recognized atmospheric compensation techniques. ELM is the most attractive choice for a trade study because perfectly calibrated reference panels can be placed in every version of the synthetic scene. Another advantage is that even with multiple view angles and multiple times of day, ELM can be easily automated. After converting the synthetic data into reflectance space, a widely accepted target detection algorithm can be applied.

3.10 Detection Algorithms

Since the focus of this work is on fusing the products from two sensors at the decision level to enhance target detection, the data from each sensor will be processed independently with some detection algorithm. Although many different target detection algorithms exist [Schott 2007], to avoid turning the project into a target detection algorithm optimization effort, representative standard algorithms will be used.

The spectral matched filter (SMF) is the most common target detection algorithm based on a stochastic description of the data, and is described by Eq. 22 [Schott 2007]

$$SMF(\mathbf{x}) = (\mathbf{t} - \mathbf{m})^T \mathbf{S}^{-1} (\mathbf{x} - \mathbf{m}) \quad (22)$$

where \mathbf{t} is the target vector and \mathbf{x} is the sample vector. Further, \mathbf{m} and \mathbf{S} are the background mean vector and covariance matrix, which are created from local values drawn from an area around the pixel of interest or global values drawn from the entire scene. Each pixel in the image is analyzed, and if the SMF score is above some user defined threshold, η , the pixel is designated a target. The threshold is used to control the false alarm rate, and so can be varied based on the particular application. Farrand and Harsanyi (1997) modified the SMF algorithm such that the output is scaled to a target

abundance value between zero and one, dubbing their method constrained energy minimization (CEM) as shown in Eq. 23:

$$CEM(\mathbf{x}) = \frac{(\mathbf{t} - \mathbf{m})^T \mathbf{S}^{-1} (\mathbf{x} - \mathbf{m})}{(\mathbf{t} - \mathbf{m})^T \mathbf{S}^{-1} (\mathbf{t} - \mathbf{m})} \quad (23)$$

The CEM technique is therefore an attractive algorithm for a multispectral or hyperspectral target detection problem.

The topological anomaly detection (TAD) algorithm has achieved a reasonable degree of success when applied to a hyperspectral target detection application [Basener and Messinger 2009]. The TAD algorithm models the background as a set of connected components of a graph, imposing a topological assumption on the data without requiring any assumptions on the geometry, linearity, or statistical distribution of the data. First, the TAD algorithm normalizes the image data so that the brightest 10% of the pixels in the image have Euclidean L2 norm equal to two, and the darkest 1% have Euclidean L2 norm equal to one. Next, a random subsample from the image of between 500 and 10,000 pixels is chosen to model the background, and the distance between every pair of pixels in the sample is computed. A graph is then constructed by adding an edge between the closest 10% of pairs of points, exploiting the idea that anomalous pixels are unlikely to have nearby neighbors. The largest groups of points are then designated as background as shown in Figure 8.

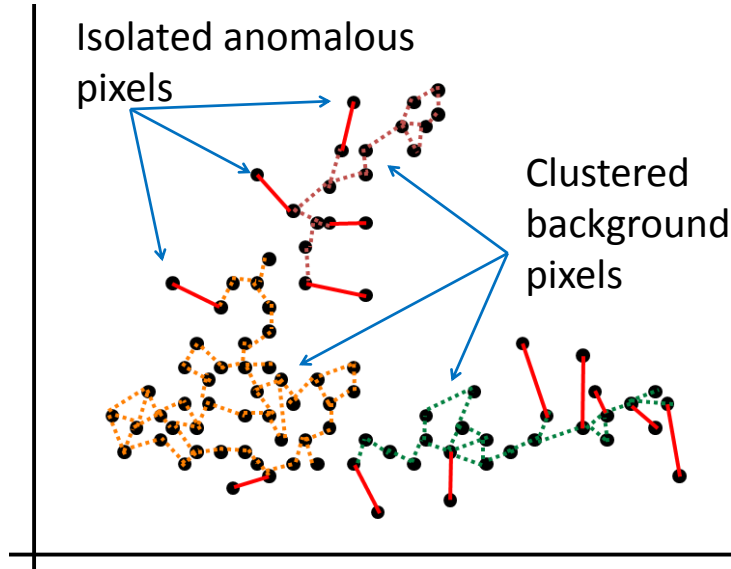


Figure 8. The TAD algorithm operating on data notionally distributed in two dimensions. The closest pairs of points are deemed background and linked by dotted lines, with different colors implying different classes. Pixels linked to their nearest neighbor by longer solid lines are likely anomalous. Image courtesy of Basener and Messinger (2009).

The percentage of pixels from the subsample that are background is then assumed to equal the percentage of the image that is background. Finally, each pixel is measured against the identified background pixels via the codensity metric, δ_k , which represents the radius of the smallest sphere enclosing k neighbors. The TAD ranking of each pixel in the image is equal to the sum of the distances to the 3rd, 4th, and 5th nearest neighbors in the background pixels, as shown by Eq. 24:

$$TAD(x) = \sum_{i=3}^5 \delta_i(x) \quad (24)$$

The TAD rankings result in level sets of arbitrary topology, allowing detection of pixels in the holes of the convex hull of the background as shown in Figure 9 [Basener and Messinger 2009].

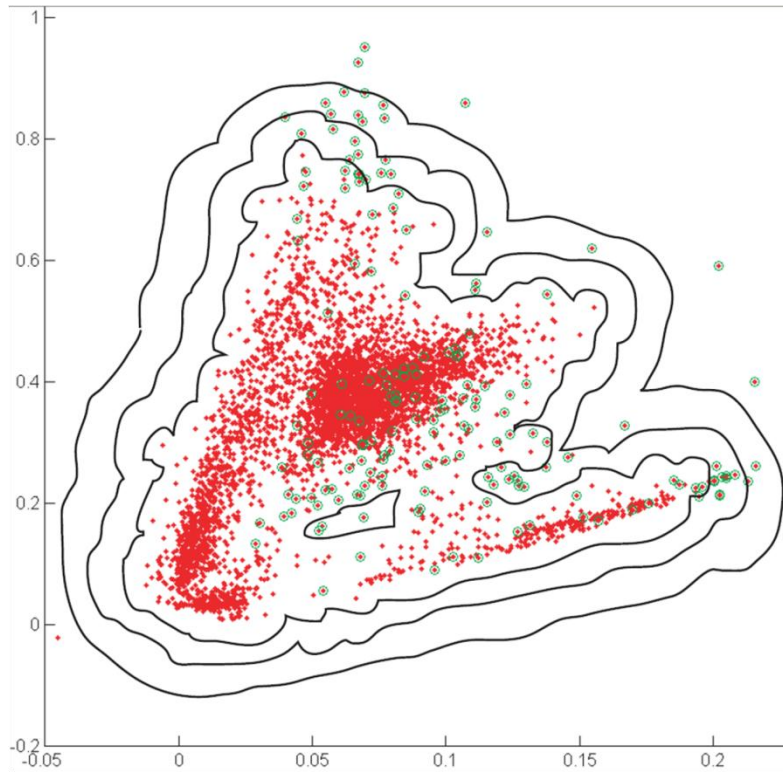


Figure 9. The TAD algorithm produces level sets of arbitrary topology (solid lines), allowing detection of anomalous members (circled) both inside and outside of the background’s convex hull. Image courtesy of Basener and Messinger (2009).

Similar to the CEM score, the TAD ranking can be normalized to fall in the range between zero and one. In their experiment, the TAD algorithm outperformed several standard spectral anomaly detection algorithms.

The TAD algorithm has also been employed for a DIRSIG polarimetric target detection application [Gartley and Basener 2009]. The specific scenario consisted of manmade targets in a natural background, observed by a multispectral polarimetric sensor. Four polarization images (S_0 , S_1 , S_2 and S_3) were captured at blue, green, red and near IR wavelengths, but the TAD algorithm performed best when the intensity, or S_0 band, was ignored. Further, the TAD algorithm outperformed the standard RX anomaly detector on panchromatic polarimetric imagery.

3.11 Section Summary

In conclusion, this section has summarized the current state of the art by describing the capabilities of existing high spatial or spectral resolution sensor systems, explaining the different data fusion levels and reviewing established decision level fusion techniques. A high level overview of existing sensing modalities and polarimetric remote sensing was conducted, identifying ongoing interest in hyperspectral and polarimetric sensors for target detection applications. Several previously published multimodal sensor fusion efforts were described, confirming that the impact of fusing hyperspectral and polarimetric data under a variety of conditions is still largely unexplored. The capabilities of several synthetic image generation models were examined for possible use in a potential trade study, and DIRSIG was identified as a useful tool. Next, the ELM was found to be a representative atmospheric compensation method that could easily be implemented for synthetic imagery under a wide variety of viewing geometries and illumination conditions. Finally, the CEM and TAD algorithms were identified as proven target detection algorithms. In a decision level multimodal fusion experiment, CEM could be used to analyze the data spectrally, while TAD could be used to identify materials with a strong polarimetric signature. With the general concepts firmly established, the specific experimental method must be determined.

4 Methodology

4.1 Section Overview

Several key tasks made up the multimodal sensor trade study proposed in Sec. 1. First, a reasonable scenario and notional sensor systems were defined. Then, a model urban scene was designed to evaluate the scenario and sensors of interest. The scene was rendered under a variety of viewing geometries, illumination conditions and SNR values before applying an atmospheric compensation technique. After introducing some degree of registration error between the images from the two sensors, the sensor outputs were fused for analysis at both the pixel and the decision levels. A method to determine the impact of incorporating data from a second modality was described by analyzing the resulting ROC curves, and a process to determine the ideal viewing conditions for each sensing modality was proposed. Finally, a procedure describing how the best-case images from each modality were geometrically registered and analyzed helped determine the maximum attainable benefit from incorporating the second modality.

4.2 Defining the Scenario

Given the intelligence community's emphasis on high value target and tracking discussed in Sec. 1 and the existing remote sensing systems described in Sec. 3.2, an unmanned sensor platform scanning an urban area for a particular vehicle described a reasonable scenario. A representative urban area contained a mix of residential houses, commercial facilities, government buildings, trees, roads, and vehicles. In this scene, the target vehicles were mixed with several other types of vehicles to provide realistic background clutter, assessing the sensor's ability to differentiate between vehicle types

rather than simply detecting any vehicle. After establishing the scene of interest, reasonable sensor parameters were defined.

4.3 Defining the Sensor

The goal of this project was to assess the utility of a particular sensor rather than to investigate the engineering challenges inherent in building that sensor. No optical aberrations were considered, such that the output represented a detector limited best-case scenario. For this project, a hyperspectral or multispectral sensor was used in conjunction with a polarimetric sensor, and the notional test sensors were assembled by merging optimum capabilities from published existing systems. As outlined in Sec. 3.2, the Predator's published operational ceiling of 7620 m provided a notional observational altitude; while the Global Hawk's published GSD (from a much higher vantage point) suggested a minimum GSD of 0.5 m is reasonable.

To achieve the desired GSD in the simulation, both sensors shared a common optic with a focal length of 300 mm. However, the hyperspectral sensor had significantly larger detector pixels than the polarimetric or multispectral sensors to account for SNR issues. The capabilities of each sensor were simulated in DIRSIG by an ideal 2-D framing array, where each pixel was capable of recording information as seen through multiple user-defined spectral bands or linear polarizers arranged at different orientations. Therefore, the resulting synthetic data appeared as an image of the scene as seen through each of the filtered bands defined by the user. It was assumed that an actual detector, sampling and interpolation scheme could be constructed to produce data equivalent to the capabilities described below.

NASA's AVIRIS instrument proves that an airborne hyperspectral system spanning the 0.4-2.2 μm region is achievable and served as the template for this scenario. The hyperspectral sensor's focal plane was simulated as a framing array of 128×128 pixels, where each pixel was $120 \mu\text{m} \times 120 \mu\text{m}$. Positioning the sensor at the Predator's flying altitude of 7500 m above ground level produced a nadir GSD of 3 m, with the intent to hunt sub-pixel vehicle targets while permitting a reasonable pixel fill factor. The hyperspectral sensor consisted of 90 channels spanning the same spectral range as AVIRIS, each with a Gaussian spectral response of FWHM 20 nm.

The spectral characteristics for the multispectral sensor in this project were based on Worldview-2's eight bands, with rectangular spectral responses of varying widths.

Table 3 shows the spectral sensitivity regions of the multispectral bands.

Table 3. Spectral bands for the multispectral sensor used in this project. The spectral bandwidths roughly match Worldview-2.

Band Name	Spectral Bandpass (nm)
Coastal	400 – 450
Blue	450 – 530
Green	520 – 610
Yellow	585 – 625
Red	640 – 690
Red Edge	705 – 745
NIR	770 – 880
NIR 2	860 - 1040

GeoEye-1's proven space-based multispectral capability has a nadir GSD of 1.65 m. If the GeoEye-1 multispectral instrument were mounted on a lower airborne platform, the change in scale would improve the GSD. Reducing the number of spectral bands, as compared to the hyperspectral instrument, allows a plausible decrease in GSD to 0.5 m. To generate this data, the multispectral focal plane was designed as a framing array of 768×768 pixels where each pixel was $20 \times 20 \mu\text{m}$.

Four of GeoEye-1's multispectral bands could easily be re-imagined as the four bands needed for polarimetric imaging. The polarimetric sensor was mounted on the same platform as the spectral sensor, and had a 768×768 pixel focal plane where each pixel was $20 \times 20 \mu\text{m}$. The resulting nadir GSD of 0.5 m ensured multiple pixels on target, capturing localized polarimetric phenomena rather than losing information by integrating over a large region. Four separate polarimetric bands were defined, each filtered with a linear polarizer oriented at 0° , 45° , 90° or -45° such that the sensor reaching radiance could be decomposed into a Stokes vector at each pixel with the modified Pickering method [Schott 2009]. The panchromatic spectral response function was a rectangle spanning 400 – 900 nm.

The combination of hyperspectral and polarimetric sensors identified in this section represented cutting-edge, but reasonably achievable, sensor capabilities. Since the goal of this project was to determine the value added by incorporating polarimetric information into spectral data analysis, plausible baseline systems were identified. This project did not address whether these were the ideal combinations of sensor hardware or tasking altitudes—rather, a future potential user could apply this method to evaluate the effect obtained with the exact systems available. Similarly, the system effectiveness was

assessed against a notional urban scene, with a future user free to exactly model a particular target of interest.

4.4 Designing a Model Scene

DIRSIG was the natural choice for use in this project, since RIT has access to the model and extensive work has been done with DIRSIG to simulate an urban environment. A subset of Megascene 1, dubbed Tile 1, provided an excellent background for a target detection application. Tile 1 was composed of a mostly residential area, a significant school complex, several trees, open fields and parking lots. Figure 10 shows both a DIRSIG rendering and a Google Earth snapshot of the region of interest, permitting vehicles in open fields, on roads, in parking lots, near houses, under trees and near large buildings while constraining the amount of data generated.



Figure 10. (Left) DIRSIG rendering of Megascene 1, Tile 1. (Right) Google Earth snapshot of the region of Rochester, NY nominally represented by Megascene 1, Tile 1.

The modeled target and decoy vehicles in the synthetic scene were faceted geometric constructions based on actual vehicles. Because polarimetric effects are highly sensitive to surface orientations, great effort was expended to accurately model small

features on the vehicle like the side view mirror, radiator grille and door handles. Figure 11 shows one of the CAD vehicle models designed for use with DIRSIG, provided by the Air Force Office of Scientific Research (AFOSR).



Figure 11. Subaru wagon model provided by AFOSR for use in DIRSIG. Notice the level of detail—a side mirror, front grille and even door handles have been incorporated.

Once the models had been converted into a format that DIRSIG could read, each facet was attributed with a material, which then linked to an emissivity curve, measured at 1 nm spectral resolution, and a pBRDF also provided by AFOSR. The target vehicles are red Subaru wagons, while the decoy vehicles are blue Ford Focus sedans, white sedans, black BMW SUVs, green VW wagons, yellow pickup trucks, and grey Volvo wagons. A total of 18 target vehicles and 108 decoy vehicles were inserted into the scene. While the facetized non-target vehicles provided polarimetric clutter, several different red materials provided spectral clutter for the target detection application. For example, measured spectra of red asphalt, roof shingles, bricks and a tennis court surface were attributed where those objects appear in the scene.

To remain plausible, the radiance reaching the sensor was affected by local weather conditions. Weather effects in the scene were modeled with MODTRAN-4P, with realistic weather inputs based on Megascene 1's location in Rochester, NY. Actual weather data from June 23, 1992 was input into MODTRAN-4P, along with the standard mid-latitude summer tape5 file adjusted to account for urban aerosols. In conjunction with MODTRAN-4P, DIRSIG predicted the sensor reaching radiance, but did not apply any sensor MTF effects. In essence, this implied that the sensor was detector limited, and represented the best-case scenario.

The synthetic scene was generated by sampling the scene at the instantaneous field of view (IFOV) in each direction, across every spectral band, leading to a GSD of 0.25 m at nadir and 2×2 spatial oversampling in general. Then, 2×2 pixel neighborhoods in each band were summed, such that the resulting DIRSIG scene captured the sensor-reaching radiance in each polarimetric or spectral band for each pixel on the high resolution focal plane described in Sec. 4.3. Since each of the high spatial resolution pixels represented the IFOV of the larger GSD hyperspectral image, one hyperspectral super pixel was created by summing the radiance values calculated from a $6 \text{ pixel} \times 6 \text{ pixel}$ area of the high spatial resolution image. Because the spectral and polarimetric information was rendered for a wide variety of viewing conditions, constructing the low spatial resolution spectral image from existing data eliminated the need to re-render each DIRSIG scene at the larger hyperspectral GSD.

4.5 Varying Relevant Parameters

For this trade study, the fundamental free parameters were solar zenith angle, sensor zenith angle, sensor relative azimuth angle, and sensor SNR. Figure 12 describes how the angles of interest could be varied while the sensor GSD values were fixed.

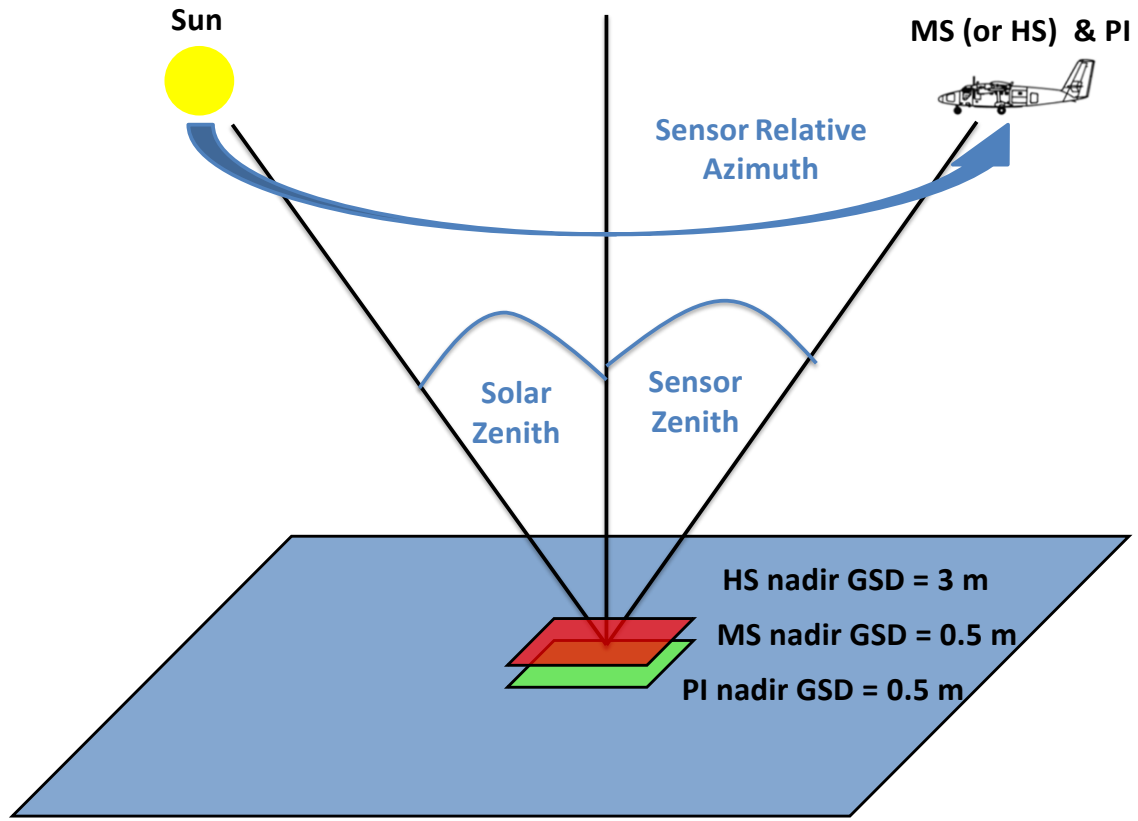


Figure 12. The solar zenith, sensor zenith and sensor relative azimuth angles were all varied for this trade study while the hyperspectral (blue rectangle), multispectral (red rectangle) and polarimetric (green rectangle) sensor GSD values were fixed.

Varying the illumination angle and sensor viewing angles affected the ability to measure spectral and polarimetric phenomena, while varying the sensor SNR helped establish sensor design requirements. The rendered scene GSD was locked to 0.5 m for nadir viewing, as described in Sec. 4.3, and was not a part of the investigated trade space. Since the scene represented an actual part of Rochester, NY, the illumination angle was varied by examining several different times of day: 0600, 0700, 0800, 1000 and 1200.

For each time, the position of the sun was set accordingly to produce five separate illumination conditions. As long as some of the targets and decoys are shadowed while some are well lit, at each time, the morning and afternoon sun positions are redundant if the images are acquired as described below.

The sensor zenith angle was varied in a uniform fashion while holding the altitude constant. As a result, the slant range to the center of the scene changed for every zenith angle. However, maintaining a constant altitude is more representative of a loitering UAV than holding the slant range constant and varying the sensor altitude. Since Megascene 1 is located about 119 m above sea level, and the operational ceiling of the Predator is 7620 m, the altitude of the sensor above ground was conveniently fixed at 7500 m. The sensor zenith angle was investigated by modeling seven concentric circular flight paths of varying radii, plus a nadir image, with the sensor focused on the center of the scene in Figure 10. The concentric flight paths represented zenith angles from $\pm 10^\circ$ to $\pm 70^\circ$ in 10° increments, spanning a range more extreme than that used by most existing systems.

The required sensor relative azimuth angles were determined by incorporating the principles discussed in Sec. 3.6. A conceptual line was drawn from the sun, through the center of the scene and on through the concentric circular flight paths described above. Then, the sensor azimuth angle was varied as shown in Figure 13 by sampling angles $\pm 20^\circ$ from the principal plane formed by the sun, the point designated as the center of the scene, and the ray representing the specular reflection.

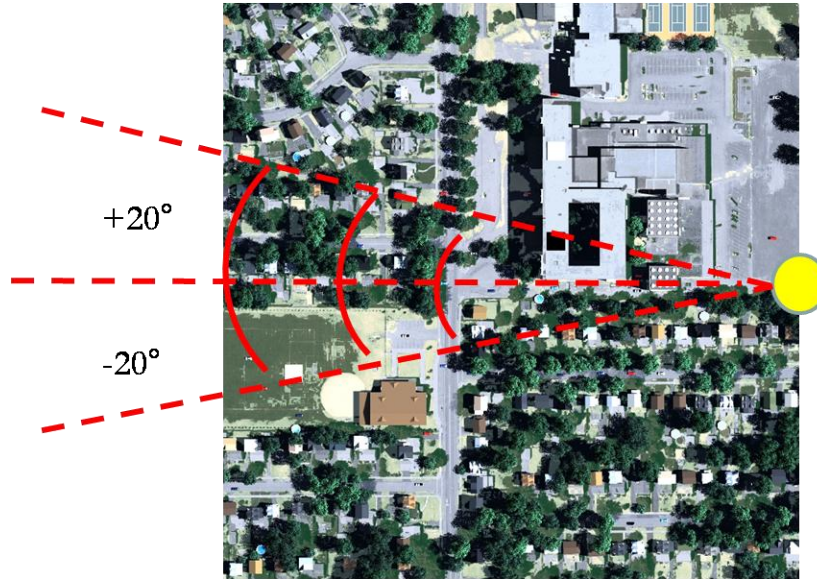


Figure 13. For each time of day, the sensor azimuth angles of interest were $\pm 20^\circ$ from the line representing the specular reflection coming from the center of the scene. The red arcs represent circular flight paths with varying sensor zenith angles.

The azimuthal angles were sampled in 5° increments, producing nine distinct images for each zenith angle. An additional control image was captured at a relative azimuth angle 90° from the sun, for a total of ten distinct azimuthal images per zenith angle. Finally, the azimuthal images for each of the seven zenith angles were captured at each time of day. Including one nadir image for each time of day required that a total of 355 DIRSIG images be rendered for analysis.

Varying the SNR of either sensor should affect the system performance. To explore the fusion best case scenarios, the polarimetric sensor will be fixed at $\text{SNR} = 200$ for all cases. The AVIRIS and Hyperion systems described in Sec. 3.2 provide an estimate of reasonable SNRs for hyperspectral systems, ranging from just under 40 to almost 200. Therefore, the spectral sensor SNR will be cycled through a series of values: 50, 60, 80, 100 and 200.

The sensor SNR was varied during post-processing by adding scaled zero-mean Gaussian noise to each band in a noise-free DIRSIG image, as shown in Eqs. 25 and 26, thereby eliminating the need to render more synthetic images in DIRSIG.

$$\sigma_N = \frac{SignalAvg}{SNR} \quad (25)$$

$$NoiseImage_{ij} = Image_{ij} + \sigma_N(randn) \quad (26)$$

In Eq. 25, *SignalAvg* represents the average radiance value across all pixels captured by a particular spectral band and σ_N represents the standard deviation of the noise in that same band. The pixel value in the *i*th row, *j*th column of the noise image was generated by appropriately scaling a random number drawn from the zero-mean standard normal distribution and adding the noise value to the radiance observed by the pixel in the *i*th row, *j*th column of the noise-free DIRSIG image.

In summary, five times of day were considered, ranging from 0600 – 1200. Also, seven sensor flight paths were modeled for each time of day, with ten images acquired during each one, sampling angles $\pm 20^\circ$ away from the specular reflection with one image acquired at a 90° azimuthal angle from the sun. In addition, one nadir image was rendered for each time of day. With the current computational resources available in the Digital Imaging and Remote Sensing Laboratory, each of the 355 DIRSIG images was rendered in about twelve hours, and fifteen instances of DIRSIG can reasonably be executed in parallel. As a result, the entire image set took roughly two weeks of constant computation to create. Finally, five SNR values for each sensor were considered at every image acquisition point. However, in addition to variations in performance induced by

illumination angle, sensor orientation and sensor SNR values, the multimodal sensor fusion process will also be impacted by registration errors.

4.6 Simulating Registration Errors

In any realistic multimodal system, the images captured by each different sensor modality will not be exactly aligned. Although a variety of techniques exist to geometrically register images obtained from different perspectives, the images generally will not align perfectly due to parallax effects [Schott 2007]. The high spatial resolution hyperspectral, multispectral and polarimetric synthetic images generated by DIRSIG for each sensor modality were exactly aligned since common platforms were assumed, so some amount of registration error was induced by shifting the polarimetric image right and down as shown in Figure 14 below.

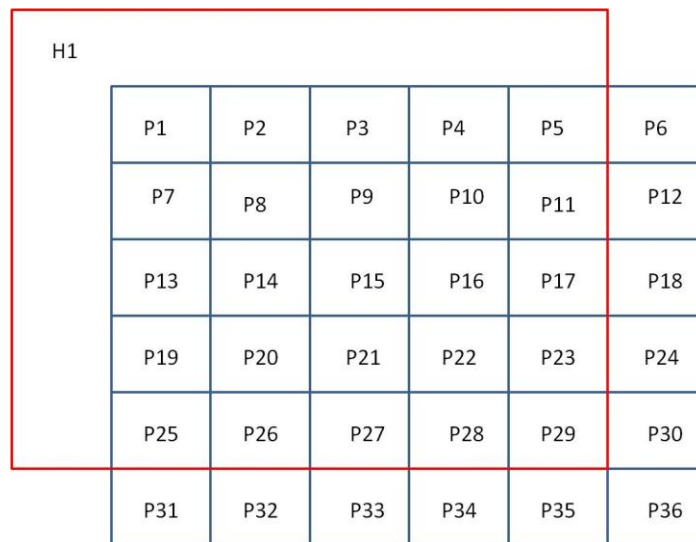


Figure 14. Registration error introduced between high spatial resolution polarimetric pixels, labeled P1 – P36, and low spatial resolution hyperspectral super pixel labeled H1.

Shifting the polarimetric image as shown in Figure 14 ensured that the polarimetric image was offset by the same distance on the ground from both the multispectral and hyperspectral images, but the pixel offset was different because of the

larger hyperspectral GSD. In the multispectral case, the registration error was a shift of one pixel in x and y, while in the hyperspectral case the polarimetric information was offset by 1/6 of a hyperspectral super pixel. The small amount of intentionally introduced registration error was meant to simulate any residual error that might be encountered after a registration routine was applied to actual data. The intent of this project was not to investigate optimal registration techniques—rather, the goal was to acknowledge that registration errors can occur and may decrease the effectiveness of data fusion to some degree.

4.7 Compensating Atmospheric Effects

This project sought to apply a nominal atmospheric compensation technique to the spectral images rather than determine the optimum atmospheric compensation technique. ELM was the most attractive atmospheric compensation technique described in Sec. 3.9, because it was easily automated and required minimal computational effort. Since part of this trade study involved varying the sensor view angle and the time of day, the atmospheric effects varied from one image to the next. Therefore, whatever atmospheric compensation technique was chosen must be executed hundreds of times, producing slightly different effects for each set of conditions, and ELM was particularly suited to this task.

To implement ELM with synthetic DIRSIG data, three 12 m × 12 m calibration panels of 0%, 50% and 100% reflectance were inserted into the image. The large size of the panels, relative to the nadir multispectral GSD, ensured the calibration panels span about 50 full pixels at a sensor declination angle of 65°. However, at the larger

hyperspectral GSD, the calibration panels may not span a full pixel at oblique angles. Since the goal of this step was to reasonably compensate the data for atmospheric effects, ELM was performed on the high spatial resolution spectral image before degrading it to the hyperspectral GSD. The calibration panels were designed to be perfectly Lambertian, with a constant reflectance across the spectral range of interest. Two of the panels were mounted on the flat roof of a large building in the center of the scene, while the third was placed in an open field. Each location had a shape factor $F \cong 1$, satisfying the assumption made for ELM in Sec. 3.9. Spatially dispersing the panels ensured at least two panels were likely to be captured in each image, regardless of sensor view angle.

When each image was analyzed in post-processing, the pixels containing calibration panels were identified via DIRSIG's truth data. The known reflectance values and corresponding sensor reaching radiances for each of those pixels were then plotted, and a best-fit line identified the slope and intercept as described in Sec. 3.9. Finally, the effective Lambertian reflectance of every other pixel in the image was determined using the slope, intercept, and sensor reaching radiance. Once the image had been converted to the reflectance domain, it was analyzed with spectral target detection algorithms.

4.8 Fusing Sensor Outputs

Two methods were explored to fuse the data in this project. First, in the case of co-located multimodal sensors, the data were fused at the pixel level using elements drawn from the pan-sharpening technique discussed in Sec. 3.3 [Price 1986] and the combination logic for hyperspectral and DOLP data described in Sec. 3.7 [Cavanaugh *et al.* 2006]. Then a different algorithm, based on the LOP and LOGP algorithms described

in Sec. 3.4, was applied to fuse the sensor data at the decision level for both the co-located and dispersed cases.

4.8.1 Pixel Level Fusion

The data from the co-located sensors was fused at the pixel level by treating the DOP described in Sec. 3.6 as an additional spectral band. Specifically, the DOP values from a given image were treated as an additional band of spectral reflectance data as shown in Figure 15.

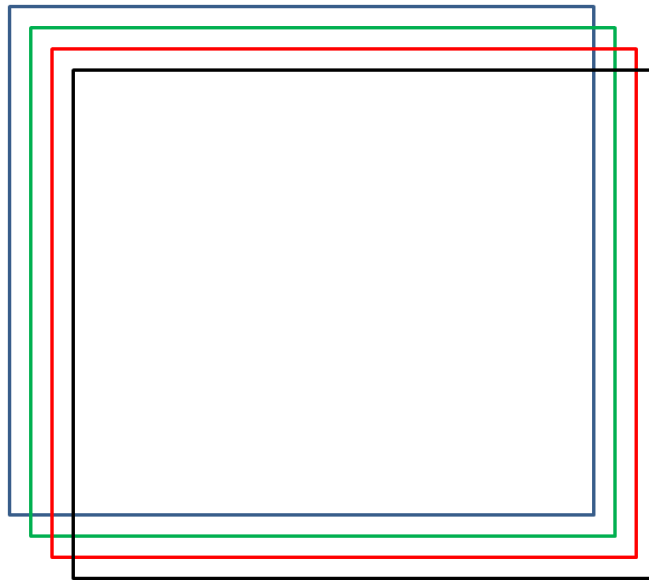


Figure 15. In the pixel fusion algorithm, DOP (black) will be conceptually stacked onto existing spectral bands (shown as RGB above) to create the fused data set.

Normalizing the DOP values by the peak value ensured that all the data from the pixel fusion case was constrained between zero and one. As a result, the Spectral Polarimetric Optimization Tool (SPOT) pixel fusion algorithm was essentially the CEM algorithm from Sec. 3.10 operating on a customized target vector as shown in Eq. 27:

$$(27)$$

where \mathbf{x} was the vector of polarimetric and reflectance values representing the target signature.

In the multispectral case, the spectral GSD exactly matched the polarimetric GSD, so the data were simply combined via stacking. In this case, \mathbf{x} was the vector of polarimetric and effective Lambertian reflectance values for each spectral band as shown in Eq. 28:

(28)

If the images were acquired under favorable polarimetric sensing conditions, man-made objects were expected to be among the most polarizing objects in the scene. Therefore, the value sought in the DOP band with the CEM algorithm were set to one—identifying objects with the highest degree of polarization as those most likely to be targets. The target reflectance values were merged with the desired DOP value to produce the target signature \mathbf{t} as shown in Eq. 29:

(29)

However, in the hyperspectral case, an additional issue had to be considered. The hyperspectral 3.0 m nadir GSD didn't match the polarimetric 0.5 m nadir GSD. Therefore, the hyperspectral super pixel will be spatially resampled by dividing it into 36 sub-pixels to match the polarimetric GSD, as shown previously in Figure 14. Although

each sub-pixel will have the same spectral reflectivity values, the sub-pixels will align with different DOP values to produce the pixel value as shown in Eq. 28.

4.8.2 Decision Level Fusion

Alternatively, since the two sensing modalities tested in this project measure different physical phenomena, their outputs were also combined via a decision level fusion technique as previously described in Sec. 3.4. As a result, data from each sensor was analyzed independently and then merged to enhance the probability of detection at a fixed probability of a false alarm. The spectral data was analyzed with the CEM algorithm described in Sec. 3.10, producing a score for each pixel between zero and one.

Conversely, each pixel of polarimetric data was evaluated using the TAD algorithm from Sec. 3.10, also producing a score between zero and one. The TAD algorithm essentially differentiated polarimetrically anomalous pixels from the urban background, rather than attempting to match pixels to a known polarimetric signature. TAD was attractive because of the relatively few targets in the scene and the difference in polarimetric signature between man-made vehicles and the background. After a TAD score was calculated for each pixel, the TAD output was mapped onto the spectral data. As Figure 14 shows, one hyperspectral pixel with GSD = 3 m represents a 6×6 pixel region of polarimetric data with GSD = 0.5 m. Therefore, each hyperspectral super pixel was divided into 36 pixels, all with the same CEM score, but with different TAD scores as described above. Since the multispectral GSD was identical to the TAD GSD, the multispectral CEM score for each pixel was combined with the corresponding TAD score. The registration error described in Sec. 4.6 meant pixels on the image boundary may not have both a TAD and CEM score. Since the TAD data was being mapped onto

the spectral data, any TAD scores shifted off the image to the right or down were ignored. The spectral pixels along the top and left edges of the images were assigned the TAD score of their nearest neighbor.

Since the focus of this project was to reduce the number of false alarms for a given spectral sensor, the outputs of the analysis above were exploited via a variation of the LOGP discussed in Sec. 3.4. Pixels were thresholded according to their overall spectral polarimetric integration (SPI) score, shown in Eq. 30:

$$SPI(\mathbf{x}) = [CEM(\mathbf{x})][CEM(\mathbf{x}) + TAD(\mathbf{x})] \quad (30)$$

where SPI scores above some user-defined threshold were declared targets. The first term in the SPI algorithm simply represented the likelihood of a given pixel being a target from the multispectral or hyperspectral sensor's point of view. The right-most term in brackets from Eq. 32 represented a way to incorporate information from the polarimetric sensor while preventing the polarimetric sensor from vetoing targets nominated by the spectral sensor.

If a target is defined as a pixel containing part of a vehicle with a particular type of paint, the spectral sensor is completely capable of identifying targets using the CEM algorithm under ideal illumination and viewing conditions, unlike the polarimetric sensor using TAD. Under poor polarimetric sensing conditions, little signal reaches the sensor and polarimetric anomalies are unlikely to be found—in which case the SPI algorithm essentially produces a squared CEM score, rather than vetoing all pixels. Alternatively, when polarimetric sensing conditions are ideal, the spectral sensor identifies pixels as potential targets, and then the polarimetric sensor modifies that ranking to place extra emphasis on anomalous pixels. The impact of incorporating additional polarimetric

information varies across a variety of viewing conditions, and ROC curves were used to quantify the effect.

4.9 Quantifying the Impact

For both pixel and decision level fusion, enhanced performance with the SPOT or SPI algorithm was defined as a higher probability of detection than that obtained with the CEM algorithm at a fixed probability of false alarm. Since a larger area under the ROC curve represented an enhanced probability of detection, the measure of merit was the percent increase in area under the ROC curve when comparing the SPOT or SPI algorithm to the CEM algorithm, for a specific viewing geometry, time of day, and SNR. However, the ROC curve behavior at extremely high false alarm rates was irrelevant for any practical application, so the integration was limited to the point where the false alarm rate met some user-defined threshold. Any viewing conditions where no targets were detected at that point were deemed not useful, noted and discarded from further analysis.

Fusing the data as described in Sec. 4.8 meant that one false alarm represented a $P_{FalseAlarm}$ of 1.7×10^{-6} , ensuring enough distinction to accurately capture behavior of the target detection algorithms being analyzed. The first step in constructing the ROC curves was to determine the number of targets actually visible in the scene. DIRSIG truth data was used to determine the dominant material present in each high-spatial resolution image pixel. Recalling from Sec. 4.4 that each pixel in the synthetic image will be generated via 2×2 oversampling, a pixel at least 50% filled with target material was flagged as a target in the DIRSIG truth data.

Once the number of targets (expressed as target pixels) present for a particular viewing geometry was determined, the pixels were separately scored by the SPOT, SPI and CEM algorithms. For both the multispectral and hyperspectral case, scoring occurred at the smaller multispectral / polarimetric GSD. The SPOT and SPI algorithms produced a score map at the desired GSD, but the hyperspectral CEM score map needed to be resampled to the same number of pixels as the SPOT and SPI scores. The scored pixels were then classified as either targets or background. By thresholding and reducing the threshold until the desired number of targets was found, the number of false alarms generated was determined. The values required for a ROC curve were calculated as shown in Eqs. 31 and 32:

$$DR = \frac{N_{TF}}{N_T} \quad (31)$$

$$FAR = \frac{N - N_{TF}}{N_{image}} \quad (32)$$

where DR , the detection rate, was the rate of correctly identifying a target, N_{TF} was the number of target pixels found, N_T was the total number of target pixels in the image, FAR , the false alarm rate, was the rate of labeling a background pixel as a target pixel, N was the number of pixels flagged as potential targets, and N_{image} was the total number of pixels in the image. With this method, a target completely obscured by buildings for a particular viewing geometry did not show up as a missed target, since no information about that target ever reached the focal plane. As a result, N_T may have changed from image to image, which should not have posed a problem given the large number of target pixels as described in Sec. 4.4.

The area under the SPOT, SPI and CEM ROC curves was calculated by numerically integrating, via the trapezoidal method, the curve's DR values from the minimum FAR value to the FAR threshold described above. In the case that no measured DR value existed for the FAR threshold value, an estimated DR value was calculated by assuming the ROC curve was piece-wise linear and interpolating between the nearest neighbor data points as shown in Figure 16.

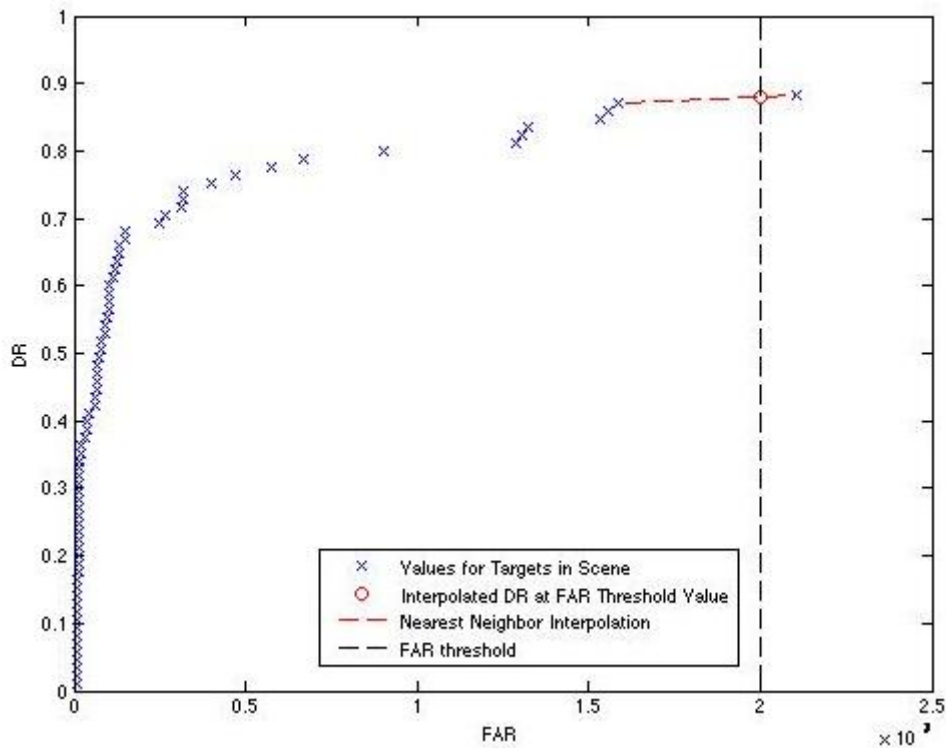


Figure 16. When the area under a ROC curve was integrated to a user-defined threshold (vertical black dashed line), an interpolated detection rate (red circle) was determined by assuming the ROC curve is piecewise linear (red dashed line) between the two nearest measured values (blue x).

The governing measure of merit for assessing the performance of the fusion algorithms was a ratio, R , of the test algorithm performance to the multispectral or hyperspectral CEM algorithm performance when applied to the nadir image for the same time of day as shown in Eq. 33:

The test area metric, A_T , was the area under the ROC curve generated for each algorithm and viewing geometry under evaluation. The nadir area metric, A_N , was the area under the ROC curve generated by applying the CEM algorithm to the nadir image for the same time of day as A_T . The ratio R therefore assessed the off-nadir multispectral (or hyperspectral) CEM, SPOT and SPI performance in terms of the nadir multispectral (or hyperspectral) CEM performance.

Since the random noise added to the image as described in Sec. 4.5 changed the appearance of pixels in the image, the target-like rankings of each pixel produced by the various algorithms varied for different patterns of random noise. As a result, the area under the ROC curves varied from one application of noise to another for the same image. The image therefore had to be analyzed several times to ensure accurate algorithm improvement metrics were reported. After each iteration, the standard deviation of the area metrics was calculated and compared to the standard deviation from the previous iteration. On the n th iteration, once the standard deviation for all area metrics changed less than 3%, each of the percent improvement metrics was defined as the median of the collection of the n individual values. The median value was used to account for possible values of infinite percentage increases resulting from zero area under a given ROC curve.

If any of the target detection algorithms failed to find a target before the FAR threshold value was reached, the area under that particular ROC curve equaled zero and that algorithm was declared not useful for the given viewing geometry. However, a

situation could arise where one algorithm was not useful, but another might have had some area under the ROC curve. The difference in performance was meaningful, but a zero value in the denominator produced an infinite difference in Eq. 33. Therefore, an infinite improvement meant that incorporating polarimetric information via a test algorithm made the algorithm useful, whereas the original nadir CEM baseline scenario was not. The metrics described above were then used to identify the ideal viewing conditions for each sensing modality.

