AFRL-OSR-VA-TR-2013-0006

(DCT FY08) DEVICE, ALGORITHM AND INTEGRATED MODELING RESEARCH FOR PERFORMANCE-DRIVEN MULTI-MODAL OPTICAL SENSORS

JOHN KEREKES

ROCHESTER INSTITUTE OF TECHNOLOGY

12/17/2012
Final Report

AIR FORCE RESEARCH LABORATORY
AF OFFICE OF SCIENTIFIC RESEARCH (AFOSR)/RSE
ARLINGTON, VIRGINIA 22203
AIR FORCE MATERIEL COMMAND

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
29-11-2012	Final	01 February 2008 - 31 August 2012

4. TITLE AND SUBTITLE
Device, Algorithm and Integrated Modeling Research for Performance-driven Multi-modal Optical Sensors

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-08-1-0028

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kerekes, John P., Jared A. Herweg, Lingfei Meng, Zoran Ninkov, Michael D. Presnar, Alan D. Raisanen, Annette O. Rivas, Tingfang Zhang, Sabino M. Gadaleta, Andrew C. Rice, Juan R. Vasquez

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Rochester Institute of Technology
1 Lomb Memorial Drive
Rochester, New York, 14623

8. PERFORMING ORGANIZATION REPORT NUMBER
DIRS Report Number 12-61-184

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
875 N. Randolpht St. Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-OSR-VA-TR-2013-0006

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for unlimited public release

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT
Basic research was conducted in technology associated with adaptive, performance-driven multi-modal optical sensing, as well as the phenomenology of hyperspectral imaging of humans. Contributions included (1) An end-to-end simulation model demonstrated multimodal adaptive sensing in the context of vehicle tracking in an urban environment. (2) A feasible design was developed for a MEMS single-pixel tunable Fabry-Perot spectrometer using novel thermal activation. (3) A novel analytical model was developed for polarimetric imaging systems and used in an adaptive target detection algorithm to optimize polarizer angles. (4) A multimodal target-tracking algorithm was developed combining spectral and polarimetric imagery. (5) A novel technique for spectral waveband selection was developed and used in a vehicle tracking demonstration. (6) A database was developed of human spectral reflectance, which found that VNIR measurements of clothing were more robust than human skin or hair in distinguishing among pedestrians. Future work is recommended in developing the tunable FP device, rotatable polarizers, and algorithms for performance-driven sensing.

15. SUBJECT TERMS
Performance-driven sensing, hyperspectral, polarimetric, multi-modal, target-tracking

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	73	John P. Kerekes
U	U	U			19b. TELEPONE NUMBER (*Include area code*) 585-475-6996

Device, Algorithm and Integrated Modeling Research for Performance-driven Multi-modal Optical Sensors

Air Force Office of Scientific Research
Award Number FA9550-08-1-0028
Final Report

1 February 2008 – 31 August 2012

DIRS Report Number 12-61-184

Authors:

John P. Kerekes, Principal Investigator
Jared A. Herweg
Lingfei Meng
Zoran Ninkov
Michael D. Presnar
Alan D. Raisanen
Annette O. Rivas
Tingfang Zhang
Digital Imaging and Remote Sensing Laboratory
Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology
Rochester, New York USA

in collaboration with

Sabino M. Gadaleta
Andrew C. Rice
Juan R. Vasquez
Numerica Corporation
Loveland, Colorado USA

Executive Summary

This report documents the accomplishments and findings of a four-year research effort conducted by researchers at the Rochester Institute of Technology (RIT) and its subcontractor, the Numerica Corporation. The primary goal of the effort was to conduct basic research into technologies associated with adaptive, performance-driven multi-modal optical sensing in the context of airborne surveillance in an urban area. An additional aspect was added midway through to explore the phenomenology associated with hyperspectral imaging of pedestrians.

The research was motivated in part by the challenges of conducting Intelligence, Surveillance, and Reconnaissance (ISR) missions in complex dynamic urban environments. While individual sensing modalities such as electro-optical (EO), polarization, and hyperspectral imaging can be effective in detecting and tracking moving vehicles, the combination of multiple methods can lead to a more robust capability. However, the additional data volume and complexity can compromise the effectiveness unless done in an intelligent manner. This has lead to the idea of performance-driven adaptive sensing where choices are adaptively made to limit the data acquisition to only what is necessary based on achieving good performance in a given task.

The primary project was built around two different concepts of adaptive sensing. The first is a Multi-Object Spectrometer (MOS) uses a Digital Micromirror Device (DMD) to selectively collect spectra at individual pixels of a scene, while simultaneously collecting an image elsewhere. The second concept uses the idea of a superpixel to adaptively collect multi-modality (spectral and polarization) imagery.

A number of contributions to understanding the science and technology of adaptive multimodal performance driven sensing were made in this research as listed here.

- An end-to-end simulation model was developed to demonstrate multimodal adaptive sensing in the context of vehicle tracking in an urban environment. This work integrated models and simulated images across the imaging chain and further identified performance characteristics of such as system.
- A feasible design was developed and prototyped for a MEMS single-pixel tunable Fabry-Perot spectrometer using a novel thermal activation technique. The design was verified through extensive model analysis of its electrical, mechanical, thermal, and optical properties.
- A novel analytical model was developed for polarimetric imaging systems and used as part of an adaptive target detection algorithm which can optimize the polarizer angles for maximum target-to-background contrast.
- A new multimodal target tracking algorithm was developed which combines spectral and polarimetric imagery to enhance target detection, followed by a novel approach to feature aided tracking using spectral information.

- A novel technique for spectral waveband selection was developed and used as part of a vehicle tracking demonstration of performance-driven sensing. This work quantified for this application the time advantages of selecting a subset of spectral information using the tunable single-pixel spectrometer concept.
- A database was developed of spectral reflectance measurements of humans and their clothing, and was used to assess the separability of pedestrians using hyperspectral images. Clothing was identified as being more robust than human skin or hair in distinguishing among pedestrians, and at least with these data, the visible through near infrared spectrum was found to be adequate for the task.

While these contributions have helped advance the basic science behind multimodal performance-driven sensing, they have in many ways just scratched the surface of understanding. Three specific areas recommended for follow-on work. One is to continue work in the promising technology emerging from this research of the single-pixel tunable Fabry-Perot spectrometer. It is recommended this device development continue as it offers the promise of a truly adaptive imaging spectrometer in a very compact configuration. A second aspect of this research that should be further pursued is the development of optimal algorithms for operating these adaptive sensor concepts. The initial efforts developed here have demonstrated the promise, but much more can be done. A third area is to explore hardware concepts to take advantage of the optimal polarizer angle algorithm to improve polarimetric target detection.

Support from this award have lead to five graduate theses (three Ph.D. and two M.S.), two journal articles in print with one in review and at least two more in preparation, three peer-reviewed conference proceeding papers, and 11 additional conference proceedings papers. The work has been accomplished with contributions from the following individuals.

Scott D. Brown
Kenneth D. Fourspring
Sabino M. Gadaleta
Jared A. Herweg
Emmett J. Ientilucci
John P. Kerekes
Zhong Lu
Robert T. MacIntyre
Michael J. Mendenhall
Lingfei Meng
Zoran Ninkov
Jeffrey P. Patel
Alan D. Raisanen
Andrew C. Rice
Annette O. Rivas
Kyle M. Tarplee
Juan R. Vasquez
Tingfang Zhang

The following publications that have appeared in print were supported at least in part under this award.

Fourspring, K., Z. Ninkov, and J. Kerekes, "Subpixel Scatter in Digital Micromirror Devices," *Proceedings of Emerging Digital Micromirror Device Based Systems and Applications II*, SPIE Vol. 7596, SPIE Photonics West Meeting, San Francisco, California, January 2010.

Fourspring, K., Z. Ninkov, and J. Kerekes, "Scattered Light in a DMD-based Multi-object Spectrometer," *Proceedings of Modern Technologies in Space- and Ground-based Telescopes and Instrumentation*, SPIE Vol. 7739, SPIE Astronomical Instrumentation, San Diego, California, June 2010.

Gadaleta, S., J. Kerekes, and K. Tarplee, "Autonomous Target Dependent Waveband Selection for Tracking in Performance-driven Hyperspectral Sensing," *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, SPIE Vol. 8390, 839023, April 2012.

Herweg, J., J. Kerekes, E. Ientilucci, and M. Eismann, "Spectral Variations in HSI Signatures of Thin Fabrics for Detecting and Tracking of Pedestrians," *Proceedings of Active and Passive Signatures II*, SPIE Vol. 8040, 8040G, SPIE Defense Sensing and Security Symposium, Orlando, Florida, April 2011.

Herweg, J., J. Kerekes, and M. Eismann, "Hyperspectral Measurements of Natural signatures: Pedestrians," *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, SPIE Vol. 8390, 83901C, April 2012.

Herweg, J., J. Kerekes, O. Weatherbee, D. Messinger, J. van Aardt, E. Ientilucci, Z. Ninkov, J. Faulring, N. Raqueno, and J. Meola, "SpecTIR Hyperspectral Airborne Rochester Experiment Data Collection Campaign," *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, SPIE Vol. 8390, 839028, April 2012.

Herweg, J., J. Kerekes, M. Eismann, "Hyperspectral Imaging Phenomenology for the Detection and Tracking of Pedestrians," *Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 5482-5485, Munich, Germany, 2012.

Lu, Z., A. Rice, J. Vasquez, and J. Kerekes, "Target Discrimination Via Optimal Wavelength Band Selection with Synthetic Hyperspectral Imagery," *Proceedings of the Second IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 14-16 June 2010.

Meng, L. and J.P. Kerekes, "Adaptive Target Detection with a Polarization-sensitive Optical System," *Applied Optics*, vol. 50, no. 13, pp. 1925-1932, 1 May 2011.

Presnar, M., J. Kerekes, and D. Pogorzala, "Dynamic Image Simulations for Adaptive Sensor Performance Predictions," *Proceedings of the Second IEEE Workshop on*

Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 14-16 June 2010.

Raisanen, A., M. Presnar, Z. Ninkov, K. Fourspring, L. Meng, and J. Kerekes., "Simulation of Practical Single-pixel Wire-grid Polarizers for Superpixel Stokes Vector Imaging Arrays," *Optical Engineering*, vol. 51, 016201, January 2012.

Rice, A., J. Vasquez, M. Mendenhall, and J. Kerekes, "Feature-Aided Tracking via Synthetic Hyperspectral Imagery," *Proceedings of the First IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 26-28 August 2009

Meng, L. and J. Kerekes, "Analytical Modeling of Optical Polarimetric Imaging Systems," *Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3998-4001, Vancouver, Canada, 2011.

Presnar, M. and J. Kerekes, "Modeling and Measurement of Optical Polarimetric Image Phenomenology in a Complex Urban Environment," *Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 4389-4392, Honolulu, HI, 2010.

Presnar, M., A. Raisanen, D. Pogorzala, J. Kerekes and A. Rice, "Dynamic Scene Generation, Multimodal Sensor Design, and Target Tracking Demonstration for Hyperspectral/Polarimetric Performance-Driven Sensing," *Proceedings of Polarization: Measurement, Analysis, and Remote Sensing IX*, SPIE Vol. 7672, 76720T, SPIE Defense Sensing and Security Symposium, Orlando, Florida, April 2010.

Rivas, A., J. Kerekes, and A. Raisanen, "Tunable Single Pixel MEMS Fabry-Perot Interferometer," *Applied Industrial Optics: Spectroscopy, Imaging and Metrology*, OSA Technical Digest, paper JWA15, July 2011.

Table of Contents

List of Figures

List of Tables

Section 1. Introduction and Objectives

This report documents the accomplishments and findings of a four-year research effort conducted by researchers at the Rochester Institute of Technology (RIT) and its subcontractor, the Numerica Corporation. The primary goal of the effort was to conduct basic research into technologies associated with adaptive, performance-driven multi-modal optical sensing in the context of airborne surveillance in an urban area. An additional aspect was added midway through to explore the phenomenology associated with hyperspectral imaging of pedestrians. The significant findings of the research conducted are reported here, with reference provided to the theses, reports, and other publications that contain the full details.

The research was motivated in part by the challenges of conducting Intelligence, Surveillance, and Reconnaissance (ISR) missions in complex dynamic urban environments. While individual sensing modalities such as electro-optical (EO), polarization, and hyperspectral imaging can be effective in detecting and tracking moving vehicles, the combination of multiple methods can lead to a more robust capability. However, the additional data volume and complexity can compromise the effectiveness unless done in an intelligent manner. This has lead to the idea of performance-driven adaptive sensing where choices are adaptively made to limit the data acquisition to only what is necessary based on achieving good performance in a given task.

The primary project was built around two different concepts of adaptive sensing. The first, shown in Figure 1, was adopted from an existing sensor designed for astronomical imaging applications. This Multi-Object Spectrometer (MOS) uses a Digital Micromirror Device (DMD) to selectively collect spectra at individual pixels of a scene, while still collecting an image elsewhere. The second concept is shown in Figure 2 and uses the idea of a superpixel to adaptively collect multi-modality (spectral and polarization) imagery.

Figure 3 depicts the architecture of the primary research project. The scene phenomenology, device modeling, and system modeling research was done primarily be researchers at RIT, while the algorithm research was conducted primarily by Numerica.

The additional research topic was added in 2010 through interest and support from the Air Force Research Laboratory's Sensors Directorate to explore the scene phenomenology associated with hyperspectral imaging of dismounts, or pedestrians. This additional work was conducted by researchers at RIT in collaboration with several organizations in the Dayton, Ohio area. That work included data collection, simulation, and analysis aspects and the significant findings are discussed further in this report as well.

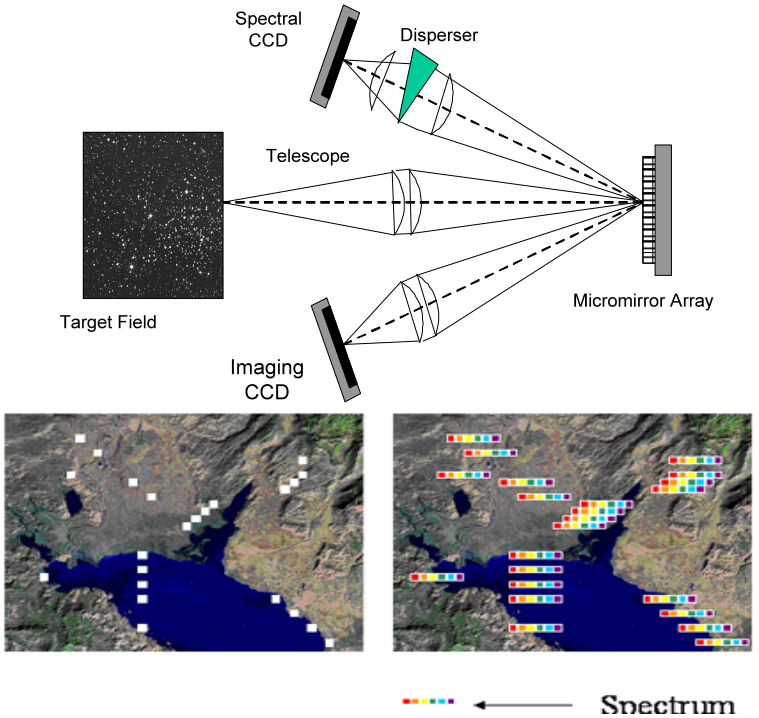


Figure 1. Concept for a DMD-based MOS (top). Illustrated use in remote sensing (bottom) where point spectra are collected at dynamically-located individual pixels.

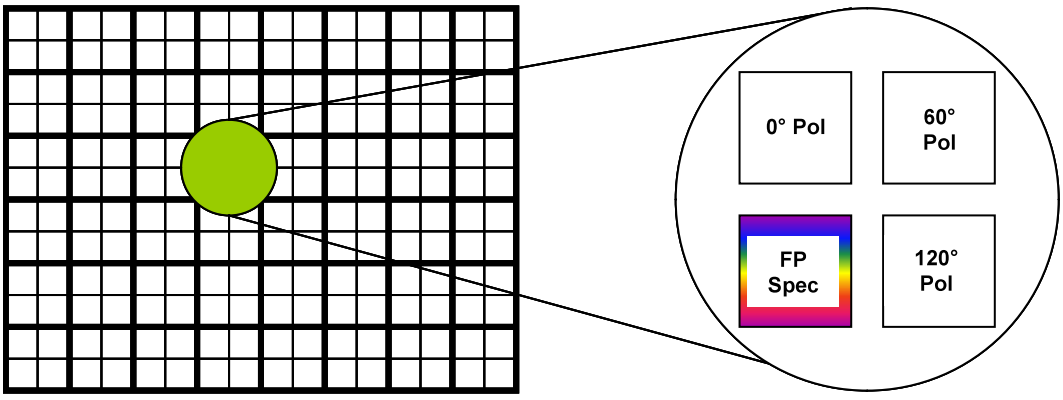


Figure 2. Integrated co-registered multimodality focal plane array concept. Each spatial pixel (thick lines on array) is actually a 2x2 array of detectors with each detector having an integrated polarization filter of a tunable Fabry-Perot spectrometer.