4.10 Independently Optimizing CEM & TAD Geometry

Next, this project assumed that a polarimetric sensor on one platform could be tasked in conjunction with a spectral sensor on another platform to simultaneously acquire imagery of the target scene from different angles as shown in Figure 17.

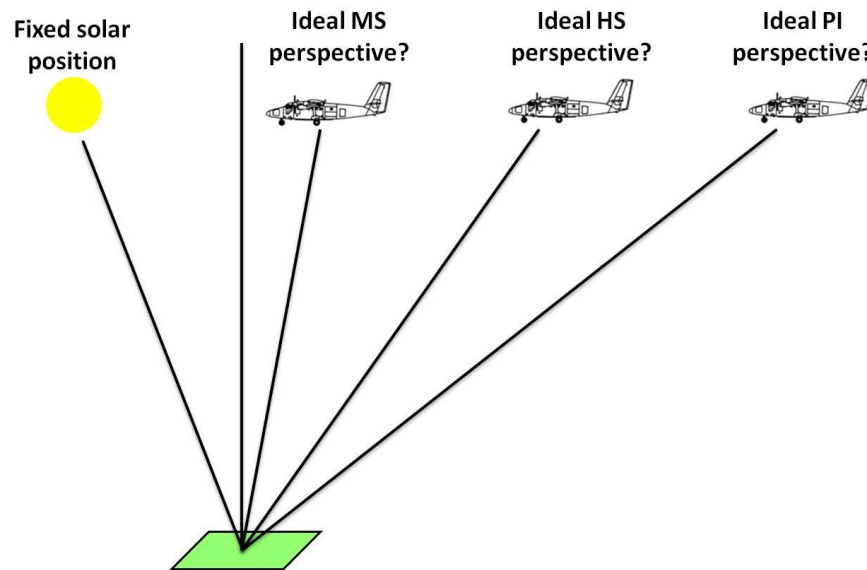


Figure 17. The ideal multispectral (MS), hyperspectral (HS) and polarimetric (PI) sensing geometries were established for each time of day.

The ideal viewing geometries for each sensing modality were determined for each of the five times of day rendered in Sec. 4.5. For the spectral sensor, the optimum viewing

geometry for each time was associated with the point where the area under the CEM ROC curve calculated in Sec. 4.9 was the greatest. This metric was chosen because the CEM algorithm is only dependent on spectral data. In contrast, the optimum viewing geometry for the polarimetric sensor was determined by treating TAD as a target detection algorithm and integrating the area under the ROC curve as described in Sec. 4.9. This metric represented the point where target pixels are most likely to be flagged as anomalies with the TAD algorithm, highlighting where the extra polarimetric information had the most impact on target detection.

4.11 Registering the Images

Once the ideal viewing geometries for each modality at a given time of day had been identified, the corresponding DIRSIG images rendered in Sec. 4.5 and atmospherically compensated in Sec. 4.7 needed to be registered. Although there were only five sets of images, the potential existed for significant amounts of scale, translation and rotation between the images. In addition, at large oblique sensor zenith angles, different elements of the scene had different slant ranges to the sensor, introducing shearing effects.

Given these challenges and a relatively few number of images to register, the images were registered by manually selecting ground control points (GCPs) in ENVI and applying the automated Delaunay triangulation warping technique. As the default option in ENVI, this method fit triangles to the irregularly spaced GCPs and interpolated values to the output grid. To best preserve data in the re-sampled image, the new radiance values for every pixel were obtained from the nearest neighbor for each band of spectral

or polarimetric information. Finally, the total root-mean-square (RMS) error from the GCPs was recorded to quantify the registration quality.

4.12 Establishing the Maximum Attainable Benefit

The spectral information in the newly registered images was atmospherically compensated using the method from Sec. 4.7 prior to applying the CEM algorithm. No atmospheric compensation was applied to the polarimetric data before applying the TAD algorithm. At small zenith angles where the GSD was roughly constant across the focal plane because of the sensor's narrow field of view ($\pm 1.5^\circ$), the polarimetric effects of upwelled radiance appeared as a constant bias. At more oblique angles, the upwelled radiance polarization effects were still smoothly varying across a neighborhood of pixels. In both cases, since TAD sought to identify anomalous pixels, small local variations due to target material and geometry had more impact than larger scale atmospheric variations.

The areas under the ROC curve for the CEM, SPI and SPOT algorithms were calculated as in Sec. 4.9, and the ratios from Eq. 33 for off-nadir CEM, SPI and SPOT algorithms were compared to assess the impact on target detection. Once ratios had been determined for each time of day, the best multimodal spectral and polarimetric sensing conditions were identified. Quantifying the benefit obtained from incorporating polarimetric information under a given set of illumination conditions will help sensor operators determine the opportunity cost of tasking the polarimetric sensor with a separate mission. Additionally, analysts can incorporate the tolerance for false alarms in a given target detection scenario and determine whether incorporating the additional polarimetric sensor is worthwhile. Finally, understanding the maximum attainable

benefit under a variety of illumination conditions will aid in designing requirements for future multimodal sensing systems.

4.13 Section Summary

This section described the essential tasks comprising the multimodal sensor trade study proposed in Sec. 1. First, a reasonable scenario and notional sensor systems were defined, and then a model urban scene was identified. A capability was established to render the scene under a variety of viewing geometries, illumination conditions and SNR values before applying the ELM atmospheric compensation technique. Registration artifacts were accounted for by shifting the polarimetric image one pixel right and down from the high resolution spectral image. Methods were described to fuse the sensor outputs for analysis via the SPOT algorithm at the pixel level and the SPI algorithm at the decision level before a performance comparison to the CEM target detection algorithm. A method to quantify any value added by incorporating polarimetric data with spectral data was described, comparing the areas under the resulting ROC curves. Further, a method to determine the ideal viewing conditions for each sensing modality was identified so that the best-case images from each modality could be geometrically registered. Finally, a process to calculate the maximum attainable benefit from incorporating the second modality via the SPI or SPOT algorithm was described.

5 Preliminary Studies

5.1 Section Overview

Several preliminary experiments were carried out to ensure the validity of the results of this research project. First, the effects of using a global versus local background mean vector and covariance matrix in the CEM algorithm were investigated, and the global values were shown to provide suitable performance while minimizing computational requirements. Next, the spectral and polarimetric sensing modalities were shown to be decorrelated to some extent, ensuring the second sensing modality was providing additional useful information. Then, the spectral signal to clutter ratio (SCR) of the synthetic scene was compared to the spectral SCR value from an actual image, verifying sufficient variation was present in the synthetic background. In addition, the most polarimetrically anomalous materials in the synthetic scene were identified to qualitatively understand the polarimetric clutter in the scene. A study was also conducted to identify a reasonable ROC curve *FAR* threshold given the range of SNR values examined in this study. Next, the effect of employing the SPI algorithm using the RX anomaly detection algorithm instead of the TAD algorithm was assessed. Then, an experiment was conducted to determine whether improvements seen with the fusion algorithms were due to value added by polarimetric information or simply the addition of high-resolution contrast information in the S_0 band. Finally, an experiment was conducted to ensure that the quality of the synthetic polarimetric data was sensitive to different viewing geometries.

5.2 Background and Covariance Matrices

In Sec. 3.10, the CEM algorithm requires a background mean vector, \mathbf{m} , and a covariance matrix, \mathbf{S} , as inputs to assign a score to the pixel of interest. For some applications, using local values instead of global values has been shown to enhance target detection performance, but the results come with an increased computational burden [Caefer *et al.* 2008].

To ensure the covariance matrix is invertible for use in the CEM algorithm, the number of pixels contributing to the covariance matrix must be greater than the number of bands in the image—ideally several times greater [Caefer *et al.* 2008]. The hyperspectral system described in Sec. 4.3 has 90 spectral bands, requiring that any local window around the pixel of interest would measure at least 10 pixels \times 10 pixels. Constantly computing the covariance of this sliding window will significantly increase the already considerable computational resources required for this project. Further, one of the primary concerns with synthetic imagery is rendering spectrally complex backgrounds [Ientilucci and Brown 2003]. The background is more likely to appear spectrally complex as more pixels are included in the covariance matrix calculation, simply because more materials are likely to be represented. Using a global covariance matrix for this project is attractive because it minimizes computation requirements, maximizes the likelihood that the covariance matrix is invertible, and maximizes the spectral variation in the background.

The mean vector, \mathbf{m} , is an estimate of the background spectral characteristics. Essentially, the CEM algorithm is scoring each pixel of interest based upon both how similar the pixel is to a known target spectrum and how different the pixel is from the

background. When a local background vector is used, pixels are averaged from within a sliding window around the pixel of interest, and the pixel of interest is masked out to avoid contaminating the background estimate in case a target is present. Using a 3×3 window around the pixel of interest has been shown to increase performance, while any attempt to include pixels farther away degraded performance [Caefer *et al.* 2008].

Although Caefer's team was focused solely on subpixel targets, the targets in this project may be subpixel size for highly oblique sensor zenith angles but will span multiple pixels at viewing angles near nadir. Therefore, the target material in the pixel of interest may also exist in nearby pixels, contaminating the estimate of the background drawn from a local sliding window.

A study was performed to examine the CEM algorithm performance using three different methods to estimate the background mean vector. First, the global mean value was extracted by considering all the pixels in the image in question. Next, pixels within a 3×3 sliding window around the pixel of interest were considered, while the pixel of interest was masked out. Finally, pixels within a 5×5 sliding window around the test pixel were considered, while all pixels within a 3×3 sliding window around the test pixel were masked out. Two viewing geometries, both at 0800 (solar zenith angle of 55°), were considered. First, the sensor was positioned viewing at nadir, while the second case was an off-nadir scenario with the sensor at a zenith angle of 50° and an azimuth of 175° from the sun—a location predicted to be favorable for polarimetric remote sensing.

First, the nadir viewing geometry was examined. Figure 18 shows the CEM ROC curves for both the multispectral and hyperspectral cases constructed with a global

covariance matrix and a background mean vector derived from the varying sized windows described above.

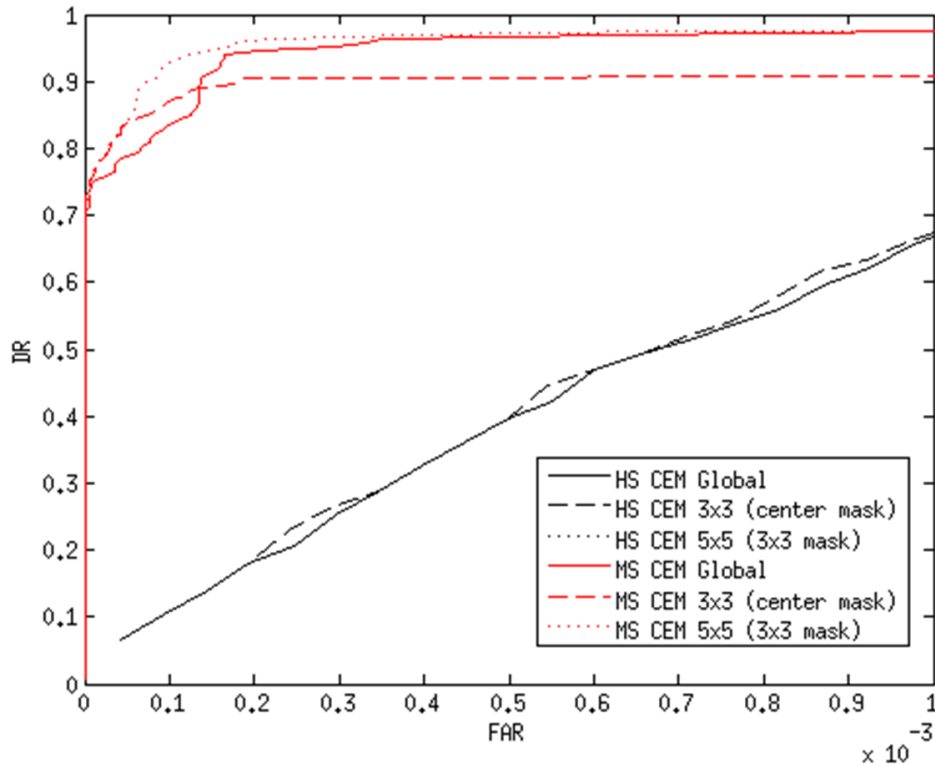


Figure 18. The larger GSD hyperspectral (black) and smaller GSD multispectral (red) CEM ROC curves were created by analyzing the nadir scenario at 0800 with a global covariance matrix. The background mean vector was constructed using the entire scene (solid lines), a 3x3 sliding window with the pixel of interest masked out (dashed lines) or a 5x5 sliding window with the pixel of interest masked out.

Slight differences were observed in the multispectral case. With the 3x3 window, the multispectral performance was degraded as compared to the global mean case or the 5x5 window. The degradation likely occurred because the targets spanned multiple pixels when viewed at nadir, meaning the estimate of the background was contaminated by the target’s spectral signature. In contrast, the hyperspectral performance was less sensitive due to the larger GSD, and the ROC curves essentially overlapped.

The other viewing geometry examined was a sensor zenith angle of 50° and a sensor azimuth angle from the sun of 175°. Figure 19 shows the CEM ROC curves for

both the multispectral and hyperspectral cases constructed with a global covariance matrix and a background mean vector derived from the varying sized windows described above.

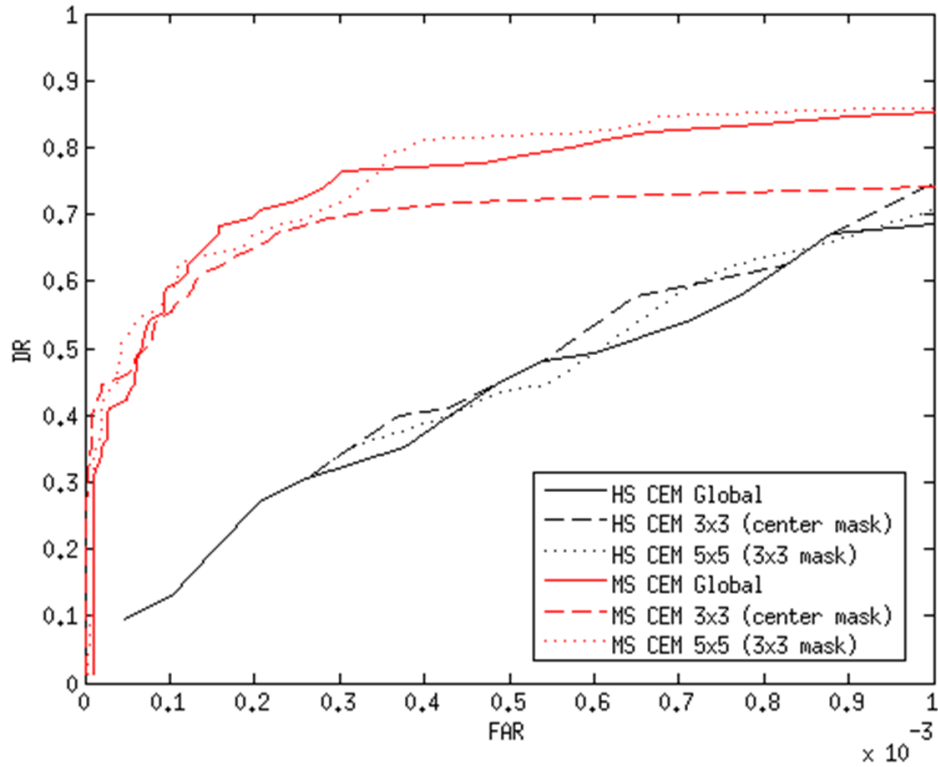


Figure 19. The larger GSD hyperspectral (black) and smaller GSD multispectral (red) CEM ROC curves were created by analyzing the off-nadir scenario at 0800 with a global covariance matrix. The background mean vector was constructed using the entire scene (solid lines), a 3x3 sliding window with the pixel of interest masked out (dashed lines) or a 5x5 sliding window with the pixel of interest masked out.

The off-nadir CEM ROC curve trends mirrored the trends observed at nadir, with the multispectral performance again degraded when using the 3x3 window as compared to the global mean or the 5x5 window. Although the GSD was increased due to the off-nadir viewing geometry, the multispectral GSD was still small enough that the target spectral signature spanned several pixels and contaminated the estimate of the background. In the hyperspectral case, the ROC curves were quite similar due to the significantly larger GSD.

Figure 20 shows the off-nadir SPI ROC curves for both the multispectral and hyperspectral cases using a global covariance matrix and a background mean vector derived from the varying sized windows described above. Again, more variation was observed in the smaller GSD multispectral case than in the larger GSD hyperspectral case—and this was expected, since the SPI algorithm leveraged the results of the CEM algorithm.

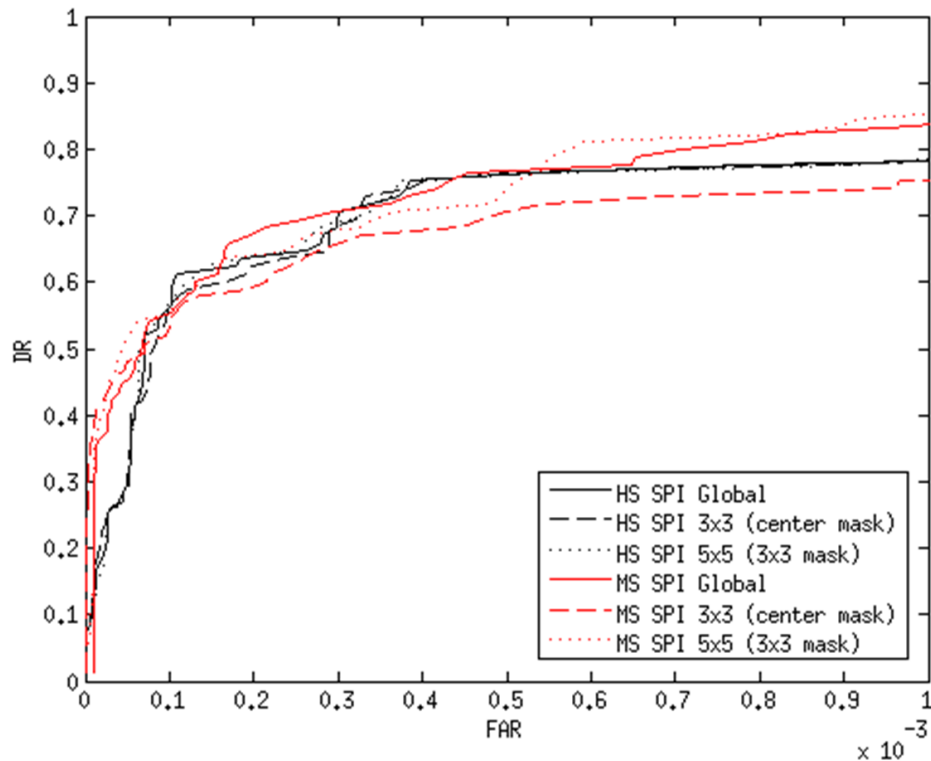


Figure 20. The larger GSD hyperspectral (black) and smaller GSD multispectral (red) SPI ROC curves were created by analyzing the nadir scenario at 0800 with a global covariance matrix. The background mean vector was constructed using the entire scene (solid lines), a 3x3 sliding window with the pixel of interest masked out (dashed lines) or a 5x5 sliding window with the pixel of interest masked out.

Figure 21 shows the off-nadir SPOT ROC curves for both the multispectral and hyperspectral cases using a global covariance matrix and a background mean vector derived from the varying sized windows described above. In stark contrast to the CEM or SPI algorithms, the SPOT pixel fusion performance proved quite sensitive to the way

the background mean vector was constructed, and using a global mean value far outperformed analyzing a limited sliding window. The dramatic difference in performance between the CEM and SPOT algorithms suggested that windowing the polarimetric information negatively impacted performance. When a sliding window was implemented, the variation in polarimetric information used to calculate the mean vector was dramatically reduced—implying that a wide-area perspective was required to accurately represent the polarimetric clutter of the urban scene.

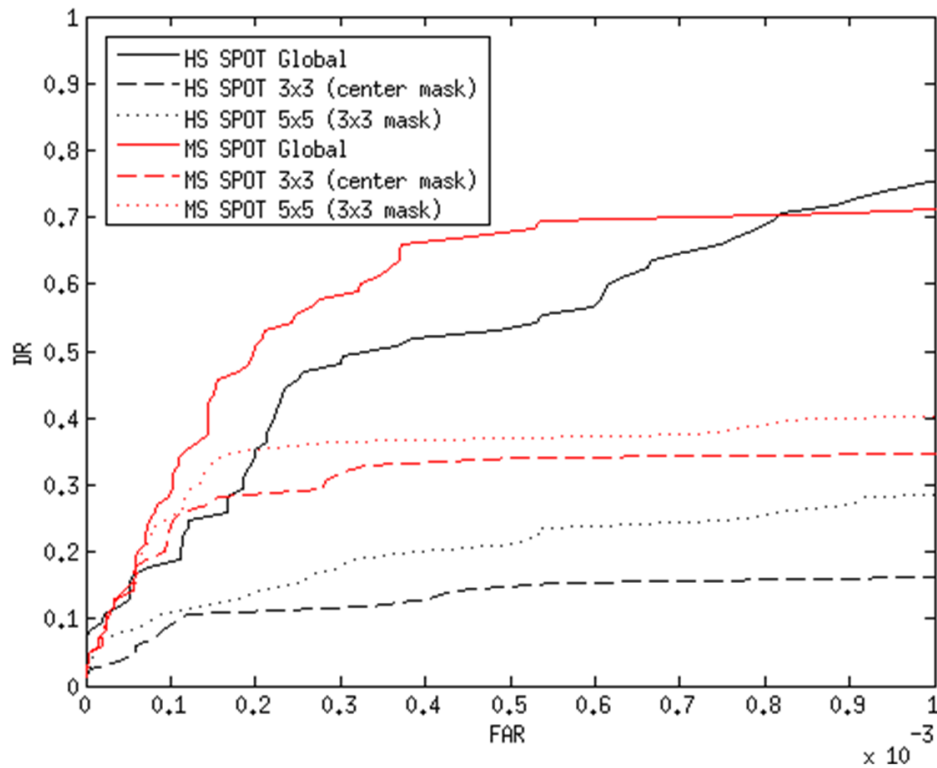


Figure 21. The larger GSD hyperspectral (black) and smaller GSD multispectral (red) ROC curves were created by analyzing a sensor zenith angle of 50° and a sensor azimuth angle of 175° at 0800 with a global covariance matrix and local background mean vector drawn from a 5×5 neighborhood with a 3×3 neighborhood around the pixel of interest masked out. The solid lines were obtained using the CEM algorithm from Sec. 3.10, while the dotted lines were obtained using the SPI algorithm from Sec. 4.8.

Although slight differences were observed in the CEM and SPI algorithm ROC curves when comparing results derived from the global and local mean estimates, the

local mean estimates did not shift the ROC curves to an order of magnitude lower false alarm rate. Further, windowing the polarimetric data for the SPOT algorithm dramatically hampered performance. In addition, the need to calculate a sliding window and different mean value for each test pixel increased the computational burden significantly. The global mean value was therefore adopted for use on the rest of this project because it provided roughly equivalent target detection rates for CEM and SPI, maximized the SPOT performance, and minimized the computational effort.

5.3 Sensing Modality Correlation

For the sensing modality correlation study, the spectral and polarimetric images were perfectly aligned to avoid any artifacts introduced by registration error. Given the covariance, C_{xy} , of two metrics x and y , a correlation coefficient r_{xy} can be calculated as shown in Eq. 34 below [Papoulis and Pillai 2002].

The multispectral CEM and TAD scores for each pixel were then be compared for a particular viewing geometry, and a correlation coefficient calculated, to determine how well the multispectral CEM score was able to predict the TAD score for each pixel. Further, the chosen viewing geometry was in a favorable location for polarimetric sensing as described in Sec. 3.6 to provide a wide variety of polarimetric signatures in the scene. Figure 22 shows a plot of the multispectral CEM score versus the TAD score for each pixel in the image acquired at 0800, with a sensor zenith angle of 40° , a sensor azimuth angle of 175° from the sun and an SNR value of 200—a case predicted to be favorable for polarimetric sensing.

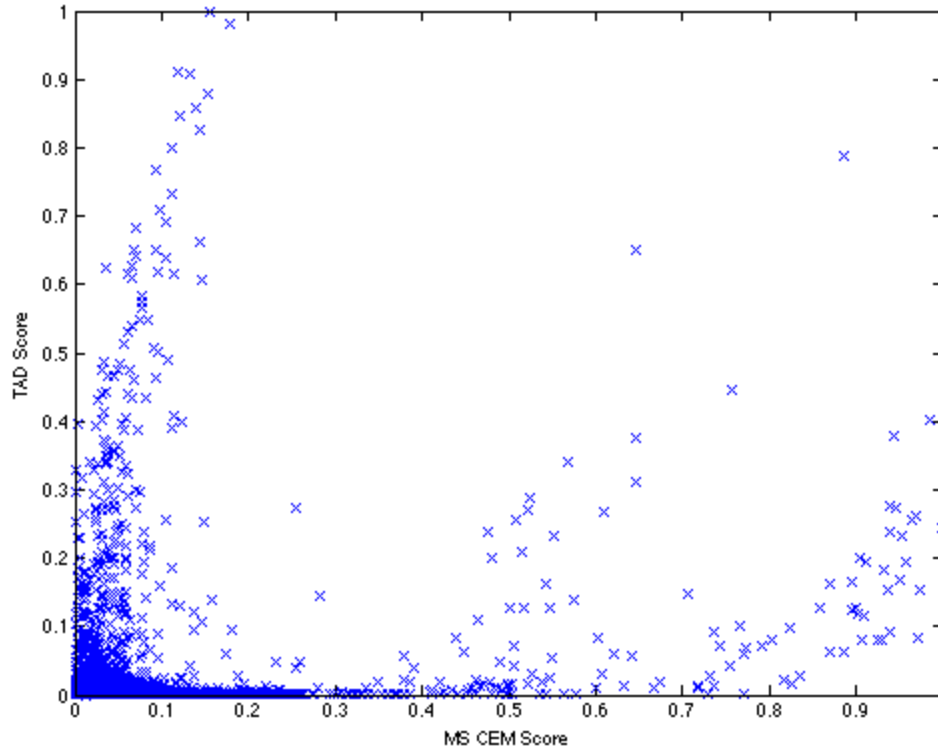


Figure 22. For an image acquired at 0800, with a sensor zenith angle of 40° and a sensor azimuth angle of 175° from the sun, the TAD score for each pixel is plotted as a function of the pixel's multispectral CEM score. Very little correlation exists, since $r = 0.1023$.

Although Figure 22 shows that a wide variety of CEM and TAD scores were assigned, the correlation coefficient for the set of the two scores is only 0.1023, meaning the two score metrics are statistically significantly uncorrelated.

Another way to assess the correlation between CEM and TAD scores is by nominating pixels with a high score as targets with both algorithms, and assessing the results with a ROC curve. As described in Sec. 3.10, CEM is a target detection algorithm, while TAD is an anomaly detection algorithm. Figure 23 shows the resulting ROC curves, confirming that CEM outperforms TAD when applied in the fashion described above.

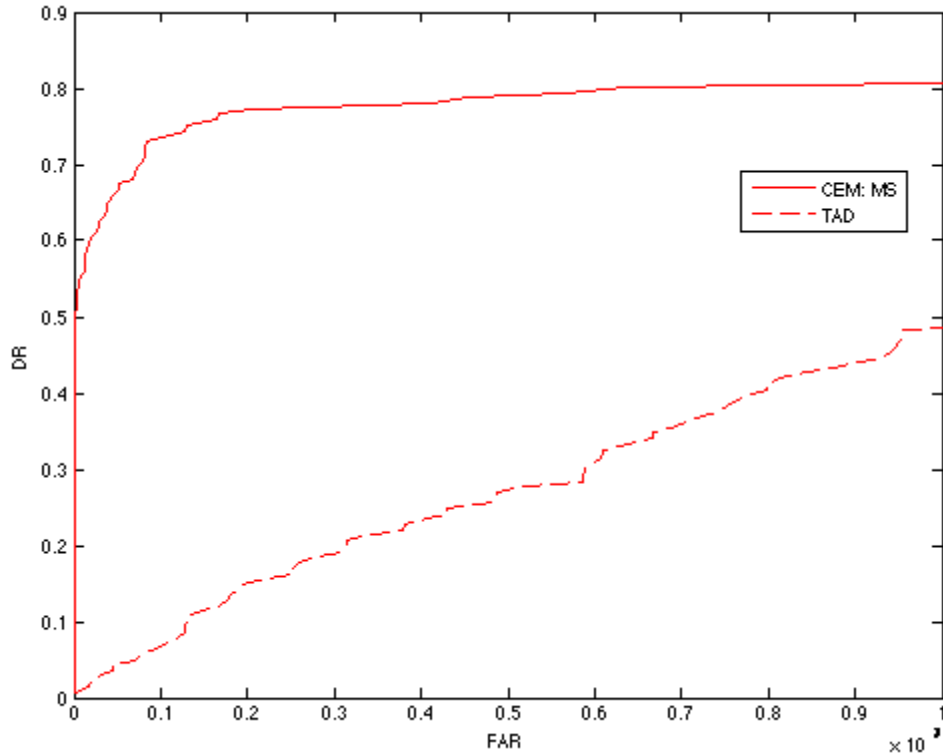


Figure 23. For an image acquired at 0800, with a sensor zenith angle of 40° and a sensor azimuth angle of 175° from the sun, ROC curves generated with either the multispectral CEM or TAD scores for each pixel are plotted, confirming very little correlation exists between the two metrics.

Further, Figure 23 demonstrates that the pixels receiving the highest TAD scores are often not target pixels, confirming that CEM and TAD scores are reasonably uncorrelated.

The uncorrelated nature of the CEM and TAD scores could be a mixed blessing for the fusion algorithms proposed in Sec. 4.8. Pixel level fusion techniques tend to rely on measurements of the same physical phenomenon, where some correlation is present between the two data sets being fused [Price 1986, Pohl and Van Genderen 1998]. Since the polarimetric and spectral measurements in this project measure different physical phenomenon, and are uncorrelated, the pixel level fusion SPOT algorithm performance may suffer. However, decision level fusion effectiveness increases when the classifiers used are uncorrelated [Petraikos *et al.* 2001]. One of the CEM algorithm’s weaknesses is

that very bright pixels will trigger false alarms, because when the brightest pixels are projected onto the target vector, they will still produce a large value. However, while target pixels are likely to be polarimetrically anomalous when compared to the background, a spectrally bright pixel with a very low TAD score most likely represents a bright false alarm. Since the TAD score is used to modify the CEM score in the SPI algorithm, under favorable polarimetric sensing conditions the SPI algorithm will increase the score of target pixels relative to spectrally bright, yet polarimetrically unremarkable false alarms. Therefore, the uncorrelated nature of the CEM and TAD algorithms should enhance the SPI algorithm's performance.

5.4 SCR Calculation

The SCR is a metric used to assess how different the target signal is from the background in an image, as shown in Eq. 35 below [Schaum 2001]:

where \mathbf{t} represents the target signature vector, \mathbf{m} represents the background mean vector, and \mathbf{S} represents the image covariance matrix. The SCR values for the synthetic images were computed for four different types of pixel vectors: 90 hyperspectral bands, eight multispectral bands, one DOP band and one TAD score band. For the 0800 time point, the sensor azimuth angle was fixed at 175° from the sun but the sensor zenith angle was varied from 10° to 70° in 10° increments.

Initially, all SCR values were calculated based on an SNR value of 200, with a synthetic image GSD of 0.5 m. Figure 24 shows the calculated SCR values as a function of the sensor zenith angle. The hyperspectral case (black) hovered around $SCR = 28$

while the multispectral case (red) decayed from SCR = 20 near nadir to SCR = 5 at the most oblique angle. The difference between the two cases demonstrated how incorporating more spectral information through finer spectral resolution helps to identify target pixels. Note that for the DOLP case (dashed blue line) the SCR remained below five for all cases—indicating that DOLP was not a useful metric to identify targets. However, in the TAD case (dotted blue line) the importance of viewing geometry in polarimetric sensing quickly became apparent. Although the SCR started low, a dramatic increase was observed as the sensor zenith angle changed resulting in a peak SCR \approx 20 at a zenith angle of 50°. This was an excellent example of how the TAD algorithm effectively separates anomalous pixels from the background under ideal polarimetric sensing conditions.

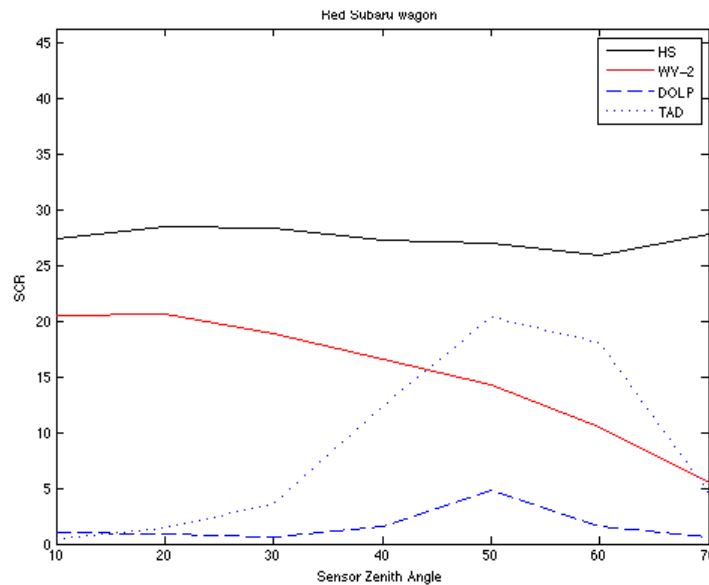


Figure 24. Signal to clutter ratio (SCR) as a function of sensor zenith angle with the sensor azimuth angle fixed at 175° from the sun, the solar zenith angle fixed at 55°, sensor SNR = 200 and GSD = 0.5 m. The hyperspectral case (black) hovered just below SCR = 30, while the multispectral case (red) decayed from SCR = 20 near nadir to SCR = 5 at the most oblique angle. Note that the DOLP case (dashed blue line) varied from about SCR = 2 to SCR = 5 near the sun’s specular reflection, while the TAD case (dotted blue line) peaked at an SCR near 20 at a zenith angle of 50°.

Next, the spectral SCR values from the synthetic imagery were compared to the spectral SCR values from an actual system to confirm that enough background variation was present in the synthetic scene. Using the COMPASS system, RIT acquired hyperspectral imagery of the Rochester, NY area notionally represented by Megascene 1 in this study [Raqueno *et al.* 2005]. Based on the COMPASS system specifications and the imagery collection parameters, the imagery was collected with a GSD ≈ 0.66 m across 256 spectral bands spanning $0.350 \mu\text{m}$ to $2.5 \mu\text{m}$ [Simi *et al.* 2001]. Figure 25 shows an RGB image extracted from the COMPASS collect, where an isolated reddish vehicle can be seen in a parking lot. Several pure vehicle paint pixels were identified, and a spectrum was extracted to serve as a target pixel.

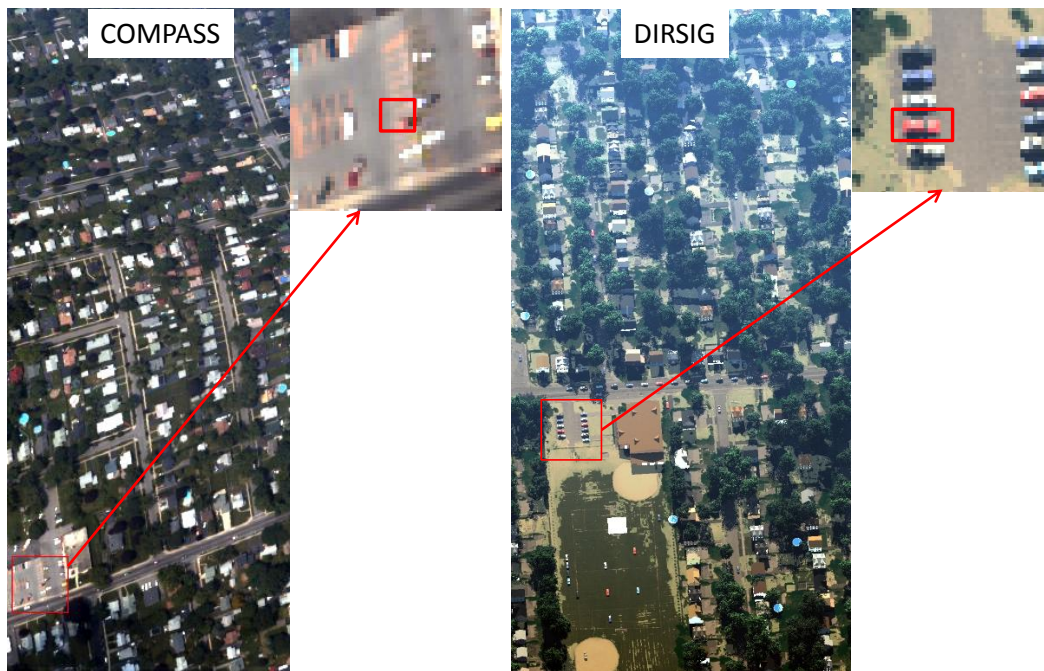


Figure 25. RGB image of 256 band hyperspectral COMPASS image collected over the region of Rochester NY similar to that represented by Megascene 1. A zoomed view shows an isolated reddish vehicle (boxed) used to identify a target pixel in the image. (Right) RGB rendering of DIRSIG's Megascene 1. Several red target vehicles can be observed along with six different models of decoy vehicles.

A spectral SCR = 50.17 was calculated for the COMPASS image by inserting the extracted target spectrum, scene mean value and scene covariance matrix into Eq. 35 without applying any atmospheric compensation.

However, the COMPASS image was acquired with 256 spectral bands, while the hyperspectral imager envisioned in this project only has 90 spectral bands—although both span a similar spectral region. The COMPASS imagery was degraded to 85 bands (to permit a more direct comparison between systems) by summing the contributions from bands 1-3 into a new band 1, summing the contributions from bands 4-6 into a new band 2, and so on, then disregarding information in the last band. The same SCR metric was calculated based on the 85 bands, resulting in $SCR \approx 45$, which was quite similar to the 256 band case. Since in both cases the SCR in the actual hyperspectral image is slightly higher than the SCR in any of the synthetic hyperspectral images examined, while the GSD of the synthetic image is smaller than that of the actual image, the background of the synthetic scenes was determined to have enough spectral variation to ensure the results of this study will be meaningful.

However, an attempt to perform a similar SCR comparison between synthetic and actual polarimetric imagery was curtailed because no data has yet been reported on the scale of the proposed scene with the capabilities of the proposed instrument. A qualitative examination of the synthetic polarimetric clutter was performed instead. The scenario with a solar zenith angle of 55° , a sensor zenith angle of 50° , and a sensor azimuth angle of 175° from the sun represents a scenario with favorable polarimetric viewing conditions. The polarimetric information for each pixel in the scene was scored with the TAD algorithm, and the materials associated with the top 0.5% anomalous pixels

(almost 3,300) were identified. Figure 26 shows a histogram of TAD scores for the target and background pixels identified as the most anomalous, confirming that the target pixels are not the most anomalous materials in the scene.

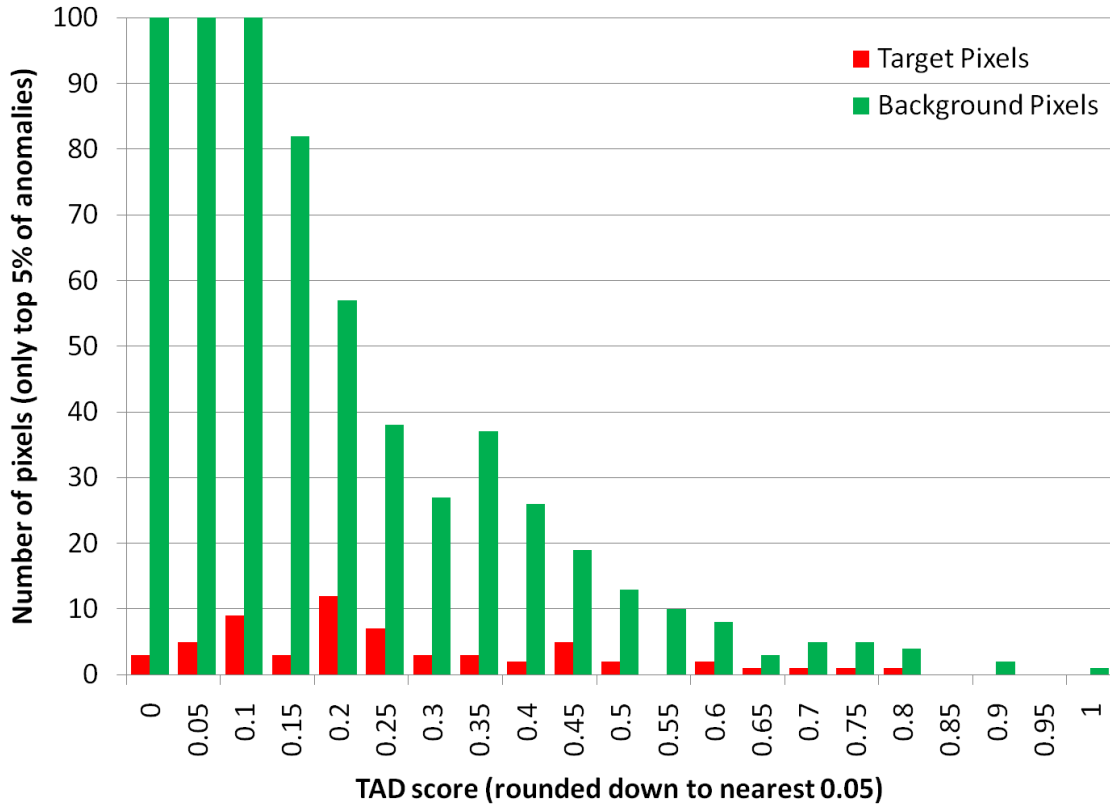


Figure 26. With the solar zenith angle fixed at 55° , the sensor zenith angle at 50° and a relative sensor azimuth angle 175° from the sun, the TAD algorithm was used to identify the most anomalous 0.5% of pixels in the urban scene. The anomalous pixels were identified as either target or background material and plotted as a histogram. The scale for number of pixels has been truncated to enable the target pixel histogram bars to be observed.

Table 4 specifically identifies the materials flagged among the most anomalous pixels, demonstrating that many of the most polarimetrically anomalous materials are background material—roadways, grass and trees. In addition to the vehicles in the scene, the TAD algorithm identifies several other man-made materials—roofing, siding and glass. The variety of materials ranked as the most anomalous therefore gives a qualitative feel of the polarimetric clutter present in the scene.

Table 4. With the solar zenith angle fixed at 55°, the sensor zenith angle at 50° and a relative sensor azimuth angle 175° from the sun, the TAD algorithm was used to identify the most anomalous 0.5% of pixels in the urban scene. The materials associated with those pixels were then tallied to determine which materials were regarded as polarimetrically anomalous. The top 20 materials are shown below.

Number of pixels	Material	Number of pixels	Material
560	Glass	70	Dark blue station wagon
489	Grass (healthy green)	67	Siding (various types)
437	Roadway surface (gray)	63	Black SUV
243	Grass (brown / dirty)	60	Red station wagon (Target)
189	Tree leaf (Silver Maple)	55	Tree leaf (Black Oak)
153	Roofing (various types)	52	Green station wagon
147	Tree leaf (Red Maple)	42	White pickup truck
129	New asphalt (black)	36	Blue sedan
114	Tree leaf (Norway Maple)	35	Swimming pool
83	Aluminum	35	White sedan

5.5 ROC Curve *FAR* Threshold

As described in Sec. 4.9, the fusion algorithm effectiveness metric depends on the area underneath two ROC curves after the user has identified a *FAR* threshold. Setting a particular *FAR* has the ability to influence the metric in two main ways. First, if the *FAR* threshold is too high, the performance of the tail of the ROC curve drives the area metric and may skew the estimate of the algorithm’s effectiveness. For all practical purposes, detections at an extremely high *FAR* are useless since it would take too long to weed out the targets from among the false alarms. Second, if the *FAR* threshold is too low, the chosen algorithm may not find any targets—resulting in a zero value (in the case of the test algorithm) or an infinite value (in the case of the reference nadir scenario). Either a zero or infinite value is of limited use when trying to establish gradations of performance increase. Given these two concerns, the ideal *FAR* threshold lies in a space bounded on

both the upper and lower ends with some room for adjustment to ensure representative behavior was captured for both the test and reference scenarios.