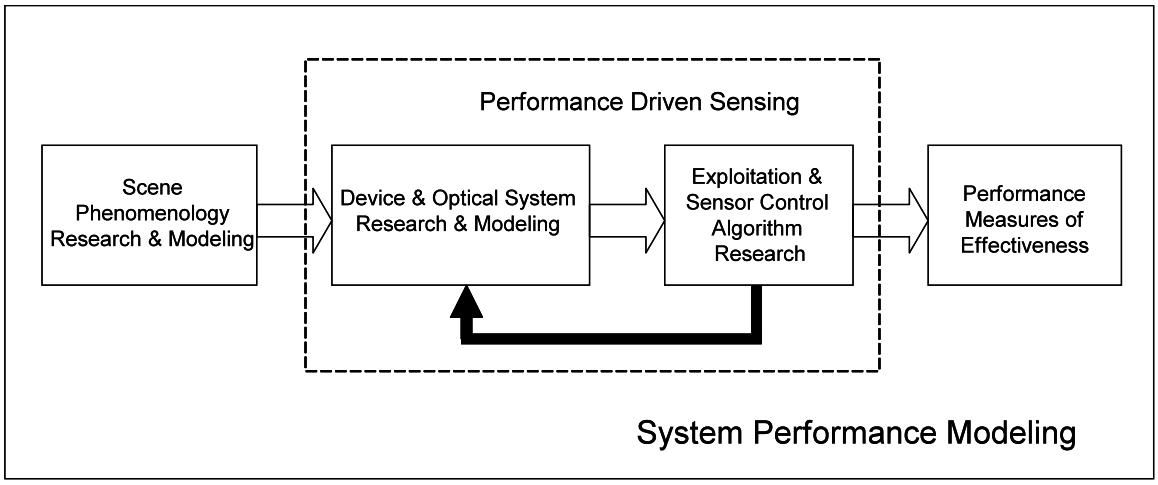


Figure 3. Primary research project architecture.

Section 1.1. Primary Research Project Objectives

The objectives of the initial research project were to conduct basic research through modeling and simulation into the challenges associated with performance-driven adaptive multimodal sensing in the context of airborne imaging and tracking of vehicles in an urban area. Topics included the following.

- Explore issues of the use of DMD's in the MOS configuration for adaptive sensing.
- Develop a system modeling capability for an adaptive sensor and demonstrate performance through simulation and analysis.
- Develop and analyze designs for single-pixel tunable spectrometers.
- Develop and analyze designs for single-pixel wire-grid polarizers.
- Develop system modeling tools for the performance of polarization imaging systems.
- Explore issues and performance capability of spectropolarimetric systems for target tracking.
- Develop and explore performance of algorithms for selecting feature measurements for adaptive sensing in target tracking applications.

Section 1.2. Additional Research Project Objectives

The additional project was incorporated to explore the basic scene phenomenology and information aspects of hyperspectral imaging of pedestrians. Topics to be explored in this project included the following.

- Establish a database of pedestrian hyperspectral signatures.
- Investigate requirements and association performance for regions of the optical spectrum used in pedestrian imaging.
- Investigate use of multiple regions on a pedestrian for association.

- Explore appropriate metrics for separating a unique pedestrian from the background and other pedestrians.
- Investigate loss in association performance due to environmental factors.

Section 2. Technical Accomplishment Summaries

The research in this project was accomplished by students, faculty, and staff at RIT, as well as by research scientists at the subcontractor Numerica Corp. The research at RIT was focused through five graduate student theses (3 Ph.D. and 2 M.S.). This section of the report provides a short summary of the significant findings of the following research thrusts, organized by the thesis and the contribution of scientists at Numerica as follows.

- Adaptive multimodal sensors – Ph.D. thesis by Presnar (2010)
- Tunable single-pixel spectrometers – M.S. thesis by Rivas (2011)
- Analytical modeling of polarimetric imaging systems – Ph.D. thesis by Meng (2012)
- Spectropolarimetric target tracking – M.S. thesis by Zhang (antic. 2013)
- Integrated multi-modal sensing, processing, and exploitation – research by Numerica Corp. (Final report in Appendix A, 2012)
- Hyperspectral imaging phenomenology of pedestrians – Ph.D. thesis by Herweg (2012)

Section 2.1. Adaptive Multimodal Sensors

This work investigated an integrated aerial remote sensor design approach to address moving target detection and tracking problems within highly cluttered, dynamic ground based scenes (Presnar, 2010). Complex modeling of novel micro-opto-electro-mechanical systems (MOEMS) devices, optical systems, and detector arrays resulted in a proof of concept for a state-of-the-art imaging spectropolarimeter sensor model and the quantification of performance in a target tracking application with varying ground scenery, flight characteristics, or sensor specifications. The research culminated an end-to-end simulated demonstration of multimodal aerial remote sensing and target tracking.

Figure 4 conveys the flow of the research. Dynamic urban scenes with moving vehicles imaged by moving platforms were simulated to produce radiometrically accurate panchromatic, color, polarization, and spectral imagery. Accurate optical and radiometric models of the sensors incorporated realistic characteristics in the imagery. These images were then analyzed using a state-of-the-art tracking algorithm and tracking performance studied as a function of scene, sensor, and algorithm characteristics. Figure 5 shows a frame of the simulated imagery as observed by a color (RGB) sensor.

In addition to the system level modeling, detailed device modeling was accomplished for innovative single-pixel wire-grid polarizers (Raisanen, et al, 2012). One significant finding from that work was that relative coarse grids (~500 nm) that can be fabricated on high NA *i*-line lithography equipment can achieve adequate performance to provide useful polarization imagery for target tracking applications.

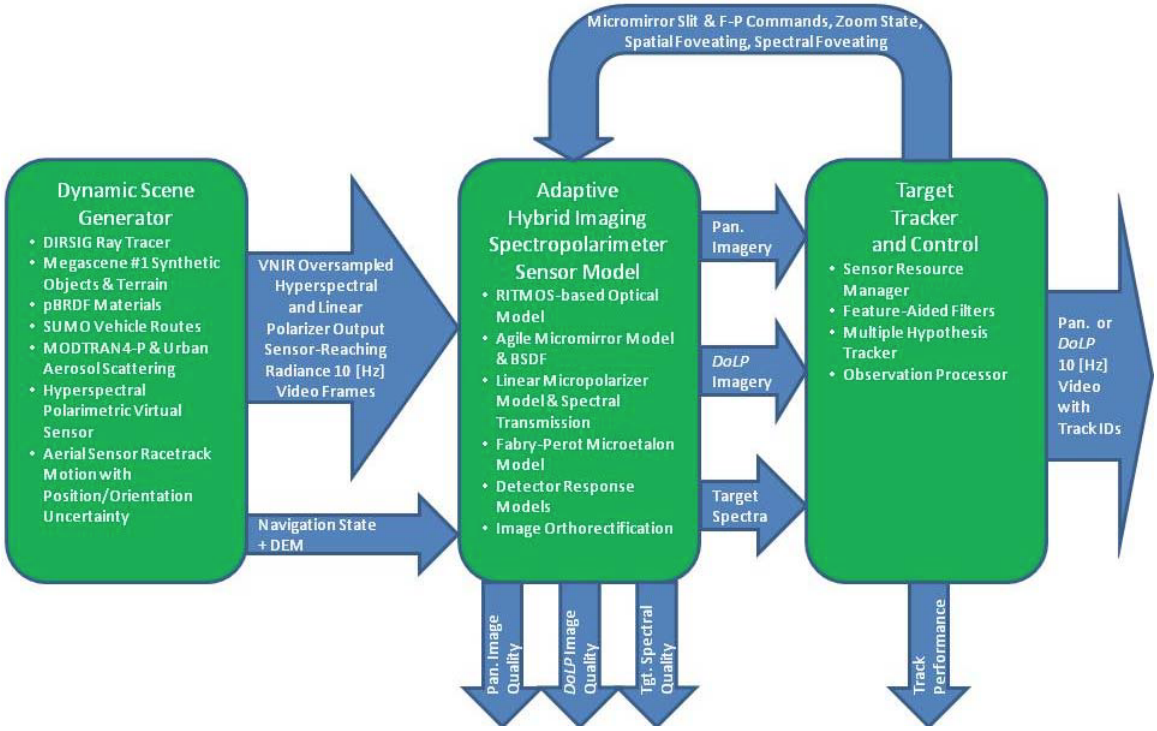


Figure 4. Adaptive multi-modal sensor research architecture.

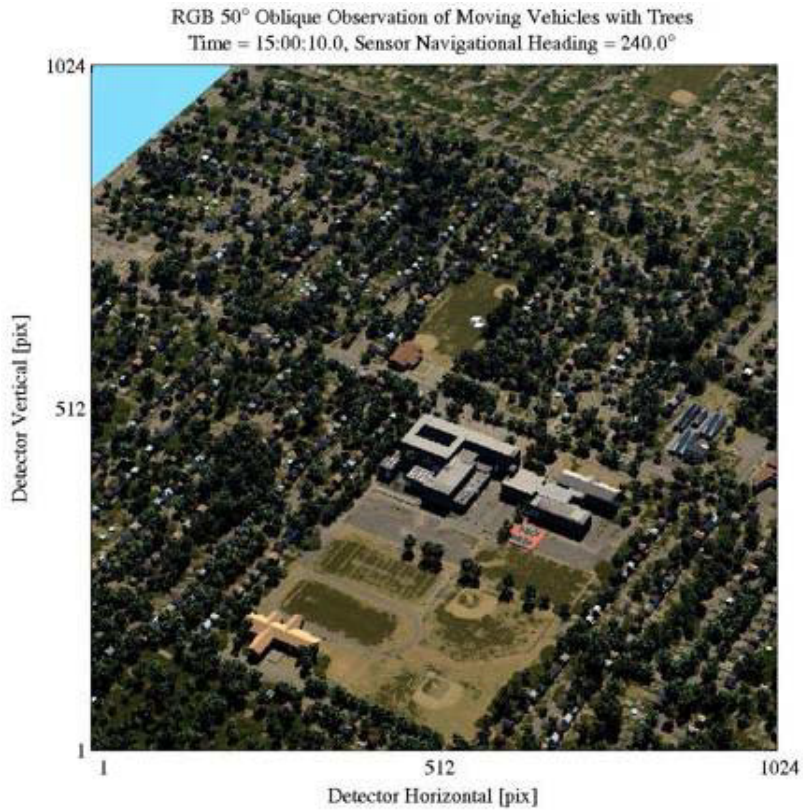


Figure 5. Simulated color oblique image from the airborne sensor.

The adaptive nature of the multi-object spectrometer sensor was demonstrated in the context of target tracking by using the Numerica Algorithm simulator for Tracking and Observations (ALTO) modified to work with the adaptive sensor. Figure 6 shows an example of how after tracks are initiated through motion detection using panchromatic imagery (or degree of linear polarization – DoLP – imagery), pixels are identified for interrogation by the spectrometer as identified by the red marks. These spectra are then used in a feature-aided-tracking scheme to improve track robustness.

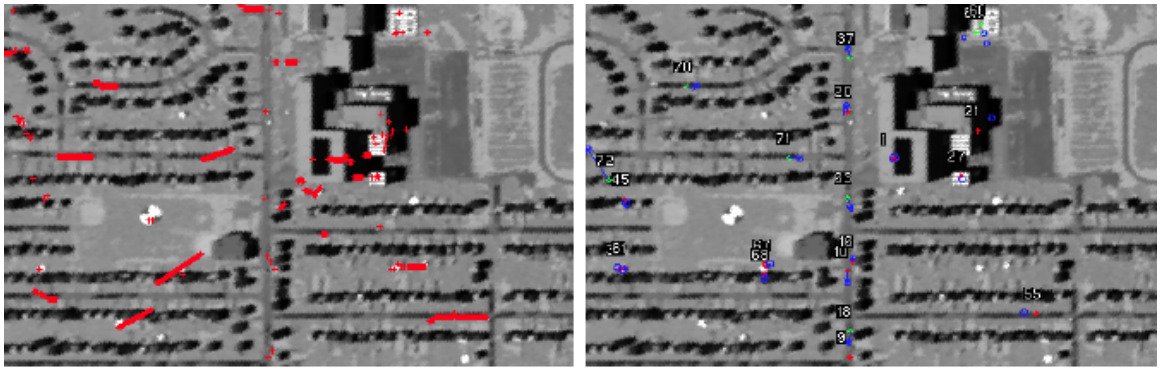


Figure 6. HSI pixel queries (red in left image) and identified tracks (right image).

Tracking performance was quantitatively assessed using the following two metrics: track completeness and track purity as defined below.

$$\text{Track Completeness} = \frac{\#\ \text{of Valid} - \text{Tracks}}{\#\ \text{of Should} - \text{Tracks}}$$

$$\text{Track Purity} = \frac{\#\ \text{of epochs a valid track maintained the same truth vehicle}}{\text{Total}\ \#\ \text{of epochs in a valid track}}$$

Overall, track performance was slightly higher using the DoLP imagery due to the enhanced polarization signature from the shiny paint and glass of the vehicles. High oblique sensor angles, clear atmospheric conditions, and facing the sensor into the sun were situations where the DoLP imagery was better than the panchromatic. Performance was also found to be higher for images with high spatial resolution (sub-meter). Degradation of performance occurred when using too many spectral bands (61 vs. 13), high levels of atmospheric aerosols were present, and when platform jitter introduced spurious motion in the imagery.

Full details on this work can be found in the Ph.D. thesis published by Michael Presnar in 2010 listed in Section 3 of this report.

Section 2.2. Tunable Single-Pixel Spectrometers

A part of the superpixel multi-modal sensor concept is a single-pixel tunable spectrometer. Basic research was undertaken in this part of the project to design and model a single-pixel tunable Fabry-Perot spectrometer using a novel thermal actuation method for the mirror movement (Rivas, 2011). The innovation here is that by having an individually tunable spectrometer at each pixel, the adaptive sensor can collect different spectral information at each spatial location independently, as opposed to current imaging spectrometers that collect the same spectra for all pixel locations.

Figure 7 shows the device design and its thermal response during actuation by thermally heating the legs by passing a current through. The spectrometer functions with a partially transmissive mirror on the top and bottom, with a detector placed below the bottom mirror. The tuning is accomplished by varying the space between the partially transmissive mirrors using a resonant cavity effect.

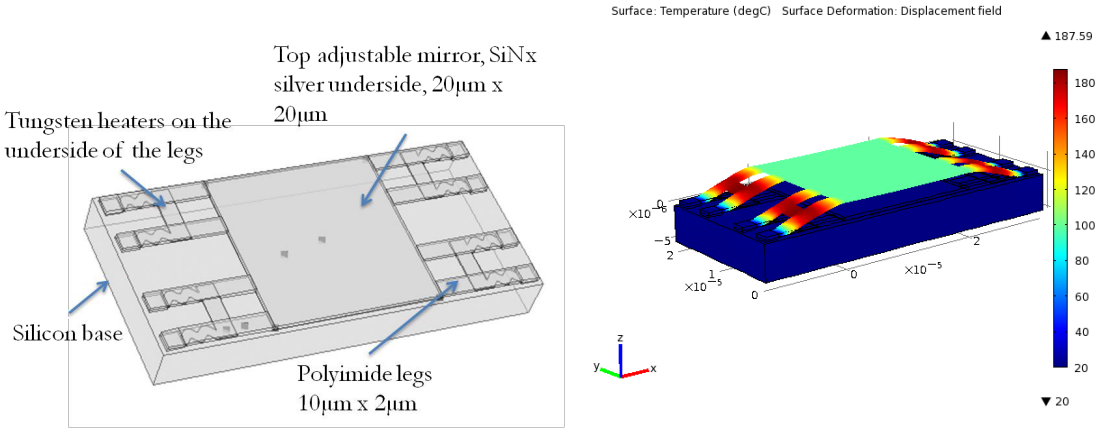


Figure 7. Tunable single-pixel FP spectrometer design (left) and thermal performance during mirror actuation from thermally heating legs (right). Linear dimensions in image on the right side are in meters.

Extensive performance modeling of the design was accomplished using the multi-physics simulation and modeling code COMSOL. Thermal modeling demonstrated the desired range of motion (~300 nm) could be achieved without excessive thermal effects, even if deployed in an array configuration. Temporal performance showed the mirror could be scanned through its range of motion at rates exceeding 100 Hz.

Optical performance was also modeled and is shown in Figure 8. These curves show the typical lower transmission associated with Fabry-Perot devices, but demonstrated the spectral selectivity possible by varying the thermally-induced gap distance.