Since all test algorithms are evaluated by comparing to the CEM algorithm at nadir, the nadir viewing scenario was chosen to begin the ROC curve *FAR* threshold study. Since the CEM algorithm performance was expected to decrease as the solar zenith angle increased, the 0600 and 0800 time points were analyzed to determine the lower bound on performance. The area under the ROC curve for both the multispectral and hyperspectral CEM algorithms was calculated for four different SNR values (200, 100, 50 and 15) with five different *FAR* thresholds (0.0003, 0.0007, 0.001, 0.0015 and 0.002). The lowest *FAR* threshold represented the condition where the number of false alarms (176) was almost equal to the total number of target pixels in the scene (166), while the highest *FAR* threshold represented more than 1,000 false alarms in the scene.

A series of bar graphs was created to examine the area under the nadir ROC curve as a function of the chosen *FAR* threshold for a particular SNR value. In the graph, the 0600 case is shown in blue while the 0800 case is shown in red. The multispectral bars are fully filled with color, while the hyperspectral bars are filled by a striped pattern. Figure 27 shows the SNR = 200 case, demonstrating that a *FAR* = 0.0003 barely captures any results from the hyperspectral sensor. However, the relatively small error bars, which represent one standard deviation of the area under the ROC curve, indicate that the displayed value is a typical value after the random noise has been applied.

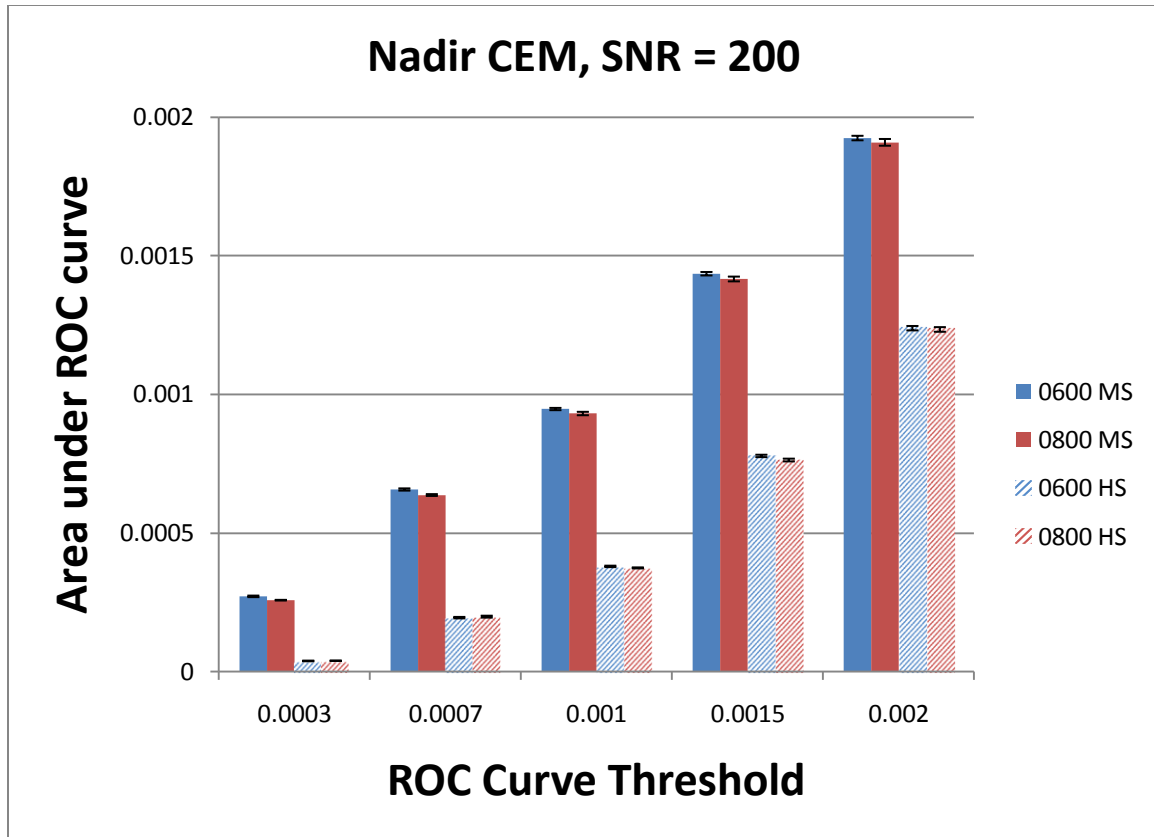


Figure 27. The area under the CEM ROC curve (SNR = 200) is shown as a function of the user-controlled *FAR* threshold for the multispectral (solid) and hyperspectral (striped) scenarios at both the 0600 (blue) and 0800 (red) times of day. Note the error bars, indicating the displayed area under the ROC curve is representative of a typical value after random noise has been applied.

Figure 28 shows the SNR = 100 case, demonstrating that a *FAR* = 0.0003 still barely captures any results from the hyperspectral sensor. The error bars have increased in magnitude, but remain relatively small, still indicating that the displayed value is a typical value after the random noise has been applied.

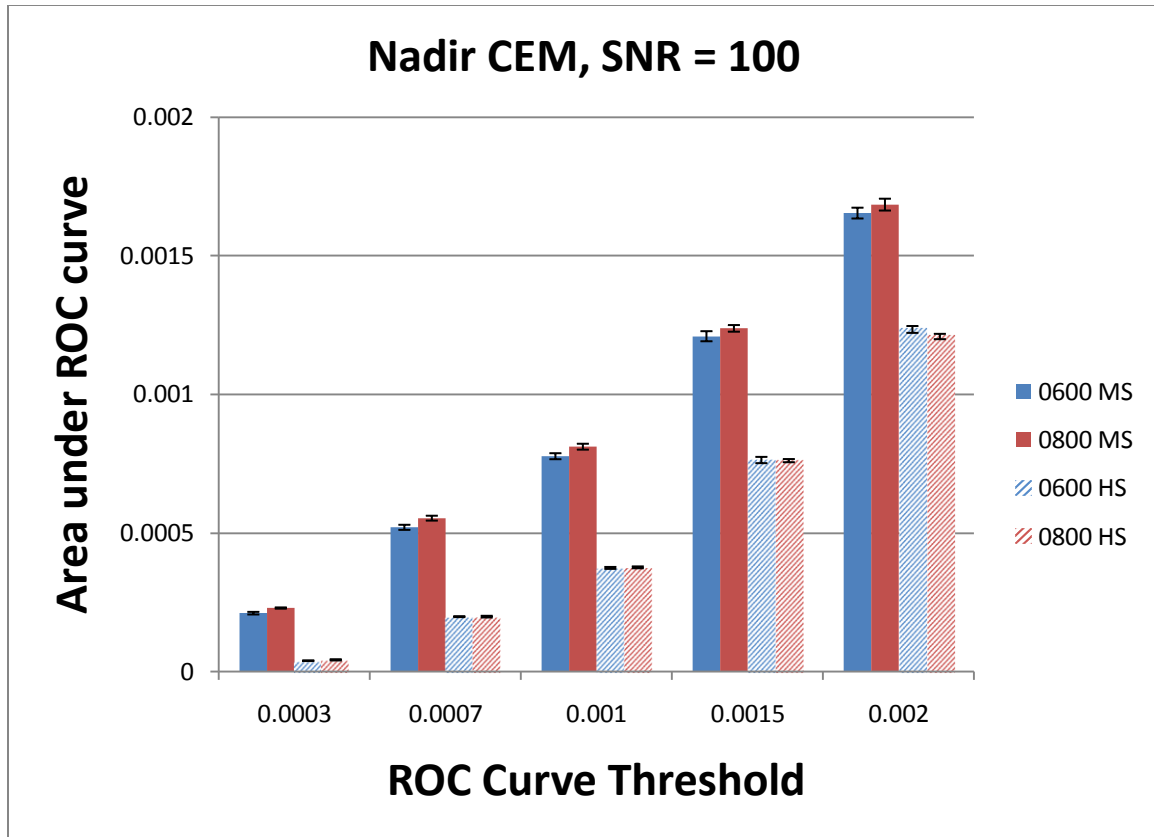


Figure 28. The area under the CEM ROC curve (SNR = 100) is shown as a function of the user-controlled *FAR* threshold for the multispectral (solid) and hyperspectral (striped) scenarios at both the 0600 (blue) and 0800 (red) times of day. Note the error bars, indicating the displayed area under the ROC curve is representative of a typical value after random noise has been applied.

Figure 29 shows the SNR = 50 case, where the area under the multispectral ROC curve hovers near zero until *FAR* = 0.001, effectively imposing a lower bound on useful *FAR* thresholds for this data. The error bars have again increased in magnitude, but remain relatively small, still indicating that the displayed value is a reasonable value after the random noise has been applied.

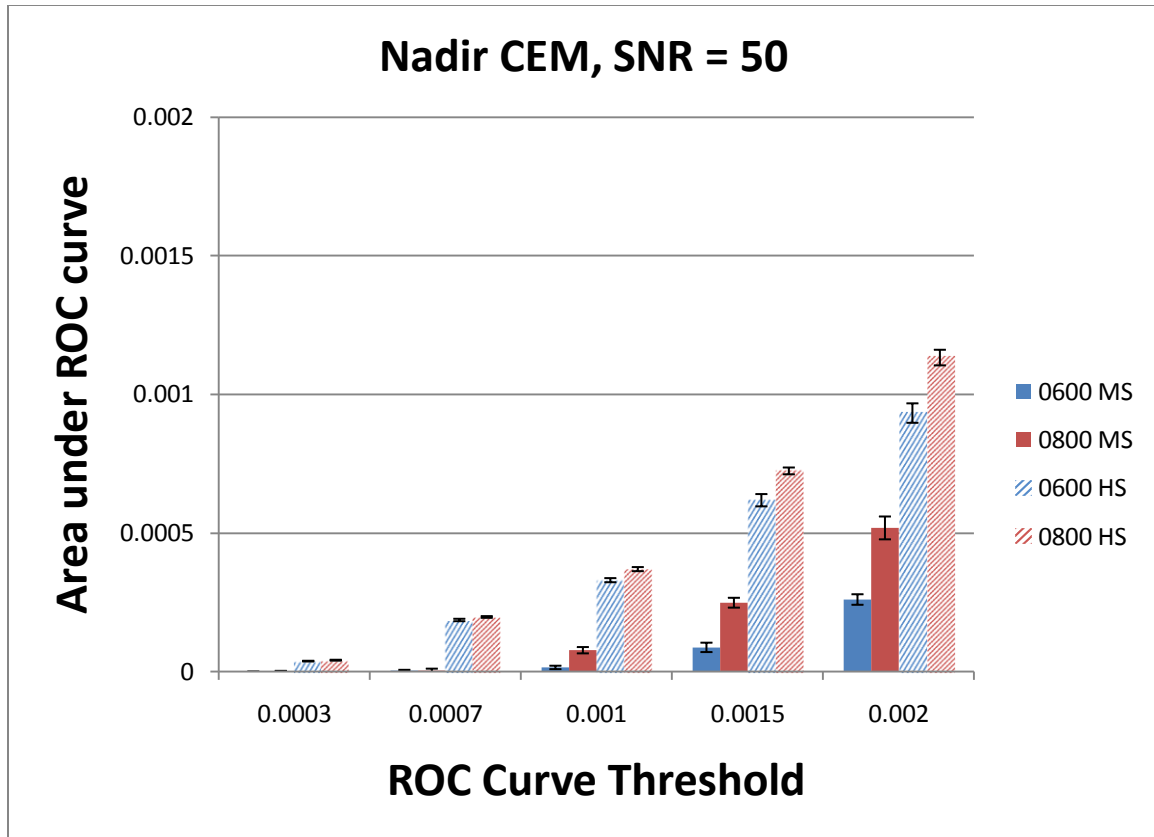


Figure 29. The area under the CEM ROC curve (SNR = 500) is shown as a function of the user-controlled *FAR* threshold for the multispectral (solid) and hyperspectral (striped) scenarios at both the 0600 (blue) and 0800 (red) times of day. Note the error bars, indicating the displayed area under the ROC curve is representative of a typical value after random noise has been applied.

Figure 30 shows the SNR = 15 case. Since the area under the multispectral ROC curve hovers near zero for all *FAR* thresholds, any test algorithm that produced a measurable area under its ROC curve would result in an infinite improvement. As a result, this SNR value is impractical and will be dropped from further analysis. The error bars also indicate significant variation in the calculated area under the ROC curve value after the random noise has been applied, showing the potential for serious error in estimating the performance of test algorithms.

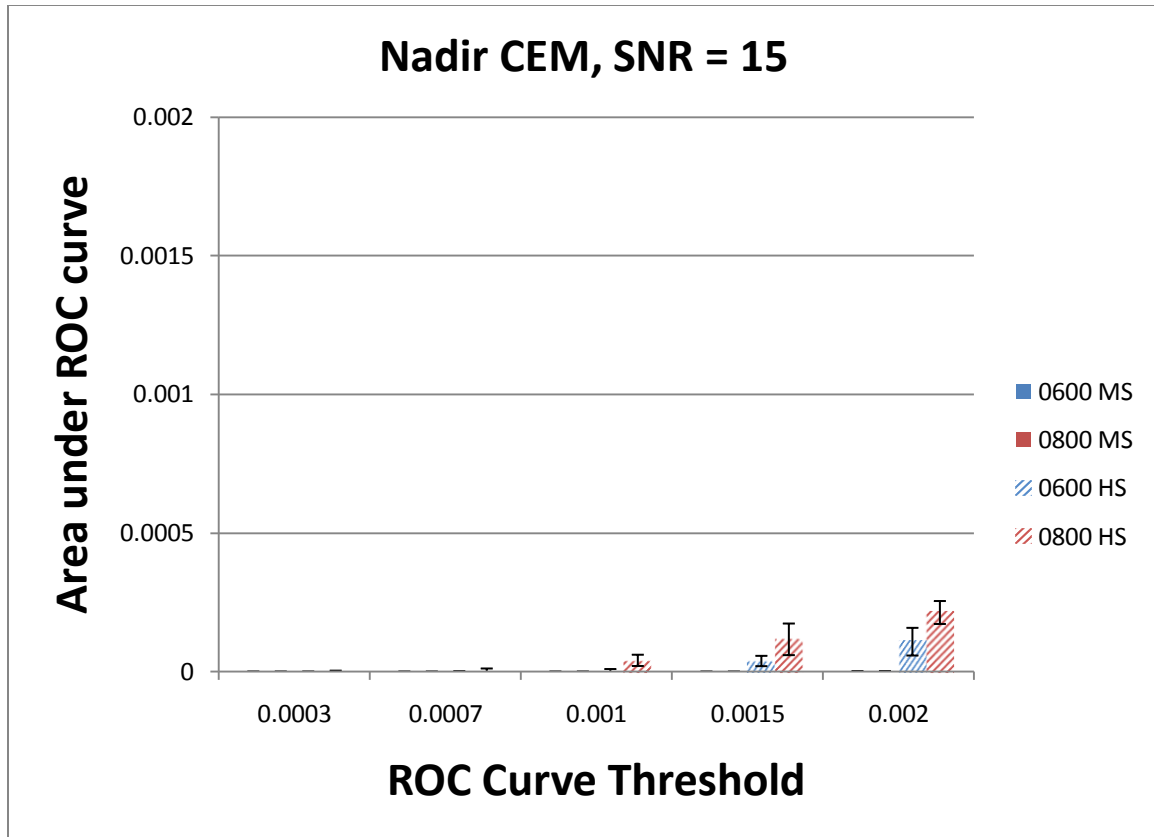


Figure 30. The area under the CEM ROC curve (SNR = 15) is shown as a function of the user-controlled *FAR* threshold for the multispectral (solid) and hyperspectral (striped) scenarios at both the 0600 (blue) and 0800 (red) times of day. Note how the multispectral area hovers near zero for all *FAR* values, making a meaningful comparison to this data point impractical. Further, the error bars indicate a high degree of variability in the area under the ROC curve after random noise has been applied.

The results of the ROC curve *FAR* threshold study indicate that useful SNR values to investigate would be SNR = 200, 100 or 50 rather than SNR = 200, 50 or 15. Further, a *FAR* = 0.001 is the minimum value that will permit a meaningful comparison between the test and nadir algorithms at all SNR values.

Once the *FAR* threshold for the project had been determined, representative multispectral (red) and hyperspectral (black) ROC curves were calculated for the CEM algorithm in the 0800 nadir scenario. The figures below show the ROC curves for SNR values of 200, 100, 75, 60 and 50.

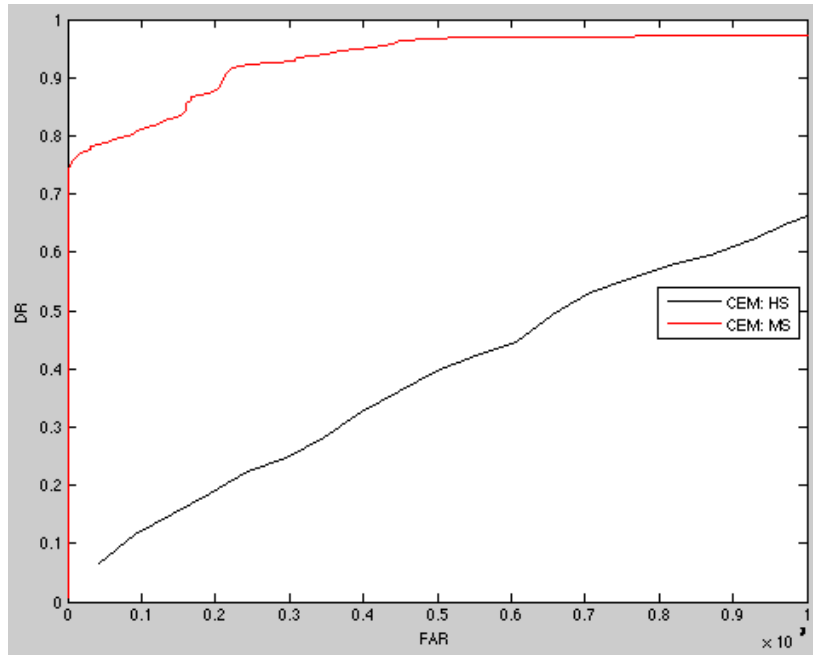


Figure 31. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 200.

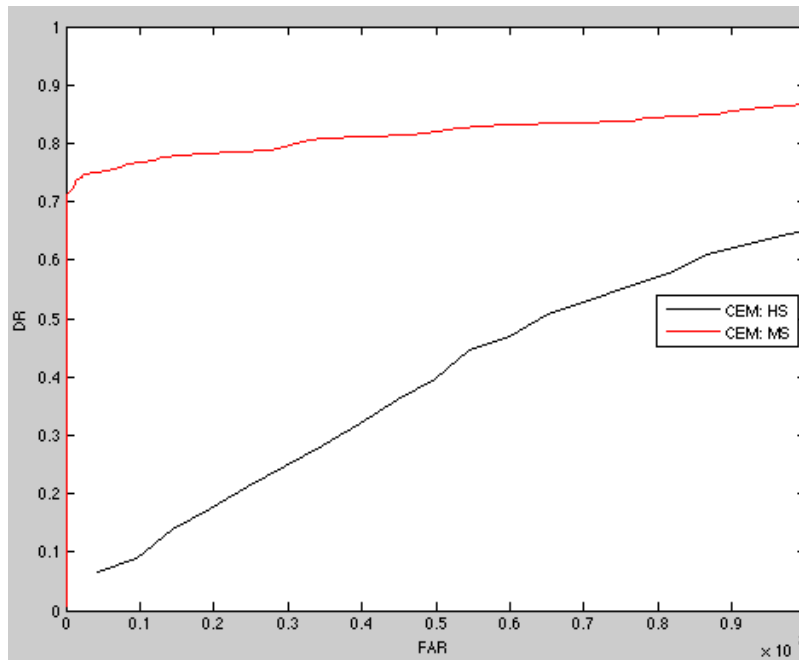


Figure 32. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 100.

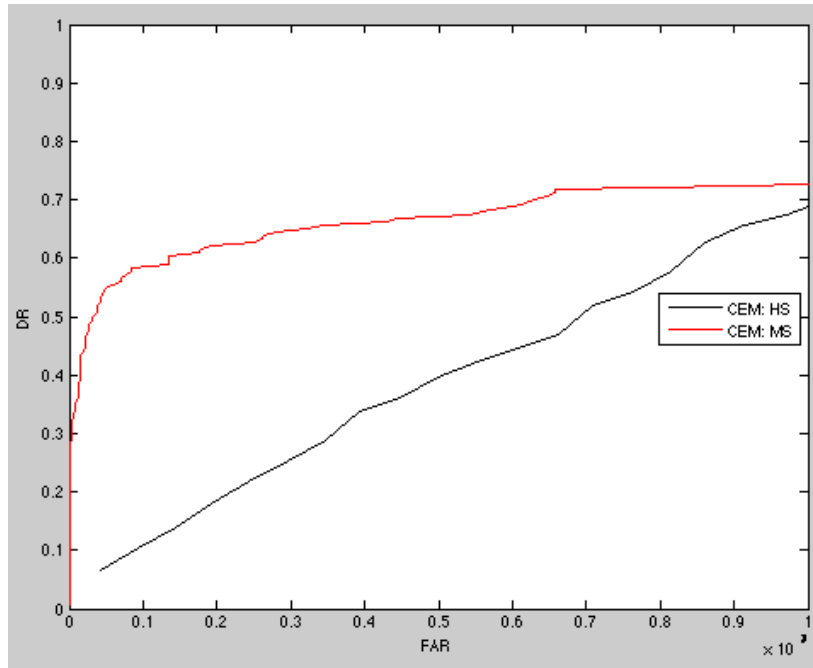


Figure 33. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 75.

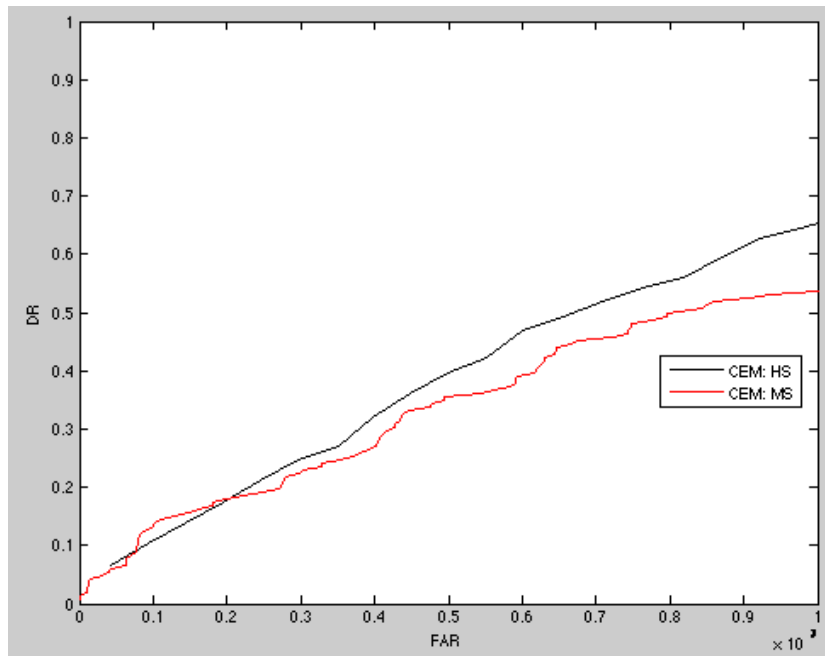


Figure 34. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 60.

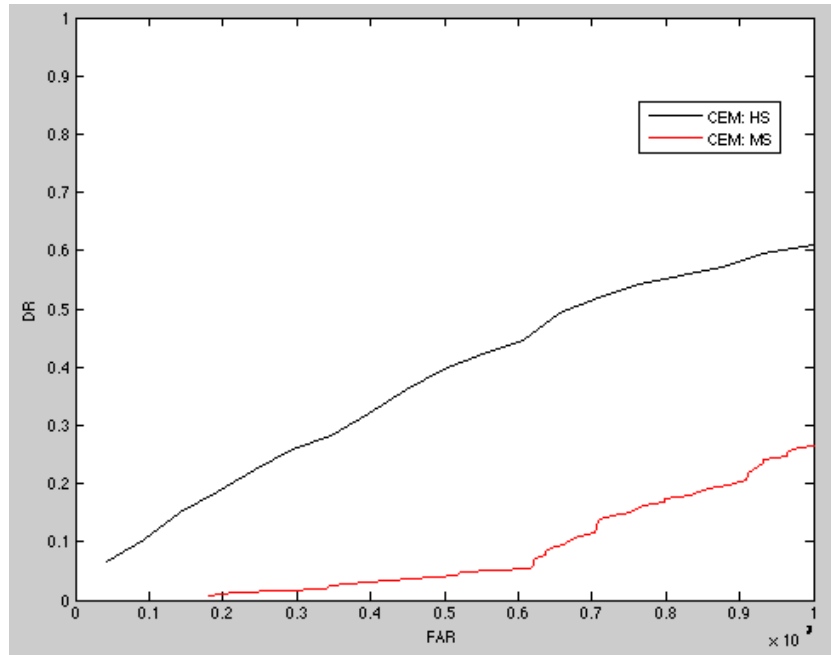


Figure 35. ROC curves for multispectral (red) and hyperspectral (black) CEM algorithms when applied to the 0800 nadir case with an SNR = 50.

The ROC curve graphs demonstrate that the hyperspectral performance decreases slightly, in a consistent manner, as the SNR decreases. Alternatively, the multispectral performance decreases gradually from SNR = 200 to SNR = 75, but performance drops precipitously between SNR = 75 to SNR = 50. The lowest SNR value examined in this project performed quite poorly at nadir, and therefore small increases in area under the ROC curve for test algorithms will produce large impacts on performance when the metric described in Sec. 4.9 is calculated.

5.6 SPI Algorithm: TAD vs. RX

A short study was carried out to assess the impact of varying the SPI algorithm from Sec. 4.8.2 by replacing the TAD algorithm from Sec. 3.10 with the more widely known RX anomaly detection algorithm [Reed and Yu 1990]. The RX algorithm, shown in Eq. 36, has been successfully employed on spectral data [Schott 2009]:

where \mathbf{x} is the spectrum of the pixel of interest, \mathbf{m} is the scene spectral mean, and \mathbf{S} is the scene covariance matrix. For this study, the RX algorithm was applied to polarimetric information by operating on the Stokes vectors described in Sec. 3.6 rather than the spectral vectors for each pixel.

The first step in determining a fair comparison between TAD and RX was to establish which Stokes bands to analyze with the TAD algorithm. Previous work has suggested that TAD performs better as an anomaly detector when only the S_1 and S_2 Stokes bands are considered, rather than when S_0 , S_1 and S_2 are all evaluated [Gartley 2009]. To test this hypothesis, the scene designed for this work was examined with polarimetric and spectral SNR values of 200, while the solar zenith angle was fixed at 34° , the sensor azimuth angle was fixed at 190° from the sun, and the sensor zenith angle was held at 30° . Figure 36 confirms that although TAD with all three Stokes bands does add value in the SPI algorithm, even better performance was obtained by only using TAD on the S_1 and S_2 Stokes bands.

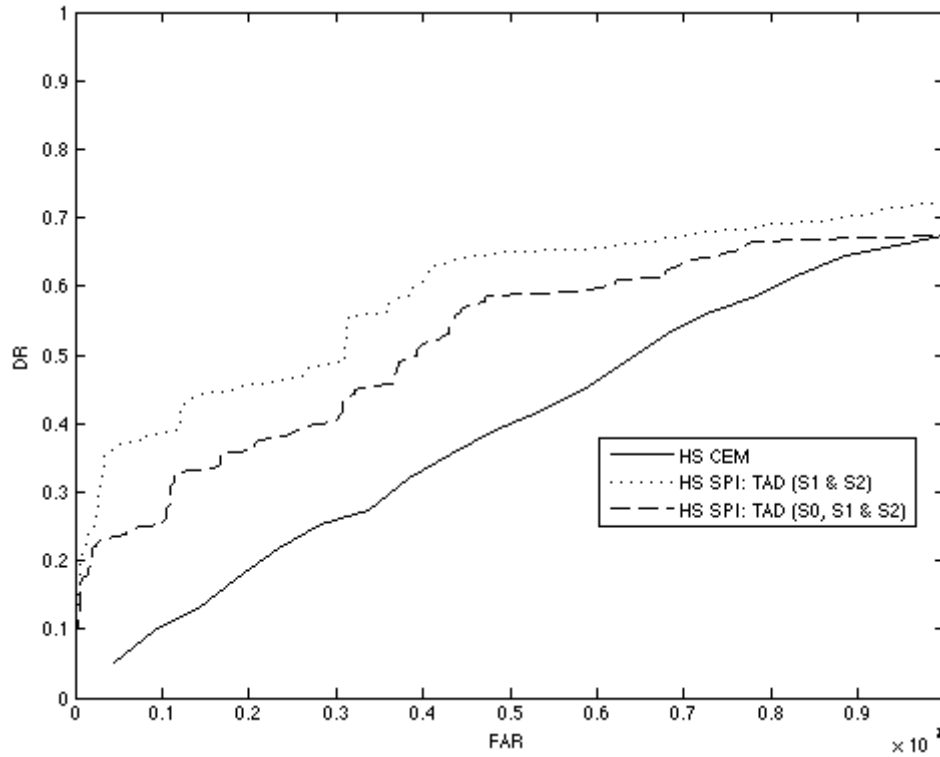


Figure 36. Although using TAD on all three Stokes bands (S_0 , S_1 and S_2) adds value with the SPI algorithm, even better performance is obtained using only two Stokes bands (S_1 and S_2).

Figure 37 plots the two and three band TAD scores for each pixel in the scene, and provides some insight as to why the two band TAD score is more useful. Since the S_0 band is contrast information, bright pixels are likely to be flagged as anomalous pixels—just as they are with the CEM algorithm. As a result, the SPI algorithm is more likely to assign a high score to a bright false alarm when relying on the three band TAD score. When the SPI scores are instead based off only the S_1 and S_2 Stokes bands, only polarimetric behavior contributes toward a pixel’s anomalous behavior.

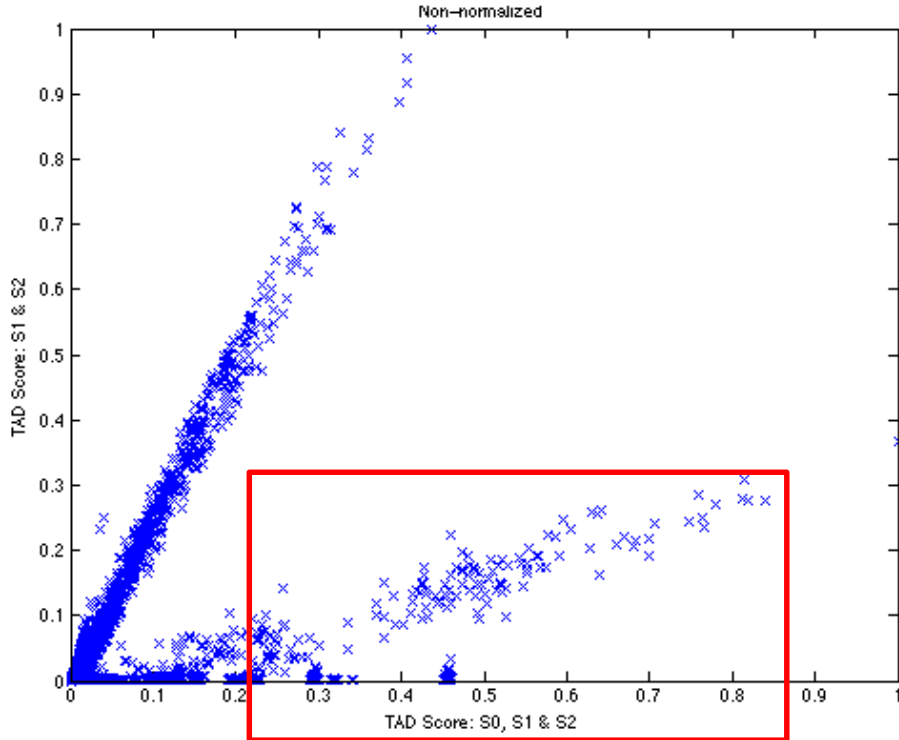


Figure 37. The TAD scores for each pixel are plotted as derived from either three Stokes bands (S_0 , S_1 and S_2) or only two Stokes bands (S_1 and S_2). The two-band case produces better results when exploited by the SPI algorithm because including the S_0 band flags bright false alarms (boxed) as polarimetrically anomalous pixels.

The RX algorithm was compared to the TAD algorithm for the same viewing geometry described above. After an RX score was determined for each pixel, the RX scores were normalized by the score of the most anomalous pixel in the scene to produce scores ranging from zero to one. The normalized RX scores were then fed into the SPI algorithm to be combined with the scene CEM scores. Figure 38 demonstrates that when a high-quality polarimetric signal is captured, the two algorithms produce different results when examined via the SPI algorithm. Interestingly, the results derived from the RX algorithm were essentially unchanged whether two or three Stokes bands were used—although the results derived from the TAD algorithm were superior to both RX cases.

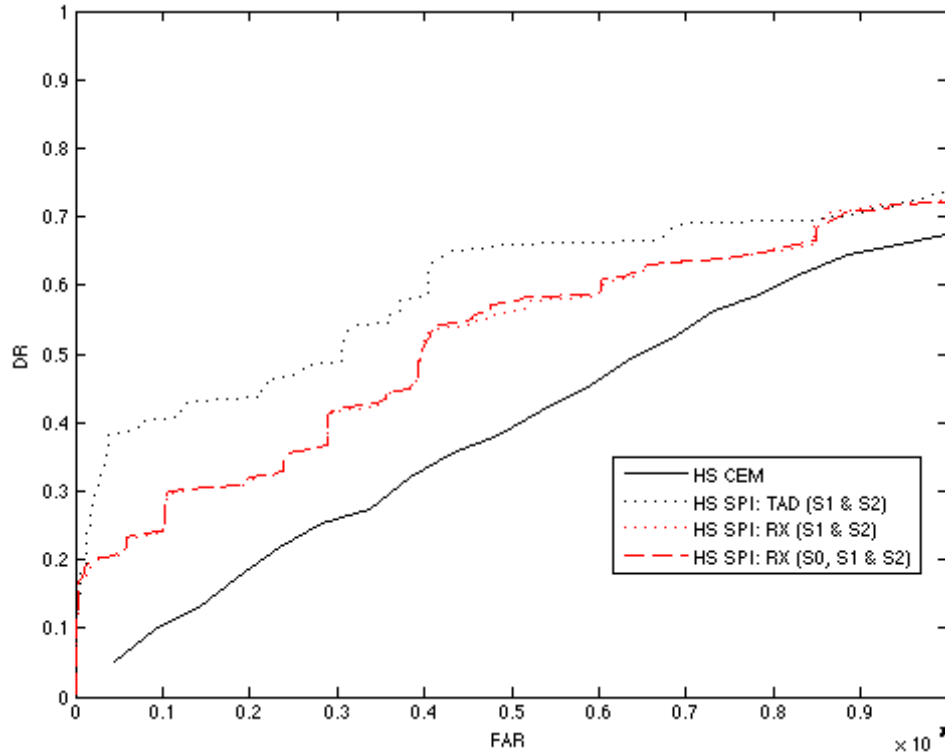


Figure 38. ROC curves were calculated for the hyperspectral SPI curves derived from applying TAD to the S_1 and S_2 bands (dotted black line), applying RX to the S_1 and S_2 bands (dotted red line), and applying RX to the S_0 , S_1 and S_2 bands (dashed red line). The hyperspectral CEM curve (black solid line) is shown for the same viewing geometry, indicating the value added by incorporating polarimetric information.

Next, the TAD and peak-normalized RX scores assigned to each pixel were examined to determine the correlation between outputs from the two algorithms. The correlation coefficient between the two scores was 0.9159, and Figure 39 shows the strong functional relationship between the TAD and RX scores. Although TAD and RX ranked anomalous pixels in a similar order, TAD weighted anomalous pixels with a higher numerical score. The polarimetric information then carried more weight when the TAD scores (as opposed to RX scores) were fed into the SPI algorithm. The increased relative value of the polarimetric information explains the significantly better performance in Figure 38 of the hyperspectral SPI ROC curve based on the TAD algorithm.

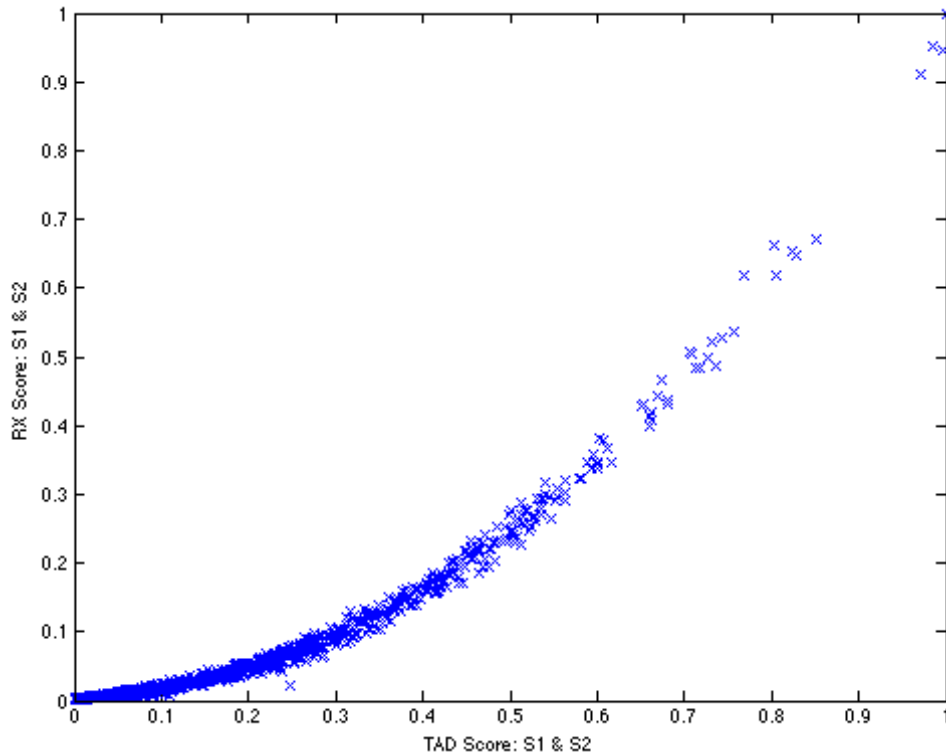


Figure 39. The TAD and RX scores used in Figure 38 are shown for each pixel in the scene. TAD and RX generally rank anomalous pixels similarly, but TAD seems to weight more anomalous pixels with a higher score than RX.

The strong functional relationship between TAD and RX scores displayed in Figure 39 raised the question of whether the TAD performance could be matched (or exceeded) by applying a nonlinear operator to the RX scores. When the TAD and square-rooted RX scores were compared for each pixel in the scene, a much more linear relationship was produced. Figure 40 shows a series of ROC curves, confirming that the TAD and the square-rooted RX scores produced similar performance when combined with the CEM scores via the SPI algorithm. This result suggested an avenue exists to further optimize the distribution of the polarimetric anomaly scores during the fusion process, but that effort will be left for future work and not further developed in this project. Because results have previously been published showing that the TAD algorithm

outperformed RX on polarimetric panchromatic imagery [Gartley and Basener 2009], and those results appear to be confirmed in this scenario, the remainder of this experiment will leverage the TAD score in the SPI algorithm.

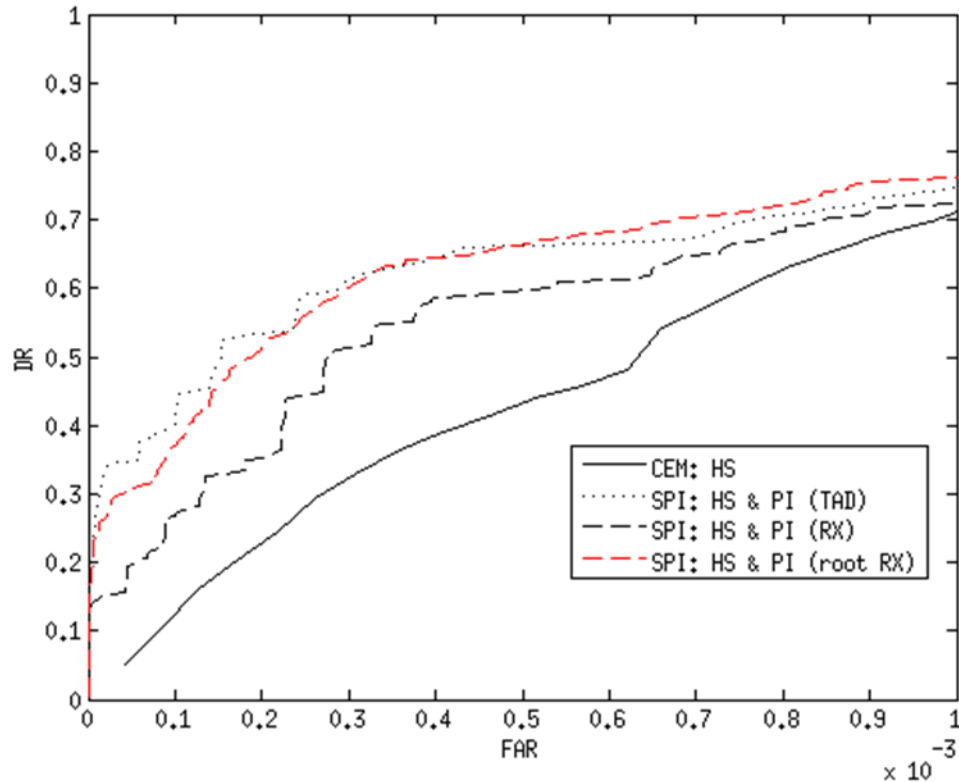


Figure 40. For a given image, ROC curves were calculated using the hyperspectral CEM score (solid black line), the SPI algorithm fusing the CEM and TAD scores (dotted black line), the SPI algorithm fusing the CEM and RX scores (dashed black line) and the SPI algorithm fusing the CEM and square-rooted RX scores (dashed red line). Re-mapping the RX scores by taking the square root of each pixel's score appears to give performance similar to using the TAD score.

5.7 Contrast vs. DOLP

A study was conducted to verify whether improvements observed with the fusion algorithms were actually due to polarimetric information or simply due to additional high-resolution contrast information. The viewing geometry chosen for analysis was a case where both of the fusion algorithms enhanced performance in the hyperspectral case. Specifically, the solar zenith angle was 34° , the sensor zenith angle was 30° and the

sensor azimuth angle was 170° from the sun. The spectral SNR was fixed at 200, and an image created from the multispectral RGB bands captured under those conditions is shown in Figure 41.



Figure 41. An RGB image of the synthetic scene was observed with a solar zenith angle of 34° , a sensor zenith angle of 30° and a sensor azimuth angle of 170° from the sun.

First, the RX anomaly detector was applied to several different combinations of Stokes band data sets to determine how they were related. Then, the S_0 band representing high-resolution contrast information was used in place of DOLP in the SPOT pixel fusion algorithm. Finally, the normalized RX scores derived from the S_0 band were used in place of the TAD scores in the SPI decision fusion algorithm.

5.7.1 RX Applied to Stokes Bands

As an initial effort, the RX algorithm was applied to a series of different unnormalized Stokes band data sets (PI SNR = 200) to evaluate the impact of the contrast

information in the S_0 band. First, the RX algorithm was applied to just the S_0 band where the contrast information was contained. Next, the RX algorithm was applied to only the S_1 and S_2 bands. Finally, the RX algorithm was applied to the entire Stokes data set (S_0 , S_1 and S_2).

Figure 42 shows the S_0 information that was analyzed to test the effect of high-resolution contrast information. As expected, the 100% reflective calibration panel on top of the building shows up brightest, the 50% reflective panel in the field has a mid-level gray value and the 0% reflective panel is dark black.



Figure 42. The S_0 information from the scene was used to test the effect of incorporating high resolution contrast information into the fusion algorithms.

After RX was applied to the data sets described above, the RX scores for each pixel in the different cases were plotted against each other to determine whether any correlation was present. Figure 43 demonstrates that although there was some correlation

($\rho = 0.6579$), applying RX to all the Stokes bands instead of just the S_0 band provided extra information about anomalous pixels in the scene.

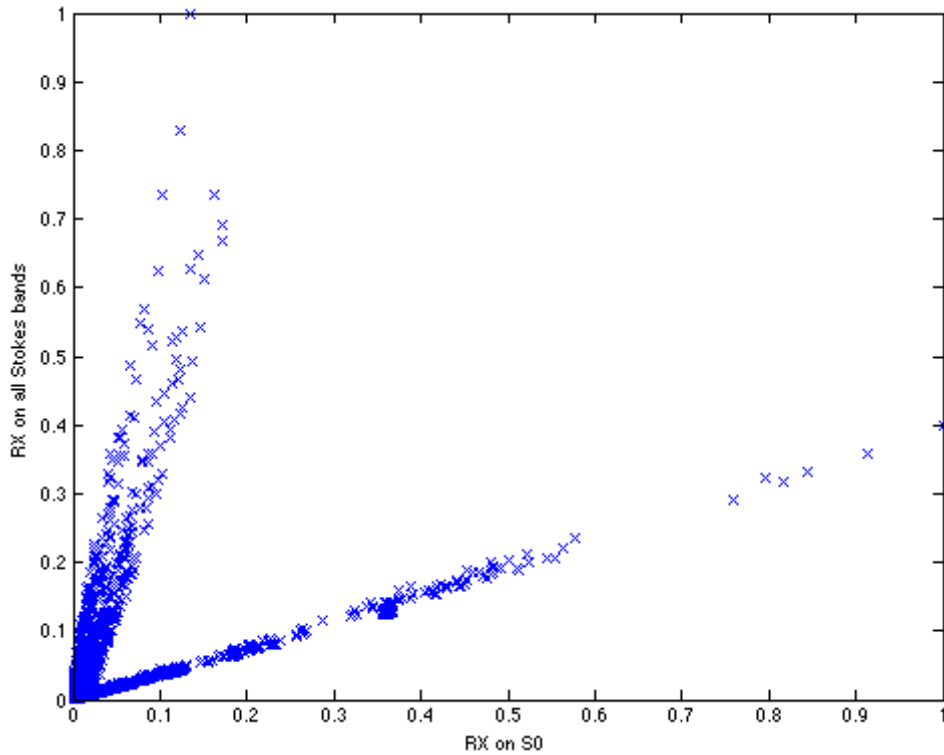


Figure 43. Each pixel in the scene was evaluated both by applying RX to just the S_0 band and then by applying RX to all the Stokes bands.

Figure 44 shows that there was very little correlation ($\rho = 0.1367$) between the RX scores derived from the S_0 band and the RX scores produced by analyzing the S_1 and S_2 bands. This discrepancy is due in part to Umov's effect, where the DOP of brighter objects will tend to be less than the DOP of darker objects if the objects are made of similar materials [Schott 2009].

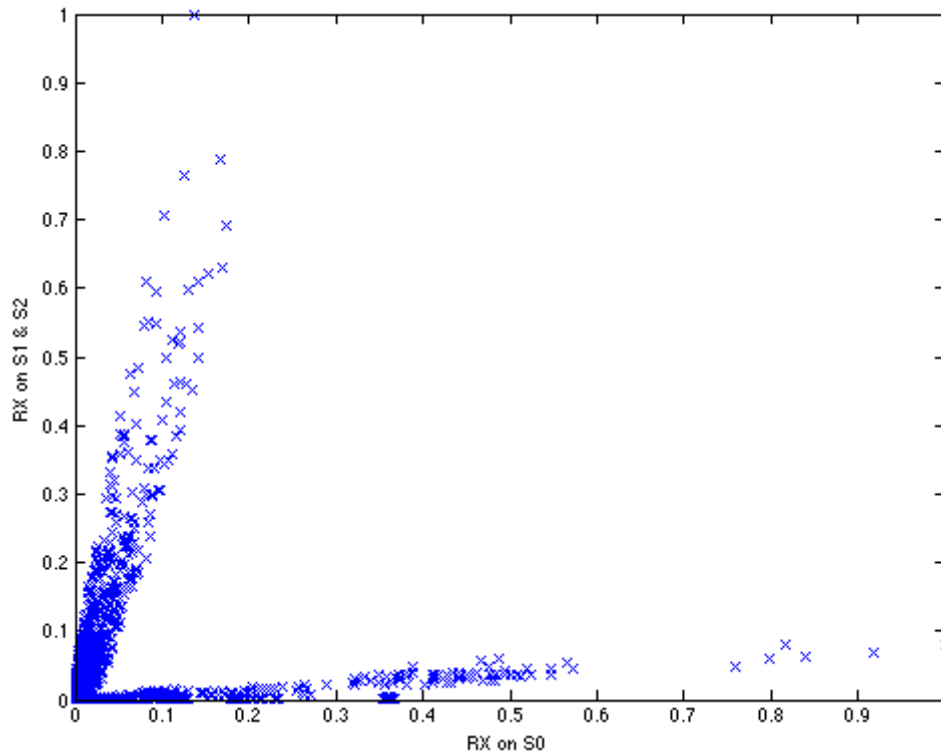


Figure 44. Each pixel in the scene was evaluated both by applying RX to just the S_0 band and then by applying RX to both the S_1 and S_2 bands.

Applying RX to the different Stokes data sets indicated that polarimetric and contrast information provided separate perspectives in the attempts to separate anomalous pixels from the background.

5.7.2 SPOT Algorithm

For the SPOT algorithm described in Sec. 4.8.1, the scene's polarimetric information was incorporated by treating the DOLP as an additional band of information. In this study, the DOLP band was replaced by the unnormalized S_0 values to instead incorporate contrast information, while the target signature was unchanged from Sec. 4.8.1—effectively assuming the target was the brightest object in the scene. Figure 45 shows the DOLP values of the scene as an intensity image. Umov's effect was highlighted by comparing the DOLP values between the 100% and 0% reflective

perfectly Lambertian calibration panels—the dark panel had some degree of polarization, while the bright panel showed almost none. In contrast, bright points representing highly polarized pixels were apparent near the vehicles in the scene.



Figure 45. The polarimetric information from the scene was incorporated as DOLP values in an additional band for the SPOT algorithm.

The SPOT algorithm was applied and a pixel fusion score was produced for each pixel in the image. Figure 46 shows the results for the multispectral case derived from S_0 on the left and DOLP on the right. In the case where the contrast information was used, variations in the pixel fusion scores were washed out. However, in the case where the polarimetric information was used, the bright spots representing target vehicles were preserved.

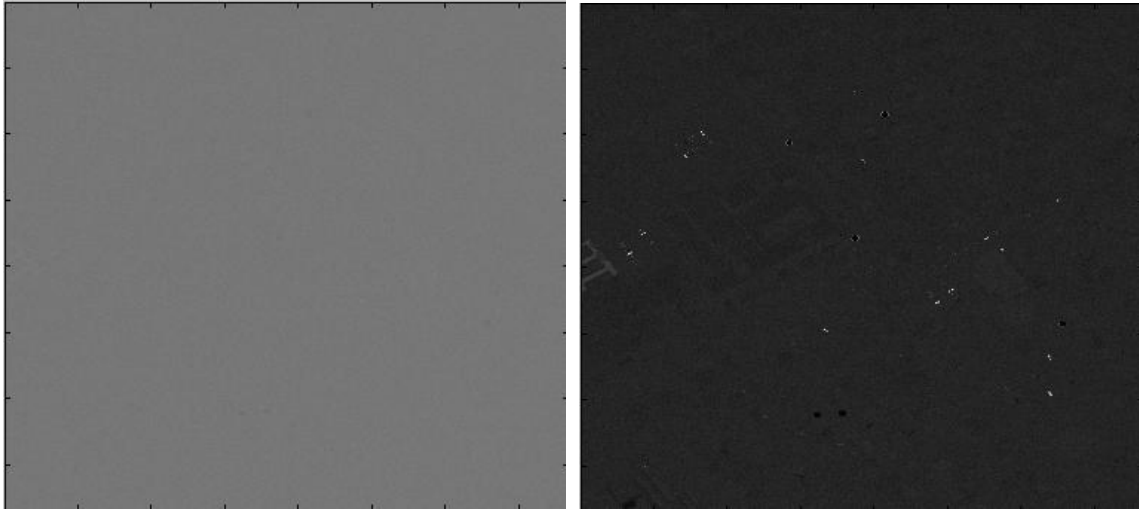


Figure 46. SPOT scores are shown for the multispectral case when using the S_0 information (left) or DOLP (right).

Figure 47 shows the results for the hyperspectral case derived from S_0 on the left and DOLP on the right. The differences between the two cases are less pronounced because the effect of varying the one band carrying polarimetric information was diluted by keeping the other 90 bands constant.

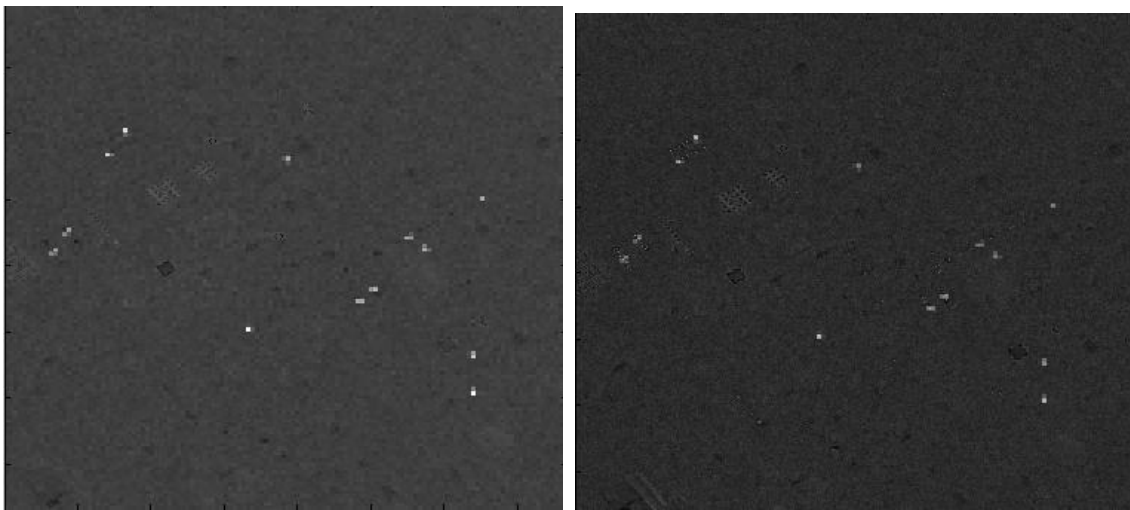


Figure 47. SPOT scores are shown for the hyperspectral case when using the S_0 information (left) or DOLP (right).

Finally, the SPOT fusion algorithm performance in both cases was evaluated via the ROC curves shown in Figure 48.

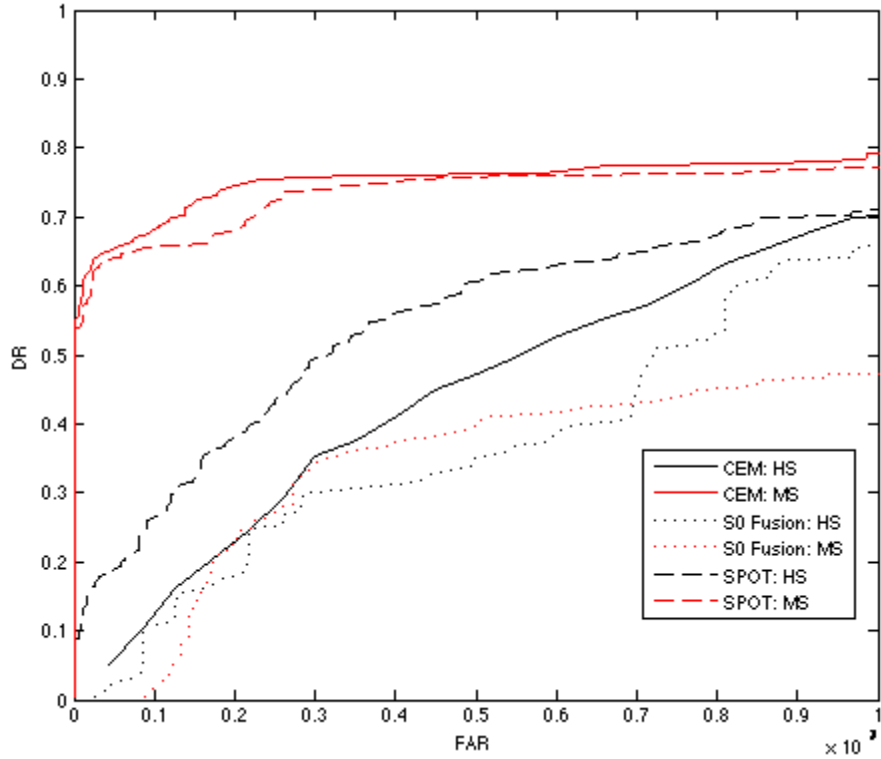


Figure 48. The multispectral (red) and hyperspectral (black) SPOT ROC curves are shown when derived from the S_0 information (dots) or DOLP (dashes). The CEM scores (solid lines) are included for comparison.

From the results presented in this section, it is apparent that the SPOT algorithm is not just evaluating high-resolution contrast information—rather, it is exploiting polarimetric information conveyed in the DOLP.

5.7.3 SPI Algorithm

For the SPI algorithm described in Sec. 4.8.2, the scene’s polarimetric information was incorporated by ranking anomalous pixels with the TAD algorithm, then combining those scores with each pixel’s CEM score. In this study, the TAD scores (shown in Figure 49) were replaced by the RX scores derived from the S_0 values (shown in Figure 50) to instead incorporate contrast information.

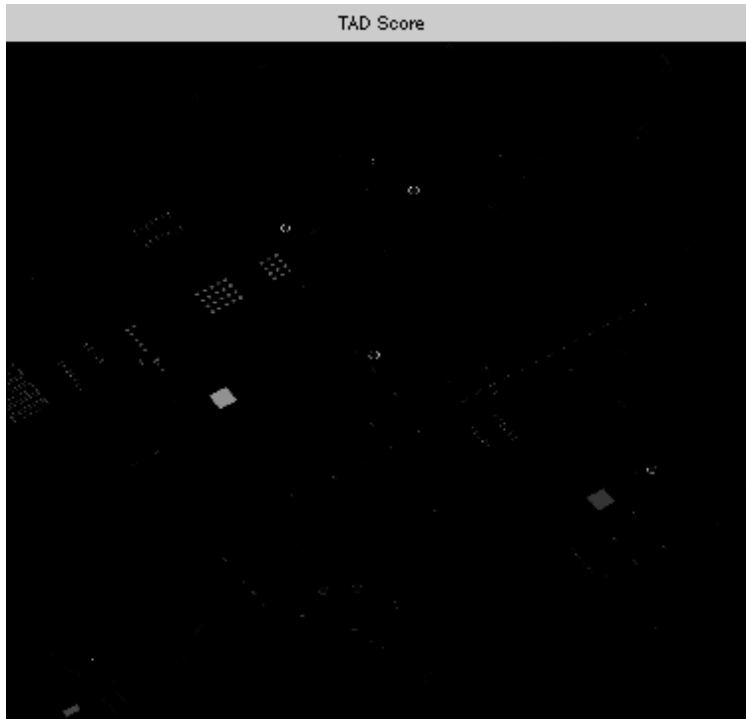


Figure 49. The TAD scores calculated by examining the S_0 , S_1 and S_2 bands are shown as an intensity image. Note how many vehicles show up as anomalous bright spots.



Figure 50. The RX scores derived from the S_0 band are displayed as an intensity image. Note that the vehicles are less visible in this image.

The SPI algorithm was applied to fuse scores from the anomaly detector with the spectral CEM scores, and the results were analyzed with the series of ROC curves shown in Figure 51.

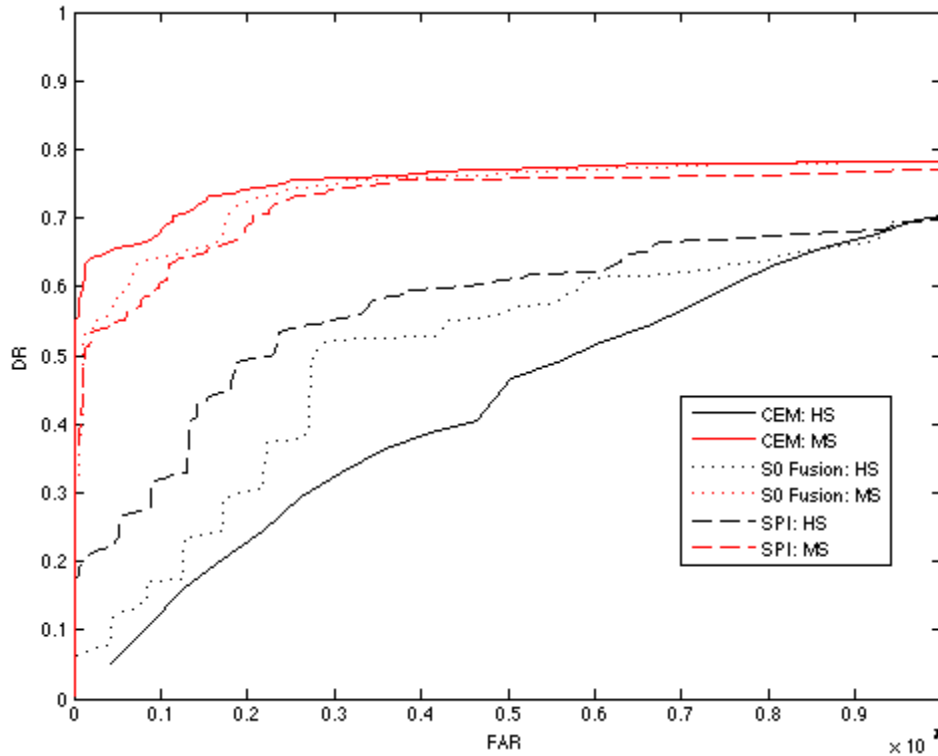


Figure 51. The multispectral (red) and hyperspectral (black) SPI ROC curves are shown when incorporating RX scores derived from the S_0 information (dots) or TAD scores derived from the S_0 , S_1 and S_2 bands (dashes). The CEM scores (solid lines) are included for comparison.

In summary, although incorporating high-resolution contrast information into the SPI algorithm did enhance performance in the higher GSD hyperspectral scenario, including polarimetric information in the other Stokes bands further increased performance.

5.8 Establishing Sensitivity to Geometry

The effectiveness of the fusion algorithms described above was expected to vary due to the sensitivity of polarimetric remote sensing to different sun-target-sensor geometries, as described in Sec. 3.6, so a small proof-of-concept experiment was carried

out to ensure this variation occurred when viewing the synthetic scene. The hyperspectral and polarimetric sensors described in Sec. 4.3 were employed at three different viewing geometries, shown in Table 5, and ROC curves were calculated using the CEM, SPI and SPOT algorithms for each off-nadir viewing geometry.

Table 5. Three different viewing geometries were considered to illustrate the variable utility of polarimetric information.

Image number	Solar zenith angle (deg)	Sensor zenith angle (deg)	Relative azimuth angle (deg)
1	55	50	180
2	55	50	90
3	55	10	180

Figure 52 shows RGB images of the sensor reaching radiance calculated by DIRSIG for the viewing geometries in Table 5.

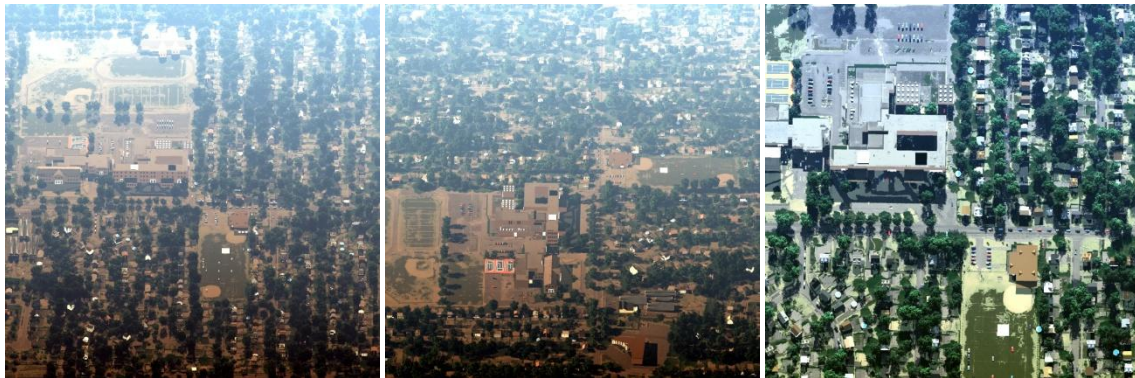


Figure 52. RGB images of sensor-reaching radiance calculated by DIRSIG for the cases described in Table 5. (Left) Image 1, captured near the sun’s specular reflection with a sensor zenith angle of 50° and relative azimuth angle of 180° . (Center) Image 2 is at the same zenith angle as Image 1, but only 90° relative azimuth from the sun. (Right) Image 3 is azimuthally in the specular reflection (180°), but at a sensor zenith angle of only 10° .

A ratio of the area under the off-nadir fusion ROC curve to the area under the off-nadir spectral CEM ROC curve was calculated, and impact of polarimetric information on performance was quantified as a percent increase as shown in Table 6.

Table 6. The area under the ROC curves for the fusion algorithms was compared to the area under the ROC curve for the CEM algorithm, producing a percent change metric.

Image #	HS SPI improvement	HS SPOT improvement	MS SPI Improvement	MS SPOT improvement	GSD at scene center: spectral (polarimetric)
1	68.4 %	35.7 %	80.4 %	61.7 %	5.13 m (0.85 m)
2	-1.5 %	-0.6 %	-0.8 %	-1.4 %	5.13 m (0.85 m)
3	1.1 %	-4.9 %	1.8 %	-8.3 %	3.05 m (0.51 m)

Representative ROC curves calculated with the CEM, SPI and SPOT algorithms for Image 1 are shown in Figure 53. Because the sensor was in the sun’s specular lobe, with both the sun and sensor at relatively large zenith angles, a significant polarimetric signal was expected. The ROC curves demonstrate that both fusion algorithms outperformed the purely spectral CEM algorithm for the viewing geometry associated with Image 1.

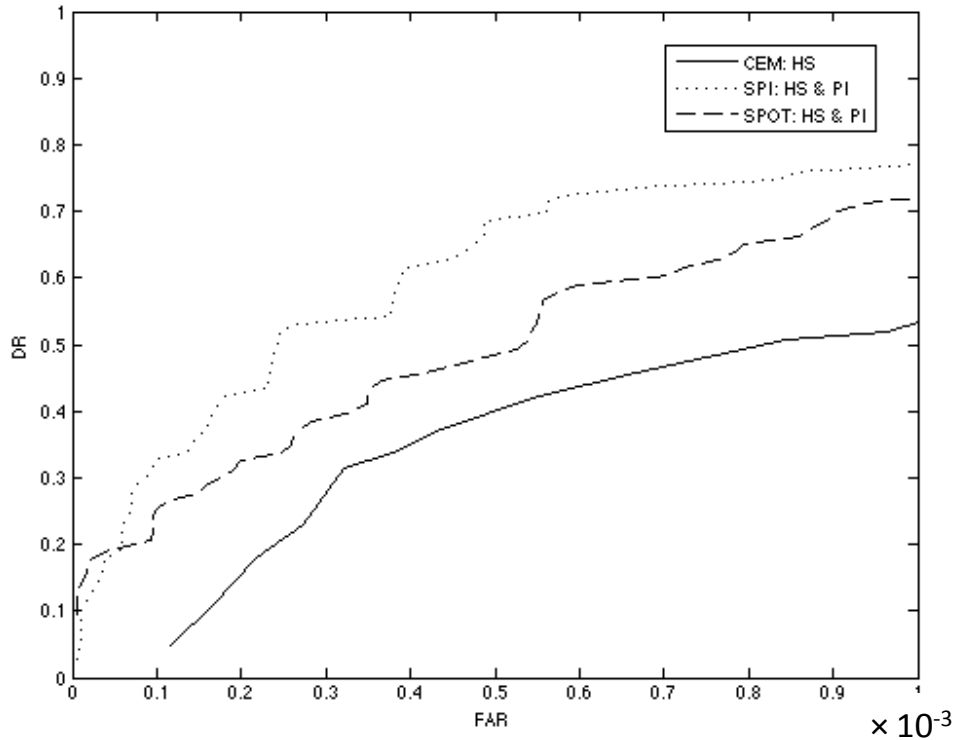


Figure 53. Hyperspectral ROC curves were generated from Image 1 (with the sensor near the sun’s specular reflection) using the CEM (solid line), SPI (dotted line) and SPOT (dashed line) algorithms.

Representative ROC curves calculated with the CEM, SPI and SPOT algorithms for Image 2 are shown in Figure 54. This viewing geometry differed from Image 1 because the sensor was only separated from the sun by a relative azimuth angle of 90° , although both the sun and sensor were maintained at the same large zenith angles. The ROC curves demonstrate that both fusion algorithms essentially performed the same as the purely spectral CEM algorithm for Image 2, meaning that incorporating polarimetric information provided no added value.

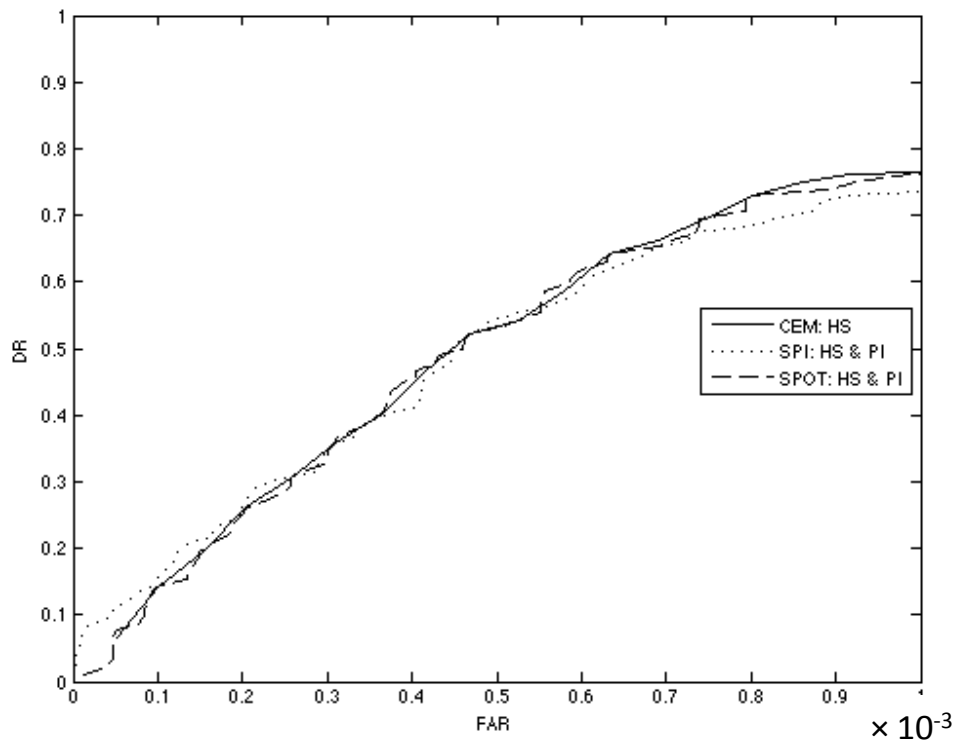


Figure 54. Hyperspectral ROC curves were generated from Image 2 (with the sensor at a 90° azimuth angle from the sun) using the CEM (solid line), SPI (dotted line) and SPOT (dashed line) algorithms.

One possible explanation was that the polarimetric contrast between the target vehicles and the background was reduced due to decreasing pBRDF values and incorporating upwelled radiance. The magnitude of the pBRDF was decreased because the target was

viewed from a perspective away from the direction of the sun's specular reflection, resulting in a decrease in reflected polarimetric radiance from the target. Further, if the car is generally flat when viewed from above, the reflected polarization is polarized horizontally [Schott 2009]. However, since the polarization of the upwelled radiance is always perpendicular to the sun-target-sensor plane [Schott 2009], the polarimetric signature of the upwelled radiance was instead oriented at $+45^\circ$ (rather than horizontal for Image 1), making the polarimetric signature of a target pixel less distinctive. When the pBRDF and atmospheric effects combined to reduce the contrast between target and background pixels, the target pixels were less likely to be ranked as highly polarimetrically anomalous by the TAD algorithm.

Representative ROC curves calculated with the CEM, SPI and SPOT algorithms for Image 3 are shown in Figure 55. This viewing geometry differed from Image 1 because the sensor zenith angle was reduced from 50° to 10° , although the solar zenith angle and relative azimuth angle were unchanged. The ROC curves demonstrate that both fusion algorithms again essentially performed the same as the purely spectral CEM algorithm for Image 2, highlighting another case where incorporating polarimetric information provided no added value. One possible explanation for the lack of polarimetric impact was that the pBRDF value again decreased as the sensor was moved further away from the sun's specular reflection.

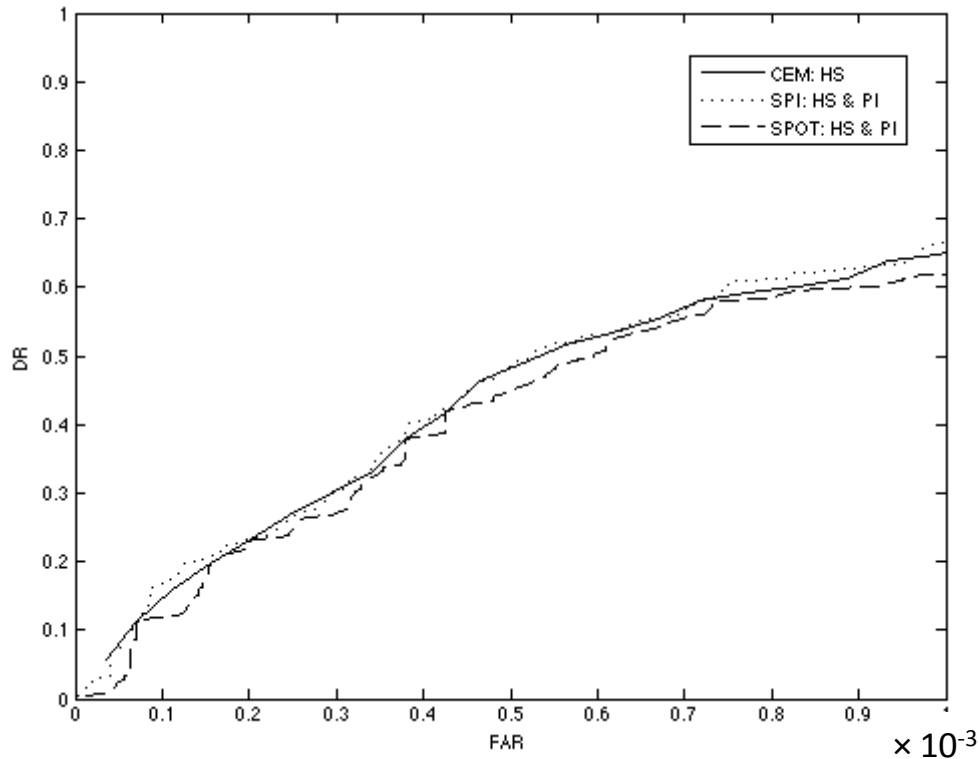


Figure 55. Hyperspectral ROC curves were generated from Image 3 (with the sensor at a 180° azimuth angle from the sun, but only a 10° zenith angle) using the CEM (solid line), SPI (dotted line) and SPOT (dashed line) algorithms.

A second explanation was that when the sensor was positioned in the sun's specular reflection lobe, the spectral sensor suffered from the spectral whitening effect, resulting in a decreased ability to distinguish the target from the background. As the sensor was moved out of the specular reflection lobe, the level of spectral whitening was decreased, and performance was improved. Some evidence of this effect may be observed by comparing the ROC curves from each of the images, as shown in Figure 56. The ROC curves demonstrate that when polarimetric information provided the most increase in performance, the purely spectral CEM performance was the poorest for all the cases examined. Further, the ROC curves indicate that the fusion algorithms also have the potential to outperform the CEM algorithm in several different viewing geometries.

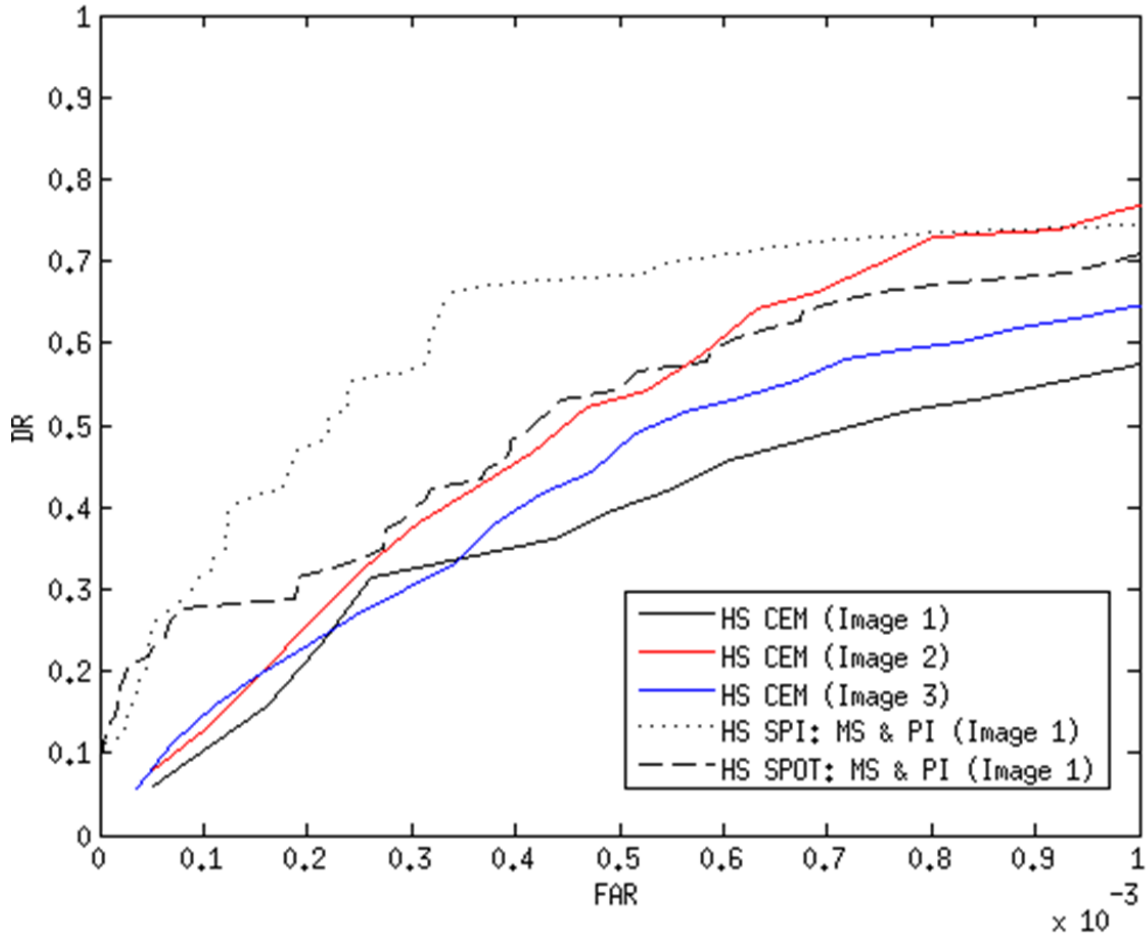


Figure 56. Hyperspectral ROC curves were generated from the series of images described in Table 5 using the CEM (solid lines), SPI (dotted line) and SPOT (dashed line) algorithms. Image 1 (black) was the scenario where polarimetric information increased performance the most, and it was also the scenario where the spectral sensor performed the poorest.

In conclusion, the three different viewing geometries investigated in this study ensured that variations in the utility of polarimetric data were observed for the synthetic scene of interest. Some part of the performance increase due to spectral and polarimetric data fusion may have been due to the spectral whitening effect reducing the capability of the spectral sensor in the sun's specular lobe. However, even though the spectral performance was worst in the sun's specular lobe, leveraging the additional polarimetric information produced better performance than exploiting solely spectral data for any of the three viewing geometries examined. The overall multimodal system performance is

therefore a complex combination spectral and polarimetric phenomenon, each of which varies in a different fashion. Further, any increased capability due to fusion may be of different value for different applications, highlighting the need for a sensor trade study to thoroughly examine a range of possible viewing and illumination conditions. The rest of this work will apply the process described in Sec. 4 to evaluate the fusion trade space.

5.9 Section Summary

This section explained several preliminary studies conducted to ensure the results of this project would be meaningful. First, the effects of using a global versus local background mean vector and covariance matrix in the CEM algorithm were investigated, and using the global values was shown to provide suitable performance while minimizing computational requirements. Next, the multispectral CEM and polarimetric TAD scores were shown to be reasonably uncorrelated, ensuring the second sensing modality was providing additional useful information.

Then, the hyperspectral synthetic scene ($GSD = 0.5$ m) was calculated to have an $SCR \approx 30$, while a red vehicle spectrum extracted from actual hyperspectral imagery of Rochester, NY ($GSD \approx 0.66$ m) produced an $SCR \approx 50$, verifying that sufficient variation was present in the synthetic background. The polarimetric SCR was also quantified by analyzing the TAD metric, where the SCR was shown to vary dramatically as a function of sensor zenith angle, with a peak value around 20. Alternatively, if the SCR was calculated using the DOP metric, the peak value was less than five for all sensor zenith angles. The top 5% of polarimetrically anomalous pixels for a particular favorable

viewing geometry were shown to be composed of trees, roadways, grass, glass, roofing material, siding and vehicle paints.

Next, the area under the ROC curve was calculated for several different SNR values, and SNR = 200, 100 or 50 were shown to be useful while SNR = 15 was shown to not be useful. Additionally, the SNR = 50 value imposed the constraint that $FAR = 0.001$ was the minimum useful FAR threshold for evaluating all cases. Separately, the TAD was shown to perform better when applied to just the S_1 and S_2 bands as opposed to when applied to the S_0 , S_1 and S_2 bands because including S_0 was likely to flag bright pixels as anomalous. The TAD algorithm (applied to the S_1 and S_2 bands) then outperformed the RX algorithm (applied to either the S_0 , S_1 and S_2 bands or just the S_1 and S_2 bands) when analyzing polarimetric information for input into the SPI algorithm. Then, the fusion algorithms were shown to perform better if anomalous pixels were determined from the complete set of polarimetric information (S_0 , S_1 and S_2 bands) rather than simply from the contrast information in the S_0 band. Finally, the CEM, SPOT and SPI algorithms' performances were evaluated at three different viewing geometries, and the impact due to fusion was determined to be highly sensitive to viewing geometry due to both pBRDF and atmospheric effects. This last study strongly motivates the rest of this study and points to the importance of a comprehensive trade study to understand the capabilities and limitations of sensors and acquisition conditions for scenarios of interest.

6 Results

6.1 Section Overview

Once the spectral and polarimetric characteristics of the synthetic scene had been validated, the *FAR* threshold had been chosen and preliminary experiments had shown promise, the full impact of incorporating polarimetric information was assessed. For the first step, the fusion algorithms were applied assuming co-located multimodal sensors with no registration error. Then, the fusion algorithms were applied assuming a co-located multimodal sensor with a single pixel of registration error. Finally, the data were analyzed assuming each modality was positioned at its ideal acquisition location for a particular time of day. The spectral and polarimetric images were registered to each other and then analyzed with the different fusion algorithms.

When a co-located sensor was assumed, at every different solar zenith angle, sensor zenith angle, sensor relative azimuth angle and SNR value outlined in Sec. 4.5, the impact on both the hyperspectral and multispectral systems was evaluated as described in Sec. 4.9. Then, performance trends were extracted by averaging the ratios from all azimuth angles (160° - 200°) for a particular solar zenith angle, sensor zenith angle and SNR values. In the spotlight charts that follow, red designates cases where neither fusion algorithm achieved the goal, green depicts cases where the SPOT pixel fusion algorithm performed best and blue represents cases where the SPI decision fusion algorithm performed best. Numbers in the cells quantify the percent improvement when the fusion algorithm was used.

6.2 Co-located Fusion with Perfect Registration

The initial effort sought to evaluate the effectiveness of the fusion algorithms before complicating the problem with registration error or spatially separated imaging systems. Fusion was defined to be useful at a particular viewing geometry if two conditions were met. First, an off-nadir fusion algorithm had to outperform the off-nadir CEM algorithm for the same viewing geometry—the polarimetric information shouldn't degrade performance. Second, the off-nadir fusion algorithm should perform better than the nadir CEM algorithm to justify capturing off-nadir images in the first place. If both conditions were met, that particular off-nadir viewing geometry was deemed useful. A percent increase in performance was then determined by comparing the average fusion algorithm ratio value to the average CEM ratio value as shown in Eq. 37:

Equation 38 shows how the $Ratio_{CEM}$ and $Ratio_{Fusion}$ terms can be rewritten to explicitly show the input values:

$$\frac{\frac{\text{Area}_{\text{Fusion}}}{\text{Area}_{\text{Nadir}}}}{\frac{\text{Area}_{\text{CEM}}}{\text{Area}_{\text{Nadir}}}}$$

Finally, the calculation was simplified, as shown in Eq. 39, so that the percent difference value relied only on ROC curve area values:

When both fusion algorithms met both of the conditions, the performance of the better fusion algorithm was reported, but the percent difference in performance was

quantified for two different scenarios. First, off-nadir fusion was compared to off-nadir CEM to determine the impact of incorporating polarimetric information if forced to image from a particular orientation. Second, off-nadir fusion was compared to nadir CEM to assess whether an off-nadir sensor orientation might make more sense if the user had the freedom to task the sensor at will, and these percent improvement values are shown in parentheses in the tables below. The results for the multispectral sensor described in Sec. 4.3 are shown in Table 7, suggesting fusion provides little benefit unless the spectral SNR value is low. One possible reason for the lack of improvement in performance is that the spectral GSD was already quite small (0.5 m at nadir) in the multispectral case, producing more spectrally pure pixels to easily differentiate targets from the background.

Table 7. The effect of both fusion algorithms is summarized for the 8 band multispectral sensor (nadir GSD = 0.5 m) across 5 solar zenith angles, 7 sensor zenith angles and 5 different SNR values with perfect registration. The data suggest fusion may be useful in the MS case at lower SNR values. Percent improvement metrics are defined vs. off-nadir CEM (or vs. nadir CEM in parentheses).

MS		Solar Zenith Angle				
SNR = 200						
Sensor Zenith						
Angle		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10						
20						
30						
40						
50						
60						
70						

MS

SNR = 100

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10					
20					
30					
40					
50					
60					
70					

MS

SNR = 80

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10					
20					
30					
40					
50					
60					
70					

MS

SNR = 60

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10					166% (148%)
20					107% (118%)
30					20.9% (10.9%)
40				9.2% (62.2%)	
50		2.5% (18.9%)	1.5% (41.7%)		
60		12.1% (1.8%)	0.7% (18.3%)		
70					

MS
SNR = 50

		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10						826% (911%)
20						345% (720%)
30						53.2% (267%)
40				0.9% (674%)	8.1% (779%)	0.9% (242%)
50			3.7% (248%)	2.6% (519%)	0.4% (683%)	0.4% (176%)
60	2.9% (29.2%)	17.4% (208%)	0.7% (423%)	0.5% (523%)		
70			3.9% (257%)			

However, in the larger GSD hyperspectral case shown in Table 8 several trends appeared. First, the theoretical sensitivity of polarimetric remote sensing to sun-target-sensor geometry was confirmed since no fusion algorithm was useful in every orientation. Second, in cases where polarimetric information enhanced capability, the SPI decision fusion algorithm outperformed the SPOT pixel fusion algorithm—reinforcing the notion that since spectral and polarimetric sensors observe different phenomena, the data from each should first be evaluated on its own merit. Third, the added polarimetric information produced the most increase in performance near the sun’s specular reflection, providing a target detection capability where spectral sensors would otherwise become less effective due to spectral whitening. Finally, for cases where the sun was lower on the horizon, as the system SNR decreased (0800 & 0700 cases) the performance difference due to fusion increased relative to off-nadir CEM but decreased relative to nadir CEM.

Table 8. The effect of both fusion algorithms is summarized for the 90 band hyperspectral sensor (nadir GSD = 3.0 m) across 5 solar zenith angles, 7 sensor zenith angles and 5 different SNR values with perfect registration. The data suggest fusion is useful when the HS sensor is positioned near the sun's specular reflection. Percent improvement metrics are defined vs. off-nadir CEM (or vs. nadir CEM in parentheses).