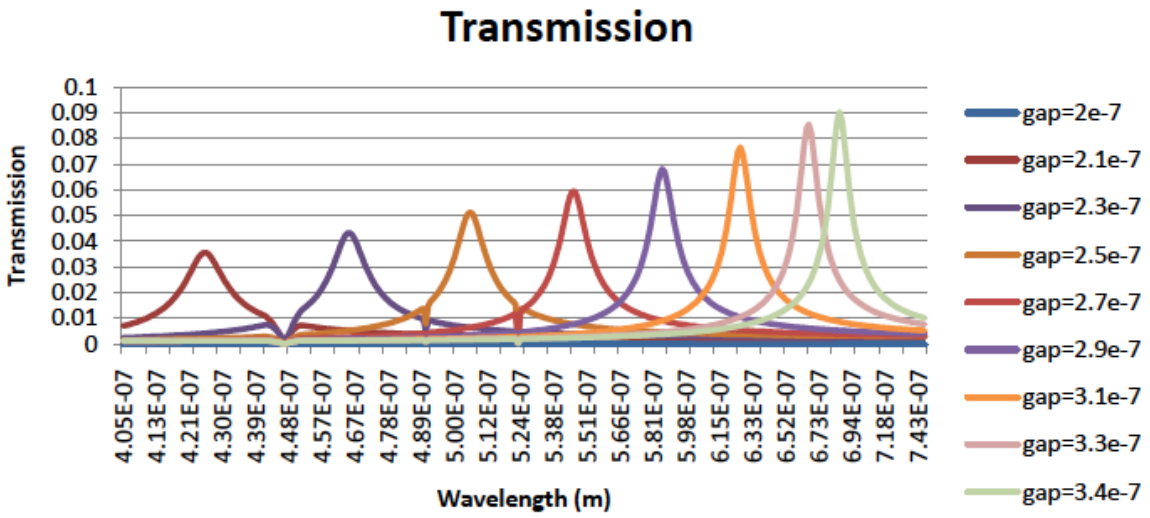


Figure 8. Optical transmission vs. wavelength of the single-pixel FP device. The units of the gap measurement are meters.

This research demonstrated the potential of a practical single-pixel tunable spectrometer to achieve useful levels of performance. Additional details are available in the M.S. thesis published by Annette Rivas in 2011 listed in Section 3 of this report.

Subsequent to the completion of the thesis by Rivas, some prototype devices were fabricated in RIT's Semiconductor & Microsystems Fabrication Laboratory (SMFL). Figure 9 shows a microphotograph of the fabricated prototype. The device is currently undergoing testing to characterize its mechanical, electrical, and optical performance.

Section 2.3. Analytical Modeling of Polarimetric Imaging Systems

In this research into multi-modal adaptive sensing it was observed that a key component was a method to characterize and predict performance of a modality (or its parameter settings) in a given situation so as to drive its adaptive setting in an optimal manner. This was observed to be particularly true for the polarimetric imaging mode since the performance of theses systems can vary significantly depending on the characteristics of the scene and the sun-target-sensor geometry.

To address this need a comprehensive end-to-end polarimetric imaging system modeling tool was developed (Meng, 2012). This end-to-end model includes models for the polarized reflectance characteristics of surface objects, the intervening atmosphere, and the polarimeter optical components, as well as the algorithms used to process the polarization imagery. Rather than pursue a discrete simulation approach, which would be computationally intensive and more limited to specific situations, a more general analytical modeling approach was pursued.

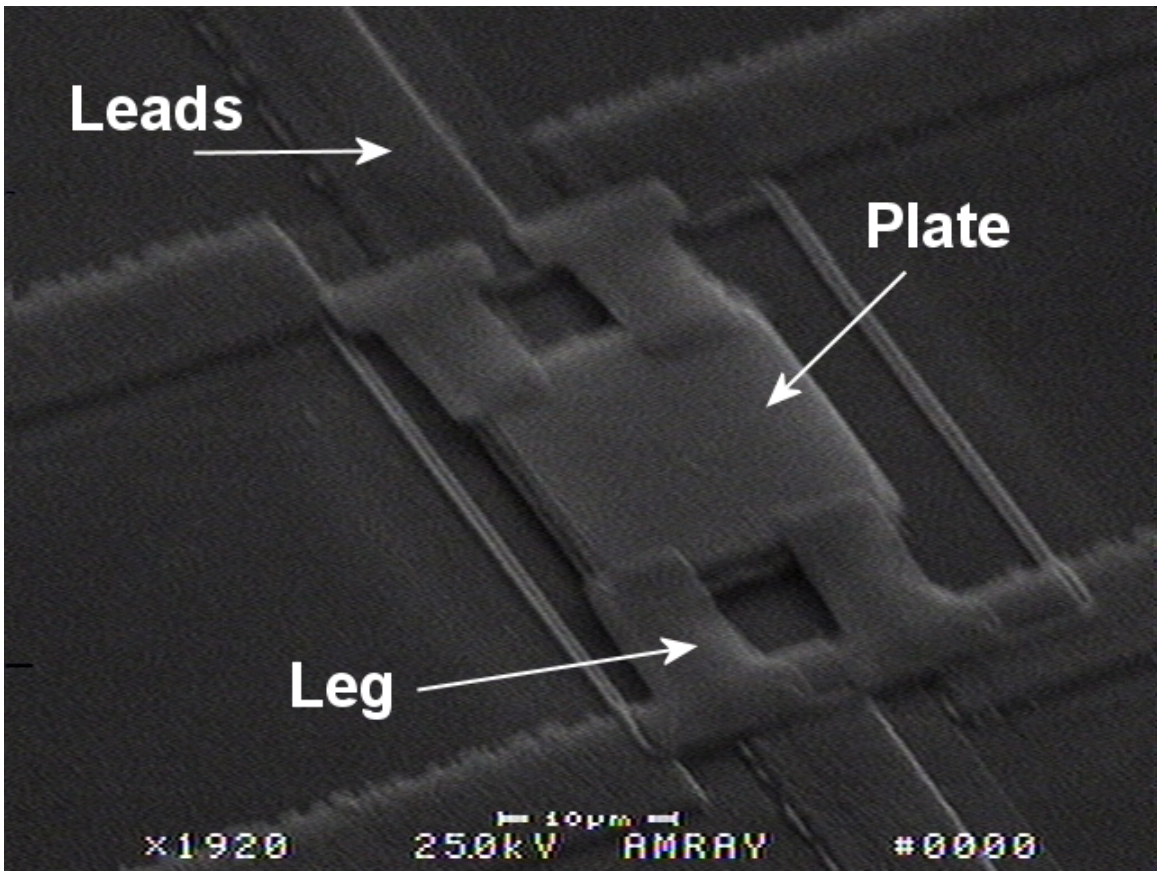


Figure 9. Prototype tunable single-pixel Fabry-Perot device. Electrical leads at top and bottom bring current to the heaters embedded in the legs. The top plate is 20 micrometers across.

Figure 10 shows an overview of the modeling approach. As can be seen the model divides the imaging process into three components: the scene, the sensor, and the processing. It is an analytical model in the sense that at each component the scene classes are characterized by statistical parameters rather than discrete spatial pixels. This analytical formulation leads to quick computation of performance.

The model uses polarized versions of the bidirectional reflectance distribution function (pBRDF) to describe the reflectance characteristics of surface materials. A polarized version of the Air Force atmospheric modeling code MODTRAN is used to model the effects of solar illumination and the atmosphere. Optical and radiometric effects of a sensor are modeled including the co- and cross-polarization characteristics of polarizing filters, the spectral response, and detector noise sources. Stokes vector means and covariances at each stage of the system are the parameters that are propagated through the imaging process. These are then used to calculate features such as the DoLP defined earlier.

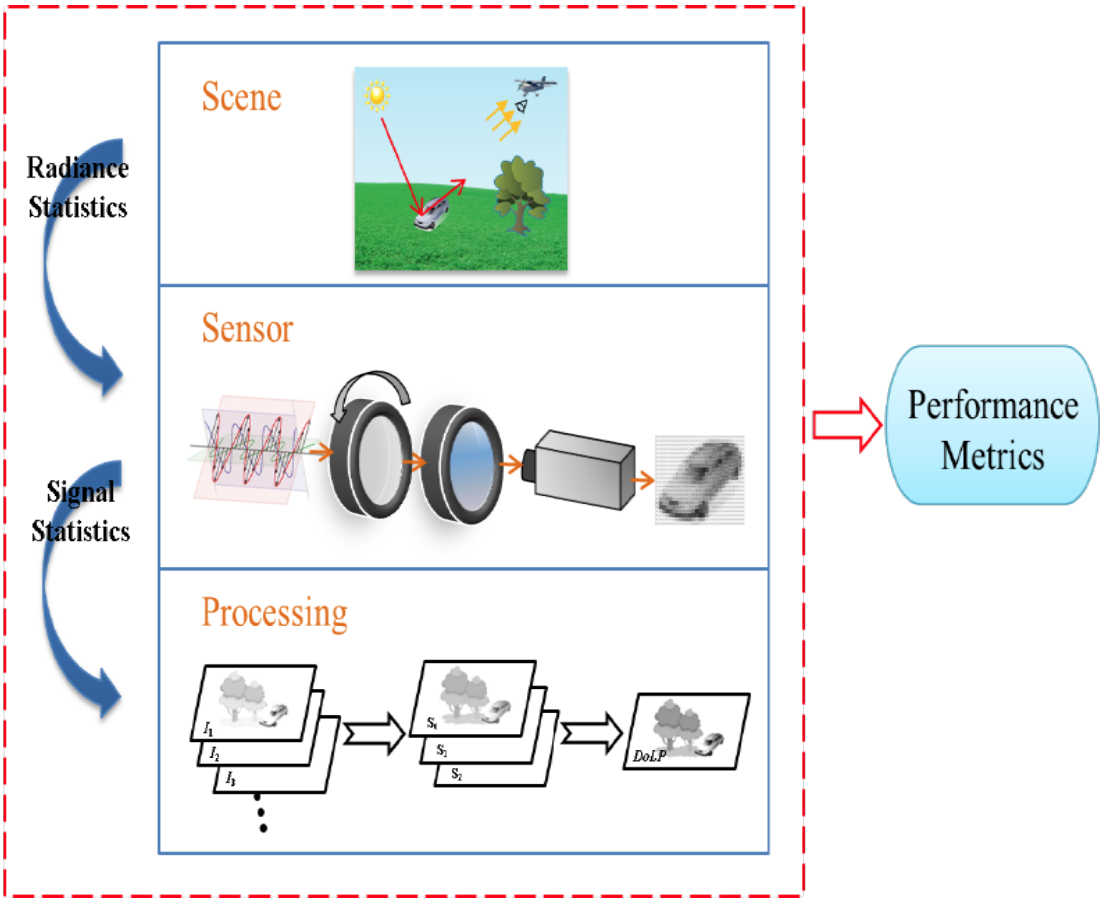


Figure 10. Framework of analytical model of polarimetric imaging systems.

Validation of the model occurred through observations of black and white painted panels with a real polarimetric imaging system and comparing the resulting observations to those predicted by the model. Figure 11 shows such a comparison demonstrating an excellent match between the real data and the model prediction. Figure 12 shows two example uses of the model to predict target detection performance (black vs. white painted panels). The left side of Figure 12 demonstrates the depolarizing effect and decreased detection probability with the increased path length for higher altitude observations, while the right side shows the higher detection probability possible with smooth (low roughness) surfaces.

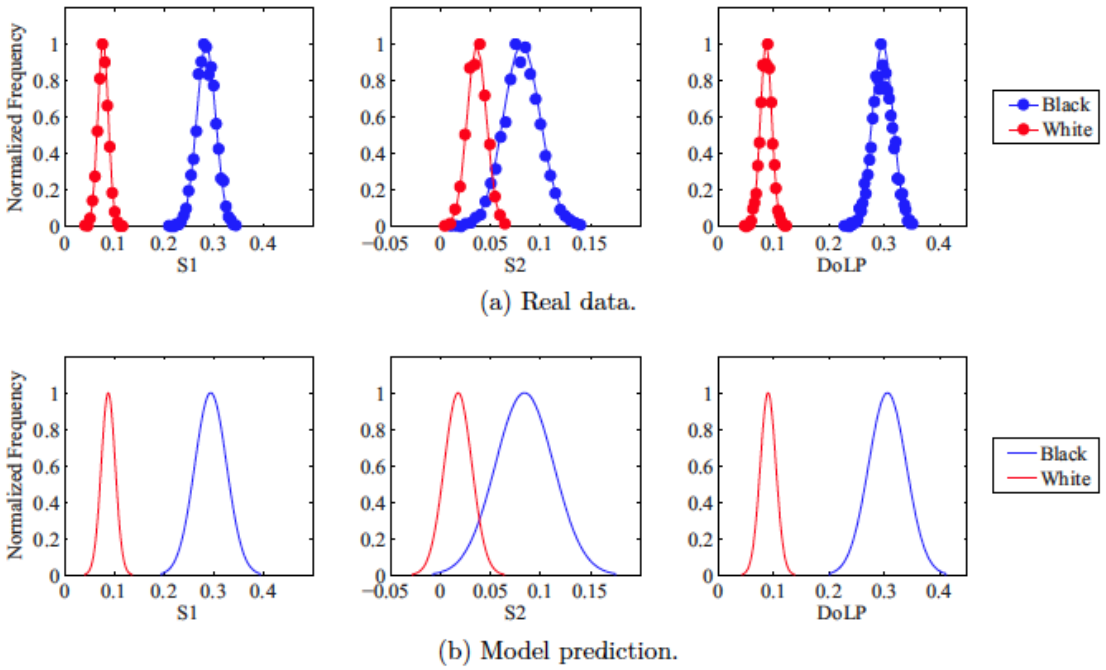


Figure 11. Polarimetric system model validation through comparison of the histograms of the Stokes parameters S1 and S2, and the DoLP.

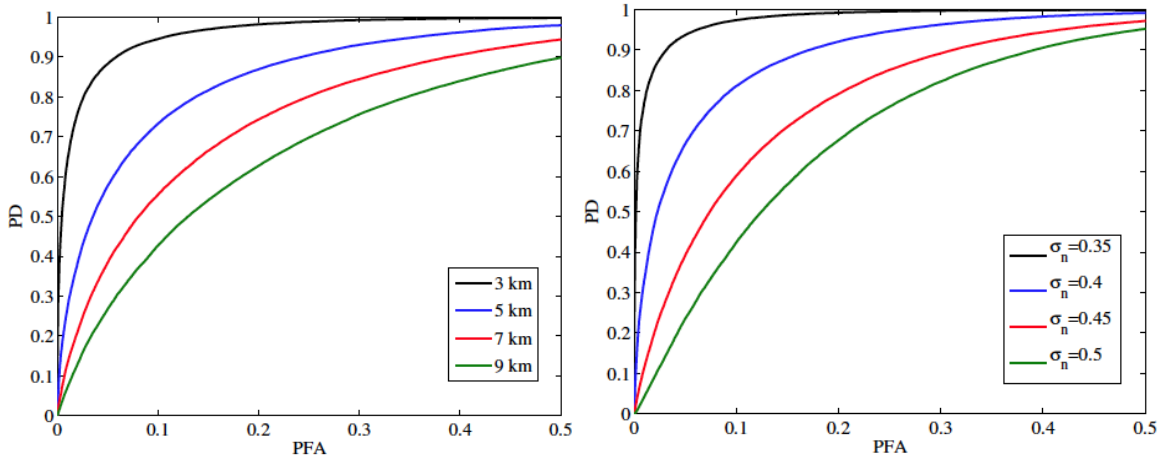


Figure 12. Model-predicted ROC curves showing the detection probability (PD) vs. false alarm rate (PFA) for sensor altitude (left) and surface roughness (right).

The system model was also used as part of a novel adaptive polarimetric target detection scheme where a metric estimated by the model (the signal-to-clutter ratio for the target vs. the background) was used to adaptively change the linear polarizing angles of the polarizers to improve target detection (Meng and Kerekes, 2011). Figure 13 below shows the flowchart for the algorithm and Figure 14 demonstrates its improved performance in a target detection scenario relative to

using the standard analyzer angle settings (indicated by their commonly used names of Pickering, Fessenkov, and Modified Pickering.)

Full details on this work can be found in the Ph.D. thesis published by Lingfei Meng in 2012 listed in Section 3 of this report.

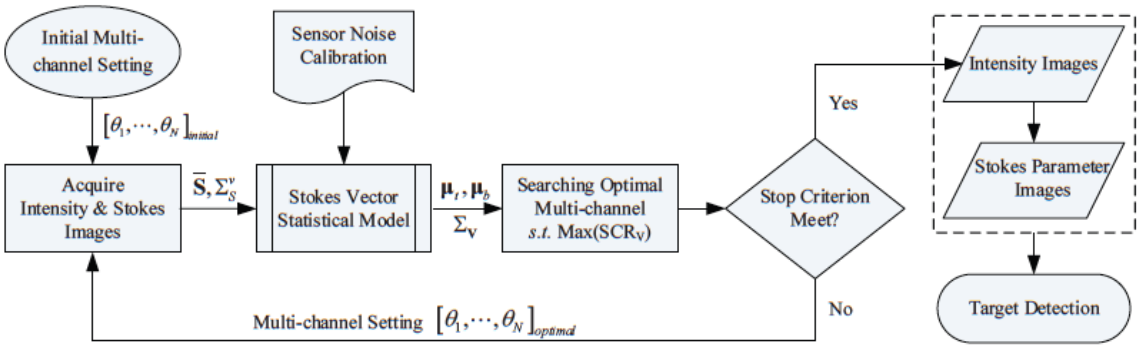


Figure 13. Flowchart for the adaptive polarimetric target detector (APTD) algorithm.