HS
SNR = 200

		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10		21.9% (30.7%)	5.8% (11.6%)	0.8% (6.0%)		
20		35.3% (40.9%)	23.9% (34.4%)	4.4% (11.0%)		
30		41.1% (54.1%)	41.3% (57.7%)	11.9% (23.2%)	1.5% (6.2%)	
40		22.3% (26.5%)	33.2% (65.8%)	24.2% (47.8%)	3.9% (19.4%)	
50		2.7% (21.1%)	31.8% (63.4%)	29.2% (46.4%)		
60		2.5% (12.7%)	4.0% (24.4%)			
70						

HS
SNR = 100

		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10		19.4% (28.2%)	5.0% (10.4%)	1.1% (6.0%)		
20		33.0% (38.6%)	21.7% (31.5%)	4.1% (10.4%)		
30		37.7% (50.7%)	38.9% (54.4%)	9.9% (20.4%)	0.4% (4.2%)	
40		20.5% (24.7%)	31.4% (63.1%)	20.5% (42.4%)	4.3% (12.8%)	
50		2.8% (21.1%)	30.9% (60.2%)	36.1% (37.5%)		
60						
70						

HS
SNR = 80

		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10		18.5% (27.6%)	5.0% (10.0%)	0.9% (5.6%)		
20		32.0% (38.0%)	21.3% (30.7%)	3.6% (9.7%)		
30		37.2% (50.5%)	37.5% (52.7%)	8.8% (18.8%)	0.4% (3.1%)	
40		20.4% (24.9%)	30.6% (61.6%)	19.8% (40.8%)	6.3% (9.5%)	
50		3.0% (21.7%)	30.6% (58.1%)	44.7% (33.5%)		
60						
70						

HS
SNR = 60

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10	18.4% (28.0%)	5.0% (10.7%)	0.5% (5.9%)		
20	31.7% (38.0%)	21.0% (31.1%)	2.9% (9.7%)		
30	36.7% (50.4%)	35.9% (51.6%)	8.0% (18.8%)	0.6% (1.7%)	
40	20.2% (25.0%)	29.3% (61.1%)	20.9% (42.6%)	8.7% (3.3%)	
50		32.5% (56.6%)	71.7% (25.2%)		
60					
70					

HS
SNR = 50

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10	17.7% (26.9%)	4.9% (11.4%)	0.2% (7.0%)		
20	30.9% (37.0%)	20.5% (31.5%)	2.5% (10.7%)		
30	36.6% (49.7%)	34.6% (51.1%)	7.6% (20.0%)	0.8% (1.4%)	
40	19.7% (24.3%)	28.7% (61.0%)	20.9% (42.6%)		
50		38.0% (55.8%)	115% (15.0%)		
60					
70					

6.3 Co-located Fusion with Single Pixel Registration Error

The next step was to introduce a single pixel of registration error as described in Sec. 4.6 and perform the same analysis demonstrated in Sec. 6.2. Table 9 summarizes the results from the multispectral case, and when they were compared to the perfect registration scenario in Table 7 the results were quite similar. Fusion still only appeared useful at low spectral SNR values. Although the numbers indicated a significant increase compared to nadir CEM performance, only a small increase was observed when compared to off-nadir CEM performance.

Table 9. The effect of both fusion algorithms is summarized for the 8 band multispectral sensor (nadir GSD = 0.5 m) across 5 solar zenith angles, 7 sensor zenith angles and 5 different SNR values with a single pixel of registration error. The data suggest fusion may be useful in the MS case at lower SNR values. Percent improvement metrics are defined vs. off-nadir CEM (or vs. nadir CEM in parentheses).

MS
SNR = 200

		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10						
20						
30						
40						
50						
60						
70						

MS
SNR = 100

		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10						
20						
30						
40						
50						
60						
70						

MS
SNR = 80

		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10						
20						
30				0.1% (9.3%)		
40						
50						
60						
70						

MS
SNR = 60

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10					74.6% (62.8%)
20	0.2% (32.7%)				52.8% (61.9%)
30	1.5% (31.8%)	1.6% (42.0%)	0.5% (2.3%)	0.4% (87.2%)	31.7% (21.2%)
40	0.5% (24.5%)	3.6% (37.3%)	2.7% (85.6%)	1.3% (51.2%)	
50		2.0% (18.5%)	2.5% (42.8%)		
60					
70					

MS
SNR = 50

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10	0.1% (242%)				434% (481%)
20	0.3% (221%)				230% (508%)
30	1.7% (220%)	1.7% (343%)	0.9% (893%)	0.7% (846%)	89.9% (359%)
40	0.6% (199%)	3.6% (319%)	3.1% (693%)	1.5% (723%)	0.8% (246%)
50	0.3% (120%)	2.3% (243%)	2.4% (518%)	1.0% (29.7%)	
60					
70			0.4% (237%)		

For the hyperspectral case, the results were generally similar to the perfect registration scenario. When the results in Table 10 (with registration error) were compared to the results in Table 8 (with perfect registration), a distinct decrease in fusion impact was observed. Although the SPI algorithm continued to enhance performance to some degree, the improvement relative to both nadir and off-nadir CEM decreased. The results also demonstrated that the SPI algorithm was reasonably robust to registration error. Further, the SPOT algorithm became the fusion algorithm of choice for sensor zenith angles of 50° or 60°. The increased utility of the SPOT algorithm for the scenario with registration error implied that the vehicle surfaces may not have been the most polarizing material in the scene. Instead, perhaps the polarimetric signature from nearby

window glass had been aligned with the spectral signature of target vehicle paint during the registration process, boosting the visibility of the target pixel. Finally, as the sensor zenith angle became more oblique, the percent increase in performance attained with fusion depended heavily on the reference scenario.

Table 10. The effect of both fusion algorithms is summarized for the 90 band hyperspectral sensor (nadir GSD = 3.0 m) across 5 solar zenith angles, 7 sensor zenith angles and 5 different SNR values with one pixel of registration error. The data suggest fusion is useful when the HS sensor is positioned near the sun’s specular reflection. Percent improvement metrics are defined vs. off-nadir CEM (or vs. nadir CEM in parentheses).

HS
SNR = 200

		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10		12.9% (21.1%)	4.1% (9.8%)	1.5% (6.8%)		
20		12.3% (16.9%)	14.5% (24.2%)	3.6% (10.2%)		
30		10.6% (20.8%)	16.7% (30.4%)	6.7% (17.4%)	0.8% (5.4%)	
40		5.3% (9.0%)	5.2% (31.0%)	9.5% (30.3%)	0.5% (15.5%)	
50		1.8% (20.1%)	2.5% (27.2%)	3.1% (16.7%)		
60		2.2% (12.1%)	1.5% (21.5%)			
70						

HS
SNR = 100

		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10		11.1% (19.3%)	3.8% (9.1%)	1.8% (6.7%)		
20		12.4% (17.1%)	13.6% (22.8%)	3.4% (9.7%)		
30		10.3% (20.6%)	15.8% (28.9%)	4.9% (14.8%)		
40		4.7% (8.3%)	3.9% (29.0%)	5.5% (24.5%)		
50		1.9% (20.1%)	2.7% (25.6%)	3.3% (4.5%)		
60						
70						

HS
SNR = 80

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10	11.2% (19.8%)	3.7% (8.6%)	1.5% (6.3%)		
20	12.2% (17.4%)	13.3% (22.0%)	3.1% (9.1%)		
30	10.0% (20.6%)	14.5% (27.0%)	3.9% (13.7%)		
40	4.7% (8.6%)	2.4% (26.7%)	5.0% (23.5%)		
50	2.0% (20.5%)	2.8% (24.6%)			
60					
70					

HS
SNR = 60

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10	10.9% (19.8%)	3.6% (9.3%)	1.0% (6.4%)		
20	11.8% (17.2%)	13.2% (22.7%)	2.4% (9.2%)		
30	9.7% (20.6%)	12.3% (25.3%)	2.9% (13.2%)	0.1% (1.4%)	
40	4.4% (8.6%)	4.2% (29.8%)	5.2% (23.1%)		
50	2.0% (20.4%)	2.8% (21.6%)			
60					
70					

HS
SNR = 50

Solar Zenith Angle

Sensor Zenith

Angle	20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
10	10.7% (19.4%)	3.5% (9.9%)	0.7% (7.6%)		
20	11.4% (16.5%)	12.8% (23.0%)	2.0% (10.2%)		
30	9.4% (20.0%)	10.6% (24.2%)	2.5% (14.4%)		
40	4.1% (8.0%)	4.7% (30.9%)	5.2% (24.2%)		
50	1.7% (18.4%)	3.3% (16.4%)			
60					
70					

When compared against off-nadir CEM, fusion generally provided a modest impact, while a more substantial impact was typically observed when compared against nadir CEM. Therefore, it appeared that simply moving the sensor off-nadir significantly enhanced performance in some cases. One possible explanation for this lies in the way

the problem was framed—since target detection was performed at the smaller PI GSD, a single hard-to-find target vehicle might represent many missed pixels at nadir. However, as the sensor moved off-nadir, the number of pixels on the focal plane remained constant but the GSD increased. As a result, the same difficult target represented fewer missed target pixels. The effects of a larger GSD and fewer difficult targets could therefore combine to increase the area under the resulting off-nadir ROC curve—while reducing the ability to exactly pinpoint the location on the ground of any targets identified.

6.4 Multiple Sensor Platforms

The first step in analyzing the multiple sensors scenario was determining the optimum viewing geometry for each sensing modality at a given solar zenith angle as described in Sec. 4.12, and Table 11 summarizes the results. In general, the polarimetric sensor was more effective at more oblique sensor zenith angles, while the multispectral sensor was most effective near nadir. Interestingly, the hyperspectral sensor was most effective at reasonably oblique sensor zenith angles rather than angles near nadir.

Table 11. For each solar zenith angle, the sensor zenith and azimuth angles where each modality was most effective are shown. The polarimetric value is shown in blue, the multispectral value is shown in red, and the hyperspectral value is shown in black.

SNR = 200		Solar Zenith Angle				
Sensor Zenith		20 (1200)	34 (1000)	55 (0800)	66 (0700)	77 (0600)
Angle						
10		170	170	160		160 / 200
20					170	
30		185 / 195				
40				190	195	190
50			185 / 175		185	
60				180		
70						

Once the ideal viewing angles had been established, the polarimetric and spectral images were registered as described in Sec. 4.11 (and shown in Figure 57) based on 36 hand-selected GCPs.



Figure 57. Several RGB DIRSIG images are displayed to illustrate the multiple sensor registration process for a solar zenith angle of 20°. (Left) The polarimetric data was drawn from an image captured at a sensor zenith angle of 30° and a sensor azimuth angle of 185°. (Center) The hyperspectral data was drawn from an image captured at a sensor zenith angle of 30° and a sensor azimuth angle of 195°. (Right) The hybrid image contains registered polarimetric and hyperspectral data from multiple sensors.

Next, the quality of registration was assessed via the total RMS error metric shown in

Table 12. A total RMS error value (in pixels) was calculated for each of the registered spectral and polarimetric data sets derived from the viewing geometries previously identified in Table 11.

Table 12. The total RMS error (in smaller GSD MS/PI pixels) is displayed for the MS/HS & PI registered data set derived from the viewing geometries described in Table 11.

Solar zenith angle (degrees)	MS fusion RMS registration error (pixels)	HS fusion RMS registration error (pixels)
77 (0600)	5.24	4.98
66 (0700)	4.46	2.20
55 (0800)	5.19	2.92
34 (1000)	5.55	1.49
20 (1200)	3.51	0.94

Finally, the hybrid spectral-polarimetric images were analyzed as described in Sec. 4.12, producing an area under the ROC curve for each off-nadir algorithm. Those areas were

then divided by the area under the ROC curve generated by applying the multispectral (or hyperspectral) CEM algorithm at nadir, producing the ratio values shown in Table 13.

Table 13. The values in the table represent the ratio of the area under the test algorithm ROC curve to the area under the multispectral (or hyperspectral) CEM ROC curve at nadir. Red means the off-nadir algorithm didn't perform as well as nadir CEM. Green and yellow both represent cases where the off-nadir algorithm was better than nadir CEM, but green represents the best choice.

Solar Zenith Angle	MS CEM	MS SPI	MS SPOT	HS CEM	HS SPI	HS SPOT
77 (0600)	0.456	0.222	0.270	1.229	0.306	1.087
66 (0700)	0.362	0.150	0.275	1.329	0.018	1.174
55 (0800)	0.450	0.392	0.372	1.339	0.717	1.276
34 (0900)	0.425	0.408	0.387	1.428	0.438	1.310
20 (1000)	0.457	0.433	0.423	1.274	1.150	1.201

The results indicated that CEM at nadir was the most powerful algorithm for the low GSD multispectral scenario, which was consistent with results shown in Sec. 6.2. In the hyperspectral scenario, the results showed that the off-nadir CEM algorithm yielded the best performance. This finding was in contrast to Sec. 6.2, which indicated that decision fusion should produce the best results. Further, in the hyperspectral case, the SPOT pixel fusion algorithm outperformed both the SPI decision fusion algorithm and the CEM algorithm at nadir. The root of this effect was the relative weighting of polarimetric information to spectral information in the fusion algorithms. The SPI decision fusion algorithm weighted the polarimetric TAD score and spectral CEM scores almost equally, while in the SPOT decision fusion algorithm the polarimetric information was weighted as only one of ninety one total bands of information. In essence, the results indicated that the polarimetric information was not being correctly aligned with the spectral information, and the total impact of the errors introduced by the registration process was proportional to the weight of the polarimetric information in each of the

fusion algorithms. Two main factors were responsible for this disconnect: registration error and target obscuration.

The most likely source of error came from applying 2-D registration techniques to images taken from different oblique perspectives that contain significant topographic variation [Schott 2007]. Since the 2-D solution calculates everything relative to a common plane, values associated with extremely high or low points will be offset spatially from their true location after the registration process. The second major source of error was obscuration, which occurred when a target was visible in one image but not in the other due to a change in the sensor's perspective. If no information about the target ever reached the focal plane for a given modality, then the required values were drawn from the nearest neighbor pixel. In the highly cluttered environment studied for this scenario, it was likely that the nearest neighbor and the pixel of interest were not composed of the same material—meaning a pure target pixel would appear as a mixed pixel in terms of spectral and polarimetric information, making it even harder to detect.

In conclusion, fusing spectral and polarimetric data from separate sensors did not produce the expected increase in performance. Rather than highlighting a phenomenological effect, the difficulty was due to inaccurately aligning correct spectral and polarimetric data for every pixel in the scene. The registration effort was hampered by differences in oblique perspective and significant topological variation in the scene. Since the quality of the fusion product seems directly tied to the registration accuracy, future efforts could revisit this scenario by applying advanced 3-D registration techniques.

6.5 Phenomenology Investigation

The results presented in Sections 6.2 and 6.3 indicated that the SPI decision fusion algorithm was the more useful fusion algorithm, so the next step was to investigate the phenomenology behind the observed effect. The viewing geometry investigated was drawn from the 0800 scenario of Sec. 6.2, with a solar zenith angle of 55° , a sensor zenith angle of 40° , a sensor azimuth angle of 180° , perfect registration, a polarimetric GSD of 0.5 m and a hyperspectral GSD of 3 m. In this configuration, SPI provided a 24% improvement in performance, and an RGB image of this scene is shown in Figure 58.



Figure 58. RGB DIRSIG image rendered with a solar zenith angle of 55° , a sensor zenith angle of 40° , and a sensor azimuth angle of 180° .

The CEM, TAD and SPI score maps were compared to the DIRSIG truth target map, highlighting the test algorithm score for every target in the scene. In this manner, high scoring, easier to find targets could be separated from low scoring overlooked targets.

6.5.1 CEM Performance

The CEM algorithm was the first algorithm tested, since it was both the basis for comparison and a critical component of the SPI algorithm. The CEM algorithm

quantified how well a test pixel matched a reference target spectrum, and Figure 59 shows the target pixels that were the easiest in the scene for the CEM algorithm to find. These targets seemed to be either well-lit target pixels or targets placed in a homogenous background. Highlighting the target in a homogenous background makes sense because the target pixels are spectrally distinct from the spectral clutter of all nearby background pixels. Similarly, well-lit target pixels are much more likely to exactly match the measured target spectrum used as an input to the CEM algorithm since the reflected target signature dominates the recorded signal and the original spectrum was presumably recorded under ideal viewing conditions.

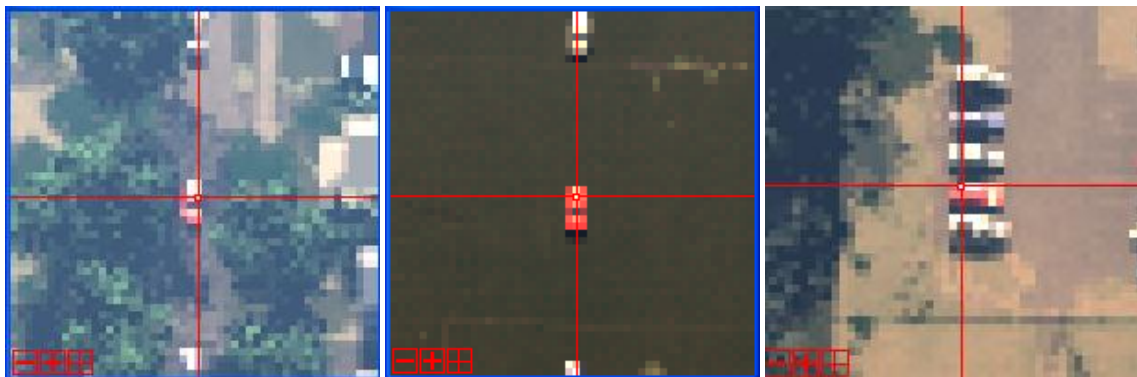


Figure 59. The easiest targets in the scene for the CEM algorithm to find were either well-lit target pixels or targets in a homogenous background. Note that these images are all shown at the lower polarimetric GSD (0.5 m) for later comparison to the SPI and TAD algorithms.

Once the more obvious targets had been noted, the focus turned to targets that were overlooked. Figure 60 shows the three most difficult targets to find in the scene. In the first case, the vehicle was in a heavily shadowed and obscured location, so the test pixel was a heavily mixed pixel with a low SNR. In the second case, the vehicle was well-lit but surrounded by relatively bright man-made features. At a large GSD, it may have been difficult to discriminate the bright mixed target pixel from the bright background pixel. Additionally, perhaps this viewing geometry, coupled with the car's

surface geometry, put the sensor in the sun's specular reflection. If so, the spectral whitening effect would mask the target signature and make the pixel harder to differentiate from other bright false alarms. Finally, the last target was well-lit but located in a cluster of decoy vehicles. The target signature was therefore likely to be mixed with a significant contribution from the other vehicles when the large spectral GSD was applied, again reducing the probability of an exact match with the desired target spectrum in the CEM algorithm.

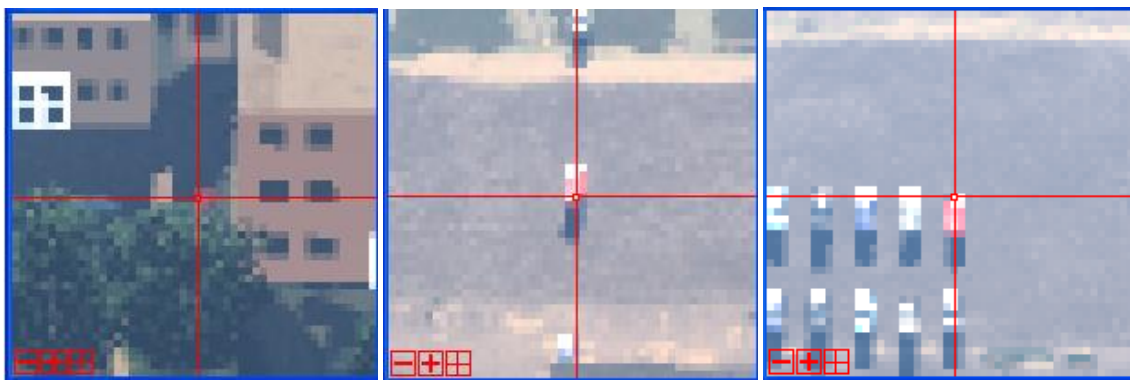


Figure 60. The hardest targets in the scene for the CEM algorithm to find were shadowed targets, targets producing a specular reflection from the sun, and targets located in a heavily cluttered environment. Note that these images are all shown at the lower polarimetric GSD (0.5 m) for later comparison to the SPI and TAD algorithms.

6.5.2 TAD Performance

The TAD algorithm examined the S_1 and S_2 bands of polarimetric information, and separated polarimetrically anomalous pixels from uninteresting background pixels. Figure 61 demonstrates that the easiest target pixels for the TAD algorithm to identify were the well-lit edges of target vehicles, representing sun-target-sensor geometries where the combined effects of the pBRDF and the polarized upwelled radiance produced a distinct polarimetric signal. Additionally, the easiest targets to identify were all in a significant amount of spectral clutter with several nearby decoy vehicles. When

compared to the CEM analysis, the small polarimetric GSD better preserved distinctive information about the target. The TAD algorithm then compared those distinctive signatures to the scene as a whole and flagged them as different.

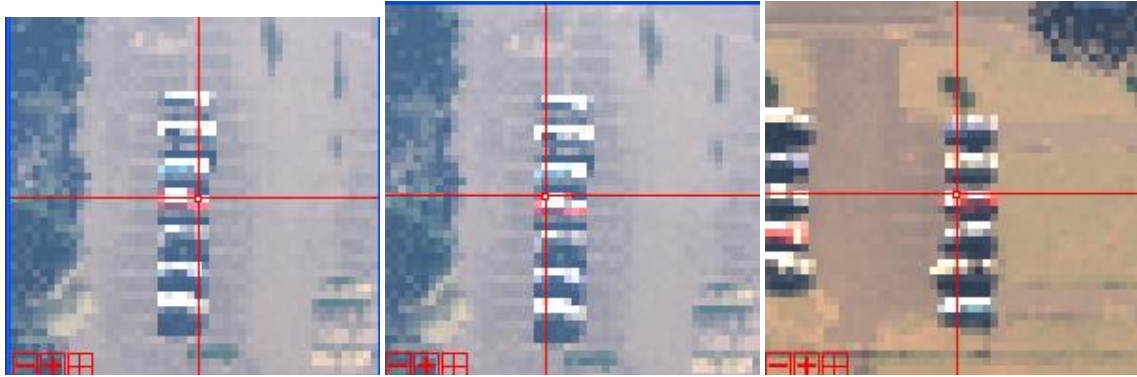


Figure 61. The easiest targets in the scene for the TAD algorithm to find were well-lit edges of target vehicles. Note that these RGB images are all shown at the polarimetric GSD (0.5 m).

However, the TAD algorithm was not successful in all cases. Figure 62 shows that shadowed and obscured targets were among the hardest target pixels to discern from the background.

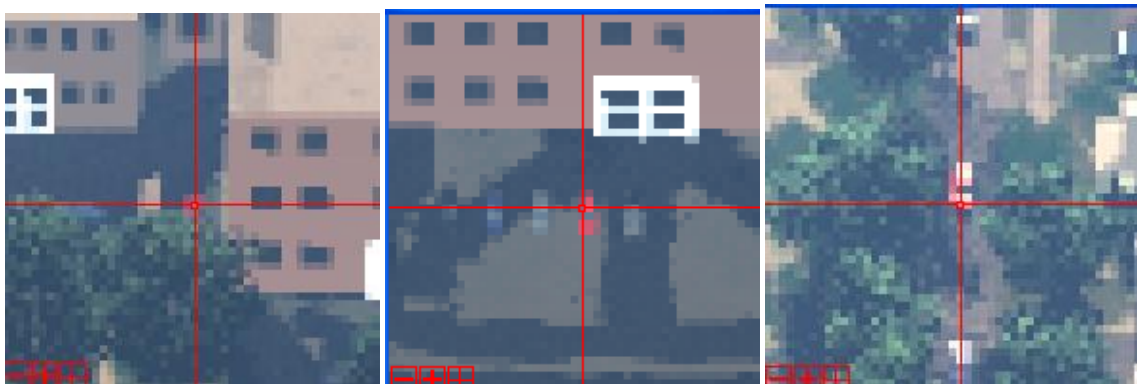


Figure 62. Shadowed and obscured targets vehicles were among the least anomalous target pixels in the scene. Note that these RGB images are all shown at the polarimetric GSD (0.5 m).

When compared to other well-lit man-made objects in the scene, shadowed and obscured targets produced a much less anomalous polarimetric signature for three main reasons. First, since the sun was blocked, there was no dominant directly reflected radiance term

contributing to the target's polarimetric signature. Rather, the downwelled radiance, with contributions from a variety of polarization states, interacted with the target material pBRDF to produce a much smaller polarized signature than that of the well-lit targets. Second, the dominant polarization state of the reflected downwelled radiance from shadowed target vehicles was not necessarily consistent with that from the well-lit vehicles; therefore this viewing geometry may not have been ideal after the upwelled radiance polarization state was taken into account. Third, the raw radiance values of shadowed areas in the scene were relatively low when seen through each linear polarizing filter. When the random noise field was applied to each pixel, the magnitude of the noise was based on a scene-wide SNR value—but since the polarimetric signature was based on differences between filtered images when expressed as a Stokes vector, even relatively small noise values had a dramatic effect. In essence, the polarimetric SNR of 200 appeared insufficient to preserve the small differences between target and background pixels for shadowed areas.

6.5.3 SPI Performance

Once the spectral and polarimetric algorithms had been assessed independently, the SPI algorithm scores were examined to determine which target scores were enhanced the most by fusing polarimetric information with existing spectral information. Eqs. 40-42 demonstrate how the SPI impact was quantified for each target pixel by comparing the SPI score to the squared CEM score:

Figure 63 shows some of the target pixels where the SPI algorithm had the most impact. Predictably, the SPI scores increased most where the TAD algorithm flagged the most anomalous pixels, and as in Sec. 6.5.2, the most anomalous target pixels tended to be well-lit edges of vehicles.

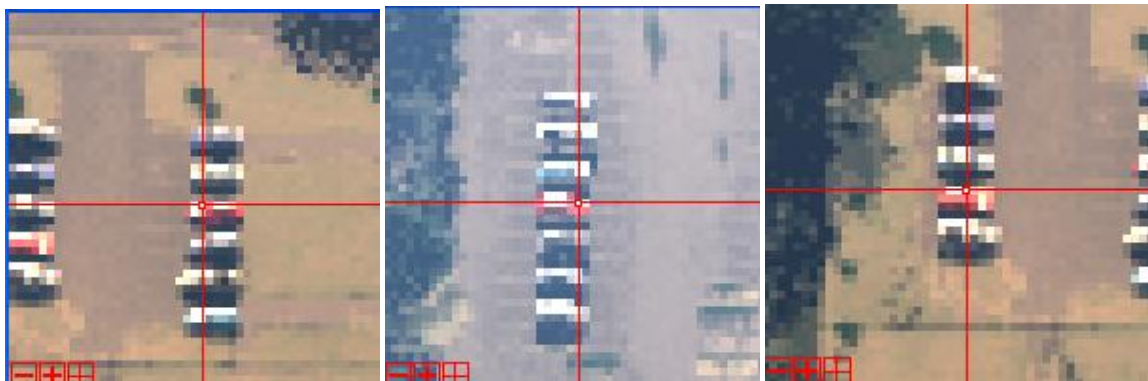


Figure 63. The SPI algorithm had the most impact on well-lit edges of target vehicles. Note that these RGB images are all shown at the polarimetric GSD (0.5 m).

Next, the SPI score map was examined to determine which target pixels were least impacted by considering the additional polarimetric information, and the results are shown in Figure 64. Two of the three least impacted targets are shadowed and obscured. Recall that Sec. 6.5.1 showed that the CEM algorithm missed shadowed targets while Sec. 6.5.2 showed that the TAD algorithm also struggled with shadowed targets. Given that both input algorithms missed this class of targets, it's not surprising that the SPI algorithm also missed the shadowed targets. The third least-impacted target was a well-lit vehicle, a slight surprise until one considers that Sec. 6.5.1 showed the CEM algorithm assigned a very low score to that vehicle. Because of the way the SPI algorithm was constructed, the polarimetric information was able to boost the pixel's score somewhat, but the score was then multiplied by a very small CEM score to produce a small final

score. This particular target pixel was missed because the CEM algorithm effectively vetoed the pixel.

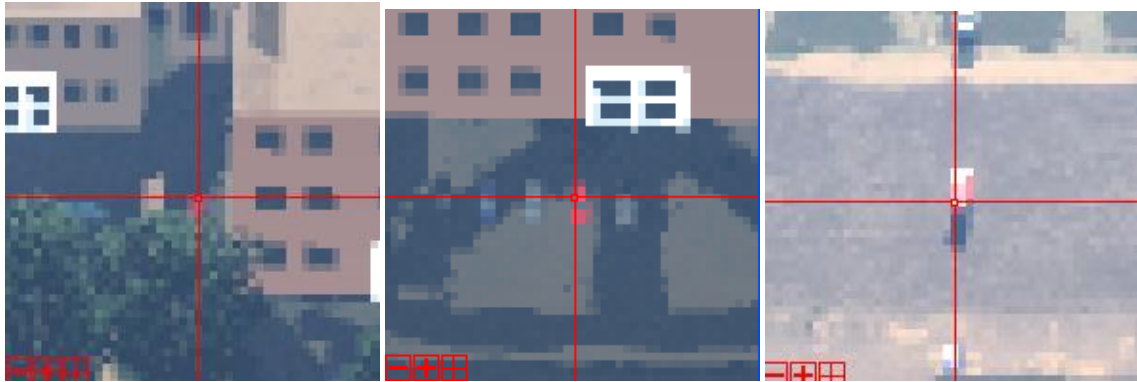


Figure 64. The SPI algorithm had the least impact on shadowed targets or bright targets with a low CEM score. Note that these RGB images are all shown at the polarimetric GSD (0.5 m).

In essence, the SPI algorithm appeared to boost the score of well-lit target pixels in clutter where a heavily mixed spectral pixel was assigned an average CEM score initially. Performance gains with the SPI algorithm occurred by weeding out bright false alarm background pixels from spectrally mixed target pixels rather than by finding targets that the CEM algorithm completely missed.

6.6 Missed Dark Targets: Sampling or SNR?

Once it became clear that both the CEM and TAD algorithms missed dark targets, the next task was to determine whether this was an SNR issue or a sampling issue. The same scene was rendered again in DIRSIG at a much higher fidelity to rule out issues with DIRSIG sampling. The hemispherical sampling of the pBRDF was increased dramatically, the ray tracing was set to a 5×5 oversample grid, and the number of bounces was increased to ensure adjacency effects were captured. These modifications increased the computation time to render the scene from about 12 hours to about a week—impractical for a full trade study with today’s computers, but not unreasonable to

test the concept. Figure 65 shows a zoomed-in image of an ROI containing several shadowed vehicles—one of which is a target—drawn from Figure 58.

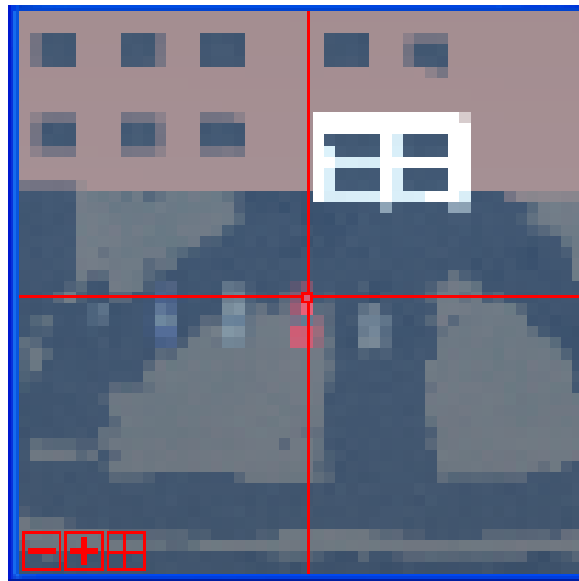


Figure 65. An ROI of Figure 58 contains several shadowed vehicles, one of which is a target. This ROI was evaluated with several different metrics for both the standard trade study scene and the high fidelity oversampled scene to determine whether SNR or sampling issues were the primary reason for overlooking shadowed targets.

6.6.1 Increased Sampling Fidelity

Once the higher fidelity scene was generated, both scenes were analyzed with the spectral and polarimetric SNR values fixed at 10,000. Although a different random noise field was applied to each scene, the variation applied to any given pixel was quite small, isolating the effects of sampling by effectively ruling out any influence due to noise.

Figure 66 shows the hyperspectral CEM ROC curves derived from scenes generated with both the standard method described in Sec. 4.4 and the higher fidelity sampling described in Sec. 6.6. Although the ROC curves do not match exactly, the trends are consistent between the two.

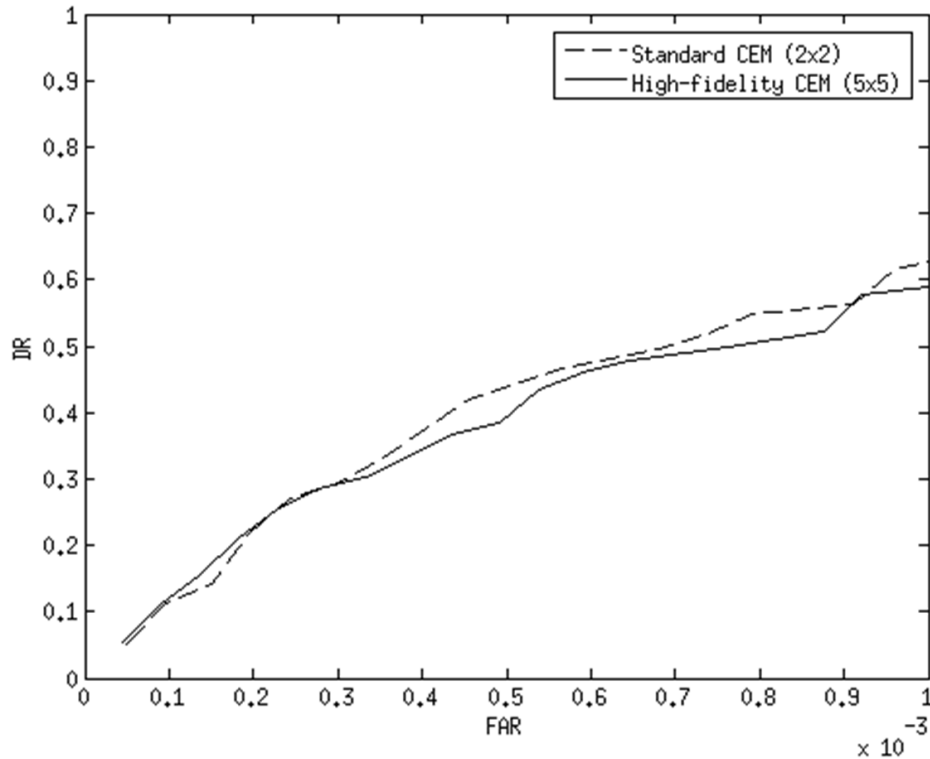


Figure 66. The HS CEM ROC curves were derived from scenes rendered with both the standard fidelity sampling (dashed line) and the higher fidelity spatial and pBRDF sampling (solid line).

Because the pBRDF was also sampled more extensively, the performance of both fusion algorithms was evaluated as well. Figure 67 shows the hyperspectral SPI ROC curves derived from the same scene with both the standard sampling method described in Sec. 4.4 and the higher fidelity sampling described in Sec. 6.6, while Figure 68. Increasing the sampling fidelity seemed to slightly decrease the performance of the SPI algorithm, while increasing SPOT performance for some targets and decreasing performance for others. However, the general shapes of the ROC curves were consistent. Additionally, since the measure of merit described in Sec. 4.9 was a percent difference between the area under the fusion and CEM ROC curves, slightly overestimating performance in both the fusion and purely spectral cases with the coarser sampling should have a minimal effect on the final answer.

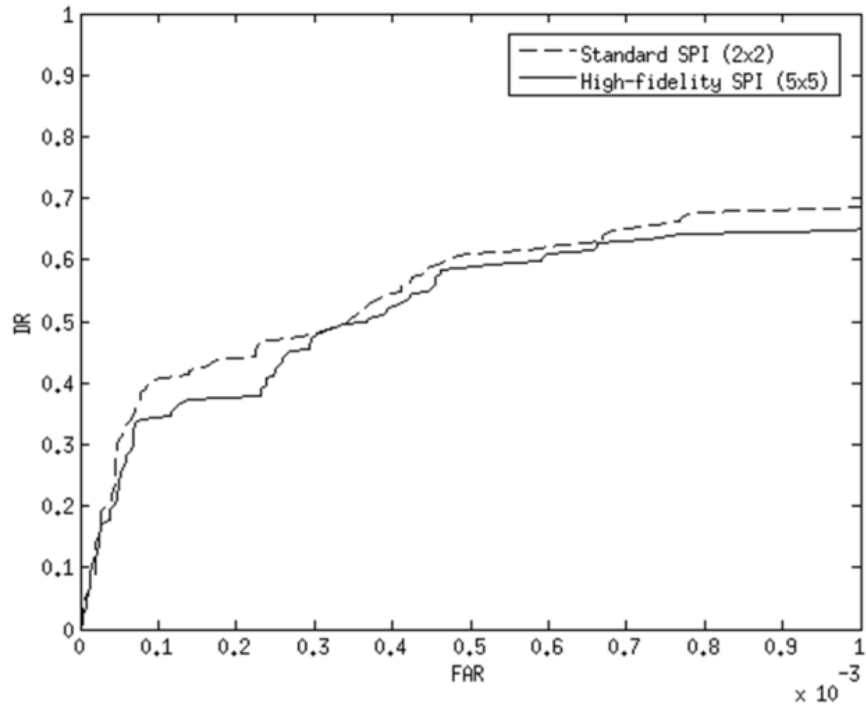


Figure 67. HS SPI ROC curves were derived from scenes rendered with both the standard fidelity sampling (dashed line) and the higher fidelity spatial and pBRDF sampling (solid line).

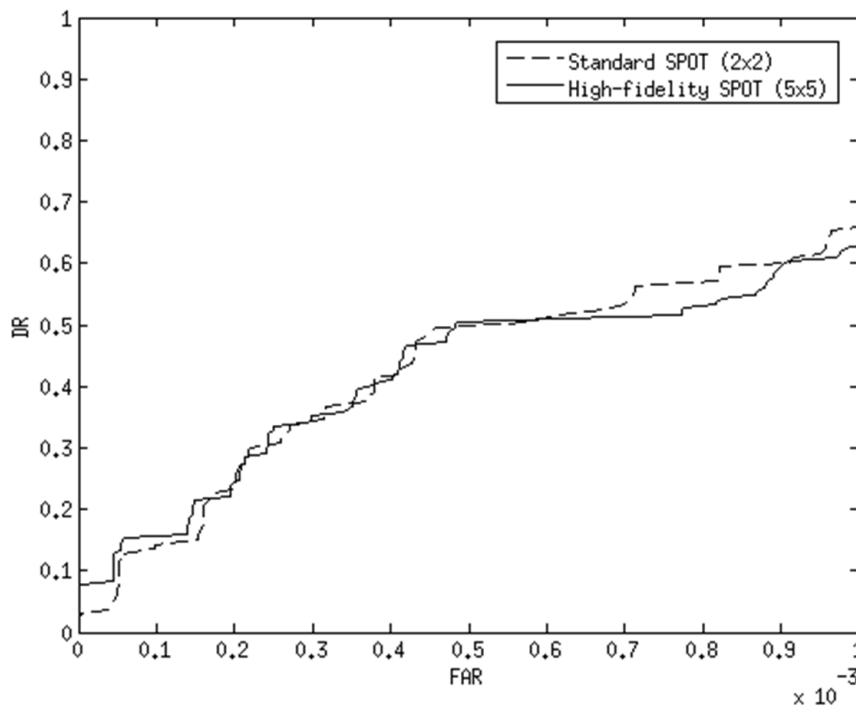


Figure 68. HS SPOT ROC curves were derived from scenes rendered with both the standard fidelity sampling (dashed line) and the higher fidelity spatial and pBRDF sampling (solid line).

Next, the ROI shown in Figure 65 was analyzed to determine the DOLP, TAD score, and hyperspectral CEM score for each pixel. Figure 69 shows the DOLP metrics generated from the ROI for both levels of sampling. In both of these images, the shadowed vehicles were barely distinguishable from the nearby shadowed background pixels. In fact, the difference in DOLP between the shadowed vehicles and background was only about 0.003 for both images, while the DOLP magnitude of the shadowed vehicles hovered around 0.138 in both cases. As a result, it appeared that increasing the sampling did not affect the DOLP metric.

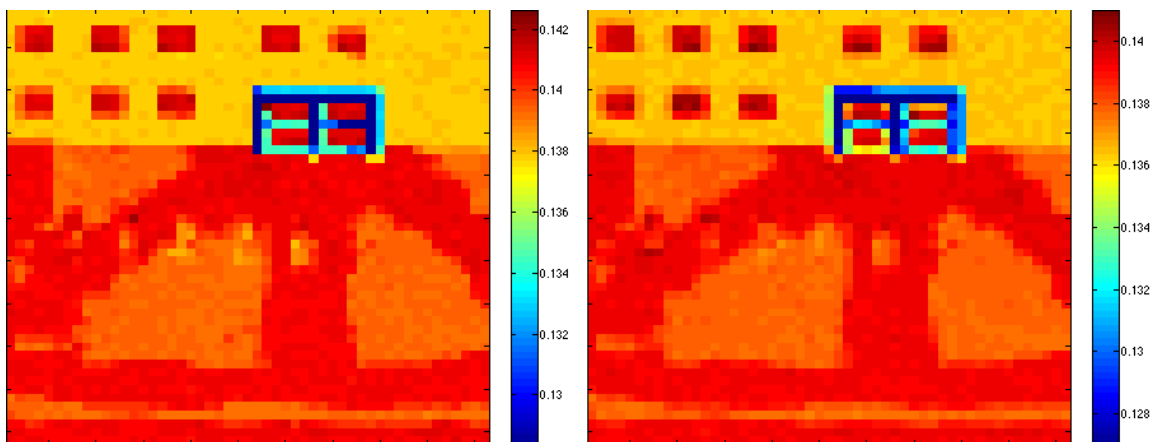


Figure 69. (Left) The DOLP metric derived from scenes rendered with the standard fidelity sampling (2x2 oversamples) for this trade study. (Right) The DOLP metric derived from scenes rendered at the higher fidelity sampling (5x5 oversamples, more extensive pBRDF) for this trade study.

Figure 70 shows the TAD score of the ROI derived from both the normally sampled and the highly oversampled images. Some differences were apparent—the edges of vehicles in the highly oversampled image were easier to distinguish from the background pixels. In the normally sampled image, the difference between the most anomalous shadowed vehicle pixels and the shadowed background pixels was about 0.003. In contrast, the difference between the vehicle edges and the background was

about 0.006 in the highly oversampled image. In both cases, the shadowed background pixels were given an almost zero by the TAD algorithm, so the difference between the background and the shadowed vehicles was essentially the magnitude of the TAD score. Although this implied almost twice as much contrast between shadowed vehicles and the shadowed background pixels was achieved by highly oversampling, recall that the TAD score could range between zero and one. When the entire scene was considered, the differences in scores assigned by the TAD algorithm due to better sampling were insignificant because of the small magnitudes of the score—a 200% improvement in performance was relatively meaningless because the initial value was so small.

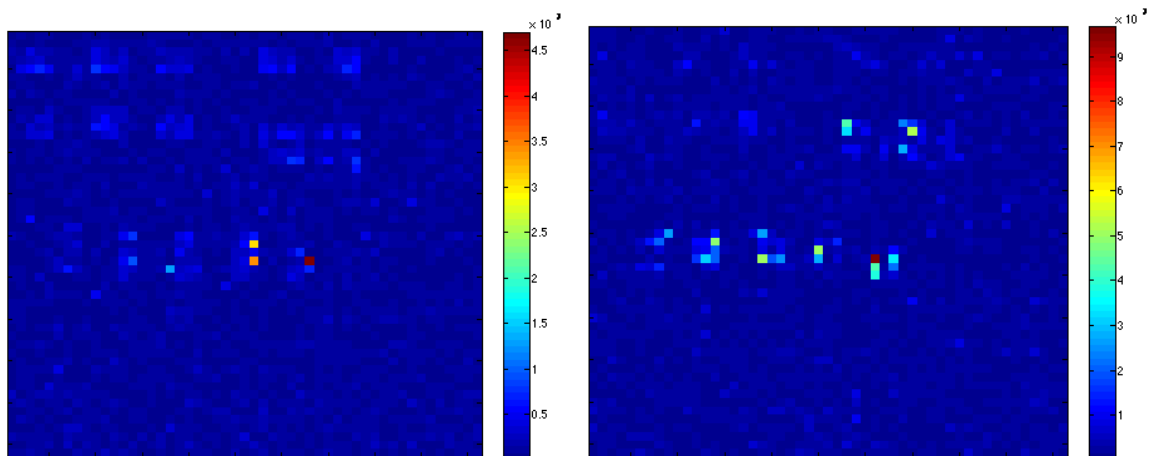


Figure 70. (Left) The TAD score derived from scenes rendered with the standard fidelity sampling (2x2 oversamples) for this trade study. (Right) The TAD score derived from scenes rendered at the higher fidelity sampling (5x5 oversamples, more extensive pBRDF) for this trade study.

Figure 71 shows the hyperspectral CEM scores derived from the ROI. In both images, the pixel containing the target vehicle was scored highest. However, the difference in score between the target pixel and the background pixels was about 0.03, and the target pixel scored only 0.035 out of one in both cases. Therefore, increasing the sampling did not significantly change the CEM scores.

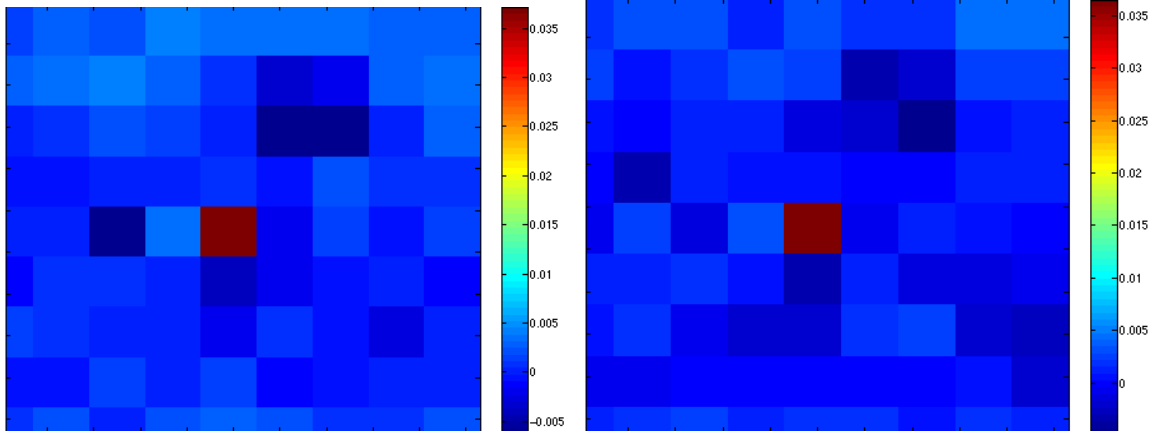


Figure 71. (Left) The HS CEM score derived from scenes rendered with the standard fidelity sampling (2x2 oversamples) for this trade study. (Right) The HS CEM score derived from scenes rendered at the higher fidelity sampling (5x5 oversamples, more extensive pBRDF) for this trade study.

In summary, examining the desired ROI with the DOLP metric, TAD score, and hyperspectral CEM scores derived from both the normally sampled and highly oversampled images demonstrated that the failure to find shadowed targets was not due to insufficient sampling in DIRSIG. Once sampling issues were ruled out, the next avenue to explore was the effect of the polarimetric SNR.

6.6.2 SNR Effects

Further investigation of the SNR required to distinguish shadowed target pixels yielded some surprising results. To verify that missing shadowed targets was an issue of SNR, this ROI was analyzed via the DOLP, S_1 and S_2 metrics first at a polarimetric SNR of 200—the value used throughout this experiment—and then at a polarimetric SNR of 10,000 to represent an upper bound on the scenario after any specialized image-processing techniques had been employed. Figure 72 shows the DOLP value for every pixel in the ROI after the noise field has been applied.

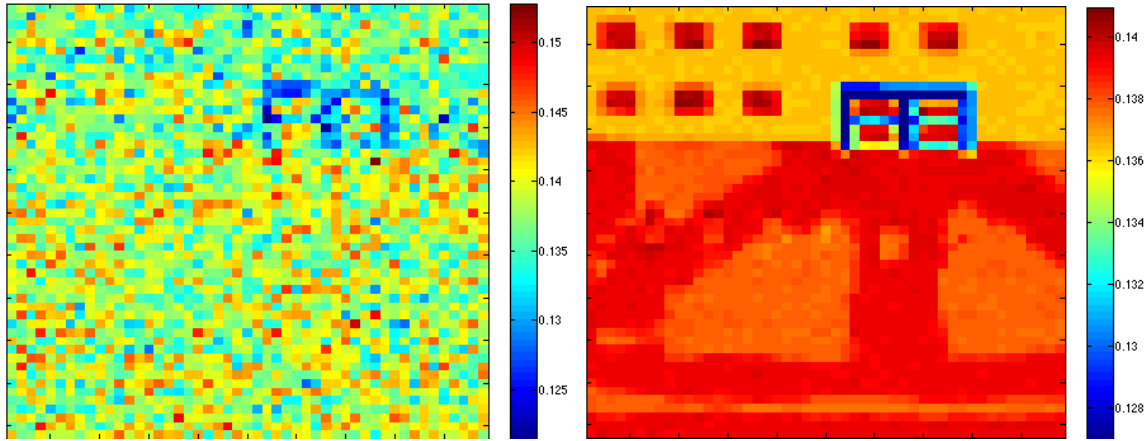


Figure 72. (Left) DOLP metric for ROI in Figure 65 at SNR = 200. (Right) DOLP metric for the same ROI at SNR = 10,000. In the lower SNR case, the added noise masked any signal from the vehicles. Even in the higher SNR case, the vehicles were barely distinguishable from the background.

In the SNR = 200 case, the vehicles were indistinguishable from the background and the entire ROI (except the doorway) appears as a random noise field. However, when the SNR = 10,000 case was examined, the vehicles were perhaps visually distinguishable, but the difference between the vehicles and the background was a DOLP value of only about 0.002.

Next, the ROI was analyzed via the S_I metric, which measures the difference between incoming horizontally and vertically polarized light. Figure 73 shows the results when the same SNR values were considered. In the SNR = 200 case, the added noise again overpowered any distinct signals from the vehicles, while in the higher SNR case, the vehicles and glass windows in the building were distinguishable from the background. However, the building windows and doorway were still more distinct than the vehicles, and the S_I value from the vehicles differed from the background by only about 0.015 .

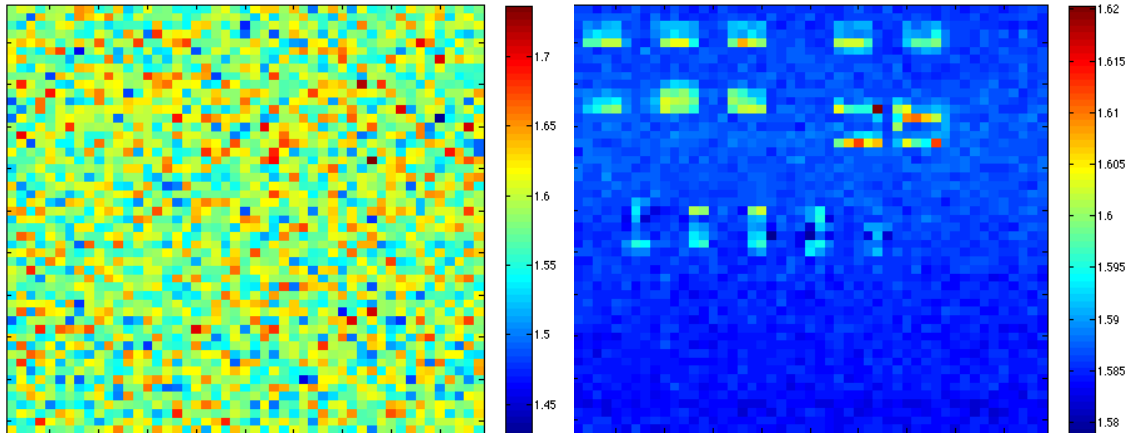


Figure 73. (Left) S_1 metric for ROI in Figure 65 at SNR = 200. (Right) S_1 metric for the same ROI at SNR = 10,000. In the lower SNR case, the added noise washed out any structure in the signal, while in the higher SNR case, the vehicles and glass windows in the building were distinguishable from the background.

Finally, the S_2 metric was computed for every pixel in the ROI for the same two SNR levels, and the results are shown in Figure 74. As with the S_1 metric, the added noise in the SNR = 200 case washed out any distinct signals, while in the higher SNR case, some vehicle pixels could be discerned from the background. However, the differences in S_2 values between the vehicles and the background were only about 0.004

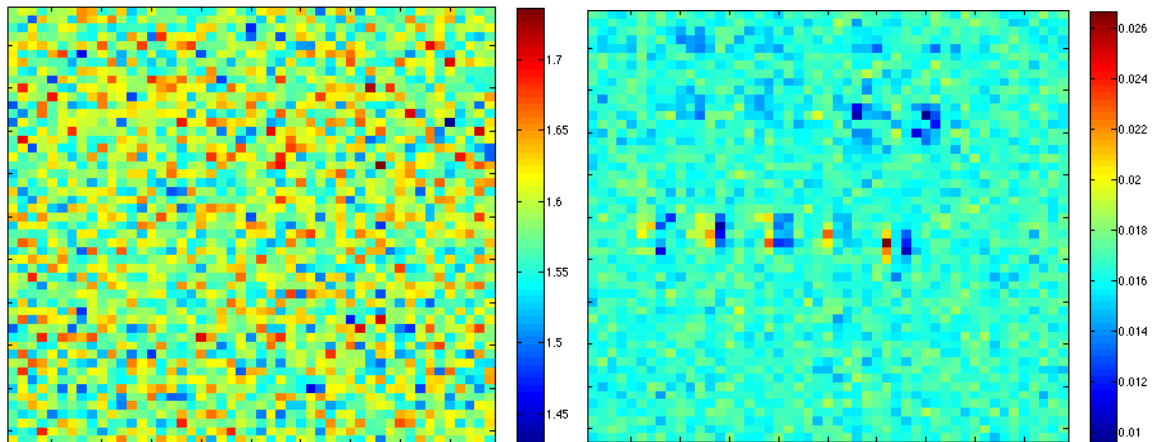


Figure 74. (Left) S_2 metric for ROI in Figure 65 at SNR = 200. (Right) S_2 metric for the same ROI at SNR = 10,000. In the lower SNR case, the added noise washed out any structure in the signal, while in the higher SNR case, the vehicles were barely distinguishable from the background.

In summary, a polarimetric SNR = 200 was simply too low to differentiate shadowed vehicles from shadowed background pixels in a scene with both shadowed and well-lit targets. Further, an SNR = 10,000 allowed slight differentiation between shadowed target vehicles and nearby shadowed background pixels in the S_1 and S_2 metrics—but not in the DOLP. This disconnect reinforces the idea that results obtained by analyzing DOLP as opposed to S_1 or S_2 are not necessarily equivalent. Finally, these experiments demonstrated that finding shadowed targets with polarimetric sensors—in a cluttered, high dynamic range scene—is a tough problem that will require very specialized hardware.

6.7 Section Summary

This section first described results from the three main experiments outlined in Sec. 2.2, then focused on the phenomenological reasons the different algorithms found (or missed) different types of targets. An additional short study examined whether sampling issues in DIRSIG or polarimetric SNR contributed more to the inability to find shadowed targets.

The first experiment assumed a spectral and polarimetric sensor mounted on the same platform with the images perfectly registered. Polarimetric information was shown to be of little value in the lower GSD multispectral scenario, but to add significant value in the higher GSD hyperspectral scenario. The impact of polarimetric information varied significantly with geometry. In general, the SPI decision fusion algorithm outperformed the SPOT pixel fusion algorithm, suggesting an enhanced capability to perform target detection with the sensor positioned in the sun's specular lobe.

The second experiment introduced one pixel of registration error between the spectral and polarimetric data. Again, very little improvement was observed in the multispectral scenario. In the hyperspectral scenario, the impact of polarimetric information was reduced but a benefit was still observed. The data fusion algorithms were shown to be reasonably robust for a more realistic scenario, with the SPI algorithm again outperforming the SPOT algorithm.

Next, spectral and polarimetric information captured from different perspectives were combined to model fusion from separate sensor platforms. No benefit was observed in either the multispectral or hyperspectral case, and the degradation in performance was attributed to limitations in registration accuracy. Future efforts could potentially increase the fusion benefit by applying advanced 3-D registration techniques, but those techniques are outside the scope of this work.

Closer inspection of the CEM, TAD and SPI score maps revealed trends about which targets were easier or harder to find. The CEM algorithm easily flagged well-lit targets or targets in a homogenous background, but missed shadowed and obscured targets or targets in significant clutter. The TAD algorithm found well-lit edges of vehicles particularly anomalous, but also missed shadowed targets. When the scores were combined via the SPI algorithm, the scores of well-lit target pixels in clutter were increased to separate bright false alarm background pixels from spectrally mixed pixels. The final experiment demonstrated that targets in the shadows were particularly hard to find due to the polarimetric SNR being too low, as opposed to insufficient sampling when the scene was rendered in DIRSIG.

Although these results provided great insight as to how spectral and polarimetric data might be combined to enhance target detection, they were derived from chasing one particular target spectrum for fixed GSD values. Further simulation work was planned to evaluate the sensitivity of the fusion algorithms to different target spectra and to a range of GSD values.

7 Simulation Supporting Multimodal Sensor System Design Requirements

7.1 Section Overview

One common approach to building new sensor systems is to recognize the reality of fiscal constraints and view cost as an independent variable [Defense Acquisition University 2009]. From this point of view, cost is treated as equally important as performance and schedule in program decisions. A primary technique employed by these programs is to perform trade-off studies analyzing the performance of a variety of different sensor configurations that might fit the need.

This work has demonstrated a process for assessing the impact of fusing spectral data with polarimetric information for an urban target detection application under a variety of different sun-target-sensor geometries, and Sec. 6 describes the impact on target detection due to spectral and polarimetric data fusion for two specific scenarios. However, the GSD, spectral SNR and spectral resolution of the sensor are all factors that might affect the performance of the SPI fusion algorithm, and these can easily be varied via simulation to conduct performance trade-offs. Further, it is expected that changing the target spectrum will affect performance, so I also sought to assess how robust the SPI algorithm is across multiple targets. This section demonstrates how DIRSIG simulations allow a design team to isolate the impact of varying one particular parameter of interest in a series of trade-off studies, serving as an integral step in considering cost as an independent variable in the sensor acquisition process.

7.2 Method

This experiment leveraged the same synthetic scene described in Sec. 4.4, but examined the performance of several different sensor combinations. Both sensing modalities were located on the same platform and for these simulations their outputs were perfectly registered to each other. The multispectral sensor employed in this study spanned eight uniform spectral response bands that roughly corresponded to DigitalGlobe's WorldView-2 satellite, described in Sec. 4.3, but was evaluated with a range of nadir GSDs. In contrast, the polarimetric imager used in this study had a fixed nadir GSD of 0.5 m and a uniform spectral response from 0.4 μm – 0.7 μm . Images of the scene were again recorded through a set of four ideal linear polarimetric filters oriented at 0°, 45°, 90° and -45°, then the polarimetric Stokes vectors were calculated as described in Sec. 3.6. The sensor viewing geometries were the same as those described in Sec. 4.5, and the measure of merit was the same as described in Sec. 4.9.

7.2.1 Varying SNR

The results in Sec 6.2 showed a dependence on spectral SNR values, where a multispectral GSD of 0.5 m was used, so the value of additional polarimetric information is again expected to vary for different spectral SNR values. For example, I hypothesized that for poorer initial spectral detection estimates, the impact of additional polarimetric information may grow. As described in Sec. 4.3, an SNR of 200 for the mean scene signal level was chosen to represent the best-case spectral SNR in this study, while lower spectral SNR values of 100, 80 and 50 were also examined. Meanwhile, the polarimetric SNR value was held constant at 200 for all cases. To isolate the impact of varying SNR, the sensor nadir GSD values in this study were fixed at 3 m for the multispectral sensor

(yielding values ranging from 3.1 m to 25.6 m for the off-nadir views) and 0.5 m for the polarimetric sensor (yielding values ranging from 0.51 m to 4.3 m for the off-nadir views).

7.2.2 Varying Spectral GSD

Besides the system SNR, another factor that should impact performance is the GSD. A variety of viewing geometries were identified in Sec. 6.2 where incorporating polarimetric information with spectral information would enhance performance, but the impact was only quantified for a single GSD value for each spectral resolution. Specifically, the scenario with a solar zenith angle of 34° , a sensor zenith angle of 30° and a relative azimuth angle of 190° was identified as one of the most promising viewing geometries for spectral and polarimetric fusion via the SPI algorithm. Therefore, an experiment was conducted to analyze that scenario repeatedly while varying the spectral nadir GSD from 0.5 m to 12 m but holding the polarimetric nadir GSD constant at 0.5 m. With these values, the quality of the spectral data ranged from several pure pixels on target to mixed pixels to significantly sub-pixel targets. At each GSD value, the areas under the off-nadir CEM and off-nadir SPI ROC curves were calculated.

7.2.3 Varying Target Spectrum

A final variable that could significantly affect the fusion results is the target spectrum, since it serves as a key input to the CEM algorithm—and thus the SPI algorithm. All inputs were known exactly with the synthetic DIRSIG imagery, so it was a straightforward process to flag green station wagons in the scene as the target vehicle instead of the original red station wagon. To isolate the spectral effects, the two vehicles were attributed with the same polarimetric BRDF model—meaning the specular (Fresnel)

components were the same but the diffuse component varied with the spectral shape of each target. Since the polarization of any diffusely reflected radiance transmitting through the first surface interface has been shown to be significant only for very bright painted surfaces [Ellis 1996]—not representative of our target vehicles—ignoring these interactions in the DIRSIG model should not impact the results. The SCR metric described in Sec. 5.4 was used to quantify how spectrally different each of the two vehicles was from the cluttered background, with the polarimetric nadir GSD of 0.5 m used, while the nadir multispectral GSD was fixed at 2 m. Further, both the spectral and polarimetric SNR values were fixed at 200. Each of the viewing geometries examined for Sec. 7.2.1 above was then re-examined while searching for the green target vehicle.

7.3 Results

When the fusion impact was assessed for the four different spectral SNR levels, two distinct trends emerged depending on the flexibility available to task the sensor. First, in the scenario where a multispectral nadir image could not be acquired, then the off-nadir SPI performance is compared to the off-nadir CEM performance. Figure 75 shows that as the spectral SNR decreased, incorporating additional polarimetric information via the SPI algorithm was the most valuable for relatively oblique sensor zenith angles (40° - 60°) until the sensor zenith angle became so high that the effect of a large GSD degraded performance. Further, this performance impact was only observed for two of the four tested solar zenith angles (34° and 55°). When the sun was extremely low in the sky (77° zenith angle), much of the scene was in shadow and the polarimetric information provided little—if any—benefit. Conversely, when the sun was high in the

sky (20° zenith angle), the benefit due to incorporating polarimetric information did not depend on the spectral SNR.

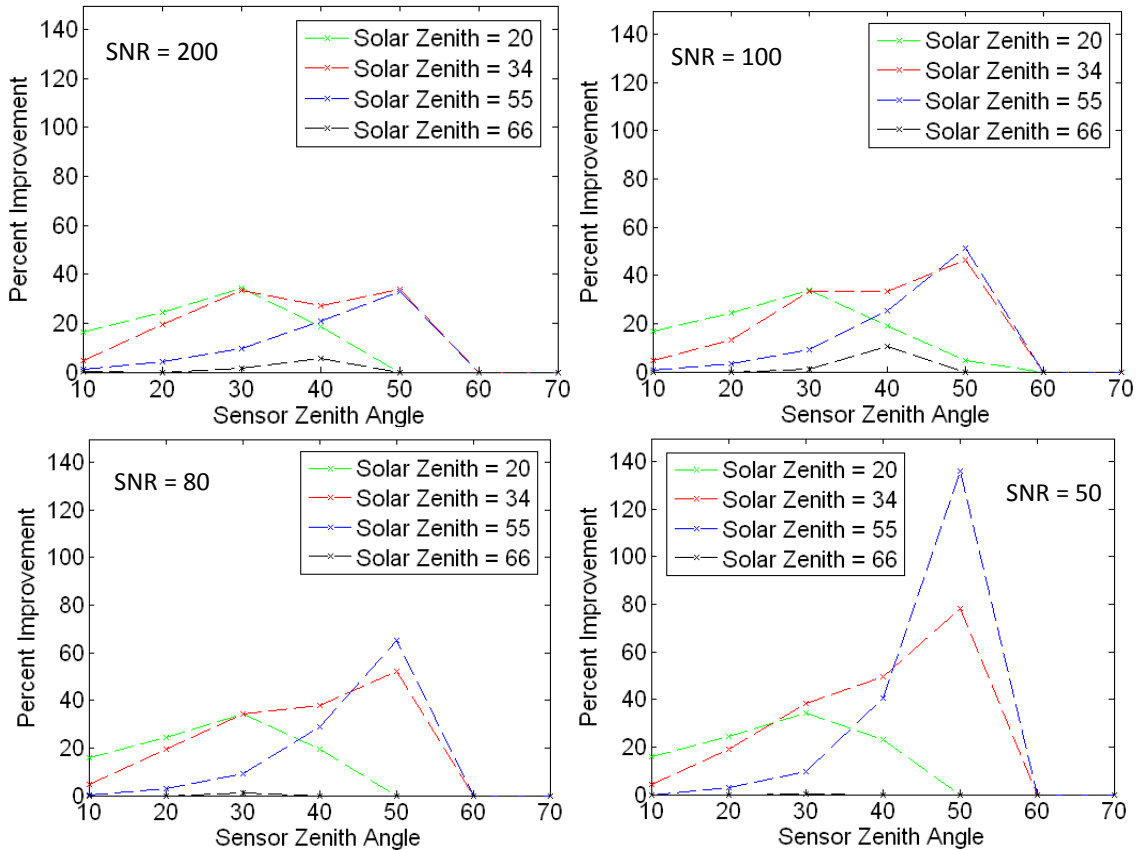


Figure 75. The percent increase in area under the ROC curve when using the SPI fusion algorithm as opposed to the CEM spectral algorithm for a given off-nadir viewing geometry is shown as a function of sensor zenith angle for solar zenith angles of 20° (green), 34° (red), 55° (blue) and 66° (black). The spectral SNR was set to 200 (top left), 100 (top right), 80 (bottom left) or 50 (bottom right).

The second scenario assumed that a multispectral nadir image could be acquired, and thus sought to determine whether the off-nadir SPI performance was superior to the nadir CEM performance for the same solar zenith angle. Figure 76 shows that as the spectral SNR decreased, incorporating additional polarimetric information via the SPI algorithm enhanced performance for moderately oblique sensor zenith angles ($20^\circ - 50^\circ$), with the most impact occurring when the sun was also at a moderately oblique zenith angle (34° or 55°).

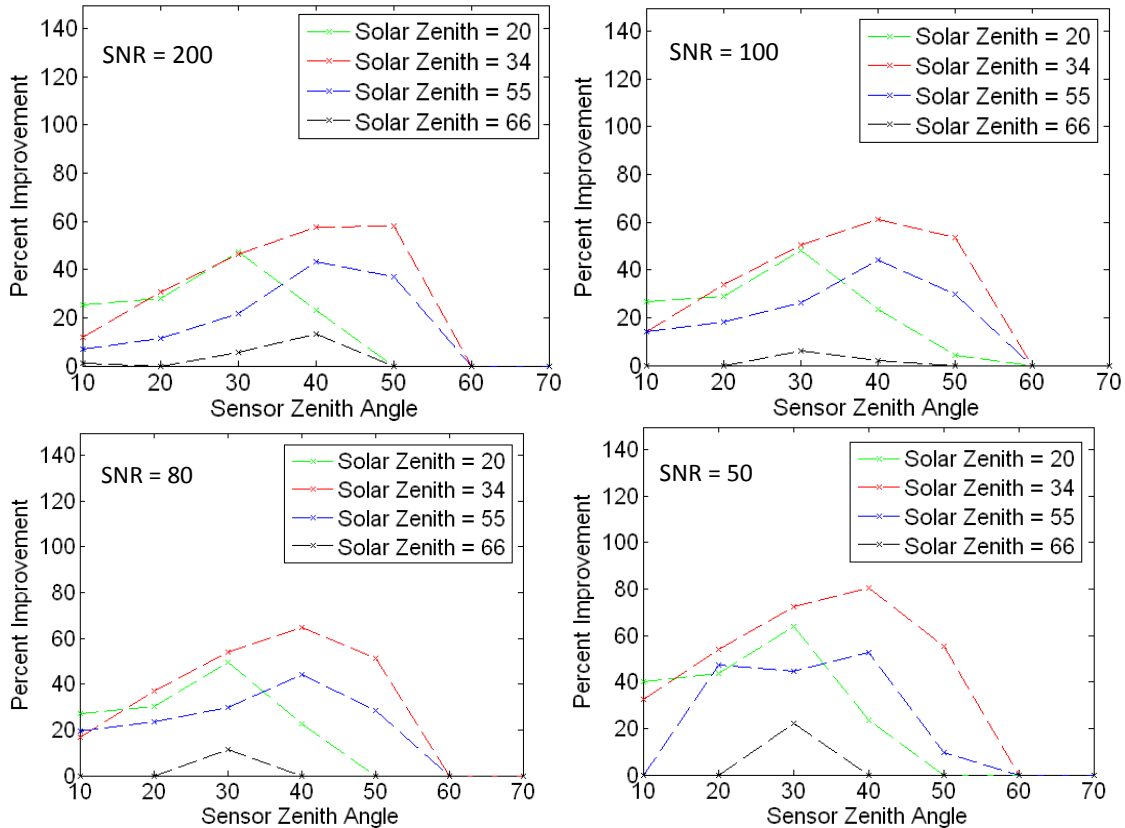


Figure 76. The percent increase in area under the ROC curve when using the SPI algorithm off-nadir as opposed to the CEM algorithm at nadir is shown as a function of sensor zenith angle for solar zenith angles of 20° (green), 34° (red), 55° (blue) and 66° (black). The spectral SNR was set to 200 (top left), 100 (top right), 80 (bottom left) or 50 (bottom right).

These two scenarios demonstrated that the viewing geometry where spectral and polarimetric data fusion provided the most impact depended on the options available for tasking the multimodal sensor. However, fusing the data via the SPI algorithm provided a benefit in both scenarios and further demonstrated the versatility of this simple decision fusion algorithm.

When the impact of varying spectral GSD for a particular off-nadir viewing geometry was assessed by examining Figure 77, several key trends emerged. First, at small GSD values, the CEM algorithm performed at least as well as, if not better, than the SPI fusion algorithm—demonstrating that even for good polarimetric sensing viewing geometries, the extra information from polarimetry doesn't always add value for target

detection. Second, although the areas under the CEM and SPI ROC curves both decreased as the spectral GSD increased, the difference between the area values grew larger. This behavior meant that as the spectral information became less precise, the relative value of polarimetric information increased dramatically.

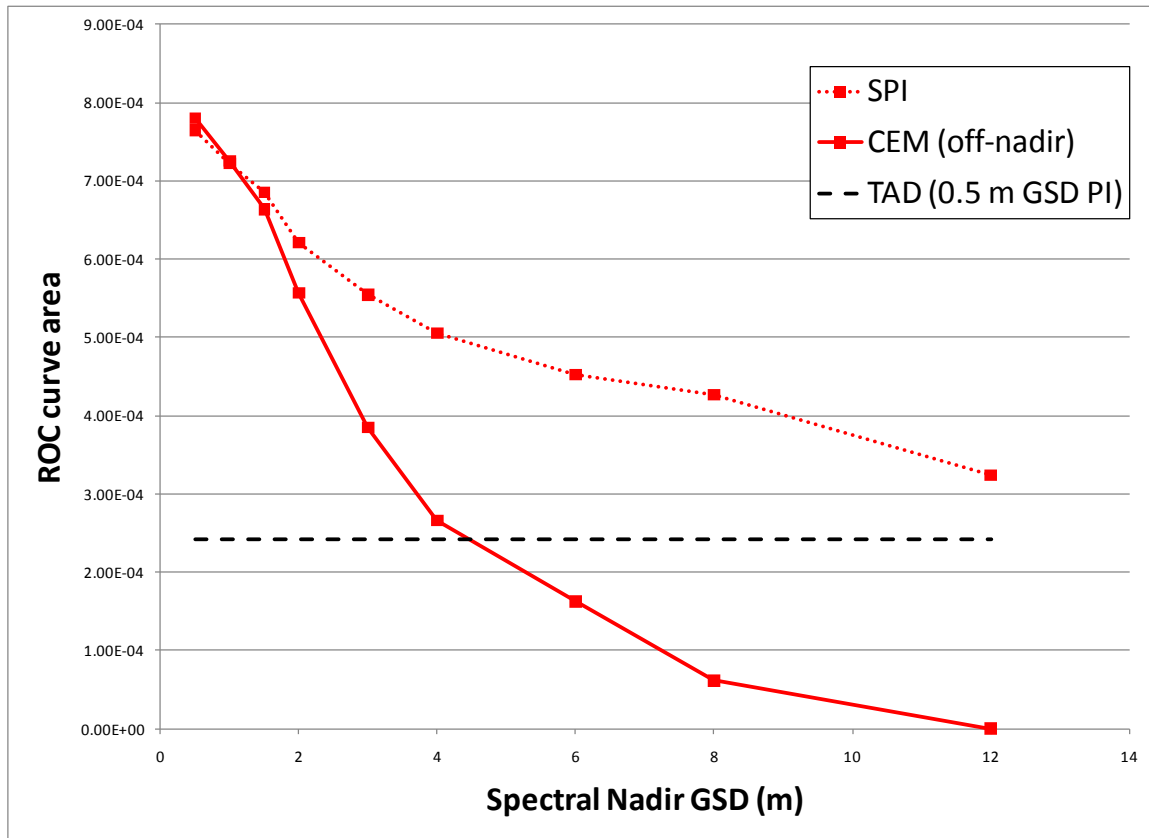


Figure 77. The area under the ROC curves generated by the off-nadir CEM (spectral data only—solid red line), SPI (spectral and polarimetric data—dotted red line) and TAD (polarimetric data only—dashed black line) algorithms was calculated for a range of spectral GSD values with the polarimetric GSD held constant at 0.5 m.

Finally, as the spectral GSD became especially large, the CEM and SPI algorithms failed differently. The CEM algorithm completely lost the ability to discern targets, while the SPI algorithm performance approached the performance obtained by simply using TAD as a target detection algorithm on the polarimetric data. Although the spectral GSD where this occurred was quite large for a target detection algorithm, one could envision similar results obtained with a smaller GSD but a much less conspicuous target. This

trend demonstrates the power of multimodal fusion via the SPI algorithm to find targets that would never be identified with spectral information alone.

When the effect of changing the target spectrum was assessed, valuable insight was gained by first calculating the SCR values for both target spectra at a nadir spectral GSD of 2 m and a nadir polarimetric GSD of 0.5 m. Figure 78 shows the SCR values for both the red and green target vehicles, based on both the spectral reflectance values and the polarimetric TAD scores derived from the scene.

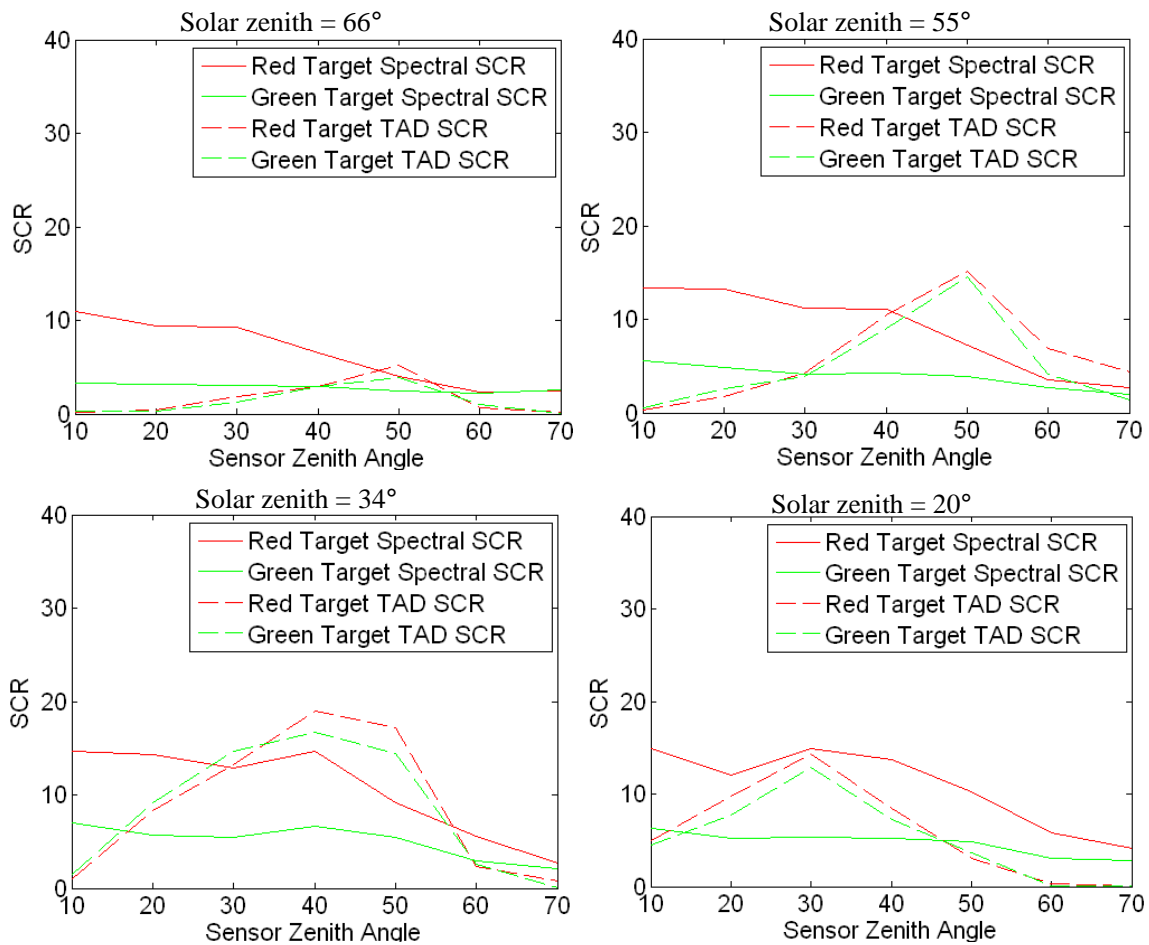


Figure 78. The SCR for the red and green target vehicles were calculated as a function of sensor zenith angle for a fixed sensor azimuth angle of 190° and four solar zenith angles: 66° (top left), 55° (top right), 34° (bottom left) and 20° (bottom right). The SCR was derived from either the spectral reflectance values (solid lines) or polarimetric TAD scores (dashed lines).

The polarimetric SCR values were quite similar for both vehicles, but the spectral SCR values for the green vehicle were about half of those for the red vehicle—meaning it was significantly harder to find the green vehicle with the CEM algorithm. Further, the peak polarimetric SCR values in Figure 78 occurred at different sensor zenith angles for different solar zenith angles. Not surprisingly, the target vehicles appeared most polarimetrically anomalous when the sensor was positioned near the sun’s specular reflection, reaffirming the trend identified in Sec. 6.2. Additionally, the differences in magnitude between the peak polarimetric SCR values at each solar zenith location suggest that incorporating polarimetric information may provide the most benefit at a solar zenith angle of 34° or 55° .

The SPI algorithm continued to enhance performance for both target vehicles by incorporating additional polarimetric information, although the performance impacts were not identical. Several trends were identified. First, Figure 79 demonstrates that a greater relative impact was observed when seeking the green target vehicle. However, fusion did not appear to improve performance for a solar zenith angle of 20° when seeking the green target. Recall that an improvement was only reported if the off-nadir SPI algorithm outperformed both the off-nadir CEM and nadir CEM algorithms—at a solar zenith angle of 20° , the off-nadir fusion performance was poorer than simply using the nadir CEM algorithm.

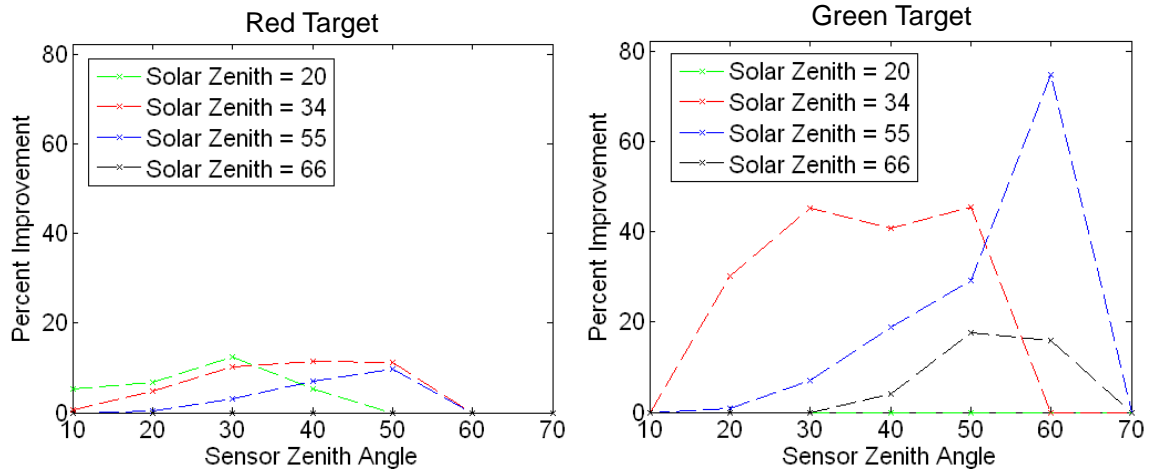


Figure 79. The percent increase in area under the ROC curve for the red target (left) and the green target (right) when using the SPI algorithm off-nadir as opposed to the CEM algorithm off-nadir is shown as a function of sensor zenith angle for solar zenith angles of 20° (green), 34° (red), 55° (blue) and 66° (black). The spectral and polarimetric SNR values were set to 200.

Second, with the spectrally easier to find red vehicle, the SPI algorithm provided a relatively constant modest boost in performance off-nadir up to a sensor zenith angle of 50° as long as the scene was well lit (solar zenith angles = 20°, 34° or 55°). In contrast, for the green target, the SPI algorithm provided a dramatic impact at most sensor zenith angles with a solar zenith angle of 34°, but as the sun moved lower in the sky (solar zenith = 55° or 66°), the percent improvement peaked at more oblique sensor zenith angles (until GSD effects dominated). Finally, the poorest viewing geometries varied depending on the target’s spectral characteristics. The SPI algorithm produced no improvement with the red target when the sun was low in the sky (66°), but aided in detecting the green target at the same solar zenith angle. Conversely, the SPI algorithm produced no improvement with the green target when the sun was high in the sky (20°), but aided in detecting the red target at the same solar zenith angle. Therefore, although the SPI algorithm is a powerful tool, the target characteristics can significantly influence

overall system performance—therefore the trade study scenario should be modeled as similar to the envisioned application (i.e. targets and backgrounds) as possible.

7.4 Section Summary

A trade study was carried out via DIRSIG simulation to assess how varying the spectral SNR, spectral GSD, or target spectrum affected the impact of spectral and polarimetric data fusion on a notional multimodal sensor via the SPI algorithm. When varying the SNR, the impact depended on the constraints placed on the sensor's tasking. When forced to image off-nadir, incorporating additional polarimetric information via the SPI algorithm was the most valuable for highly oblique sensor zenith angles (40° - 60°). However, when free to image at nadir if desired, incorporating additional polarimetric information via the SPI algorithm enhanced performance for moderately oblique sensor zenith angles (20° - 50°).

When GSD was varied, polarimetric information helped in some cases but not others. At small GSD values, the CEM spectral algorithm performed at least as well as, if not better, than the SPI fusion algorithm, but as the spectral GSD increased, the polarimetric information became more valuable. As the spectral GSD became especially large, the CEM algorithm failed to find any targets while the SPI algorithm did, reflecting a seamless transition from multimodal fusion to polarimetric target detection. Changing the target spectrum from a red vehicle to a green vehicle, reducing spectral contrast, produced variations in the impact due to fusion. When the spectral SCR decreased, incorporating additional polarimetric information became even more valuable at more oblique solar zenith angles. Although the SPI algorithm produced a general increase in

performance in both cases, correctly modeling the most likely target(s) for any future multimodal system was shown to be an integral part of an accurate design trade-off study. In summary, the SNR, GSD and target spectrum trade studies confirmed that as the initial spectral information became less precise, incorporating additional polarimetric information produced a more significant impact on performance. Therefore, fusing polarimetric data with lower quality multispectral data may provide an alternative to constructing a higher quality multispectral sensor for urban target detection scenarios. Additionally, the SPI decision fusion algorithm was shown to be robust across a number of variations possibly encountered in the multimodal sensor design process. Finally, the results highlight the importance of including simulation efforts in the sensor design process. By exploiting DIRSIG's capabilities, a team can accurately model a variety of complex scene, target and sensor characteristics—hopefully identifying system limitations early in the design stages rather than during hardware testing or operational missions. After assessing potential performance via simulation, the next step was to perform a small field experiment and apply the fusion algorithms to actual data.

8 Field Verification

8.1 Section Overview

This section describes an effort to build on the DIRSIG model results from Sec. 6 with a modest field experiment where the SPOT and SPI fusion algorithms described in Sec. 4.8 could be applied to real data. Since an airborne asset with both multispectral and polarimetric capabilities wasn't available, WASP-Lite and a separate camera with polarimetric filters in place were used to image a scene of model cars and bits of urban clutter. Two sets of images were acquired—one under favorable polarimetric viewing conditions, and the second when much of the scene was in shadows. The images were then registered to each other and analyzed with both the SPI and SPOT algorithms for several different target spectra.

Two main objectives were identified for this experiment. First, I sought to determine whether incorporating additional polarimetric information enhanced multispectral target detection for a viewing geometry identified as favorable by DIRSIG in Sec. 6. Second, I wanted to determine whether spectral and polarimetric data fusion could be used to find shadowed targets in real data, since the DIRSIG simulation results discussed in Sec. 6.5 suggested that the SPI algorithm would miss shadowed targets.

8.2 Scenario Design

The DIRSIG results from Sec. 6 indicated that spectral and polarimetric data fusion should enhance target detection when the sensor was located in the sun's specular reflection lobe. Further, performance was enhanced most at reasonably oblique sensor zenith angles (30°-50°), where the reflected polarimetric signature was increased but the

GSD was not too large. Additionally, polarimetric sensing theory suggests that the reflected polarimetric signature for a flat target should be minimized when the sensor is at a relative azimuth angle of 90° to the sun [Schott 2009]. Figure 80 shows a site on the RIT campus that was identified where both of these conditions could be met (at different times of day) without adjusting the scene.



Figure 80. The ROI for the multispectral / polarimetric data fusion experiment is outlined in red. This image shows the conditions for the morning data collect, with a solar zenith angle of 26° , a sensor zenith angle of 48° , and a relative azimuth angle between the sensor and the sun of 160° .

Good polarimetric sensing conditions were represented by a morning data collect where all targets were sunlit, with a solar zenith angle of 26° , a sensor zenith angle of 48° , and a relative azimuth angle between the sensor and the sun of 160° . Poor polarimetric sensing conditions were created with an afternoon scenario with a solar zenith angle of 35° , a sensor zenith angle of 48° , and a relative azimuth angle between the sensor and the sun of 100° . In contrast to the morning collect, many of the vehicles were shadowed in the afternoon collect.



Figure 81. (Left) The morning data collect represented favorable polarimetric sensing conditions. All vehicles in the scene were sunlit. (Right) The afternoon data collect represented poor polarimetric sensing conditions. Additionally, many of the target vehicles were shadowed.

This experiment was intended as a simple proof-of-concept experiment, so I was limited to using available RIT assets and a representative model environment. The vehicles in the scene were red, blue, silver, gray and yellow 1:32 scale die cast metal model cars shown in Figure 82.



Figure 82. The vehicles staged throughout the scene were 1:32 die cast metal model cars. Reflectance measurements of the vehicles were made under direct sunlight before the imagery was collected. Next, sample sections of siding material, roofing material,

asphalt, brick and smooth wooden boxes were also inserted in the scene to create clutter representative of an urban environment. Finally, a black felt panel and a measured spray-painted white panel were inserted into the scene for atmospheric compensation via ELM as in Sec. 4.7.