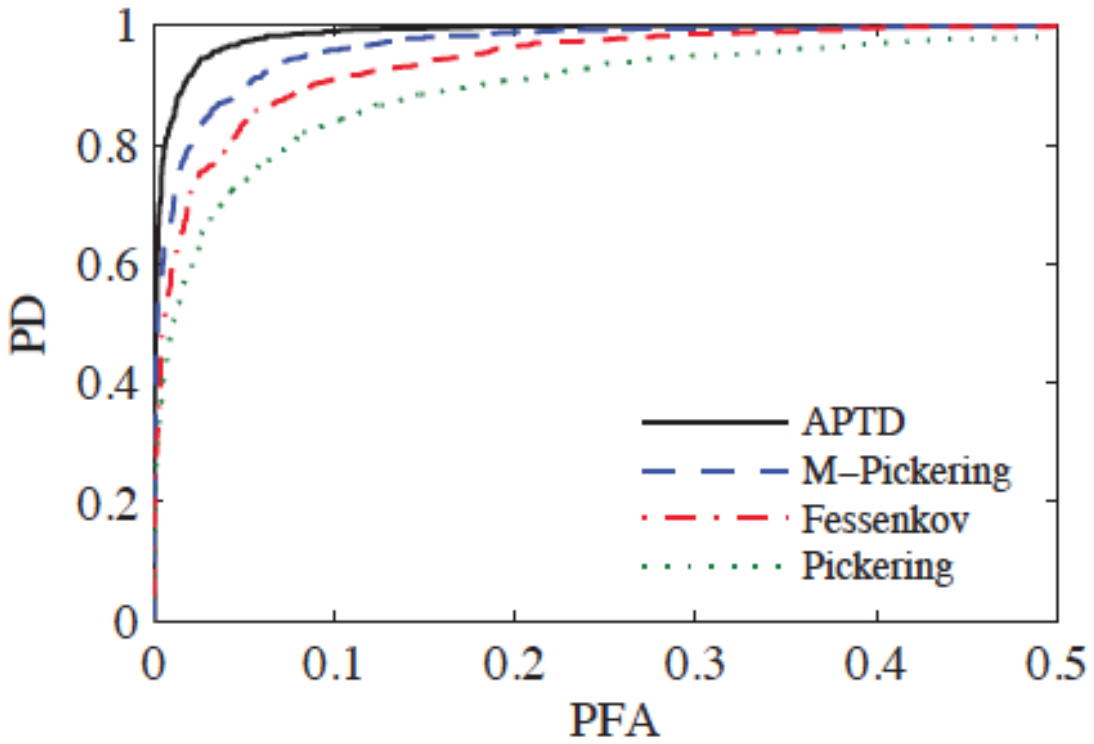


Figure 14. ROC curves demonstrating the improved detection performance possible using the APTD algorithm relative to the existing schemes for setting the linear polarizer angles.

Section 2.4. Spectropolarimetric Target Tracking

This aspect of the research was focused on exploring the utility of spectral and polarimetric information to help with the vehicle tracking application. Through the use of both empirical and simulated multi-modality imagery, combinations of previously developed algorithms were applied and the enhancements to performance studied. In particular, the combined use of spectral and polarimetric information was found to enhance the moving target detection step while spectral information alone was found to be helpful when used in a feature aided tracking approach. Polarization information was found to be not helpful in the tracking aspect (Zhang, 2013).

Figure 15 shows the flowchart of the approach found to be successful in detecting moving vehicles by the combined RX anomaly detection with change detection approach. Figure 16 shows the flowchart for the feature aided tracking algorithm developed in this work.

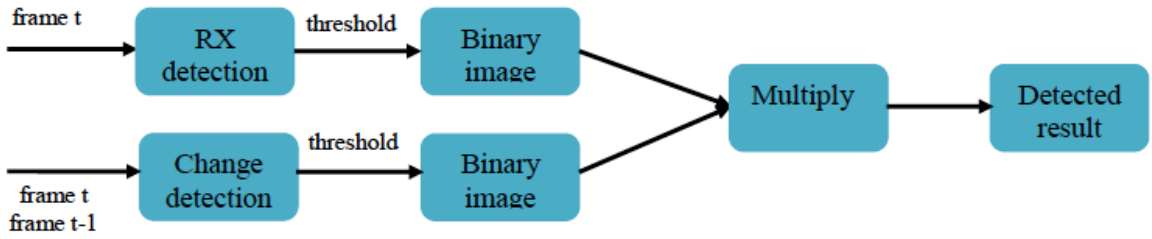


Figure 15. Flowchart for the combined detection method.

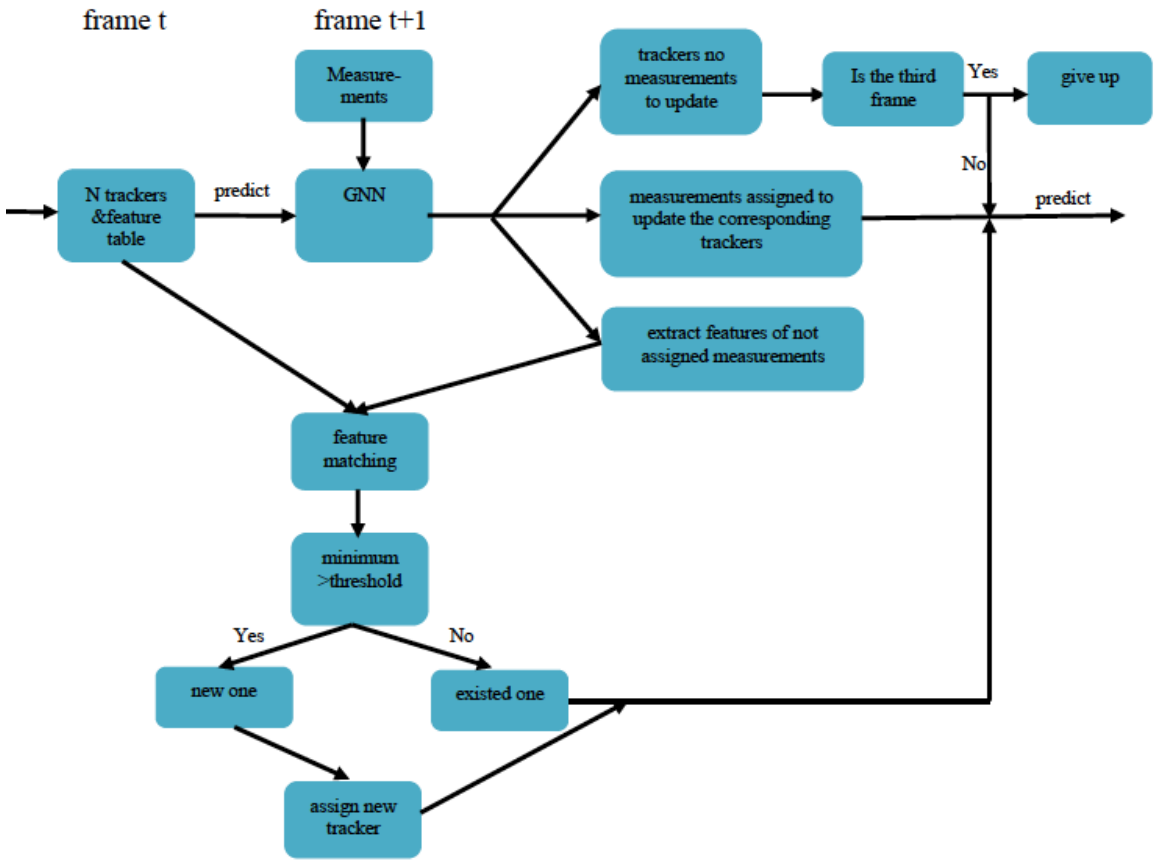
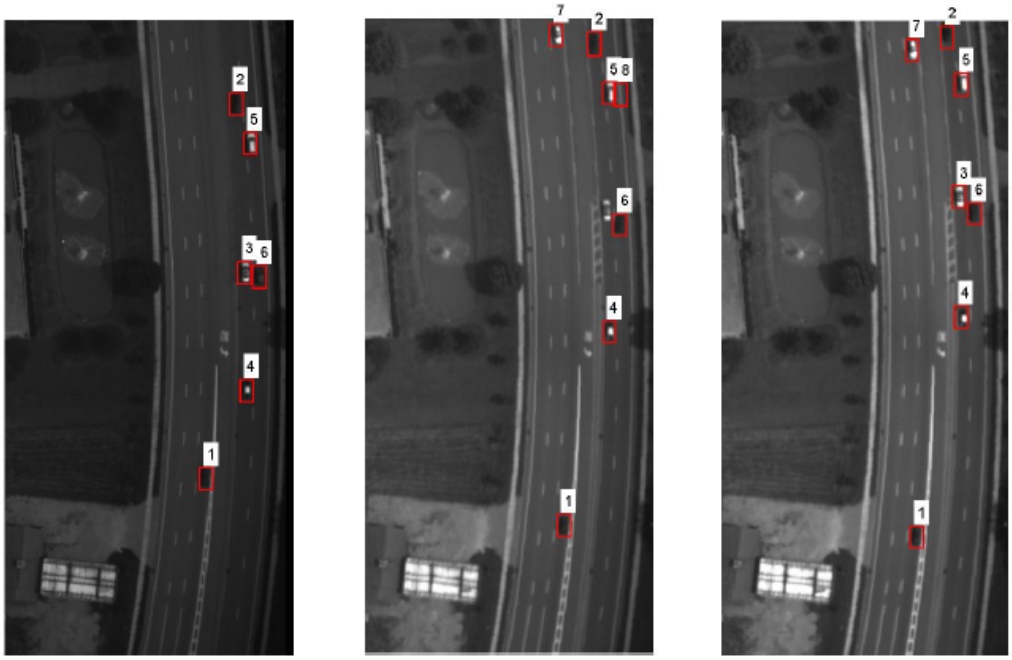


Figure 16. Flowchart for the spectral feature aided tracking method.

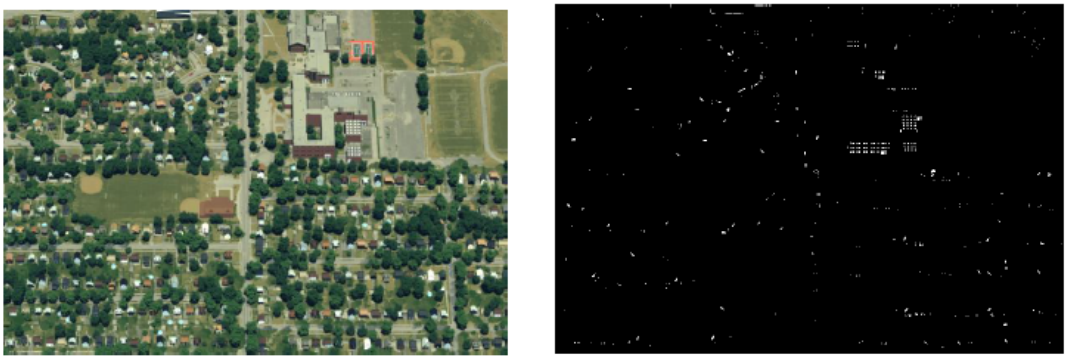
Figure 17 shows an example result for the feature aided tracking using empirical airborne spectral imagery. In this case the use of the spectral information lead to the successful tracking of all vehicles while the Kalman filter motion-only algorithm could not properly track one of the vehicles. Figure 18 demonstrates the improved detection (reduced false alarms) achieved using the combined detection method.

Full details on this work can be found in the M.S. thesis to be published in early 2013 by Tingfang Zhang as listed in Section 3 of this report.

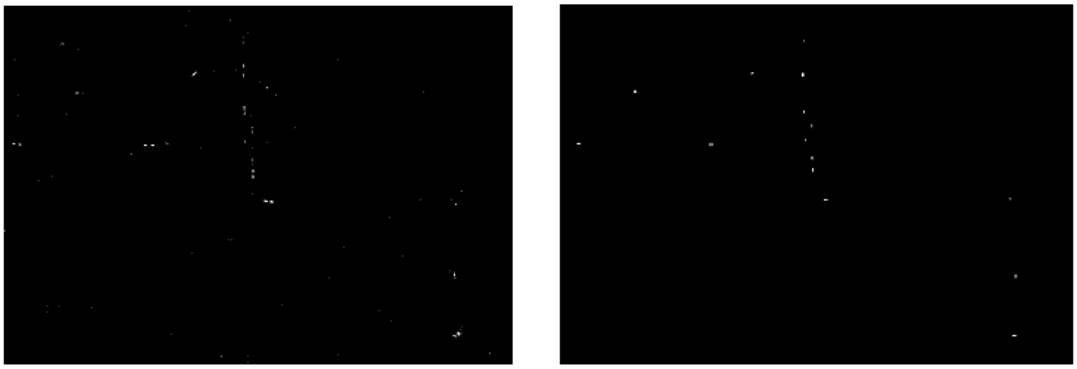


(a) Frame 1 (b) Frame 5 (c) Frame 6

Figure 17. Example result obtained using the spectral feature aided tracking method applied to airborne imagery.



(a) RGB image (b) Detected result of RX detection



(c) Detected result of change detection (d) Final result using combined method

Figure 18. Example result showing improved detection with combined method.

Section 2.5. Integrated Multi-Modal Sensing, Processing, and Exploitation

The research performed by collaborator Numerica, Inc., focused on the algorithmic aspects of performance-driven sensing. While the complete final report is included in the Appendix of this report, the main contributions are summarized in this section.

The main contribution of this work was in the development and testing through application to simulated data of a complete integrated tracking algorithm in the context of performance driven sensing. In particular, the work used the single-pixel tunable spectrometer concept discussed earlier in the report to allow the collection of selected spectral wavebands on a per-pixel basis for use in feature aided tracking. The algorithm was tested on simulated spectral data produced by RIT's DIRSIG image simulation tool.

Figure 19 shows the architecture of this adaptive tracking algorithm. Figure 20 shows an RGB of the DIRSIG- simulated scene used in the analysis.

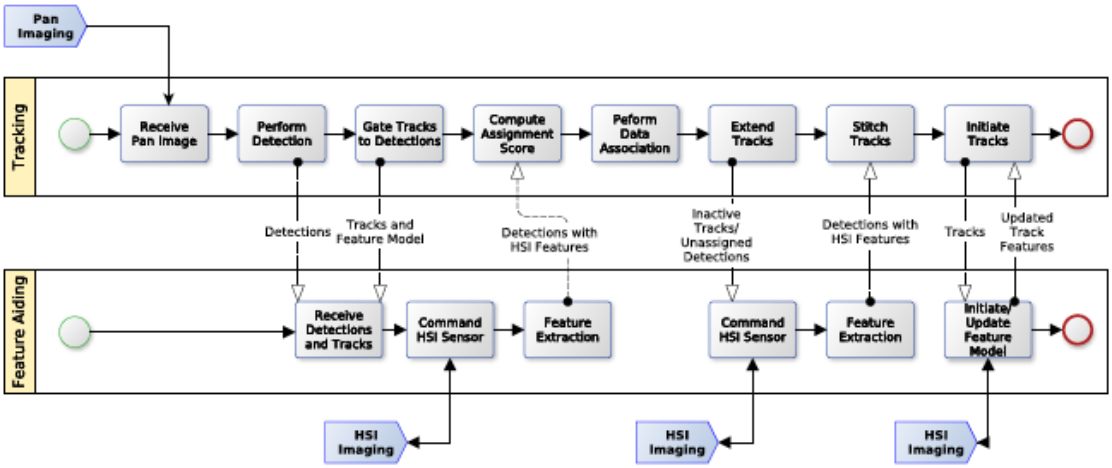


Figure 19. Adaptive track-based feature-aided tracking architecture using spectral feature information that is updated dynamically from local information.

Figure 20. RGB of simulated urban area used in the track performance analysis.

For comparison purposes, several tracking algorithms were investigated as listed below. The results of these algorithms are reported in Table 1.

Kinematic Only Tracking (KOT) – only motion information is used in the algorithm, no spectral features.

Track Derived Feature Only Tracking (TDFOT) – no motion information is used in this case, only the dynamically selected spectral features.

Track Derived Feature Aided Tracking (TDFAT) – this is the fully adaptive algorithm using both motion and selected spectral features.