8.3 Image Acquisition

Multispectral imagery was acquired by placing different spectral filters over five of the individual cameras in the WASP-Lite sensor. The resulting spectral bands were Gaussians with a FWHM of 10 nm centered at 436 nm, 550 nm, 650 nm, 750 nm and 870 nm. The WASP-Lite sensor was mounted on the roof of Building 76, and a digital level was used to record the sensor zenith angle. Several test images were taken to calibrate the instrument gain so that no pixels in the scene were saturated, and this gain value was fixed for both the morning and afternoon scenarios. The WASP-Lite images were acquired simultaneously, but the different cameras were inherently mis-registered and had different fields of view. To create a multispectral image cube, the individual images were registered to the green band by manually selecting ground control points in ENVI, then applying a rotation-scale-translation calculation with nearest neighbor interpolation. Figure 83 shows the RGB representations of the registered multispectral imagery for both the morning and afternoon data collections. Some artifacts are apparent where the edge of the roof clipped the field of view (FOV) for certain spectral channels but not others. Additionally, the sensor was repositioned between the two imaging scenarios, so the morning and afternoon images are not perfectly aligned with each other.



Figure 83. (Left) Registered RGB WASP-Lite bands captured during the morning scenario. (Right) Registered RGB WASP-Lite bands captured during the afternoon scenario.

Polarimetric imagery was acquired by placing a linear polarizer in front of an IR cutoff filter to create a filtered panchromatic image. Several test images were taken to calibrate the instrument gain so that no pixels in the scene were saturated, and this gain value was fixed for both the morning and afternoon scenarios. The linear polarizer was rotated from 0° to 135° in 45° increments, with an image captured after each rotation. After the four images were captured, the filtered intensity images were converted to a Stokes vector for each pixel using the modified Pickering method described in Sec. 3.6. The division-of-time approach ensured the maximum polarimetric spatial resolution, but potentially introduced artifacts due to registration error (from adjusting the polarizer between images) or due to changes in the atmosphere between images. Figure 84 shows the processed S_0 images for both the morning and afternoon data collections. As with the multispectral sensor, the polarimetric sensor was repositioned between the two imaging scenarios, so the morning and afternoon images are not perfectly aligned with each other.



Figure 84. (Left) Processed S_0 image derived from the morning scenario. (Right) Processed S_0 image derived from the afternoon scenario.

Once the registered multispectral and polarimetric image cubes were assembled, the multispectral image was registered to the polarimetric data to create a fused dataset and ensure target detection could be done at the higher spatial resolution. Figure 85 shows RGB images of the fused datasets.

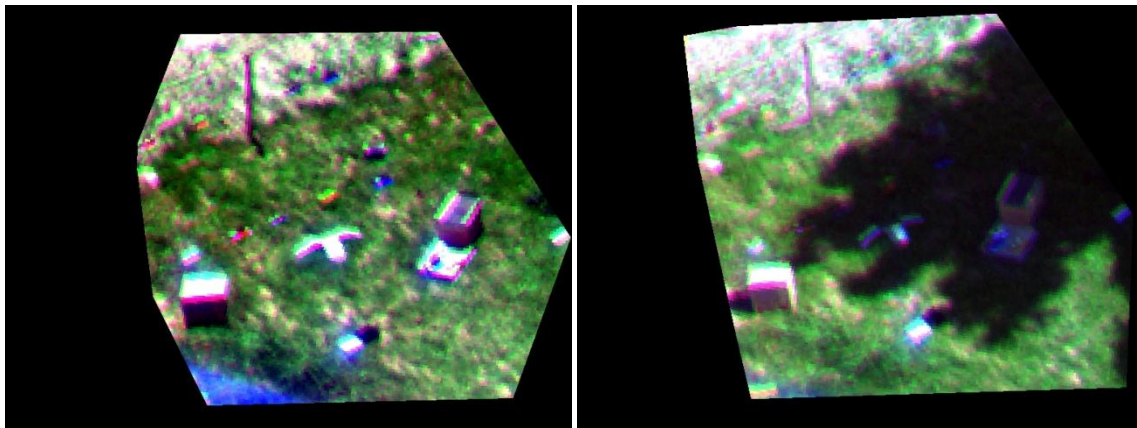


Figure 85. (Left) RGB image of multispectral data registered to the polarimetric imagery derived from the morning scenario. (Right) RGB image of multispectral data registered to the polarimetric imagery derived from the afternoon scenario.

The multispectral imagery had a GSD roughly three times as large as the polarimetric imagery, so using a nearest-neighbor interpolation produced a somewhat blocky final image. Also, since the multispectral FOV was different than the polarimetric FOV, some objects (like the side of the building) were present in one image, but not the other.

However, this was anticipated—four specific siding tiles placed in the scene served as the boundary of the scene ROI, and the multispectral imagery was able to capture the entire ROI. Areas outside the ROI were disregarded during registration. Finally, a rectangular spatial subset of the fused dataset was extracted to ensure every pixel tested had valid multispectral and polarimetric data, and the multispectral data was converted to the reflectance domain using ELM as shown in Figure 86.



Figure 86. (Left) The green band of the cropped fused morning data subset is shown as an intensity image after converting to reflectance values. (Right) The green band of the fused afternoon data subset is shown as an intensity image after converting to reflectance values.

8.4 Scene Clutter Assessment

Once the fused dataset had been constructed, the first step was to analyze the scene using the SCR metric defined in Sec. 5.4 to determine whether enough spectral and polarimetric clutter was present. Target ROIs were defined for each of the five different colored vehicles, and SCR values based on the multispectral imagery, TAD score and DOLP were computed for both the morning and afternoon cases, and the results are displayed in Table 14 and Table 15.

Table 14. The SCR metric was computed for the morning scene based on the multispectral data, TAD score, and DOLP value.

Model Car	MS SCR (am)	TAD SCR (am)	DOLP SCR (am)
Blue	4.51	1.31	0.39
Grey	1.58	1.50	0.20
Red	2.53	3.93	0.51
Silver	5.09	6.09	0.39
Yellow	3.18	3.68	0.46

Table 15. The SCR metric was computed for the afternoon scene based on the multispectral data, TAD score, and DOLP value.

Model Car	MS SCR (pm)	TAD SCR (pm)	DOLP SCR (pm)
Blue	2.52	0.48	0.58
Grey	0.98	0.07	0.84
Red	0.66	1.73	0.25
Silver	2.96	3.33	0.15
Yellow	1.12	0.17	0.16

The SCR values of the real data highlight three main points. First, the SCR values from this field collect were much lower than the SCR values in either the DIRSIG or COMPASS data described in Sec. 5.4. However, recall that Sec. 7.2.3 showed the SCR metric was sensitive to the target vehicle spectrum, so the different SCR values for each target were not unexpected. Since the SCR values demonstrated that the model cars in the field collect were harder to find than the vehicles in the DIRSIG data, this experiment perhaps underestimated the effectiveness of data fusion in an actual urban scenario. Second, the TAD SCR values were generally higher than the DOLP SCR values, confirming the DIRSIG trend that identified TAD as a better choice to find vehicles. One possible explanation was that for a fixed DOLP value, the TAD score could vary based on the differences between the S1 and S2 values—in effect, TAD provided more degrees of freedom to analyze the data. Third, the SCR values generally decreased from morning

to afternoon. On the spectral side, the decrease was likely due to the signatures from vehicles in shadow becoming more similar to nearby background pixels.

Polarimetrically, the decrease implied either a reduction in contrast between the reflected polarimetric signature and the upwelled radiance or the increased influence of the pBRDF (due to the downwelled radiation) as opposed to simply relying on the specular reflection—regardless, these SCR values predicted that the shadowed targets would be harder to find with polarimetric remote sensing. Once the scene was determined to have a reasonable amount of clutter, the fused datasets were analyzed with the CEM, SPOT and SPI algorithms.

8.5 Fusion Algorithm Evaluation

The first step in evaluating target detection performance was to define a target mask ROI for each of the five model vehicle types in the scene. Then, the fused datasets were evaluated for each of the different target vehicles as described in Sec. 4.9, except the ROC curve *FAR* threshold was increased from 0.001 to 0.01 because of the generally weak performance of all the test algorithms. A percent improvement in target detection was calculated based on the areas under the test algorithm ROC curves for each different target vehicle as shown in Eq. 43:

The fusion algorithms produced the most dramatic impact when the silver vehicle was identified as the target, so all graphics presented in this section reflect that scenario. Figure 87 shows the multispectral CEM scores for both the morning and afternoon scenes as intensity images, with brighter pixels being ranked as more target-like. Besides the

silver vehicles in the scene, several bright man-made objects were accidentally flagged as targets. Additionally, the contrast between high and low CEM scores was much lower in the afternoon than in the morning scenario.

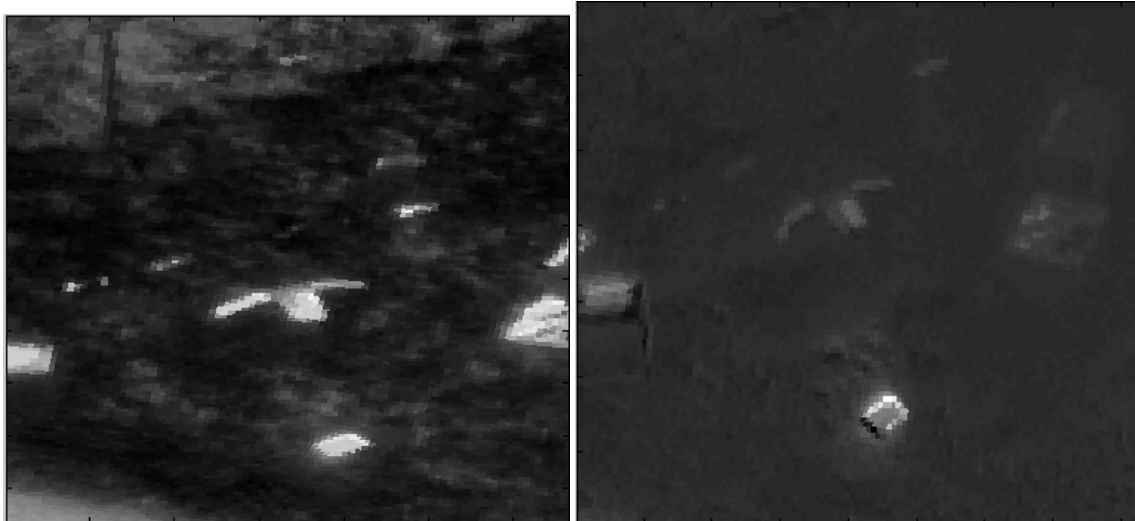


Figure 87. The MS CEM score maps are shown as intensity images, with brighter pixels being ranked as more target-like. (Left) Morning (Right) Afternoon

Figure 88 shows the DOLP for each pixel in the morning and afternoon scenes as intensity images, with brighter pixels representing more polarizing materials. In the morning scene, some vehicles contained pixels with relatively high DOLP, but vehicles were less polarizing in the afternoon scene. The difference in polarimetric signature was due to the combination of a shift in the sun-target-sensor geometry and the targets changing from sunlit to shadowed. Additionally, the DOLP in the shadowed grass for the afternoon scene appeared higher than the DOLP for well-lit grass. This enhanced DOLP was likely an artifact resulting from noise in each of the original filtered polarimetric images. The DC values recorded for the scene as seen through each polarimetric filter were about six times lower for the shadowed grass than for the sunlit grass. If the thermal noise in the detector was a roughly constant value, then the SNR value changed

dramatically between the shadowed and sunlit pixels. Recall that the DOLP metric was derived from the Stokes vectors, which were in turn derived from the polarimetrically filtered images. DC variations due to noise in pixels with low SNR values could result in significant differences when the Stokes vectors are created, and then dividing by a small number to create the DOLP further magnified the effect of those differences. The image with significant shadowed area represents one of the main challenges with polarimetric imaging—if the signal is integrated long enough that noise in the shadowed pixels isn't an issue, there is a significant risk of saturating well-lit pixels unless a very high dynamic range detector is used.

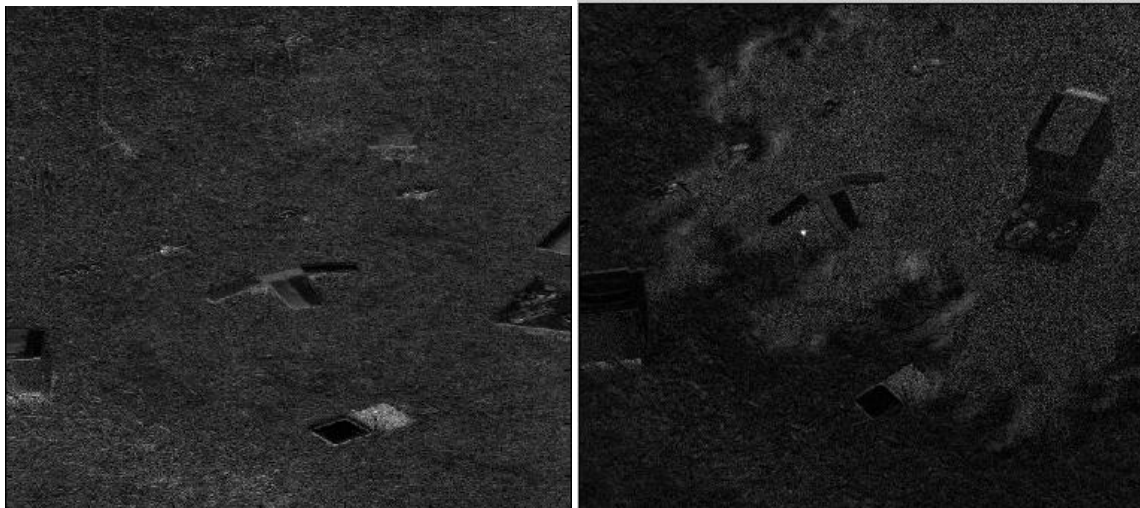


Figure 88. The DOLP values for each pixel are shown as intensity images, with brighter pixels being ranked as more polarizing. (Left) Morning (Right) Afternoon

Figure 89 shows the TAD score for each pixel in the morning and afternoon scenes as intensity images, with brighter pixels representing more anomalous pixels. In the morning scenario, under favorable polarimetric viewing conditions, the vehicles were highlighted along with other man-made polarimetric clutter in the scene. However, in the afternoon, the shadowed vehicles were scored much less anomalous than the sunlit

vehicles or urban clutter. This difference in scoring for shadowed vehicles was consistent with the DIRSIG results discussed in Sec. 6.5.2.

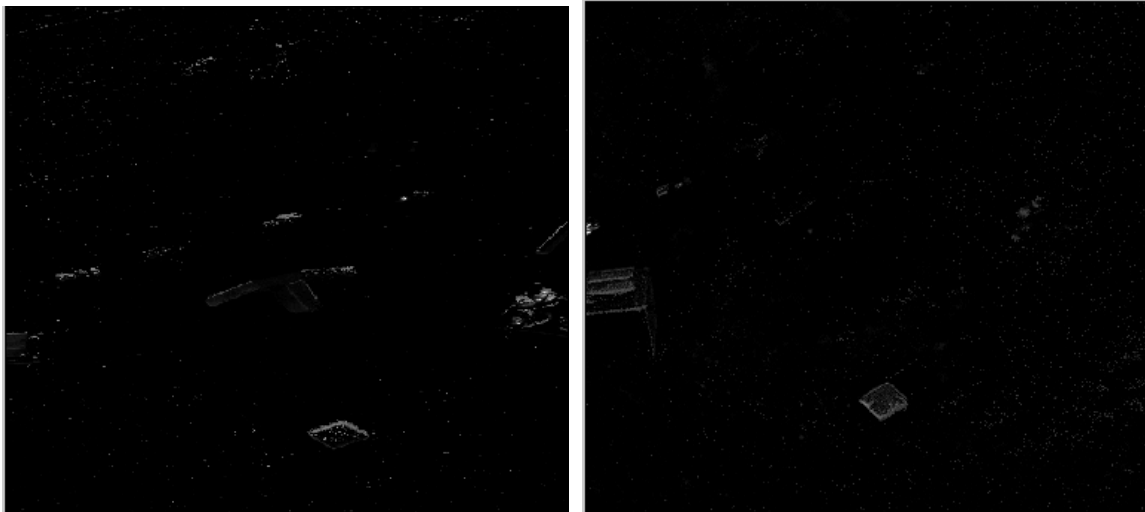


Figure 89. The TAD scores for each pixel are shown as intensity images, with brighter pixels being ranked as more anomalous. (Left) Morning (Right) Afternoon

Figure 90 shows the SPI decision fusion algorithm scores for each pixel in the morning and afternoon scenes as intensity images, with brighter pixels representing more target-like pixels. In the morning image, the target vehicles were some of the brightest pixels. However, some bright man-made clutter pixels were still flagged as potential targets, because bright false alarms shine through the CEM algorithm and the SPI algorithm weights the CEM score more than the TAD score. Note that with the SPI algorithm the nominated target pixels were dramatically separated from the natural background materials, and less spectrally obvious, but polarimetrically anomalous, target pixels were boosted above some of the bright false alarms. Further, the shadowed targets were still missed with the SPI algorithm—also confirming the performance observed with the DIRSIG imagery that was discussed in Sec. 6.5.3.

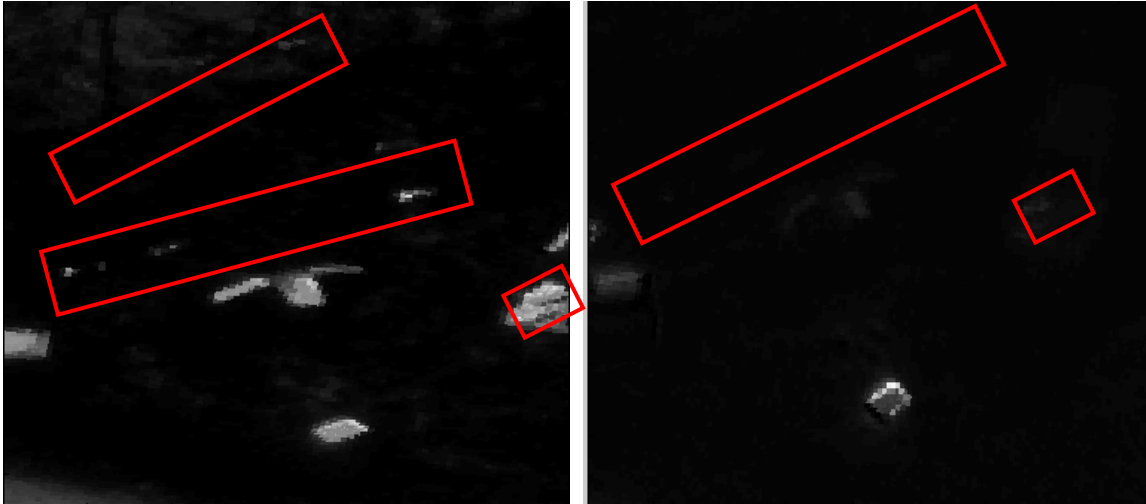


Figure 90. The SPI decision fusion algorithm scores for each pixel are shown as intensity images, with brighter pixels being ranked as more target-like. Red boxes indicate regions containing the vehicles in the scene. (Left) Morning (Right) Afternoon

Figure 91 shows the SPOT pixel fusion algorithm scores for each pixel in the morning and afternoon scenes as intensity images, with brighter pixels representing more target-like pixels. In the morning scene, using the DOLP metric produced mixed results because of Umov's effect. The scores of bright man-made false alarms were reduced due to their small polarimetric signature, but the scores of dark man-made false alarms were dramatically increased. The overall effect seemed to be a reduction of contrast in the scene between target vehicles and background materials. In the afternoon, a similar effect occurred, but the magnitude of the impact was non-uniform because of the sunlit versus shadowed DOLP calculation discussed above.

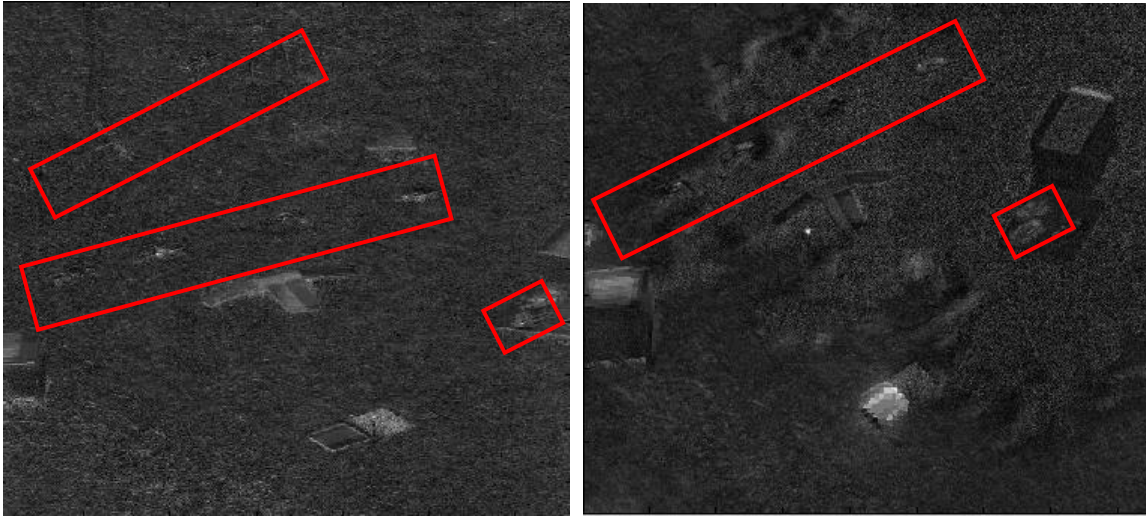


Figure 91. The SPOT pixel fusion algorithm scores for each pixel are shown as intensity images, with brighter pixels being ranked as more target-like. Red boxes indicate regions containing the vehicles in the scene. (Left) Morning (Right) Afternoon

Next, the multispectral and polarimetric SNR values were calculated for the 80% reflector reference panel in both the morning and afternoon scenes. The SNR values for the different polarimetric filter orientations were essentially unchanged, while the multispectral SNR values were band dependent. Table 16 summarizes the results, and emphasizes that both the polarimetric and multispectral SNR values captured in this experiment were far lower than the values applied to the synthetic data described previously. The low multispectral SNR values were due in part to the very narrow spectral filter bandpass window (10 nm FWHM) described in Sec. 8.3.

Table 16. The SNR values were calculated for the 80% reflector reference panel in the scene. The PI filter orientation affected SNR only minimally, so the average is reported here, while the MS SNR varied significantly from band to band.

Time of Day	PI SNR	Blue SNR	Green SNR	Red SNR	NIR 1 SNR	NIR 2 SNR
Morning	13	10	4	4	4	4
Afternoon	10	4	8	2	3	5

Finally, the effectiveness of the CEM, SPI and SPOT algorithms for both scenarios was assessed via ROC curves. Table 17 shows the results from the morning scenario as a percent difference in area under the fusion ROC curves when compared to the area under the CEM ROC curve with each different colored car identified as the target vehicle. Upon initial inspection, the grey vehicle appeared to buck the trend displayed by all the other vehicles. However, a closer examination of the individual ROC curves revealed that the grey vehicle was the hardest target to find, and due to the FAR threshold defined for this experiment, the percent differences were based on very small differences in very small areas under the ROC curve.

Table 17. The fusion algorithm effectiveness was assessed in terms of a percent difference in area under the ROC curve when compared to the CEM algorithm for each differently colored car in the morning scenario.

Model Car	SPI % Improvement	SPOT % Improvement
Blue	1.02	-29.1
Grey	-3.47	105.99
Red	9.17	-49.42
Silver	10.81	-70.49
Yellow	3.58	-14.44

Figure 92 shows the ROC curves generated when seeking the silver car in the morning scenario, which demonstrated that the SPI algorithm dramatically enhanced performance at low false alarm rates. In contrast, incorporating polarimetric information via the SPOT algorithm significantly degraded performance.

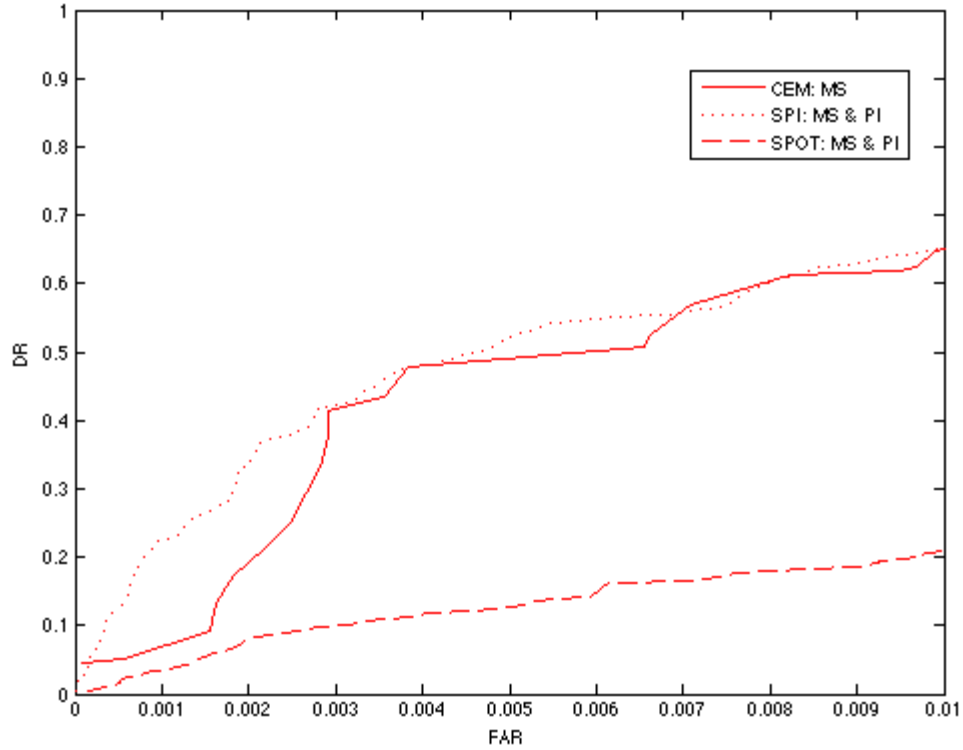


Figure 92. The CEM, SPI and SPOT ROC curves were derived from the morning scenario.

Table 18 shows the results from the afternoon scenario as a percent difference in area under the fusion ROC curves when compared to the area under the CEM ROC curve with each different colored car was identified as the target vehicle. If neither the fusion algorithm nor the CEM algorithm found any targets before reaching the *FAR* threshold, the difference in performance was undetermined. Also, if the fusion algorithm found any targets but the CEM algorithm did not, the increase in performance was infinite.

Table 18. The fusion algorithm effectiveness was assessed in terms of a percent difference in area under the ROC curve when compared to the CEM algorithm for each differently colored car in the afternoon scenario.

Model Car	SPI % Improvement	SPOT % Improvement
Blue	-20.16	-3.32
Grey	Undetermined	Infinite
Red	Undetermined	Infinite
Silver	14.02	-61.7
Yellow	Undetermined	Infinite

Although the SPOT algorithm produced an infinite improvement for three of the targets, the overall performance was still quite poor since only a few target pixels were identified before the *FAR* threshold, and they were only found along with a significant number of false alarms. Figure 93 shows the ROC curves generated when seeking the silver car in the afternoon scenario.

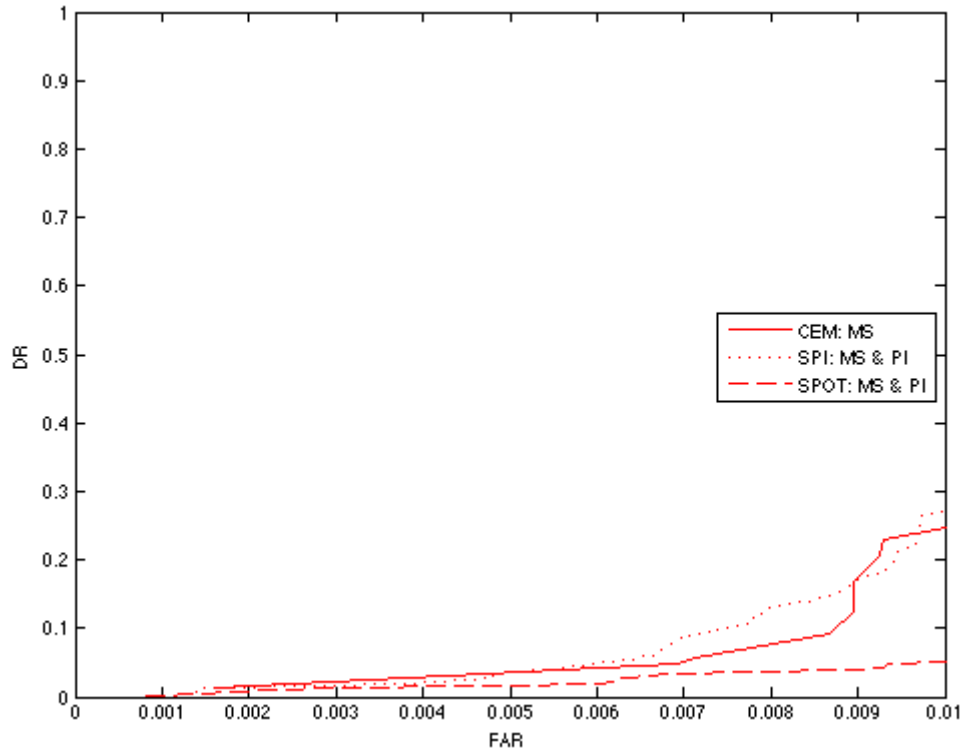


Figure 93. The CEM, SPI and SPOT ROC curves were derived from the afternoon scenario.

In general, a severe degradation in performance was seen when compared to the morning scenario, since the 14% improvement with the SPI algorithm was based on small variations in the small ROC curve areas. The poor performance was due to a combination of poor sun-target-sensor geometry, which reduced the quality of polarimetric info from sunlit targets, and shadows in the scene that minimized the spectral and polarimetric signatures from targets. As in the morning scenario,

incorporating polarimetric information via the SPOT algorithm significantly degraded performance.

8.6 Section Summary

In conclusion, this section described a modest field experiment designed to exploit the DIRSIG model results from Sec. 6 and apply the SPOT and SPI fusion algorithms to real data. A small-scale scene was constructed of model cars and bits of urban clutter. Five-band multispectral imagery was captured with WASP-Lite, while simultaneous polarimetric imagery was captured by rotating a linear polarimetric filter in front of a separate camera using the division of time method. Two sets of images were acquired—one under favorable polarimetric viewing conditions, and the second when much of the scene was in shadows. The images were then registered to each other and analyzed with both the SPI and SPOT algorithms for several different target spectra. Incorporating additional polarimetric information reasonably enhanced multispectral target detection for a favorable viewing geometry, while spectral and polarimetric data fusion was ineffective in finding shadowed targets in this experiment with real data. In executing this experiment, it quickly became apparent that the target pose could vary dramatically within a scene. Recent work at RIT suggested that acquiring multiple polarimetric images might maximize the ability to discriminate the target from the background [Devaraj 2010], so a series of studies were planned to assess the impact of incorporating multiple polarimetric images into the spectral/polarimetric fusion process.

9 Multi-View Polarimetric Fusion

9.1 Section Overview

Previous portions of this work have shown the potential to enhance target detection performance by fusing spectral and polarimetric data under a variety of different sun-target-sensor geometries. However, the value added by polarimetric data was demonstrated to be quite sensitive to the sun-target-sensor geometry. Further, since the exact target orientation (pose) in any given scene may be uncertain, the results may be somewhat pose dependent. In scenes where wildly different target orientations are possible—extreme elevation changes, irregularly shaped targets, etc—the potential exists to incorporate polarimetric images obtained from many different sensor zenith angles to increase the chances of capturing meaningful polarimetric data [Devaraj 2010].

Operational constraints often realistically limit the number of image acquisitions possible, so the multiple polarimetric image approach is only viable if a significant impact can be realized from a few carefully chosen viewing geometries. This section defines a process to model a particular situation, determines which polarimetric images will produce the most impact, quantifies the impact of incorporating polarimetric information from additional viewing geometries and evaluates the performance degradation introduced by a reasonable degree of registration error. The process described here could be adopted by a simulation team, who could render radiometrically accurate synthetic imagery of any scene of interest using the DIRSIG model. The scenario could be further customized by incorporating sensor parameters of interest and

any particular target in a variety of likely poses. After analyzing the synthetic data, meaningful image acquisition tasks could then be shared with sensor designers/operators.

9.2 Multiple PI Image Theory

Previous results showed that analyzing multiple polarimetric images of the same scene, taken at different sensor zenith angles, could enhance anomaly detection performance [Devaraj 2010]. A further expansion was to adapt the multiple polarimetric image technique to a more realistic target detection scenario including multispectral data by developing a way to incorporate polarimetric information from more than one image into the SPI algorithm discussed in Sec. 4.8.2. One conceivable scenario, shown in Figure 94, would be a polarimetric sensor flying at a constant altitude towards a point above the region of interest, capturing multiple images of the same scene from different sensor zenith angles (but constant sensor azimuth angle).

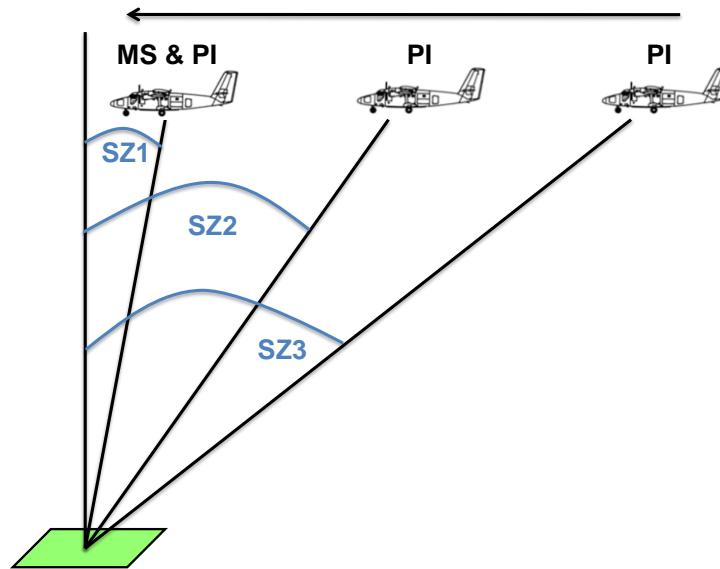


Figure 94. One sensor platform could potentially capture multiple polarimetric (PI) images of a scene in addition to a multispectral (MS) image. Although the images are obtained from different sensor zenith angles $SZ1$, $SZ2$ and $SZ3$, the aircraft altitude is held constant.

The resulting images would then each have a different FOV, and a pixel in each image would represent a different GSD. Since we desire to later fuse the polarimetric data with high quality spectral data, the fusion process begins by registering the polarimetric image data obtained at the more extreme sensor zenith angles to the polarimetric image data obtained closest to nadir. Next, the registered images are spatially resampled, and their FOVs are cropped such that only pixels containing polarimetric data from all available sensor zenith angles are retained. Based on the registered, resampled data, the Stokes vector for each pixel is calculated as observed from each different set of polarimetric images [Schott 2009], and the S_1 and S_2 elements of these Stokes vectors are concatenated into a stacked Stokes vector as shown in Eq. 44:

where SZl represents the sensor zenith angle where polarimetric data was obtained that was closest to nadir and SZn represents the most extreme sensor zenith angle where polarimetric data was obtained. If SZl is also the sensor zenith angle where spectral imagery was collected, such that polarimetric and spectral information is present for each pixel in the scene, then the new super Stokes vector shown in Eq. 44 can be used as the input into the TAD algorithm and fused via the SPI algorithm with the spectral data as described in Sec. 4.8.2. Note that the S_0 element has been neglected for this work, since it carries contrast information as opposed to information about surface roughness or orientation—we already capture brightness in the spectral information, and want to avoid flagging pixels as polarimetrically anomalous simply because of brightness.

9.3 Multiple PI Assessment Method

9.3.1 Scenario Modeling

This work continued to assess the SPI algorithm when applied to the urban target detection scenario, where several instances of either a red or green station wagon target vehicle were hidden (in varying orientations) amidst a crowded urban background with many differently colored decoy vehicles in the synthetic scene as described in Sec. 4.4. Once the scene parameters were set, the spectral and polarimetric sensors of interest were designed. Previous work described in Sec. 7.2.2 suggested that the impact of fusing spectral and polarimetric information depended in part on the spectral resolution of the spectral sensor under evaluation. Therefore, both multispectral and hyperspectral sensors were considered in this effort, with the spectral responses of those notional sensors described in Sec. 4.3. Besides spectral response, the sensor GSD and SNR values must be defined. Both spectral sensors were modeled to produce a nadir GSD of 3 m, while the polarimetric sensor was modeled to produce a nadir GSD of 0.5 m. As before, the sensor-reaching radiance was then altered by applying randomly generated zero-mean Gaussian noise to each band of the noise-free DIRSIG image to achieve a notional detector-limited sensor $\text{SNR} = 200$ for both the spectral and polarimetric data.

9.3.2 Multiple polarimetric image acquisition and registration

For this scenario, polarimetric images of the scene were rendered from sensor zenith angles of 10° to 70° in 10° increments, while fixing the sensor altitude at 7500 m above ground level and the sensor azimuth angle at 190° relative to the sun. The spectral images of interest were rendered at a 10° sensor zenith angle and sensor azimuth angle of 190° , corresponding to the polarimetric imagery closest to nadir (but at different GSD

values). Each set of sensor viewing geometries was then evaluated at four different times of day, where the solar zenith angles corresponded to sun locations for the modeled region of Rochester, NY on 23 June 1992 at times of 0700 (66°), 0800 (55°), 1000 (34°) and 1200 (20°).

Each of the polarimetric images were registered and resampled to align with the polarimetric image obtained at a sensor zenith angle of 10°. Because DIRSIG enables the user to track the exact scene coordinate represented by a given pixel in an image, the resulting registration was close to geometrically perfect. However, since the GSD increased as the sensor zenith angle increased, some small error was introduced as mixed pixels from large GSDs were resampled spatially to align with pure pixels at small GSDs. Further, at high sensor zenith angles, pixels in the scene (as viewed near nadir) were more likely to be occluded by surrounding objects—resulting in no valid polarimetric information being recorded for that target from the extreme sensor zenith angle, a condition that would occur in an operational setting.

9.3.3 Polarimetric image quality metric

The measure of merit for this work was the area under the receiver operating characteristic (ROC) curve as integrated to some user defined false alarm rate (FAR), set to 0.001 for the entirety of this work. The first part of this work sought to determine whether the adapted multiple polarimetric image analysis technique described in Eq. 44 produced better results with the complete set of polarimetric imagery described above, as opposed to simply using the single polarimetric image obtained simultaneously with the given spectral imagery in the SPI fusion algorithm or the solely spectral CEM algorithm.

Since we hypothesized that having more polarimetric perspectives should improved performance, the next logical step was to determine how many different perspectives were required to maximize performance and whether some perspectives were more valuable than others. First, the TAD algorithm described in Sec. 3.10 was applied to the stacked Stokes vector defined in Eq. 44, and the TAD scores for every target pixel in the image (available from DIRSIG truth data) were arranged in a vector, \mathbf{T} , as shown in Eq. 45:

where the subscript $10...70$ indicates that the stacked Stokes vector analyzed by the TAD algorithm for this vector was derived from the complete set of rendered polarimetric imagery: sensor zenith angles of 10° , 20° , 30° , 40° , 50° , 60° , and 70° . Then, target score vectors \mathbf{T}_{10} , \mathbf{T}_{20} , \mathbf{T}_{30} , \mathbf{T}_{40} , \mathbf{T}_{50} , \mathbf{T}_{60} , and \mathbf{T}_{70} were calculated—each produced by analyzing only the Stokes vectors derived from polarimetric imagery obtained at one particular sensor zenith angle. The process for deriving the vector of target TAD scores based on the complete set of polarimetric imagery, $\mathbf{T}_{10...70}$, is depicted graphically in Figure 95. In contrast, Figure 96 demonstrates the process for deriving the vector of target TAD scores based on polarimetric imagery obtained from one particular sensor zenith angle, \mathbf{T}_i .

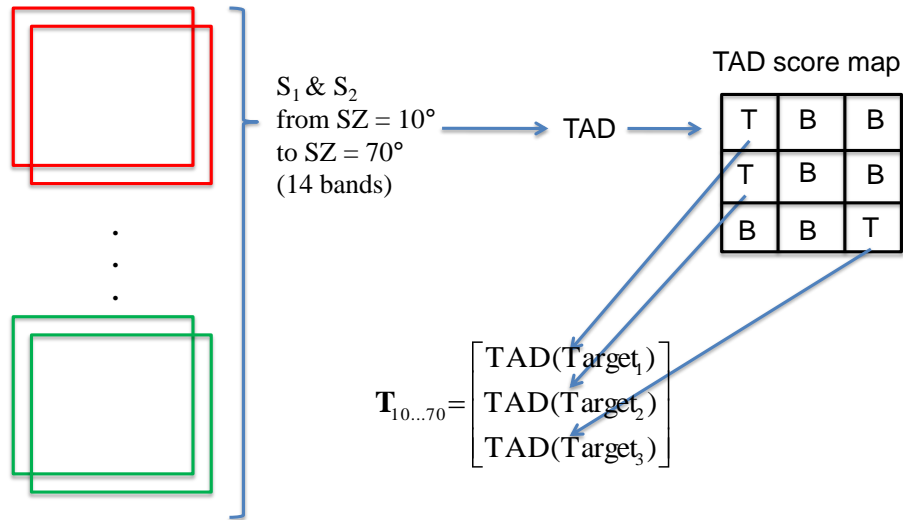


Figure 95. The S_1 and S_2 values for each pixel in the scene, as derived from sensor zenith (SZ) angles of 10, 20, 30, 40, 50, 60 and 70 degrees, were combined into a 14 band stack. The TAD algorithm was applied to the stacked data cube to identify anomalous pixels, and the $\mathbf{T}_{10...70}$ vector was constructed by collecting the TAD scores for each target pixel (T) in the image, while background pixels (B) were ignored.

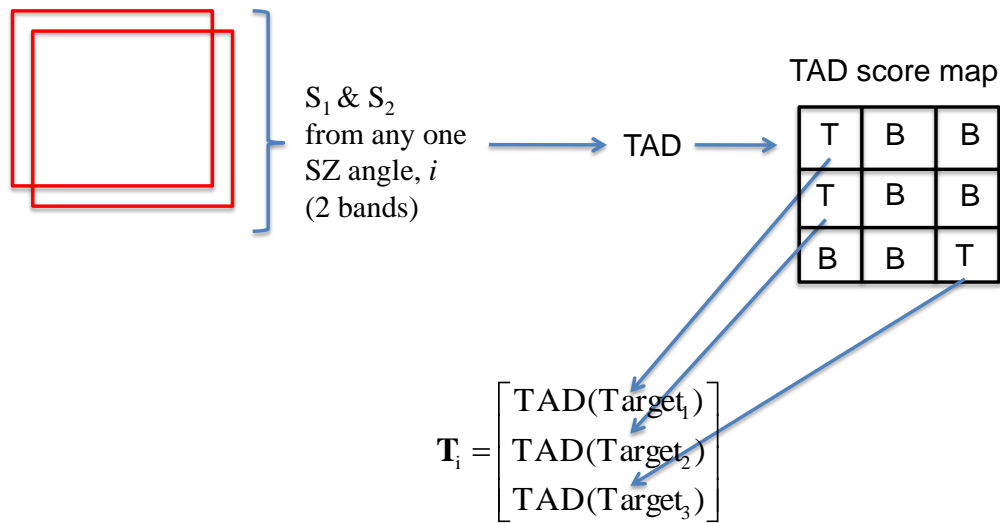


Figure 96. The S_1 and S_2 values for each pixel in the scene, as derived from one particular sensor zenith (SZ) angle, i , of interest were combined into a 2 band stack. The TAD algorithm was applied to the stacked data cube to identify anomalous pixels, and the \mathbf{T}_i vector was constructed by collecting the TAD scores for each target pixel (T) in the image, while background pixels (B) were ignored.

The quality of polarimetric information obtained from a particular sensor zenith angle was defined as the ability to replicate the target score vector $\mathbf{T}_{10...70}$, shown in Eq. 45, solely from the polarimetric information obtained at one particular sensor zenith

angle. The quality metric, Q , is derived by projecting a given target score vector derived from one polarimetric image, \mathbf{T}_i , in the direction of the complete target score vector $\mathbf{T}_{10...70}$ via a dot product as shown in Eq. 46:

where avoiding normalizing \mathbf{T}_i takes into account both the magnitude and direction of the individual score vectors, as opposed to a tool like the spectral angle mapper (SAM), which only accounts for direction [Schott 2007].

Once the quality metric was defined, an experiment was performed to assess its effectiveness by comparing the Q rankings from Eq. 46 for each different sensor zenith angle to the rank order of the area under the ROC curves produced by the SPI algorithm from Sec. 4.8.2. In this configuration, the spectral imagery input to the CEM algorithm was the same for each evaluation (sensor zenith angle of 10°), while the polarimetric information input to the TAD algorithm was derived from one of the various sensor zenith angles (10° - 70°). Because the only differences in each case came from the TAD scores, any variation in area under the ROC curve was due to the different polarimetric information incorporated from a particular sensor zenith angle.

9.3.4 Optimizing sensor image acquisition tasks

In a perfect world, as many images as desired could be captured for every scene of interest—however, operational considerations often limit the number of images that can realistically be acquired. The next step was to show how the simulation results could be translated into two meaningful sensor collection parameters, and determine whether one or two carefully chosen polarimetric collects could produce a benefit similar to the

complete set. To do so, the sensor zenith angles were rank ordered by their Q values to identify which produced the most useful polarimetric information. Then, the polarimetric information from the sensor zenith angles producing the highest and the two highest Q values was used to construct either a two or four element stacked Stokes vector (S_1 and S_2 derived from each image) as shown in Eq. 47 and Eq. 48:

The SPI decision fusion algorithm was then applied to the spectral imagery (sensor zenith angle of 10°) and two (or four) element stacked Stokes vector as shown in Eq. 49 and Eq. 50:

The area under the ROC curve was then evaluated for the CEM, SPI_{10} (\mathbf{x}_s based solely on a sensor zenith angle of 10°), SPI_{Full} (\mathbf{x}_s based on the entire set of polarimetric information sensor zenith angles from 10° to 70°), SPI_{Top1} and SPI_{Top2} algorithms.

9.3.5 Registration error

This work on multiple polarimetric images has represented a best-case scenario thus far. However, some amount of registration error is likely in any application where images from different perspectives are combined without detailed ground truth data.

Therefore, the next step was to incorporate varying degrees of registration error into the

SPI_{Top2} process described above and determine whether any performance improvement could still be attained.

After the polarimetric images had been registered and spatially resampled as described above, a desired number of pixels of registration error ($n = 1, 2$ or 3) was defined by the user. Since the images could be shifted left (or right) and up (or down) from each other, with potentially different effects, an iterative process was begun. First, a random number generator was used to produce a value $i = -1, 0$ or 1 . Then, the complete set of polarimetric images from the first sensor zenith angle was shifted to the right by $i \times n$ pixels, with negative values representing a shift to the left. Next, the random number generator was used to produce a value $j = -1, 0$ or 1 . Then, the same complete set of polarimetric images from the first sensor zenith angle was shifted downward by $j \times n$ pixels, with negative values representing an upward shift. The process was then repeated for the complete set of polarimetric images from the second sensor zenith angle. Since the spectral imagery was left unchanged, this process effectively introduced registration error between the spectral imagery, the polarimetric imagery from the first sensor zenith angle, and the polarimetric imagery from the second sensor zenith angle. Note that this process did not introduce registration error between images acquired with differently oriented linear polarimetric filters for a specific sensor zenith angle, because those effects are highly sensor dependent. Further, although the registration error introduced was an integer number of pixels for the polarimetric imagery, it represented anywhere from one sixth to one half of a pixel of error in the spectral imagery because of the difference in GSD values.

After the imagery had been shifted as described above, the SPI_{Top2} algorithm from Eq. (50) was applied to produce a ROC curve. Since the randomly directed registration error could produce a number of different final imagery combinations, the potential to skew the ROC curve results had to be addressed. Therefore, the image set was analyzed multiple times with different seeds in the random number generator, and the area under the generated ROC curve was recorded each time. The standard deviation of the ROC curve area values was calculated after each iteration (i.e. for a different seed in the random number generator) and once the standard deviation of the ROC curve areas changed by less than 3%, the median ROC curve area was identified as a representative value.

9.4 Results

To assess the quality metric, Q , used to derive the information for SPI_{Top2} above, the SPI algorithm from Sec. 4.8.2 was adapted to fix the spectral input to the CEM algorithm (obtained from a sensor zenith angle of 10°) while the polarimetric information was derived from any one sensor zenith angle (10° - 70°). The results were rank ordered to determine which single sensor zenith angles produced the most useful (and 2nd most useful) polarimetric image as determined by these different methods. Table 19 summarizes the results for the red target, while Table 20 shows the results for the green target.

Table 19. Top 2 polarimetric sensor zenith angles via different measures of merit (red target).

Solar zenith angle	Polarimetric Q value	MS SPI ROC curve area	HS SPI ROC curve area
66°	50° & 30°	50° & 30°	20° & 30°
55°	50° & 40°	40° & 50°	40° & 50°
34°	30° & 40°	30° & 40°	30° & 40°
20°	30° & 20°	20° & 30°	20° & 30°

Table 20. Top 2 polarimetric sensor zenith angles via different measures of merit (green target)

Solar zenith angle	Polarimetric Q value	MS SPI ROC curve area	HS SPI ROC curve area
66°	60° & 50°	60° & 40°	20° & 10°
55°	50° & 40°	50° & 40°	10° & 20°
34°	30° & 40°	30° & 40°	30° & 20°
20°	30° & 20°	30° & 20°	30° & 20°

For the red target, the two sensor zenith angles with the highest Q value almost always matched the two sensor zenith angles whose polarimetric data resulted in the most area under the ROC curve. However, results for the green target scenario were more complex. In the multispectral case, the two sensor zenith angles with the highest Q value again almost always matched the two sensor zenith angles whose polarimetric data resulted in the most area under the ROC curve, yet in the hyperspectral case there was poor agreement between the Q value and the ROC curve areas. This inconsistency was explained by considering the impact of polarimetric information as shown in Figure 97 and Figure 98—cases where the angles identified by the Q metric generally matched the angles used to produce the largest ROC curve areas were those where incorporating polarimetric information improved performance. Alternatively, when the angles identified by the Q metric didn't match the angles used to produce the largest ROC curve areas, incorporating any polarimetric information reduced performance. This result emphasized the need to perform a sensor trade study as described earlier in this work to isolate generally favorable polarimetric viewing conditions first, and then leverage the Q metric to determine the optimal combination of polarimetric imagery.

When the areas under the ROC curves produced by the CEM, SPI (polarimetric imagery obtained simultaneously with the spectral imagery at a sensor zenith angle of 10°), SPI_{Full}, (using the complete set of polarimetric imagery), SPI_{Top2} (smartly choosing

only two polarimetric images to include) and SPI_{Top1} (using only the best polarimetric image, often obtained at a different sensor zenith angle than the multispectral image) algorithms were compared, several trends were observed.

Figure 97 shows the results for the multispectral scenario (i.e. using 8 bands). Incorporating simultaneous polarimetric information via the SPI algorithm essentially left the multimodal system performance unchanged, although a slight increase in performance was observed for the red target vehicle as the solar zenith angle was decreased.

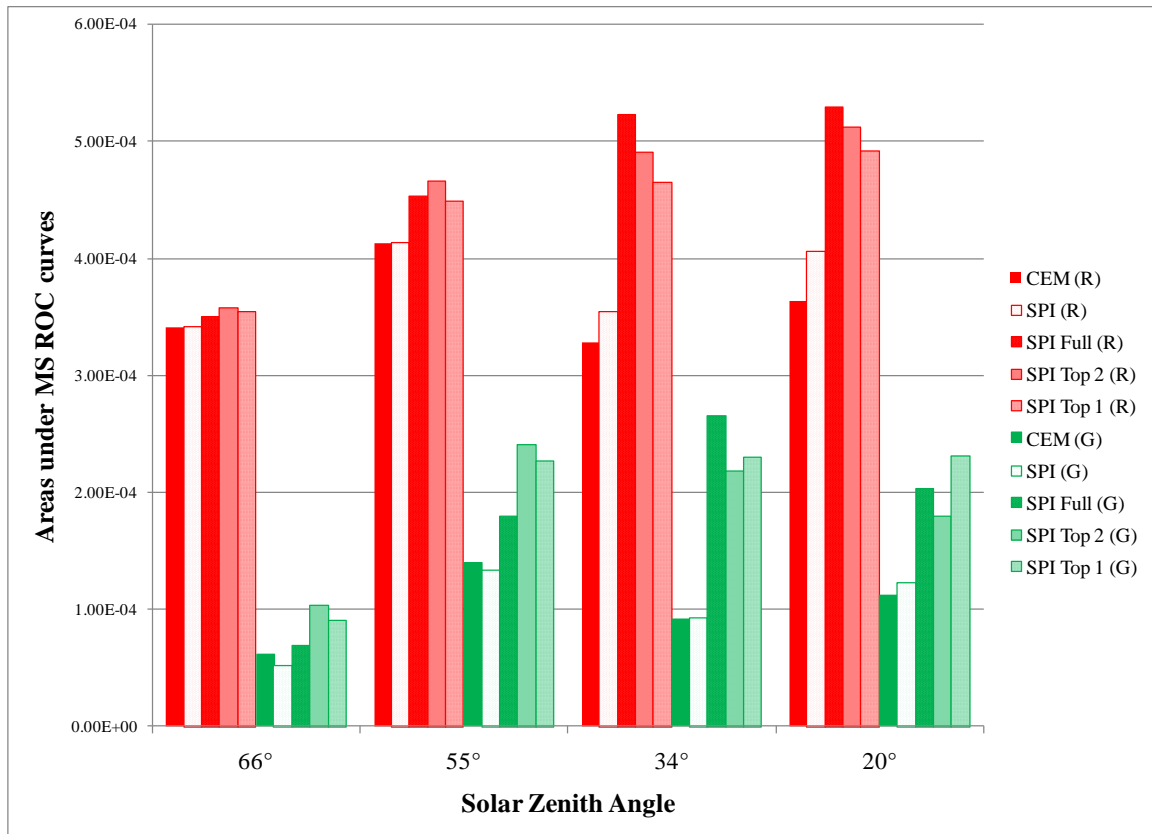


Figure 97. The areas under the multispectral ROC curves for the CEM, SPI, SPI_{Full} , SPI_{Top2} and SPI_{Top1} algorithms were calculated for both the red and green target vehicles. Acquiring multiple polarimetric images generally improved target detection performance, but most of the impact could be attained by logically choosing the single best polarimetric image.

The performance increase with SPI coincided with sensor more likely to be in the sun's specular lobe, consistent with previous findings described in Sec. 6. In contrast,

exploiting the complete set of polarimetric imagery via the SPI_{Full} algorithm produced a dramatic increase in performance for both the red and green target vehicle at most solar zenith angles. When using the best two polarimetric viewing geometries (as determined by the quality metric, Q) the SPI_{Top2} algorithm generally performed almost as well as the SPI_{Full} algorithm. Finally, simply using the best polarimetric image (SPI_{Top1}) produced most of the observed increase in performance. These results demonstrate that although multiple-look polarimetric imaging can play a decisive role in enhancing target detection performance, and a significant impact can be achieved with only two well chosen viewing geometries, there appears to be a diminishing return for the extra effort expended to capture the complete series of polarimetric images. In fact, for some cases, the SPI_{Top2} approach outperforms the SPI_{Full} approach—indicating we may be adding more noise than information by incorporating too many angles. Finally, incorporating high quality polarimetric information seems to help more as the solar zenith angle decreases.

Although a somewhat counterintuitive finding, the result is partly explained by the fact that as the solar zenith angle decreases, the multispectral data acquired in this scenario at a sensor zenith angle of 10° is more likely to be obtained near the sun's specular lobe—and thus degraded due to spectral whitening. In those cases, the additional high quality polarimetric information is much more likely to improve target detection performance because the baseline spectral target detection performance is decreased.

Figure 98 shows the results for the hyperspectral scenario, where the impact of polarimetric information was more complicated. Note that the hyperspectral CEM algorithm performed quite similar to the multispectral CEM algorithm for the red target, yet when searching for the green target, the hyperspectral CEM algorithm performed far

better than the multispectral CEM algorithm due to the low multispectral contrast of the green target.

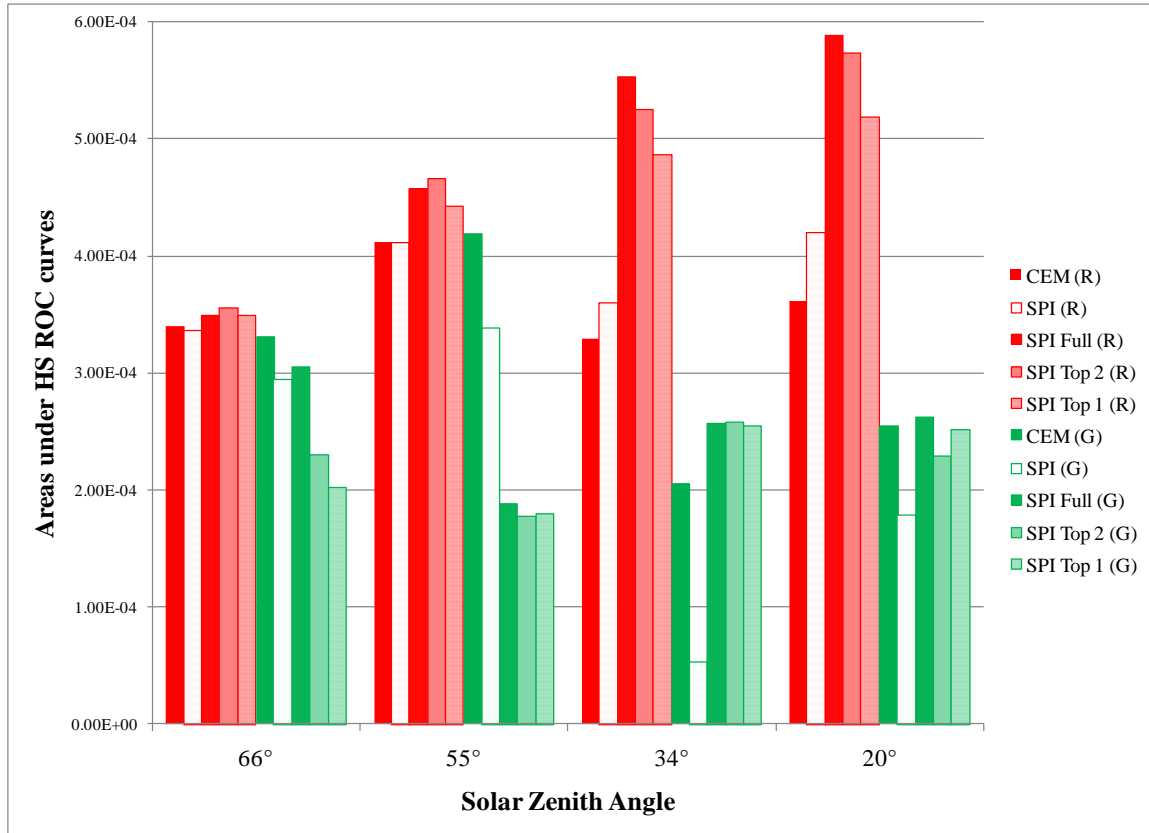


Figure 98. The areas under the hyperspectral ROC curves for the CEM, SPI, SPI_{Full}, SPI_{Top2} and SPI_{Top1} algorithms were calculated for both the red and green target vehicles. Acquiring multiple polarimetric images generally improved target detection performance for the red target, while incorporating polarimetric information generally decreased performance for the green target. In all cases, choosing the best two polarimetric images produced essentially the same impact as using the complete polarimetric image set.

Further, when the CEM performance was compared for the red and green target vehicles, the areas under the ROC curves were much more similar in the hyperspectral case than in the multispectral case. As expected, the overlap (or lack thereof) between the target features and sensor spectral response directly influenced the CEM algorithm performance. When searching for the red target, the results essentially mirrored the trends in the multispectral scenario. However, when searching for the green target, incorporating simultaneous polarimetric information via the SPI algorithm decreased

performance. Then, when the complete polarimetric image set was incorporated with SPI_{Full} , the performance was only increased for solar zenith angles of 20° and 34° , which were also the viewing geometries where the purely spectral CEM algorithm performed worst. Finally, choosing the best two polarimetric geometries using SPI_{Top2} again produced generally the same performance as using the complete set. These results demonstrate that the effectiveness of spectral / polarimetric data fusion depends not only on the sun-target-sensor geometry, but also on the spectral characteristics of both the sensor and the target of interest—further highlighting the utility of a customizable simulation process to evaluate performance before building or operationally tasking the sensor.

Finally, the SPI_{Top2} algorithm performance was analyzed to determine the impact of multiple polarimetric images when some degree of registration error was incorporated. Recall that although the registration error introduced was an integer number of pixels for the polarimetric imagery, it represented anywhere from one sixth to one half of a pixel of error in the spectral imagery because of the difference in GSD values. Figure 99 shows the results for the multispectral case, demonstrating that incorporating registration error degraded fusion performance but still outperformed simply relying on spectral information.

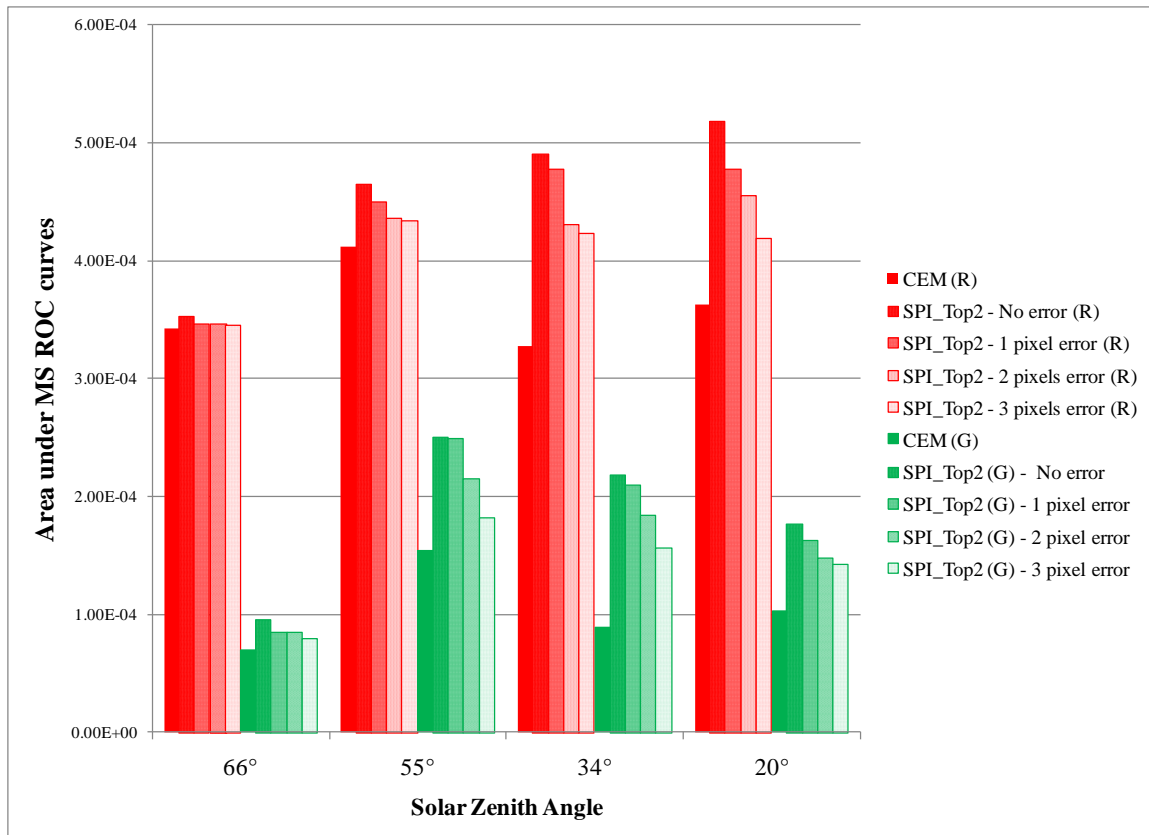


Figure 99. The areas under the multispectral ROC curves were calculated for the SPI_{Top2} fusion algorithm, with a varying degree of registration error in the polarimetric images, and compared to the purely spectral CEM algorithm performance for both the red and green target. Incorporating registration error produced some decrease in fusion performance, but still outperformed simply relying on spectral information.

Figure 100 shows the results for the hyperspectral case, again demonstrating that the fusion algorithm performance was somewhat target and sensor dependent. Although incorporating registration error produced some decrease in fusion performance for all scenarios, fusion still outperformed simply relying on spectral information for the red target.

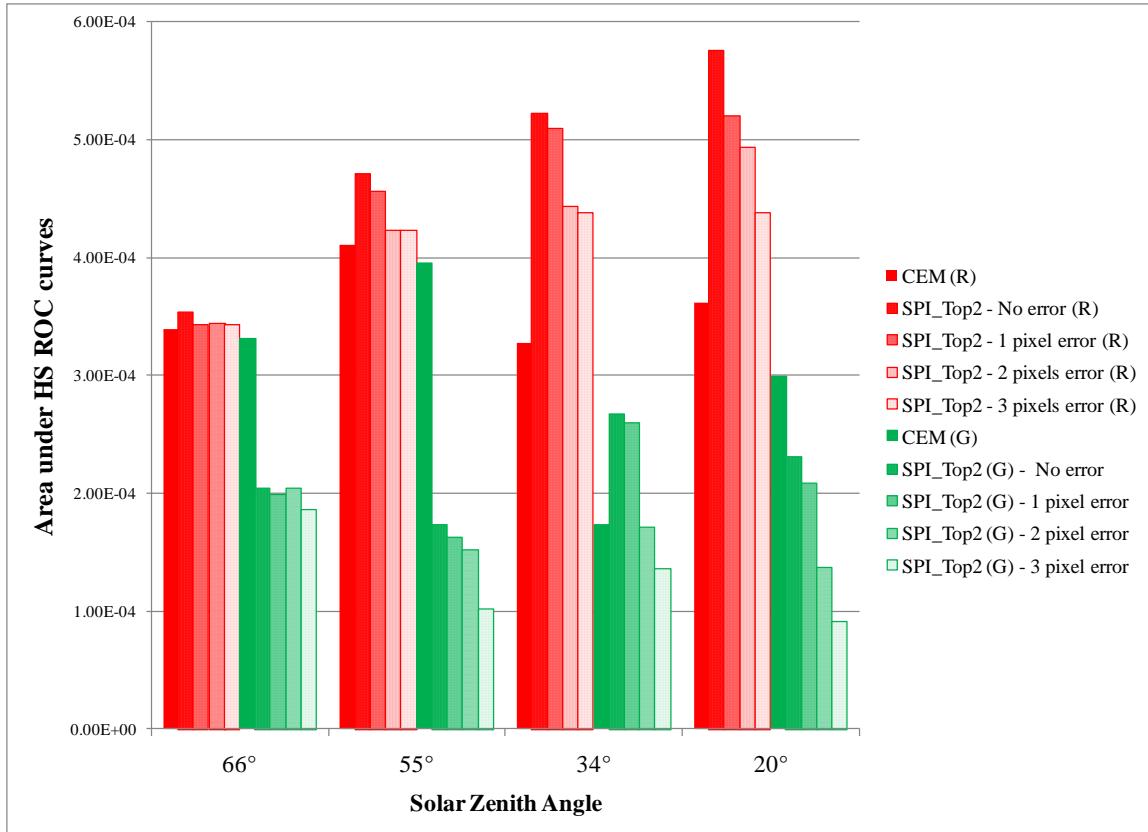


Figure 100. The areas under the hyperspectral ROC curves were calculated for the SPI_{Top2} fusion algorithm, with a varying degree of registration error in the polarimetric images, and compared to the purely spectral CEM algorithm performance for both the red and green target. Incorporating registration error produced some decrease in fusion performance, but still outperformed simply relying on spectral information.

For the green target, incorporating polarimetric information generally degraded performance even with perfect registration, so the fusion algorithm’s sensitivity to registration error was mostly irrelevant. This fact highlights exactly how simulated trade studies can be useful—rather than devoting resources to minimizing registration errors in the scenario with the green target, a design team faced with these results would likely opt to use a single-modality hyperspectral sensor.

9.5 Section Summary

Although results from Sec. 6 have shown the potential to enhance target detection performance by fusing spectral and polarimetric data, the value added by polarimetric

data was demonstrated to be quite sensitive to the sun-target-sensor geometry. Additionally, since the exact target orientation (pose) in any given scene may be uncertain, the results become somewhat scene dependent. This section described an effort to incorporate polarimetric images obtained from multiple sensor zenith angles and thereby increase the chances of capturing meaningful polarimetric data in scenes where different target orientations are possible. However, operational constraints often realistically limit the possible number of image acquisitions, so the multiple polarimetric image approach is only viable if a significant impact can be realized from a few carefully chosen viewing geometries.

This section defined a process to model a conceptual scenario, to determine which polarimetric images will produce the most impact, to quantify the impact of incorporating polarimetric information from additional viewing geometries and to evaluate the performance degradation introduced by a reasonable degree of registration error. Variations of the SPI decision fusion algorithm were described, adapting the decision fusion process to accept information from multiple images as inputs, with each image taken from a different viewing geometry. A metric was introduced to assess the quality of off-nadir polarimetric information—balancing the increased ability to detect polarimetric target signatures with the degradation in performance due to increasing the sensor GSD. After logically identifying two sensor zenith angles for capturing polarimetric imagery, a minimal return on investment was observed when additional polarimetric imagery was captured from a full set of sensor zenith angles. Finally, incorporating a modest amount of registration error between each of the polarimetric

images and the spectral image degraded fusion performance to some degree, but often still outperformed simply relying on spectral information.

This process was developed for use by sensor design/tasking teams, who could render radiometrically accurate synthetic imagery of any target/background scenario of interest using the DIRSIG model. The scenario could be further customized by incorporating sensor parameters of interest and any particular target in a variety of likely poses. After analyzing the synthetic data, meaningful sensor design and/or acquisition tasks could then be shared with sensor developers/operators.

10 Conclusion

The studies presented in this dissertation leveraged DIRSIG's radiometrically sound simulation capabilities to explore the impact of fusing spectral information with polarimetric information for an urban target detection application. A synthetic scene was designed and then validated both quantitatively and qualitatively. Two novel methods of fusing the data types were developed, and the effectiveness of each was assessed under a wide variety of specific scenario conditions. Several general trends were observed, with performance depending on sensor GSD, SNR and spectral resolution while also varying with the chosen target spectrum. A modest field experiment was carried out to apply the fusion algorithms in a more realistic setting. Finally, a method leveraging simulation was defined to assess the impact of incorporating multiple polarimetric images into the multimodal fusion process. The results from both synthetic and real data suggest that a niche mission might be available for emerging polarimetric remote sensing capabilities.

When fusing the data at the pixel level, the SPOT algorithm effectively treated the DOLP as another spectral band in a scaled matched filter target detection algorithm. Although useful when information from only a few spectral bands was available, the polarimetric information was weighted significantly lower than the spectral information. Further, the SPOT algorithm required the assumption that the target of interest was the most polarizing object in the scene. In contrast, the SPI algorithm was designed to fuse polarimetric and spectral information at the decision level. The spectral information was analyzed via a scaled matched filter algorithm, while the polarimetric information was analyzed with an anomaly detection algorithm based on graph theory. Each pixel was then scored by combining the results of both the spectral and polarimetric algorithms,

weighted roughly evenly, while permitting the spectral information veto power. The SPI algorithm assumed the target pixel was among the most polarimetrically anomalous pixels in the scene, which seemed reasonable if searching for a civilian vehicle in an urban environment. Typically, the SPI decision-level fusion algorithm outperformed the SPOT pixel-level fusion algorithm.

A series of trade studies was carried out to assess how varying the spectral SNR, spectral GSD, or target spectrum affected the impact of spectral and polarimetric data fusion via the SPI algorithm for a notional multimodal sensor. When varying the SNR, the impact depended on the constraints placed on the sensor's tasking. When spectral GSD was varied, the benefit of incorporating polarimetric information increased as the GSD increased. However, a threshold GSD was identified beyond which no benefit was observed. Reducing the target/background spectral contrast by changing the target spectrum from a red vehicle to a green vehicle produced variations in the impact due to fusion, although the SPI algorithm produced a general increase in performance in both cases. The trade studies demonstrated that incorporating additional polarimetric information may enable suitable performance with a less capable multispectral sensor.

A field experiment was designed to exploit the DIRSIG simulation results and apply the SPOT and SPI fusion algorithms to real data by constructing a small-scale scene of model cars and bits of urban clutter. Five-band multispectral imagery was captured with WASP-Lite, while simultaneous polarimetric imagery was captured by rotating a linear polarimetric filter in front of a separate camera using the division of time method. Two sets of images were acquired—one under favorable polarimetric viewing conditions, and the second when much of the scene was in shadows. The images were

then registered to each other and analyzed with both the SPI and SPOT algorithms for several different target spectra. Incorporating additional polarimetric information reasonably enhanced multispectral target detection for a favorable viewing geometry, while spectral and polarimetric data fusion was ineffective in finding shadowed targets in the real data.

Next, a process was defined to model a particular set of image acquisition scenarios, to determine which polarimetric image (from a set of many) will produce the most impact on target detection performance, to quantify the impact of incorporating polarimetric information from multiple viewing geometries and to evaluate the performance degradation introduced by a reasonable degree of registration error. Variations of the SPI decision fusion algorithm were described, adapting the decision fusion process to accept information from multiple images as inputs—with each image taken from a different viewing geometry. A metric was introduced to assess the quality of off-nadir polarimetric information—balancing the increased ability to detect polarimetric target signatures with the degradation in performance due to the increased sensor GSD. After logically identifying the best two sensor zenith angles for capturing polarimetric imagery, performance improved by exploiting those angles but a minimal return on investment was observed when additional polarimetric imagery was captured from more than two sensor zenith angles. Finally, incorporating a modest amount of registration error between each of the polarimetric images and the spectral image was shown to degrade fusion performance slightly, but often still outperform spectral information alone.

Through the evaluations described above, this dissertation demonstrated a generalized approach to performing trade space evaluations via synthetic image generation tools. In the remote sensing community, polarimetric capabilities are often seen as a tool without a widely accepted mission. The results in this work essentially suggest that polarimetric information may be leveraged to restore the capabilities of a spectral sensor when forced to image under less than ideal circumstances. Separately, this work makes a compelling case for simulation as an attractive tool in designing cutting-edge systems because of the sheer volume of data required for a reasonable trade study, especially when multimodal sensors are considered. The generalized analysis method presented here will allow system designers to tailor future target and sensor parameters to their particular scenarios of interest, ensuring that resources are best allocated without the delays associated with constructing a variety of different prototypes.

References

- Alouini, M., Goudail, F., Grisard, A., Bourderionnet, J., Dolfi, D., Beniere, A., Baarstad, I., Loke, T., Kaspersen, P., Normandin, X., and Berginc, G., "Near-infrared active polarimetric and multispectral laboratory demonstrator for target detection", *Applied Optics*, Vol. 48, No. 8, 10 March 2009
- Barcomb, K., Schott, J., Brown, S. and Hattenberger, T., "High-Resolution, Slant-Angle Scene Generation and Validation of Concealed Targets in DIRSIG", *Proceedings of SPIE*, Vol. 5546, 2004
- Basener, W. F. and Messinger, D. W., "Enhanced Detection and Visualization of Anomalies in Spectral Imagery", *Proceedings of SPIE*, Vol. 7334, 2009
- Benediktsson, J. A., Palmason, J. A., Sveinsson, J. R., and Chanussot, J., "Decision Level Fusion in Classification of Hyperspectral Data from Urban Areas", *IEEE International Geoscience and Remote Sensing Symposium 2004, IGARSS '04*, Vol. 1, 20-24, September 2004
- Benediktsson, J. A. and Kanellopoulos, I., "Classification of Multisource and Hyperspectral Data Based on Decision Fusion", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, No. 3, May 1999
- Benediktsson, J. A. and Swain, P. H., "Consensus Theoretic Classification Methods", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 22, No. 4, July/August 1992
- Berk, A., Bernstein, L. S., and Robertson, D. C. "MODTRAN: A moderate resolution model for LOWTRAN 7", GL-TR-89-1022, Spectral Sciences, Boston MA, 1989
- Briottet, X., Boucher, Y., Dimmeler, A., Malaplate, A., Cini, A., Diani, M., Bekman, H., Schwering, P., Skauli, T., Kasen, I., Renhorn, I., Klasen, L., Gilmore, M., and Oxford, D., "Military applications of Hyperspectral Imagery", *Proceedings of SPIE*, Vol. 6239, 2006
- Cafer, C., Stefanou, M., Nielsen, E., Rizzuto, A., Raviv, O., and Rotman, S., "Analysis of false alarm distributions in the development and evaluation of hyperspectral point target detection algorithms", *Optical Engineering*, Vol. 46, No. 7, pp. 1-15, July 2007
- Cafer, C., Silverman, J., Orthal, O., Antonelli, D., and Sharoni, Y. "Improved covariance matrices for point target detection in hyperspectral data", *Optical Engineering*, Vol. 47, No. 7, pp. 4-13, July 2008

- Cavanaugh, D. B., Castle, K. R. and Davenport, W. "Anomaly detection using the hyperspectral polarimetric imaging testbed", *Proceedings of SPIE*, Vol. 6233, 2006
- Chen, M., Jia, Y., Li, D., and Qin, Q., "Image Fusion Algorithm for Multispectral and Panchromatic Data Based on Multiple Criteria", *Proceedings of SPIE*, Vol. 6043, 2005
- Cheriyadat, A., Bruce, L. M. and Mathur, A., "Decision level fusion with best-bases for hyperspectral classification", *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, 399-406, October 2003
- Defense Acquisition University, *Defense Acquisition Guidebook*, Section 3.2.4, <https://acc.dau.mil/dag>, July 2009
- Devaraj, C., "Polarimetric remote sensing system analysis: Digital Imaging and Remote Sensing Image Generation (DIRSIG) model validation and impact of polarization on material discriminability", PhD dissertation, Rochester Institute of Technology, 2010
- Devaraj, C., Brown, S., Messinger, D., Goodenough, A., and Pogorzala, D., "A framework for polarized radiance signature prediction for natural scenes", *Proceedings of SPIE*, Vol. 6565, 2007
- DigitalGlobe QuickBird satellite datasheet, <http://www.digitalglobe.com/index.php/85/QuickBird>, 2009
- DigitalGlobe WorldView-1 satellite datasheet, <http://www.digitalglobe.com/index.php/86/WorldView-1>, 2009
- DigitalGlobe WorldView-2 satellite datasheet, <http://www.digitalglobe.com/index.php/88/WorldView-2>, 2009
- DIRSIG web page, Digital Imaging and Remote Sensing Laboratory, Rochester Institute of Technology, <http://www.dirsig.org/>, 2009
- Douglas, J., Burke, M., and Ettinger, G. J., "High-resolution SAR ATR performance analysis", *Proceedings SPIE*, Vol. 5427, 293-301, (2004)
- Duggin, M. J., and Loe, R. S., "Calibration and exploitation of a narrow-band imaging polarimeter", *Opt. Eng.*, Vol. 41, No. 5, 1039-1047, May 2002
- Eismann, M. T., "Strategies for Hyperspectral Target Detection in Complex Background Environments", *Proceedings of the IEEE Aerospace Conference*, Paper #1626, 2006

- Ellis, K. K., "Polarimetric bidirectional reflectance distribution function of glossy coatings," *J. Opt. Soc. Am. A.* **13**, 1758-1762, 1996.
- Farrand, W. H. and Harsanyi, J. C., "Mapping the distribution of mine tailings in the Coeur d'Alene River Valley, Idaho, through the use of a constrained energy minimization technique", *Remote Sensing of Environment*, Vol. 59, No. 1, 64-76, 1997
- Fauvel, M., Benediktsson, J. A., Chanussot, J., and Sveinsson, J. R., "Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profiles", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No. 11, Nov. 2008
- Gartley, M. G., Brown, S. D., Goodenough, A. D., Sanders, N. J. and Schott, J. R., "Polarimetric Scene Modeling in the Thermal Infrared", *Proceedings of SPIE*, Vol. 6682, 2007
- Gartley, M. G. and Basener, W., "Topological Anomaly Detection Performance with Multispectral Polarimetric Imagery", *Proceedings of SPIE*, Vol. 7334, 2009
- GeoEye IKONOS satellite specification sheet,
<http://www.geoeye.com/CorpSite/products/imagery-sources/Default.aspx>, 2009
- GeoEye GeoEye-1 satellite specification sheet,
<http://www.geoeye.com/CorpSite/products/imagery-sources/Default.aspx>, 2009
- Google 3D warehouse, "Subaru Legacy WRX",
<http://sketchup.google.com/3dwarehouse/details?mid=b3f43a055e109d9d2b9d9d193d34262e&prevstart=12>, 2009
- Harvey, C., Wood, J., Randall, P., Watson, G., and Smith, G., "Simulation of a new 3D imaging sensor for identifying difficult military targets", *Proceedings of SPIE*, 2008
- Huang, X. and Zhang, L., "An Adaptive Mean-Shift Analysis Approach for Object Extraction and Classification from Urban Hyperspectral Imagery", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No. 12, Dec. 2008
- Iannarilli, F. J. and Conant, J. A., "Quantifying key trade-off between IR polarimetric discriminability versus pixel resolution against complex targets", *Proceedings of SPIE*, Vol. 3699, 1999
- Iannarilli, F. J., Shaw, J. A., Jones, S. H. and Scott, H. E., "Snapshot LWIR hyperspectral polarimetric imager for ocean surface sensing", *Proceedings of SPIE*, Vol. 4133, 2000

- Ientilucci, E. J. and Brown, S. D., "Advances in Wide Area Spectral Image Simulation", *Proceedings of SPIE*, Vol. 5075, 2003
- Kruse, F. A., Boardman, J. W., and Huntington, J. F., "Comparison of Airborne Hyperspectral Data and EO-1 Hyperion for Mineral Mapping", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 6, June 2003
- Kwan, Y., Sawtelle, S., Bernstein, U., Pereira, W., and Less, D., "A simulation for hyperspectral thermal IR imaging sensors", *Proceedings of SPIE*, Vol. 6966, 2008
- Lach, S. R., Kerekes, J. P. and Fan, X., "Fusion of multiple image types for the creation of radiometrically-accurate synthetic scenes", *Journal of Applied Remote Sensing*, Vol. 3, 12 January 2009
- Leachtenauer, J. C., and Driggers, R. G., *Surveillance and Reconnaissance Imaging Systems*, Artech House, Inc., Norwood, MA, 2001
- Lockheed Martin UK CameoSim product webpage,
<http://www.lockheedmartin.co.uk/products/cameosim.html>, 2009
- Macon, C., Wozencraft, J., Park, J. Y., and Tuell, G., "Seafloor and land cover classification through airborne LIDAR and hyperspectral data fusion", *IEEE International Geoscience and Remote Sensing Symposium 2008, IGARSS '08*, Vol. 2, 77-80, July 2008
- Mangolini, M., "Apport de la fusion d'images satellitaires multicapteurs au niveau pixel en teledetection et photo-interpretation", Dissertation published at the University of Nice-Sophia Antipolis, France, 15 November 1994
- Meyers, J. P., Schott, J. R. and Brown, S. D., "Incorporation of polarization into the DIRSIG synthetic image generation model", *Proceedings of SPIE*, Vol. 4816, 2002
- Mitchell, A. A., Brewster, J. M., Haynes, A. W., and Richardson, P. A., "Modeling reflections from exact surfaces in CameoSim", *Proceedings of SPIE*, Vol. 6543, 2007
- Papoulis, A. and Pillai, S. U., *Probability, Random Variables and Stochastic Processes*, The McGraw-Hill Companies, Inc., New York, NY 2002
- Pearlman, J. S., "Hyperion Validation Report", *Boeing Report Number 03-ANCOS-001*, Phantom Works, The Boeing Company, Kent WA, July 2003

- Petrakos, M., Benediktsson, J. A., and Kanellopoulos, I., "The Effect of Classifier Agreement on the Accuracy of the Combined Classifier in Decision Level Fusion", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 11, November 2001
- Piech, K. R. and Walker, J. E., "Interpretation of soils", *Photogrammetric Engin. Remote Sensing*, Vol. 40, 87-94, 1974
- Piech, K. R., Schott, J. R. and Stewart, K. M., "The blue-to-green reflectance ratio and lake water quality", *Photogrammetric Engin. Remote Sensing*, Vol. 44, No. 10, 1303-1319.
- Pogorzala, D., Brown, S., Messinger, D. and Devaraj, C., "Recreation of a nominal polarimetric scene using synthetic modeling tools", *Proceedings of SPIE*, Vol. 6565, 2007
- Pohl, C. and Van Genderen, J. L., "Multisensor image fusion in remote sensing: concepts, methods and applications", *Int. J. Remote Sensing*, Vol. 19, No. 5, 828-834, 1998
- Prasad, S. and Bruce, L. M., "Decision Fusion with Confidence-Based Weight Assignment for Hyperspectral Target Recognition", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No. 5, 1448-1456, May 2008
- Prasad, S., Bruce, L. M., and Kalluri, H., "A robust multi-classifier decision fusion framework for hyperspectral, multi-temporal classification", *IEEE International Geoscience and Remote Sensing Symposium 2008, IGARSS '08*, Vol. 2, 7-11, July 2008
- Price, J. C., "Combining panchromatic and multispectral imagery from dual resolution satellite instruments", *Remote Sensing of Environment*, Vol. 21, 119-128, 1986
- Priest, R. G. and Germer, T. A., "Polarimetric BRDF in the microfacet model: Theory and measurements", *Proceedings of the 2000 Meeting of the Military Sensing Symposia Specialty Group on Passive Sensors*, 169-181, 2002
- Raqueno, N. G., Smith, L. E., Messinger, D. W., Salvaggio, C., Raqueno, R. V. and Schott, J. R., "Megacollect 2004: Hyperspectral Collection Experiment of Terrestrial Targets and Backgrounds of the RIT Megascene and Surrounding Area (Rochester, New York)", *Proceedings of SPIE*, Vol. 5806, 554-565, 2005
- Rhinoceros software, "Rhinoceros: NURBS Modeling for Windows", <http://www.rhino3d.com/>, 2010

- Reed, I. S. and Yu, X., "CFAR detection of the optical pattern with unknown spectral distribution", *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 38, No. 10, 1760-1770, 1990
- Rustan, P., "Building an Integrated Intelligence Network: Challenges and Opportunities", *High Frontier*, Vol. 4, No. 4, 10-14, Aug 2008
- Sadjadi, F. A. and Chun, C. S. L., "Remote sensing using passive infrared Stokes parameters", *Opt. Eng.*, Vol. 43, No. 10, 2283-2291, October 2004
- Savage, J., Coker, C., Thai, B., Aboutalib, O., and Pau, J., "IRMA 5.2 multi-sensor signature prediction model", *Proceedings of SPIE*, Vol. 6965, 2008
- Schott, J. R., *Remote Sensing*, Oxford University Press, New York, NY 2007
- Schott, J. R., *Fundamentals of Polarimetric Remote Sensing*, SPIE, Bellingham, WA, 2009
- Schaum, A. P. "Spectral subspace matched filtering", *Proceedings of SPIE*, Vol. 4381, 114-117, 2001
- Shell, J. R., *Polarimetric Remote Sensing in the Visible to Near Infrared*, PhD thesis, Rochester Institute of Technology, Rochester, NY, 2005
- Simi, C., Winter, E., Williams, M., and Driscoll, D., "Compact Airborne Spectral Sensor (COMPASS)", *Proceedings of SPIE*, Vol. 4381, 129-136, 2001
- Tree Professional software, <http://www.onyxtree.com/>, 2010
- US Air Force MQ-1 Predator factsheet, <http://www.af.mil/information/factsheets/factsheet.asp?id=122>, 2009
- USGS EO-1 pointing mode description, <http://edcsns17.cr.usgs.gov/eo1/lookAngles.php>, 2009
- USGS Hyperion sensor datasheet, <http://edcsns17.cr.usgs.gov/eo1/hyperion.php>, 2009
- Van den Heuvel, J. C., Van Eijk, A. M. J., Bekman, H. H. P. T., Van Putten, F. J. M., Cohen, L. H., Pace, P. W., "Laser applications in the littoral: search lidar and ship identification", *Proceedings of SPIE*, Vol. 7090, 2008
- WASP Lite sensor system, Laboratory for Imaging Algorithms and Systems', Center for Imaging Science, Rochester Institute of Technology, <http://www.cis.rit.edu/~lias/wasplite/>, 2009

Yang, M. and Moon, W. M., "Decision Level Fusion of Multi-Frequency Polarimetric SAR and Optical Data with Dempster-Shafer Evidence Theory" *IEEE International Geoscience and Remote Sensing Symposium 2003, IGARSS '03*, Vol. 6, 3668-3670, July 2003

Zabuawala, S., Nguyen, H., Wei, H. and Yadegar, J., "Fusion of LIDAR and Aerial Imagery for Accurate Building Footprint Extraction", *Proceedings of SPIE*, Vol. 7251, 2009