Detection Derived Feature Aided Tracking (DDFAT) – this algorithm uses both motion and spectral features, but uses all spectral information available.

Track Derived Feature Aided Tracking and No Track Stitching (TDFAT_NTS) – this is similar to TDFAT but does not use track stiching.

Table 1 shows the results of applying these algorithms to 100 frames of simulated DIRSIG imagery.

Table 1. Track metrics for different algorithm configurations.

Metric	**KOT**	**TDFOT**	**TDFAT**	**DDFAT**	**TDFAT_NTS**
# Truth Tracks	47	47	47	47	47
# Tracks Initiated	59	56	56	56	95
Max. # Redundant Tracks	1	1	1	1	1
Max. # Missing Tracks	4	4	3	3	13
Total # Swaps	6	7	4	4	3
Mean # Tracks per Truth Object	1.38	1.32	1.28	1.28	2.09
# Truth Objects Tracked by Multiple Tracks	16	13	12	12	28

The results of Table 1 demonstrate that the fully adaptive TDFAT algorithm achieves the same performance as the DDFAT algorithm, which uses all spectral information. Figure 21 shows the reduction in data volume (percentage of cube interrogated) as well as theoretical increase in frame rate possible using the adaptive TDFAT algorithm. These results demonstrate the significant improvement in collection efficiency possible using the adaptive approach with no loss in tracking performance.

Full details on these algorithms are available in the Numerica report included in the Appendix.

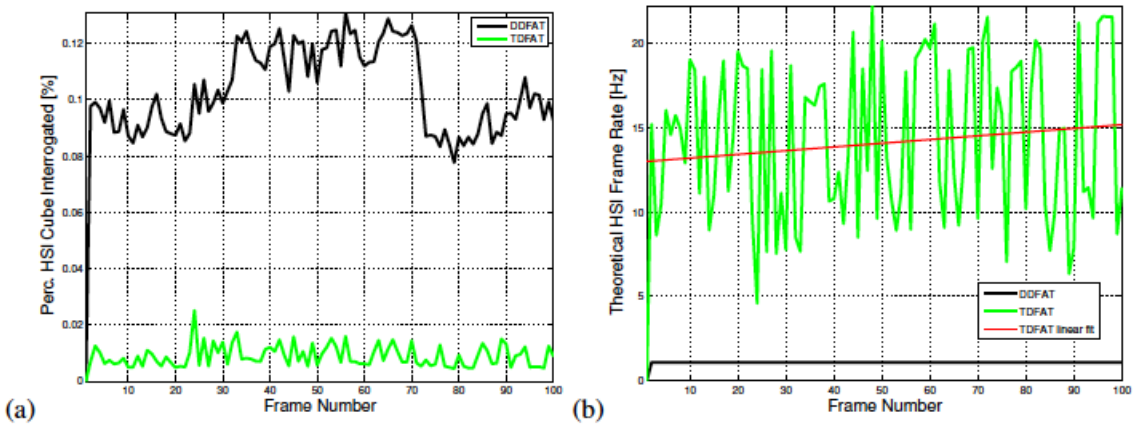


Figure 21. Bandwidth metrics for tracking algorithms: (a) Percentage of full HSI cube interrogated, and (b) HSI frame rate possible assuming the tunable spectrometer sensor concept.

Section 2.6. Hyperspectral Imaging of Pedestrians

The research tasks under this award were expanded to include investigation of the phenomenology associated with hyperspectral imaging of pedestrians (dismounts). Thus, in addition to the tracking of vehicles in a cluttered, urban environment, the use of spectral imagery to distinguish amongst pedestrians in a scene was investigated.

The research included field experiments measuring the spectral reflectance of the hair, skin and clothes of volunteers as well as the collection of hyperspectral imagery. Simulation using RIT's DIRSIG tool was also used to expand the range of conditions studied. Figure 22 shows an example comparison of the spectral reflectance measured of different clothing materials. Note the differences between 400 and 700 nm are primarily due to color while the differences between 700 and 2500 nm are primarily due to the material type.

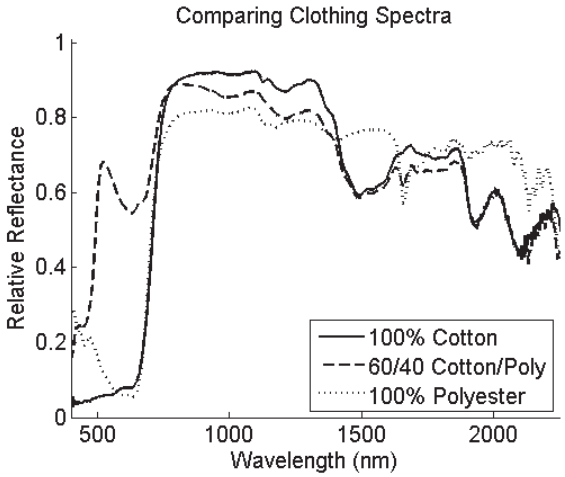


Figure 22. Comparison of spectral reflectance of different clothing materials.

Figure 23 shows an RGB of the scene imaged by the hyperspectral imager. Table 2 presents the empirically calculated probability of error using a spectral distance metric in a binary classification (one against all) using the mean spectrum of each subregion of each pedestrian as the prototype for the given class, for several spectral subsets.

Figure 23. RGB of the hyperspectral image collected for this study.

Table 2. Probability of error for one vs. all binary classification of pedestrians using subregions in hyperspectral image.

	Torso		Skin		T rousers		Hair	
Spectral Range	mean	σ	mean	σ	mean	σ	mean	σ
450 – 2250 nm	0.377	0.335	0.906	0.205	0.372	0.309	0.595	0.361
480, 550, 650 nm	0.788	0.310	0.969	0.097	0.792	0.346	0.875	0.244
450 – 700 nm	0.553	0.375	0.975	0.060	0.596	0.358	0.762	0.337
450 – 1000 nm	0.360	0.337	0.901	0.183	0.348	0.305	0.612	0.324
1000 – 1700 nm	0.832	0.314	0.949	0.142	0.787	0.344	0.792	0.339
1800 – 2250 nm	0.998	0.006	0.990	0.040	0.962	0.141	0.820	0.242

The results of Table 2 demonstrate that the clothing (Torso and Trousers) constitute the subregions that lead to the lowest error probability. They also demonstrate that the use of the VNIR region (400 – 1000 nm) only leads to similar performance as the entire full spectral range (400 – 2500 nm), suggesting the shortwave infrared may not be critical for this problem. However, further experiments should be done before making this a general conclusion.

Full details on this work can be found in the Ph.D. thesis published by Jared Herweg in 2012 listed in Section 3 of this report.

Section 3. List of Theses and Publications

The following theses and publications were supported at least in part under this research award.

Theses[1]

Herweg, Jared X., "Hyperspectral Imaging and Association Phenomenology of Pedestrians in a Cluttered Urban Environment," Ph.D. Thesis, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, 17 August 2012, *ProQuest Dissertations & Theses (PQDT)*.

Meng, Lingfei X., "Analytical Modeling, Performance Analysis, and Optimization of Polarimetric Imaging System," Ph.D. Thesis, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, 9 November 2012, *ProQuest Dissertations & Theses (PQDT)*.

Presnar, Michael D., "Modeling and Simulation of Adaptive Multimodal Optical Sensors for Target Tracking in the Visible to Near Infrared," Ph.D. Thesis, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, 4 August 2010, *ProQuest Dissertations & Theses (PQDT)*.

Rivas, Annette O., "Tunable Micro-electro Mechanical Fabry Perot Etalon," M.S. Thesis, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, 1 August 2011, *ProQuest Dissertations & Theses (PQDT)*.

Zhang, Tingfang, "Multiple-target Tracking using Spectropolarimetric Imagery," M.S Thesis, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, anticipated February 2013.

Journal Publications (peer reviewed)

Raisanen, A., J. Kerekes, A. Rivas, and S. O'Brien, "Thermally Driven Fabry-Perot Single-Pixel filter for Hyperspectral Imaging Applications," in preparation for submission to *SPIE Optical Engineering* anticipated in March 2013.

Meng, L. and J. Kerekes, "An Analytical Model for Polarimetric Imaging Systems," in preparation for submission to the *IEEE Transactions on Geoscience and Remote Sensing*, November, 2012.

[1] In addition to being available through ProQuest, electronic versions of these theses can be accessed through RIT's Digital Media Library at http://ritdml.rit.edu.

Herweg, J., J. Kerekes, and M. Eismann, “Separability Between Pedestrians in Hyperspectral Imagery,” submitted for review to *Applied Optics*, October, 2012.

Meng, L. and J.P. Kerekes, "Adaptive Target Detection with a Polarization-sensitive Optical System," *Applied Optics*, vol. 50, no. 13, pp. 1925-1932, 1 May 2011.

Raisanen, A., M. Presnar, Z. Ninkov, K. Fourspring, L. Meng, and J. Kerekes., “Simulation of Practical Single-pixel Wire-grid Polarizers for Superpixel Stokes Vector Imaging Arrays,” *Optical Engineering*, vol. 51, 016201, January 2012.

Conference Proceedings (peer reviewed)

Lu, Z., A. Rice, J. Vasquez, and J. Kerekes, "Target Discrimination Via Optimal Wavelength Band Selection with Synthetic Hyperspectral Imagery," *Proceedings of the Second IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 14-16 June 2010.

Presnar, M., J. Kerekes, and D. Pogorzala, "Dynamic Image Simulations for Adaptive Sensor Performance Predictions," *Proceedings of the Second IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 14-16 June 2010.

Rice, A., J. Vasquez, M. Mendenhall, and J. Kerekes, "Feature-Aided Tracking via Synthetic Hyperspectral Imagery," *Proceedings of the First IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 26-28 August 2009

Conference Proceedings (non-peer reviewed)

Fourspring, K., Z. Ninkov, and J. Kerekes, "Subpixel Scatter in Digital Micromirror Devices," *Proceedings of Emerging Digital Micromirror Device Based Systems and Applications II*, SPIE Vol. 7596, SPIE Photonics West Meeting, San Francisco, California, January 2010.

Fourspring, K., Z. Ninkov, and J. Kerekes, "Scattered Light in a DMD-based Multi-object Spectrometer," *Proceedings of Modern Technologies in Space- and Ground-based Telescopes and Instrumentation*, SPIE Vol. 7739, SPIE Astronomical Instrumentation, San Diego, California, June 2010.

Gadaleta, S., J. Kerekes, and K. Tarplee, "Autonomous Target Dependent Waveband Selection for Tracking in Performance-driven Hyperspectral Sensing," *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, SPIE Vol. 8390, 839023, April 2012.

Herweg, J., J. Kerekes, E. Ientilucci, and M. Eismann, "Spectral Variations in HSI Signatures of Thin Fabrics for Detecting and Tracking of Pedestrians," *Proceedings of Active and Passive Signatures II*, SPIE Vol. 8040, 8040G, SPIE Defense Sensing and Security Symposium, Orlando, Florida, April 2011.

Herweg, J., J. Kerekes, and M. Eismann, "Hyperspectral Measurements of Natural signatures: Pedestrians," *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, SPIE Vol. 8390, 83901C, April 2012.

Herweg, J., J. Kerekes, O. Weatherbee, D. Messinger, J. van Aardt, E. Ientilucci, Z. Ninkov, J. Faulring, N. Raqueno, and J. Meola, "SpecTIR Hyperspectral Airborne Rochester Experiment Data Collection Campaign," *Proceedings of Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, SPIE Vol. 8390, 839028, April 2012.

Herweg, J., J. Kerekes, M. Eismann, "Hyperspectral Imaging Phenomenology for the Detection and Tracking of Pedestrians," *Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 5482-5485, Munich, Germany, 2012.

Meng, L. and J. Kerekes, "Analytical Modeling of Optical Polarimetric Imaging Systems," *Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3998-4001, Vancouver, Canada, 2011.

Presnar, M. and J. Kerekes, "Modeling and Measurement of Optical Polarimetric Image Phenomenology in a Complex Urban Environment," *Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 4389-4392, Honolulu, HI, 2010.

Presnar, M., A. Raisanen, D. Pogorzala, J. Kerekes and A. Rice, "Dynamic Scene Generation, Multimodal Sensor Design, and Target Tracking Demonstration for Hyperspectral/Polarimetric Performance-Driven Sensing," *Proceedings of Polarization: Measurement, Analysis, and Remote Sensing IX*, SPIE Vol. 7672, 76720T, SPIE Defense Sensing and Security Symposium, Orlando, Florida, April 2010.

Rivas, A., J. Kerekes, and A. Raisanen, "Tunable Single Pixel MEMS Fabry-Perot Interferometer," *Applied Industrial Optics: Spectroscopy, Imaging and Metrology*, OSA Technical Digest, paper JWA15, July 2011.

Section 4. Summary and Suggestions for Future Work

This project had as its broad objective to conduct basic research in adaptive multimodal performance driven sensing. To this end, a number of contributions can be identified.

- An end-to-end simulation model was developed to demonstrate multimodal adaptive sensing in the context of vehicle tracking in an urban environment. This work integrated models and simulated images across the imaging chain and further identified performance characteristics of such as system.
- A feasible design was developed and prototyped for a MEMS single-pixel tunable Fabry-Perot spectrometer using a novel thermal activation technique. The design was verified through extensive model analysis of its electrical, mechanical, thermal, and optical properties.
- A novel analytical model was developed for polarimetric imaging systems and used as part of an adaptive target detection algorithm which can optimize the polarizer angles for maximum target-to-background contrast.
- A new multimodal target tracking algorithm was developed which combines spectral and polarimetric imagery to enhance target detection, followed by a novel approach to feature aided tracking using spectral information.
- A novel technique for spectral waveband selection was developed and used as part of a vehicle tracking demonstration of performance-driven sensing. This work quantified for this application the time advantages of selecting a subset of spectral information using the tunable single-pixel spectrometer concept.
- A database was developed of spectral reflectance measurements of humans and their clothing, and was used to assess the separability of pedestrians using hyperspectral images. Clothing was identified as being more robust than human skin or hair in distinguishing among pedestrians, and for these data, the visible through near infrared spectrum was found to be adequate for the task.

While these contributions have helped advance the basic science behind multimodal performance-driven sensing, they have in many ways just scratched the surface of understanding. One of the most promising technologies emerging from this research is the design for the single-pixel tunable Fabry-Perot spectrometer. It is recommended this device development continue as it offers the promise of a truly adaptive imaging spectrometer in a very compact configuration. Since the spectral data are captured over time in a tunable manner, it can reduce the volume of data collected by only sampling the spectrum as needed, rather than collecting a full spectrum everywhere. Also, the single-pixel nature allows a different spectral sampling to occur pixel-to-pixel, furthermore enhancing the efficiency of the data collection. Another aspect of this research that should be further pursued is the development of optimal algorithms for operating these adaptive sensor concepts. The initial efforts developed here have demonstrated the promise, but much more can be done. A third area is to explore hardware concepts to take advantage of the optimal polarizer angle algorithm to improve polarimetric target detection.

Appendix: Numerica, Inc. Final Report

Numerica Corporation
4850 Hahns Peak Drive, Suite 200
Loveland, CO 80538

Office: (970) 461-2000
Fax: (970) 461-2004
E-mail: info@numerica.us

www.numerica.us

FINAL REPORT:

DISCOVERY CHALLENGE THRUST FINAL REPORT: Integrated Multi-modal Sensing, Processing, and Exploitation

Contract Number: Subaward Agrement No.: 30788-01 under Prime Award No. FA9550-08-1-0028

Reporting Period: 02/01/2008 – 08/31/2012

Contract State Date: 1 February 2008

Contract Expiration Date: 31 August 2012

Government Sponsor: Air Force Office of Scientific Research

Technical Monitor:

Dr. John Kerekes
Associate Professor
Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology
54 Lomb Memorial Drive
Rochester, New York 14623 USA

Date: November 29, 2012

Prepared by:

Dr. Sabino M. Gadaleta
(970) 461-2000, ext. 206

Dr. Randy Paffenroth
(970) 461-2000, ext. 233

Table of Contents

List of Figures

List of Tables

1 Executive Summary

Persistent tracking of objects, i.e., the tracking of objects over long time in realistic environments, subject to occlusions and dense clutter, is an important problem for ground and aerial video surveillance. The use of multi-modal, hyperspectral, or polarimetric sensors for the problem of target tracking remains an active area of research and adaptive and tunable sensor concepts have emerged in order to address problems of detecting, tracking, and identifying targets in highly cluttered, dynamic scenes. These sensor concepts support performance-driven sensing that is a new concept that relies on sensing, processing, and exploiting only the most "decision-relevant" sets of target data for the purpose of reducing requirements on data recording, processing, and communications. This work presents tracking architectures in support of the goals of performance-driven sensing by providing sensor adaptation based on exploitation results. In this report we will describe two sensor designs that are capable of collecting hyperspectral data in a commanded manner on a per-pixel and per-band level and thus support hyperspectral data collection in performance-driven sensing systems. In this report we introduce three different tracking configurations: kinematic only, detection derived feature aided, and target derived feature aided. The latter configuration represents a fully adaptive tracking configuration where spectral feature data is managed and optimized on a per track level using local information. In support of the adaptive feature collection, we introduce a waveband selection algorithm that minimized the spectral wavebands to be collected per target. The report describes the integrated algorithm components and presents simulation results, using DIRSIG simulated data, that compare the different configurations on a simulated scenario that represent a typical urban scene. The simulation results and performance metrics demonstrate that the adaptive track-dependent feature tracking configuration realizes the goals of the performance-driven sensing paradigm: By only collecting a subset of features as needed, the data collection and bandwidth requirements are significantly reduced, while obtaining the same target tracking performance as obtained from feature aided tracking when using the full set of spectral features. We further motivate that by managing a small set of features, the update rate for feature data collected on individual targets is greatly increased which in turn should offer improved tracking and object identification performance. Further improvements are expected by applying techniques from the area of compressed sensing to extend the present work. To this end we review the area of compressed sensing as it relates to the area of hyperspectral imaging.

2 Introduction

2.1 Problem Identification

A fundamental problem in ground target tracking using airborne EO/IR sensors is maintaining a "persistent track" on targets, i.e., a track of long duration (more than a few minutes). Using panchromatic or color video data, traditional motion tracking methods may fail in typical urban settings when objects have similar sizes, shapes and colors, or when foreground object and background color are similar [1]. The feasibility of using *hyperspectral imaging sensors* for vehicle tracking has been demonstrated [2, 3] and the use of *multi-modal, hyperspectral, or polarimetric sensors* for the problem of target tracking remains an active area of research.

Adaptive and tunable multi-modal or hyperspectral sensor concepts have emerged in order to address problems of detecting, tracking, and identifying targets in highly cluttered, dynamic scenes [4]. These sensor concepts are investigated in the context of *performance-driven sensing*. Performance-driven sensing is a promising new concept that relies on sensing, processing, and exploiting only the most "decision-relevant" sets of target data for the purpose of reducing requirements on data recording, processing, and communications [5]. Our main interest in performance-driven sensing is the *sensor adaptation* based on exploitation results.

Rochester Institute of Technology (RIT) has developed different sensor designs for adaptive sensing including the RIT Multi-Object Spectrometer (RITMOS) [6] and a tunable micro-electro mechanical Fabry

Perot Etalon [7]. These sensors are notionally capable of providing imaging data with support for an adaptive per-pixel interrogation of full spectral and polarimetric target information. Due to various inherit communication bandwidth and data recording limitations, it is advantageous to carefully select the bands to capture per target in order to maintain a desired number of targets under track with a sufficiently high update rate.

This report presents tracking architectures in support of the goals of performance-driven sensing systems. In particular we present work that combines kinematic and spectral feature data for Feature-Aided Tracking (FAT) and present new results that demonstrate an active, performance-driven, management of the feature data on a per-track basis. This provides a systematic integration of dynamic feature selection algorithms based on waveband selection with a target tracking system for potential "in-the-loop" application with an adaptive sensor.

2.2 Related Prior Work

We briefly review prior related work on the topic of HSI and performance driven sensing. A bio-inspired multi-modal sensor design for efficient hyperspectral sensing for tracking moving targets was presented by Wang and Zhu [1]. An architecture and implementation regarding persistent, hyperspectral, adaptive, multi-modal, Feature Aided Tracking (FAT) within the urban context was presented by Rice *et al.* [8]. The paper formulated a utility function for Sensor Resource Management (SRM) including control of per-pixel hyperspectral sensing. Furthermore, a cost function for FAT with hyperspectral data within an MHT tracking system was presented. The work was extended towards an adaptive modification of track costs and tracking parameters [9]. Nguyen *et al.* [10] presented a framework that uses the mean shift algorithm to track objects in hyperspectral images. To reduce data dimensionality, the full Hyperspectral Imaging (HSI) spectrum is reduced using a random projection. HSI waveband selection algorithms were presented by Lu [11] and Nakariyakul [12]. In these works, waveband selection is applied globally to find bands that distinguish a given set of target signatures. More recently, Vodacek *et al.* [13] has extended the previous modeling approached by Rice towards modeling of a system for optical tracking in complex environments, with a focus on integrating an adaptive imaging sensor within the system framework. The approach builds upon the Dynamic Data Driven Applications Systems (DDDAS) paradigm. Gadaleta *et al.* [14] has presented an autonomous target-dependent waveband selection approach for performance-driven sensing with an adaptive hyperspectral imaging sensor and demonstrated it on a short tracking scenario. The work [14] considered target dependent dynamic waveband selection in which a set of HSI bands is selected for each target depending on local target and background spectrum and updated over time to address potential changes in the respective spectra.

2.3 Performance Driven Sensing and Dynamic Data Driven Applications Systems

To motivate the objectives of this work, stated in the following subsection, we include formal definitions of Performance Driven Sensing (PDS) and the related concept of Dynamic Data Driven Applications Systems (DDDAS).

Performance Driven Sensing (PDS) is a concept being actively pursued by the Air Force Office of Scientific Research (AFOSR) [5]:

> A very interesting and promising approach [...] is "performance-driven sensing," which relies on sensing, processing, and exploiting only the most "decision-relevant" sets of target data in order to reduce by orders-of-magnitude requirements on image data processing-throughput and communications bandwidth. The key to this approach is the ability to autonomously, dynamically, and in near-real-time select and process data from the most judicious sets of same-platform sensor pixels (spatial locations) and pixel photon modes (wavelength, polarization, and perhaps phase information), as well as the fusion and exploitation of data collected across multiple sensor platforms. It's a well known fact that the fusion and exploitation of optimum sets

> of multi-modal target spectra data can exponentially quicken target ID, dramatically improve ID fidelity, and reduce false alarms. Today, however, two capabilities essential for decision-relevant sensing don't yet exist: adaptive (pixel & mode tune or reconfigure) multimode-pixel (spatial, spectral, polarization, etc.) sensing capabilities, and autonomous data processing and exploitation algorithms for closed-loop sensor mode control.

Related to the PDS concept is the DDDAS framework that was originally funded through the National Science Foundation. The following definition of DDDAS is taken from [15]:

> DDDAS is a paradigm whereby application (or simulations) and measurements become a symbiotic feedback control system. DDDAS entails the ability to dynamically incorporate additional data into an executing application, and in reverse, the ability of an application to dynamically steer the measurement process. Such capabilities promise more accurate analysis and prediction, more precise controls, and more reliable outcomes. The ability of an application to control and guide the measurement process and determine when, where, and how it is best to gather additional data has itself the potential of enabling more effective measurement methodologies. Furthermore, the incorporation of dynamic inputs into an executing application invokes new system modalities and helps create application software systems that can more accurately describe real world, complex systems. This enables the development of applications that intelligently adapt to evolving conditions and that infer new knowledge in ways that are not predetermined by the initialization parameters and initial static data.

From the definitions above it is apparent that DDDAS and PDS are similar paradigms. Key to both concepts is to "close-the-loop" between the sensing process and the execution of the desired application such that the performance of the current task is optimally enhanced through the measurement process while minimizing system resource constraints such as computational resources or bandwidth needs. The purpose of this work is to develop a set of algorithms for target tracking that implement and demonstrate these concepts for a specific HSI application.

2.4 Objectives

The objective of this work are as follows:

- Develop, implement, and demonstrate a feature aided tracking system that combines kinematic and spectral feature costs for improved tracking performance.
- Systematically integrate a target dependent dynamic waveband selection algorithm with the target tracking system and an adaptive hyperspectral imaging sensor to manage a dynamically updated and minimal set of target dependent spectral features.
- Demonstrate the performance driven sensing concept within a vehicle tracking application using simulated HSI data.
- Present tracking performance metrics and demonstrate that the waveband selection approach does not degrade the tracking performance compared to using the full HSI feature data.
- Quantify the reduction in bandwidth and improvement in theoretical feature update rate obtained by using a target dependent waveband selection algorithm for managing the feature data to be collected in support of target tracking.

The resulting tracking system serves as a prototype PDS system for autonomous data processing and exploitation supporting closed-loop sensor mode control.

2.5 Organization of the Report

The report is organized as follows: Section 3 describes adaptive sensor designs that are the motivation for this work. The data set used for simulation and performance metrics used to quantify simulation results are descibed in Sec. 5. The HSI tracking system with several different architectures and it's algorithm components are described in Sec. 4. The algorithm for dynamic feature updating is described in Sec. 6.8. Performance simulation results are presented in Sec. 7. Section 8 surveys the area of compressed sensing as it applies to hyperspectral sensing applications and motivates it's application to extend the current work. Section 9 concludes the report and Sec. 10 provides recommendations for future work.

3 Adaptive Sensing Devices

Hyperspectral or polarimetric data can be a helpful feature to aid target tracking and surveillance applications [2, 10]. Many different devices exist that are capable of spectral data collection. Several require dispersing incoming light with a grating or prism. Other devices employ sensors that only detect light in a specific range. Another type of device used for spectrum collection is the interferometer. The current array designs collect the entire hyperspectral cube. At times this data is excessive in that only a small fraction of the data is actually needed in support of the particular surveillance task. Furthermore, collection is costly in bandwidth and memory [16] and the collection of a full set of spectral data may decrease the feature update rate compared to an approach that only collects the spectral subset that is most useful. Performance-driven sensing integrates sensing, processing, and exploiting to collect only the most decision-relevant sets of target data. To achieve these objectives, adaptive sensing devices that can work "in-the-loop" with signal processing algorithms are needed. Two examples of such sensor devices are described in the following two subsections.

3.1 RITMOS Sensor

An example of an adaptive sensor supporting performance-driven sensing is the Rochester Institute of Technology Multi-Object Spectrometer (RITMOS), shown in Fig. 1. This sensor utilizes a digital micro-mirror array at the focal plane of its fore-optics. Nominally, the micro-mirrors reflect light along an imaging path that produces a fully framed image (e.g., panchromatic or RGB). Individual micro-mirrors can be commanded to flip, directing that portion of the scene along a hyperspectral imaging path. A spectrometer and high resolution focal plane array measure the per-pixel hyperspectral signature of these pixels. Ideally, the dispersed pixels should not be allowed to overlap along their axis of dispersion; this manifests as a constrained resource optimization problem via micro-mirror tasking [4]. The RITMOS sensor does not allow simultaneous independent collection of HSI data accross the focal plane. This would require individual pixel capable of recording HSI data. Such a sensor design is described in Sec. 3.2.

It is interesting to note that the micro-mirror array in principle allows recording of ramdomized samples by using a randomized micro-mirror array control that randomly sets subimage mirror elements to on or off. In this manner, the RITMOS device effectively represents a compressive sampling sensor that can be used to record HSI image data similiar to the compressive HSI camera described in [17].

3.2 Tunable Micro-Electro Mechanical Fabry-Perot Etalon

As part of the AFOSR Discovery Challenge Thrust, the requirement to have a sensor that collects different types of information as commanded has been identified. To address this requirement, RIT has proposed a device that consists of an array of sensors that are individually tunable over the visible range and can be commanded to collect only the desired data. The potential exists in the design to overlay polarizers on individual pixels to collect polarized spectral data [7]. The device is a tunable MEMS Fabry-Perot Etalon

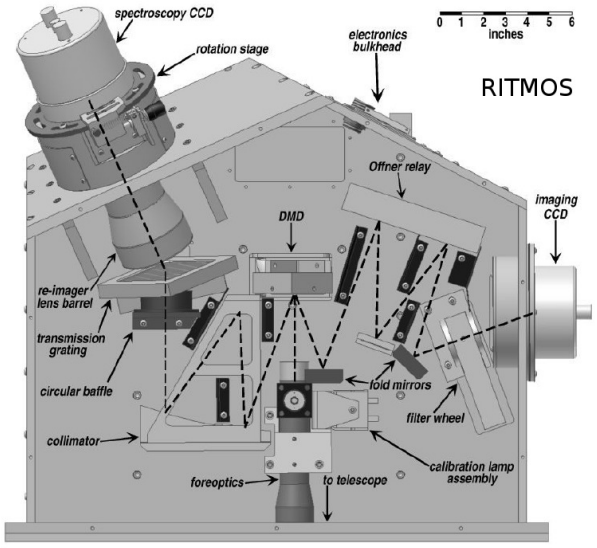


Figure 1: RITMOS sensor.

where, in the first design of [7], each 10 μm pixel is designed to scan in the visible range from 400 nm to 750 nm. The device design allows capturing of the individual spectrum of a target object on subregions of interest on command, as the situation dictates. Compared to a standard approach that collects the full HSI cube, this greatly decreases the total data collected and thus increases the ease with which the data is transferred, stored and manipulated, without loss of tactical capability [7]. Figure 2 provides an illustration of a Tunable Single Pixel (TSP) Fabry-Perot Etalon Interferometer. The sensor propsed in [7] combined an array of these pixel into a Focal Plane Array (FPA).

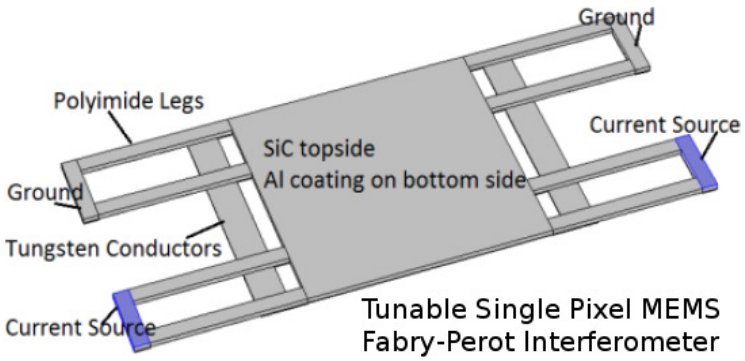


Figure 2: Individual pixel of a tunable Fabry-Perot Etalon sensor as shown in [16].

As discussed by Rivas [7], a Fabry Perot interferometer works by creating a resonant cavity for a specific wavelength of light as determined by the optical path length inside the cavity. The Fabry Perot is known for having the ability to select very narrow segments of the spectrum as determined by the precision of the physical geometry of the device and the order of the resonant mode selected. Most tunable MEMS designs use an electrostatic pull-in method to control the gap size. The device shown in Fig. 2 is thermally actuated and has been modeled [7] to show theoretical scan rates of approximately 66 different narrow wavelength bands per seconds per pixel. Furthermore, no significant thermal cross-talk between neighboring pixel was identified in the simulations which shows the theoretical feasibility of an FPA consisting of individually tunable pixel that can record high time-frequency spectral information for surveillance applications.

The following sections describe a tracking system that is designed to work "in-the-loop" with an adaptive HSI sensor such as the TSP device described above and motivate some of the benefits derived from such a sensor design when used for a vehicle tracking surveillance application.

4 Tracking Architectures for Adaptive Sensing

Throughout this report, we distinguish between three different tracking architectures. The first architecture represents a Kinematic-Only Tracking (KOT) architecture that uses only kinematic information for tracking, i.e., no spectral information. The other two tracking architectures exploit feature data and present two possible Feature Aided Tracking (FAT) architectures for adaptive sensing. In particular we distinguish between a Detection Derived FAT (DDFAT) architecture, where a full set of spectral data is recorded for subimages that cover the extent of each detection, and a Track Derived FAT (TDFAT) architecture that maintains a minimal set of features per track. These architectures are described in more detail below.

4.1 Kinematic Only Tracking

The first tracking architecture is illustrated in Fig. 3 and represents a kinematic image-based tracking architecture that relies only on recorded panchromatic imaging data. The main components of this tracking system are detection, gating, assignment score computation for solving the data association problem, track extension, track stitching, and track initiation. These individual components are common to all tracking architectures and will be discussed in more detail in Sec. 6. Since this tracking architecture uses only kinematic data we refer to it as Kinematic Only Tracking (KOT).

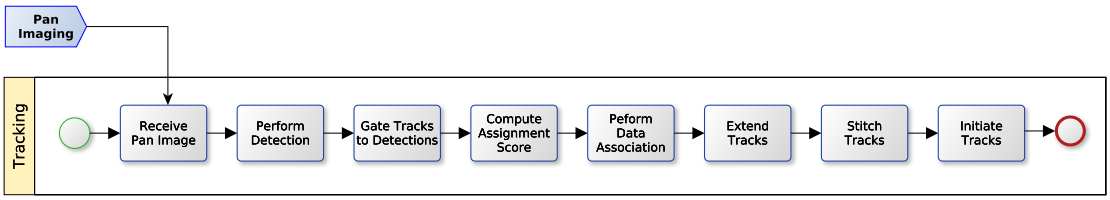


Figure 3: Kinematic tracking architecture for image-based tracking relying only on panchromatic imaging.

A main task of the tracking system is to perform data association. This relies on an assignment score that values the update of a particular track with a measurement. The assignment score that is needed to formulate this *assignment problem* between existing tracks and new detections can be based on pure kinematic information, i.e., track and detection position, or include feature components, if available. In the later case, i.e., if the panchromatic tracking system is aided by feature data, we refer to a *Feature-Aided Tracking (FAT) system*.

4.2 Detection Derived Feature Aided Tracking

A notional architecture for a tracking system that exploits an adaptive HSI sensor is shown in Fig. 4. In this architecture, a full set of spectral data is recorded for subimages that cover the extent of each detection. Since we only use detection data to control the sensor, we refer to this architecture as Detection Derived FAT (DDFAT).

An example showing the recorded subimages is illustrated in Fig. 5. The sensor records HSI data in the target region corresponding to detected pixel and includes a small background region to support waveband selection (not used in the DDFAT architecture but used in the TDFAT architecture discussed next). In the DDFAT architecture, we do not maintain a target-dependent feature set over time. Thus, the sensor is tasked to record the full HSI spectrum in the subimage region at each scan time.

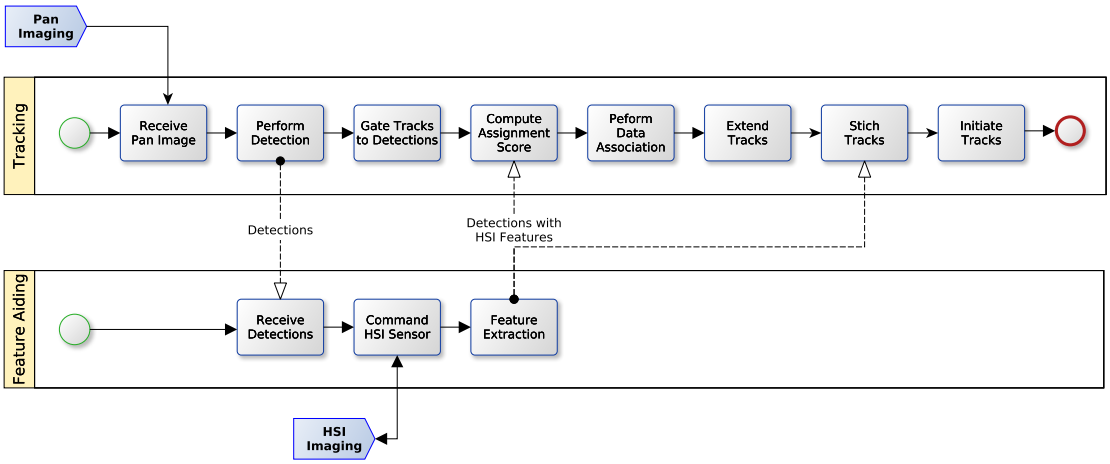


Figure 4: Adaptive detection-based feature-aided tracking architecture for feature-aided image-based tracking recording HSI feature data for each detection.

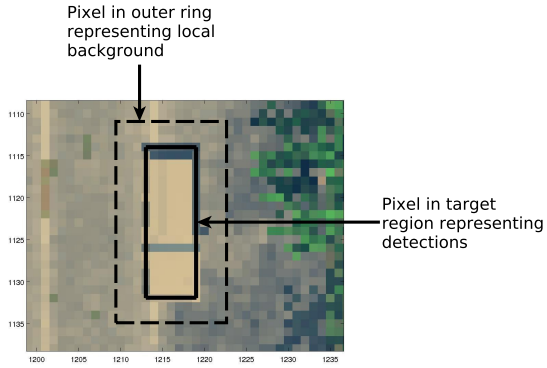


Figure 5: Illustration of HSI subimage recorded through adaptive control. The sensor records HSI data in the target region corresponding to detected pixel and includes a small background region to support waveband selection.

Feature extraction is used to derive spectral feature data from the HSI subimage that can be used in the assignment score computation. As discussed in Sec. 6.4, we use a mean spectral vector where the mean for each band is computed over all target pixels. Given the availability of target characteristic spectral feature data, we include the feature data in the track stitching component as shown in Fig. 4. Track stitching is a technique that reconnects broken track segments for the purpose of improving track continuity over time. In this work we set tracks that are not updated within a few frames to inactive. The purpose of track stitching in this work is avoiding track breakage due to temporary target occlusion. Before an unassigned detection (detection that could not extend any active track) initiates a new track, kinematic and feature data are used to attempt to reconnect it with another track segment. the track stitching component attempts to use kinematic and/or feature data to match those detections to inactive tracks in the neighborhood of the detection. Note that an unassigned detection will initiate a new track in this work. Thus, the track stitching component attempts to stitch new track initiation candidates to existing inactive tracks. The track stitching is discussed in more detail in Sec. 6.6. There are other techniques within a tracker that mitigate track breakage. For example, one could allow the tracker to coast through outages (longer than normal) when occlusions are expected.

4.3 Track Derived Feature Aided Tracking

The third tracking architecture is illustrated in Fig. 6. In this architecture, sensor control is based both on detection and track data. As above, HSI subimages are recorded in the neighborhood of detections. However, instead of recording a full spectrum, data is recorded only in a subset of bands that is derived from a target dependent feature set that is maintained for each track. This feature set is initially computed when initiating a new track using a *waveband selection* algorithm that is discussed in more detail in Sec. 6.8. In addition, the waveband set is updated over time under certain conditions, e.g., if a track assignment score starts degrading or if a new target appears in the vicinity of an established track that may lead to ambiguity. In those cases we re-apply a waveband selection process to dynamically update the feature data maintained for a track. To distinguish this third architecture from the second architecture we refer to this architecture as Track Derived FAT (TDFAT).

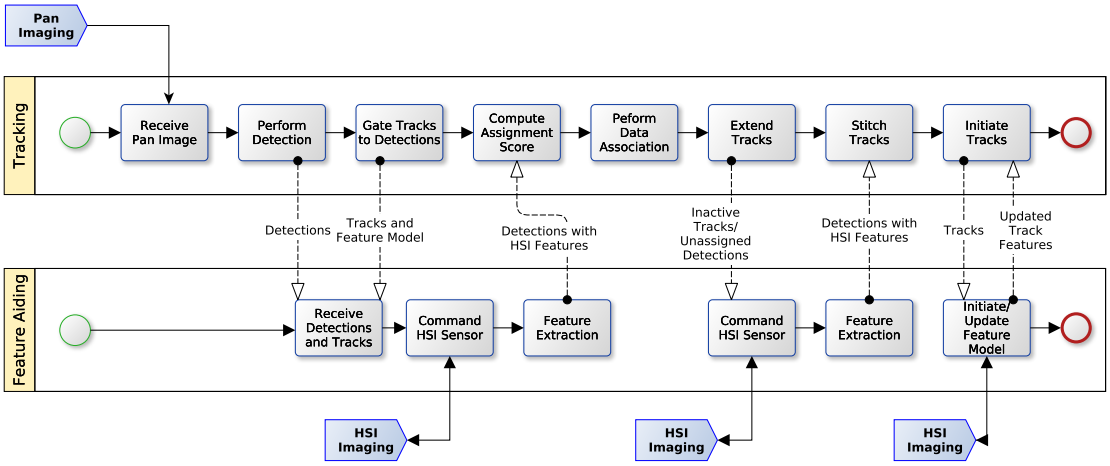


Figure 6: Adaptive track-based feature-aided tracking architecture recording HSI feature based on a track-based feature model that is updated dynamically from local information.

The waveband selection process is based on local information and has two main goals: (1) reduce the amount of spectral information that needs to be recorded, while (2) finding a feature set that improves tracker performance by emphasizing track identity purity. It is important to note that reducing the amount of spectral information that needs to be recorded for a target greatly improves real-time run-time performance not only because the reduction in required bandwidth but in particular because the practical frame rate for updating a target's feature data is greatly increased. Since the recording of a single band requires a certain amount of time, the scanning of a large set of bands for a target may take a significant amount of time making real-world application of HSI sensors for surveillance applications challenging. By reducing the set of bands, the feature set update rate is increased. To measure this performance criterion, we include in our performance metrics a measure of the HSI spectral feature update rate.

In previous work [18], *Context Aided Tracking* (CAT) was used to form background statistics from observed data to determine a mapping from classes of background objects, e.g., trees, roads, buildings. This background information was then used to adjust tracking parameters that affect the score. In this work we use a different approach and maintain target dependent features that are designed to improve target discrimination from local background and other nearby targets. The features are adapted over time as local background changes or different targets enter the vicinity of a target under track.

While hyperspectral imaging provides rich feature information about an object's material composition through light interaction in the bulk material, polarization imaging tends to provide additional information regarding the top surface characteristics [19] and it has been observed that man-made objects have a higher

degree of polarization [20]. For this project, Numerica did not have any polarimetric data set available and thus only used hyperspectral feature data. However, in the future, the developed algorithms could be used for polarimetric data by extending the feature vectors with the polarized data sets.

5 Datesets and Metrics for Performance Evaluation

This section describes the dataset used for simulation studies and the process of deriving detections and truth tracks for performance evaluation from the available data.

5.1 Dataset

A series of synthetic data frames, generated with the Digital Imaging and Remote Sensing Image Generation (DIRSIG) tool [21], was used to evaluate the performance of algorithms developed during this project. The DIRSIG tool has the ability to produce imagery in a variety of modalities, including multispectral, hyperspectral, polarimetric, and LIDAR in the visible through thermal infrared regions of the electromagnetic spectrum [4]. The base scene used was part of Tile #1 of DIRSIG Megascene #1. This scene represents a high-fidelity recreation of objects comprising a vast region of the Rochester, NY metro area [22]. The video data simulates a hyperspectral sensor operating at 10 Hz mounted to an airborne platform and oriented towards nadir. Platform motion has been excluded for simplicity, but is more generally resolved with registration techniques. Each frame was rendered as a hyperspectral cube of 61 bands synthesized from 0.4 to 1.0 μm at a resolution of 0.01 μm. This is representative of a realizable silicon-based visible-light MOS instrument.

The sensor array size is 880 x 560 pixels at 17 μm with 3 times spatial oversampling. Each image is sub-sampled in a 3x3 fashion, such that 9 independent spectral radiance values are computed and linearly mixed. This approximates the spectral mixing that is common to real HSI data. The resulting image is 2640 x 1680 pixels and the size of an individual HSI data cube is 2640 x 1680 x 61. The sensor platform operated at 3000 m altitude, stationary and nadir looking to provide an overall field-of-view of 660 x 420 m with a ground sample distance of 0.75 m [8]. Figure 7 shows different three-band combination images of the scene. Figure 7(a) shows the RGB Bands 26, 16, and 6 as a regular color image. The bands are scaled to maintain the relative image intensity within the selected band combination. Figure 7(b) shows the Bands 47, 39, and 32 displayed as a regular color image with Band 47 being used for the “Red” band, Band 39 for the “Blue” band, and Band 32 for the “Green” band.

(a) (b)

Figure 7: Part of DIRSIG Megascene #1 tile. (a) RGB Bands 26, 16, 6 displayed as a color image. (b) Bands 47, 39, 32 displayed as a color image.

5.1.1 Simulation of Noise

To simulate image noise in each spectral band, we add Gaussian white noise to the images with a standard deviation

$$\sigma_{\text{noise}}(b) = \frac{m(b)}{\text{SNR}},$$

where b denotes the spectral band, and $m(b)$ the mean pixel value for the respective band, averaged over all pixel in that band. Figure 8 shows part of the simulated scene as a color image constructed from the bands 26, 16, 6 with simulated noise using (a) SNR = 400, and (b) SNR = 10. We assessed performance for noise levels with SNR$\geq$ 10 in this work but did not notice any degradation in performance due to simulated noise for these levels. Thus, results presented in this report do not include a separate study for different noise levels.

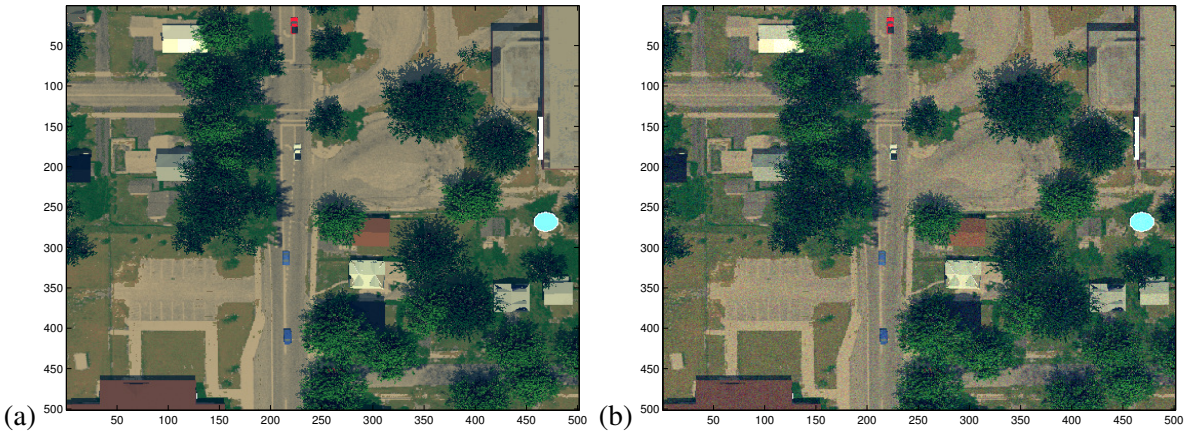


Figure 8: Part of the simulated scene. RGB Bands 26, 16, 6 displayed as a color image with simulated noise using (a) SNR = 400, and (b) SNR = 10.

5.1.2 Simulation of Vehicles

Different vehicles were simulated in DIRSIG with different vehicle paint models. Vehicle motion was introduced in the scene as described by Kerekes *et al.* [4]. Some of the different vehicles, used for performance simulations in this project, are shown in Fig. 9. On average around 40 vehicles are within the FOV of the sensor with vehicles entering and leaving the scene along the borders of the image.

5.2 Deriving Truth Detections from Data

To support performance metrics and simulation studies it is necessary to have *truth tracks* available for the targets of interest, i.e., vehicles. Truth tracks are correct data associations of vehicles over time. This data is necessary to assess the performance of *estimated tracks* obtained from the tracking system. The data available to Numerica for this study did not come with a predetermined set of truth tracks. However, the data available allowed identifying the pixels in an image that correspond to detections of vehicles. Note that the tracking algorithm under test is not using any truth information in its algorithms for data association. Truth is used only for purposes of metrics and to obtain detections passed into the tracking system. In addition, the available data allowed us to identify which truth detections correspond to detections of targets under trees and which to unobstructed targets. In this work we will use the full set of truth detections (unobstructed + under tree) to generate *truth tracks*. The HSI tracking system on the other hand is fed only by the unobstructed truth detections. Performance metrics assess the performance of the tracks from the

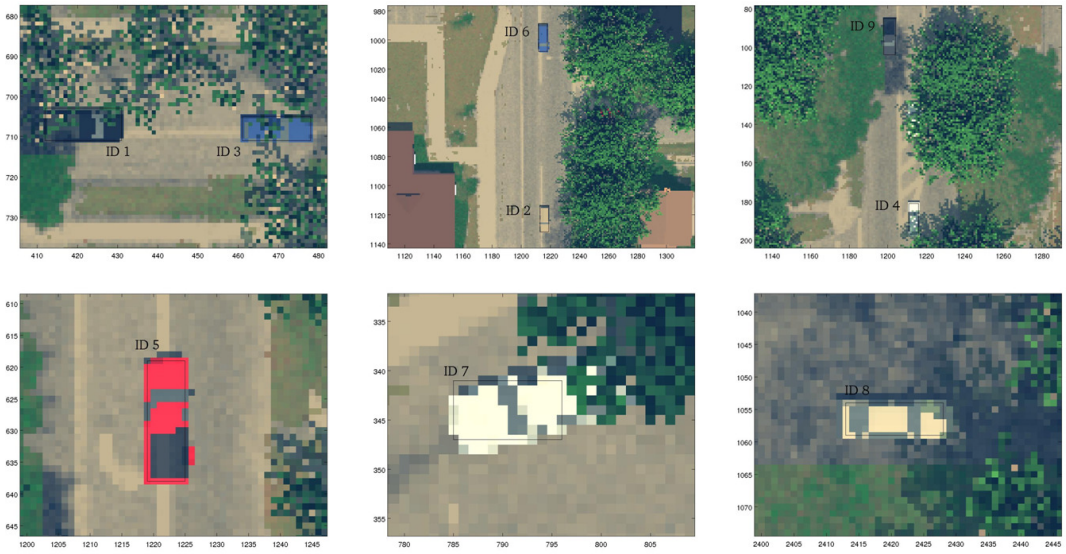


Figure 9: RGB images of nine vehicles simulated in DIRSIG.

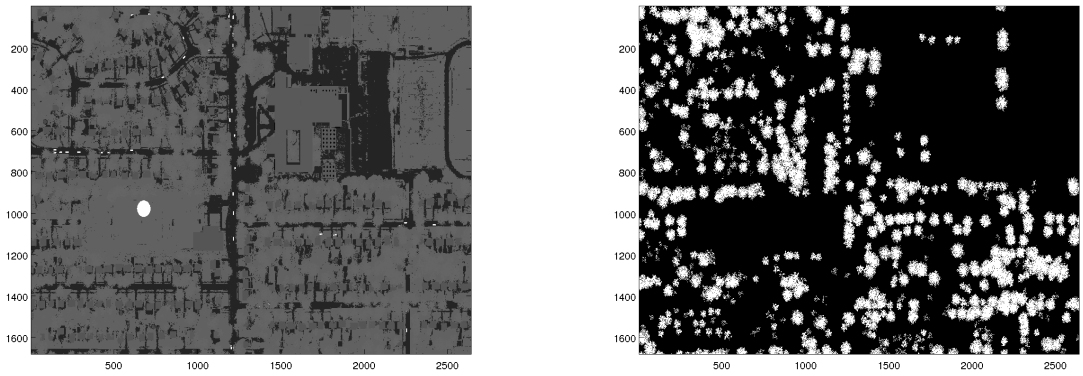
HSI tracking system against the truth tracks. The following describes how we derived these *truth detections* from the available DIRSIG data. The rationale for using the unobstructed truth detections (as opposed to detections derived from a regular detection algorithm) is to limit scope of this work to validation of the FAT technique and the demonstration of a dynamic feature updating algorithm. A full end-to-end solution would include an actual detection algorithm but this is out of the scope of this work.

DIRSIG generates each image into a dedicated folder entitled with a job number, e.g., `job_00700`. Each such folder contains two image files in the ENVI file format [23]. One file, called `afosr-t0000-c0000` contains the HSI image data of the simulated bands. The second file, called `afosr_truth` contains truth information about each pixel, such as material properties. This truth information is used to extract truth detections where each individual detection contains all pixel in an image generated from an individual vehicle. To identify vehicle targets as truth, we use the first band (First Material) and fourth band (First Opaque Mat.) of the truth image file. Figure 10(a) shows the first band for the `afosr_truth` image file of `job_00700`. Using thresholding for each of the two material images we obtain two image masks that mark all vehicle pixels. We combine the two masks and use a connected component analysis to group all neighboring pixels that belong to individual vehicles to generate a set of *truth detections* per HSI image file.

We distinguish between unobstructed truth detections and truth detections that are *under a tree*. To identify vehicle truth detections that are under a tree, we use the seventh band (First Trans. Mat. (Top ID)) and ninth band (First Trans. Mat. (Top Fraction)) to identify pixels with tree foliage. Figure 10(b) shows the seventh band for the `afosr_truth` image file of `job_00700`. As an example, Fig. 11 shows truth detections in an RGB image of `job_00700`. Detections that have pixel subsets of the vehicle under a tree are shown in red while completely unobstructed truth detections are shown in green.

5.3 Deriving Truth Tracks from Data

As noted above, the data available to Numerica for this study did not come with truth tracks. Thus, truth tracks were generated from the full set of truth detections by associating the truth detections produced across the entire sequence of frames. For association we use a spectral correlation procedure, which uses the full set of bands to compute a spectral distance measure between the signature vectors of detections between frames (see Sec. 6.4.1). To determine correlation candidates between tracks of a previous frame and detections in



(a) (b)

Figure 10: (a) Band 1 and (b) Band 4 of the `afosr_truth` image file for `job_00700`.

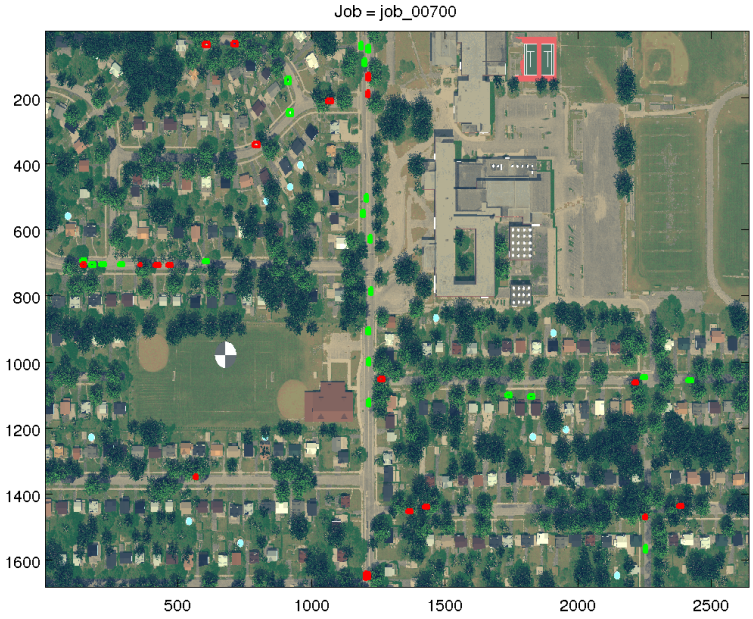
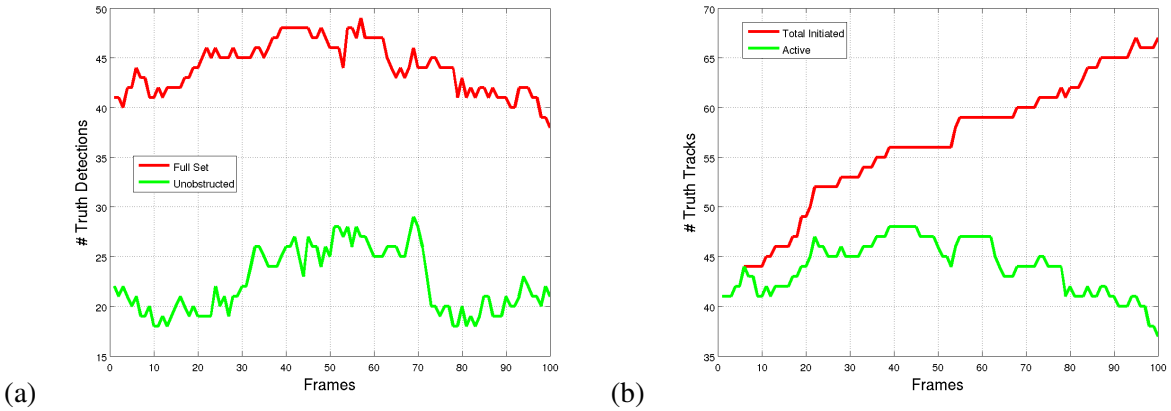


Figure 11: Truth detections in an RGB image of `job_00700`. Green: unobstructed truth detections; Red: Truth detections of vehicles that are under a tree.

the current frame, we use a gate region that is computed from the tracks in the previous frame. If multiple truth detections are feasible with a truth track, we correlate the detection to the track that corresponds to minimal spectral distance. In some rare cases, we found that multiple truth detections are generated on a single target (i.e., different detections for front and back of car). These cases are identified manually and the respective detections subsequently merged.

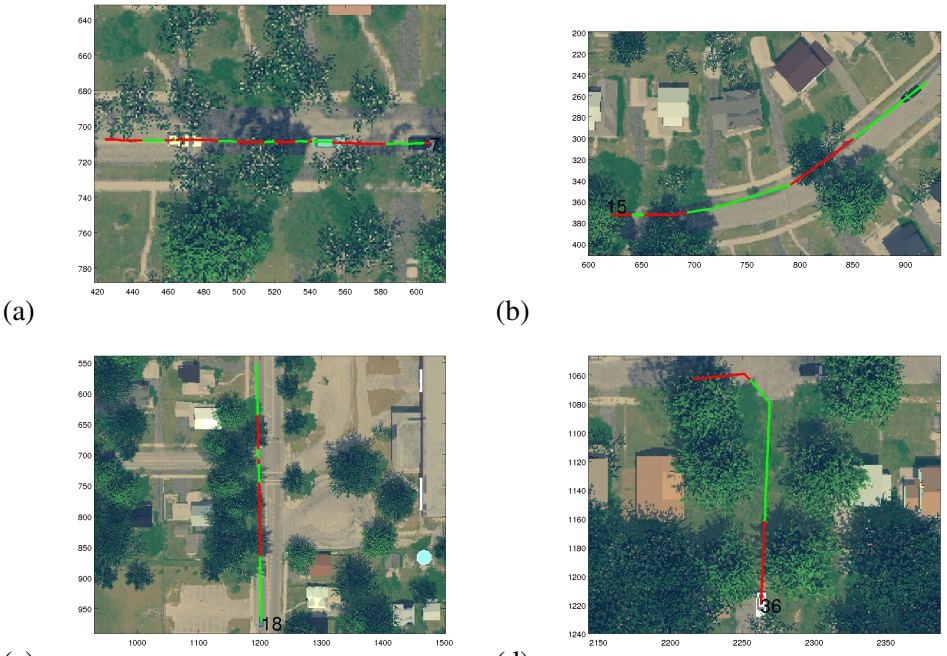
Figure 12(a) shows the number of truth detections for the first 100 images of the simulated scenario. Both the number of detections in the full set and the number of unobstructed detections are shown. We note that a significant number of the truth detections are under foliage and thus are not provided to the actual tracker during our simulations. Figure 12(b) shows the number of active truth tracks for the first 100 images of the simulated scenario where a track is counted as active if it was updated by a truth detection either at the current time or before or after the current time. Figure 12(b) also shows the total number of initiated truth tracks over scenario time. The difference increases as vehicles leave and enter the surveillance region.



(a) (b)

Figure 12: Number of (a) truth detections and (b) truth tracks over the first 100 images.

Figure 13 shows four different truth tracks maintained during the first 100 frames of the scenario. Shown in red are parts of the truth track where the truth detections are obscured by foliage. These detections are used to generate the truth tracks but not provided as input to the tracking system. Thus, one of the challenges for the tracking system is to maintain track through the temporary target occlusions.



(a) (b)

(c) (d)

Figure 13: Four different truth tracks maintained during the first 100 frames of the scenario. Red corresponds to portions of the truth track obscured by foliage.

5.4 Performance Metrics

Within an urban environment the main challenge is to maintain tracks with pure identity through ambiguity, e.g., merging targets, and severe occlusion. The failure to preserve a single, consistent identity for an object has negative consequences, e.g., in forensic applications or target engagement. Track identity loss occurs when the observations of an object no longer support the tracker's ability to maintain the track,

and it is deleted. Subsequently, the measurements resume and a new track with a new identity is formed. Track swaps occur when closely spaced objects meet and exchange track identities. Here, the chain of custody is polluted for both tracks. Track divergence occurs when the observations of an object no longer support tracking, yet the tracker does not delete the track quickly enough. Instead, the track coasts until it encounters the measurements of a different object, which it locks onto. Thus, *Track Continuity* is a critical metric that is of interest. It represents whether a track is maintained on an object over the time when it is observable to the sensor. Loss of continuity is a serious problem because then most of the historical information about a target is lost. This eliminates the opportunity using long-term behavior analytics to perform target identification or classification. These types of critical track errors are assessed in Sec. 5.4.1 through *Track Picture Metrics* [24].

In this project we are interested in an adaptive, track based, control of HSI sensor resource to acquire feature data as and when needed. The goal of this adaptive management is to minimize bandwidth and hardware resources to improve the potential run-time capability of HSI based tracking systems. To demonstrate the benefit of the adaptive feature-based tracking methods discussed in this report, we present *Bandwidth Metrics* in Sec. 5.4.2.

5.4.1 Track Picture Metrics

The process of generating performance metrics is illustrated in Figure 14. The HSI truth data is used to generate truth detections as discussed in Sec. 5.2. The truth detections are separated into unobstructed truth detections and truth detections that correspond to detections under foliage. The full set of detections is passed to a tracking process that generates truth tracks. In addition, the unobstructed truth detections are tagged with *truth IDs* that correspond to the ID of the truth track to which a respective truth detection was correlated in the truth tracking process. The unobstructed truth detections, tagged with truth ID, are passed to a tracking process that generates a set of *estimated tracks*. Performance metrics are computed from the set of estimated tracks and use detections' truth ID tags to compute track picture and track purity metrics over the time of the scenario. Note that the truth IDs are not available to the algorithm under test. They are just generated and passed through the tracking process for the purposes of computing metrics.

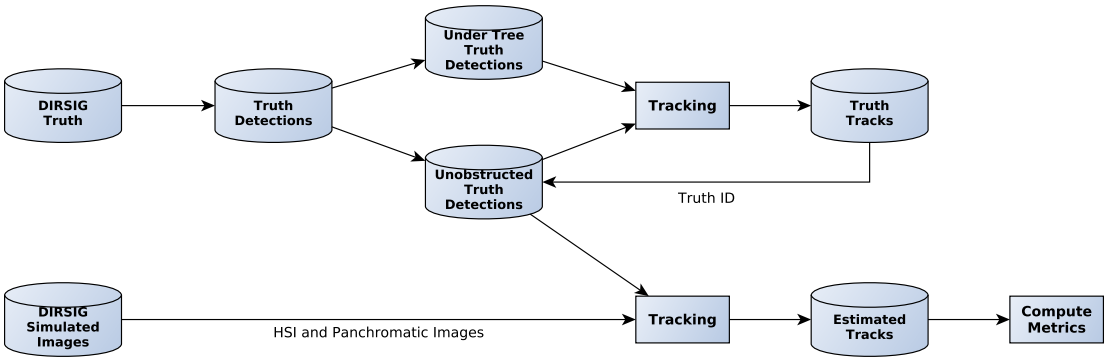


Figure 14: Process of generating performance metrics.

Let $\mathcal{D}$ denote the set of detections received over the time of a scenario. We assume that each detection is received with a truth ID. The number n_{truth} denotes the unique number of truth IDs on which detections were received over the scenario. This corresponds to the theoretical number of tracks the tracking system should generate. This process ensures that the tracking system is only assessed based on the data actually received. In particular, if truth tracks exist within the scenario time that were generated only from obstructed detections, the tracking system would not have received any data on this track and should not be penalized if a track on this target is not generated. Further, let $\mathcal{T}_{\text{est}}$ denote the set of estimated tracks generated from the

tracking system for the simulated scenario and denote with $n_{\text{est}} = |\mathcal{T}_{\text{est}}|$ the number of individual estimated tracks in this set.

A global scenario *track number* performance metric is obtained by comparing n_{truth} with n_{est}. This is also referred to as track completeness. Ideally we have $n_{\text{est}} = n_{\text{truth}}$. If $n_{\text{est}} < n_{\text{truth}}$ this indicates that the tracking system has initiated fewer tracks than truth. This can indicate that multiple vehicles are tracked by the same track. If $n_{\text{est}} > n_{\text{truth}}$ this indicates that the tracking system has initiated more tracks than truth (redundant tracks exist). This can indicate that tracks on a single vehicle are broken into multiple tracks.

More refined track picture metrics are computed in the form of *redundant* and *missing* track metrics. Sometimes these two metrics are combined into a single track completeness metric [24]. These metrics are computed for each simulated image over time and are computed using the *truth ID* that is attached to each detection used in the tracking system. For each received truth ID, we compute the first and last time that a detection with this truth ID is received. A truth track is considered active at a certain time t_i, if the tracking system has received an observation with the respective truth ID either at time t_i or before and after time t_i. Similarly, we consider an estimated track to be active at time t_i, if the tracking system has updated the respective track with an observation either at time t_i or before and after time t_i.

The number of missing tracks at time t_i is the maximum of zero and the difference between the active number of truth tracks at time t_i and the active number of estimated tracks at time t_i. The number of redundant tracks at time t_i is the maximum of zero and the difference between the active number of estimated tracks at time t_i and the active number of truth tracks at time t_i. Missing tracks at time t_i denote that fewer estimated tracks are active at this time than there are active truth tracks. Redundant tracks at time t_i denote that more estimated tracks are active at this time than there are active truth tracks. The ideal number for both of these metrics is zero.

Another important metric is the *cumulative number of swaps* metric. This metric assess the purity of a track. A track is considered pure if it only associates detections from the same object. We use the truth ID associated with associated detections to identify swaps. A swap occurs for a track at time t_i if the truth ID changes between the association made at time t_i and the previous correlation that was made to the track. The cumulative number of swaps metric counts the total number of swaps that occurred over the associations of all estimated tracks over the scenario time. The ideal number for this metric is zero.

Two additional track metrics are computed. The *mean number of tracks per target* metric checks if multiple tracks exist on the same truth target. For example, two tracks exist on a target, if the same truth ID appears in the associations of the two tracks. This indicates duplicate tracks on the same vehicle over time as would occur if a track is discontinued due to temporary target obstruction and a new track initiated on the target when the object reappears. The ideal number for this metric is one. The *number truth tracked by more than one track* metric is the total number of truth tracks for which more than one estimated track exists. This metric is similar to a traditional track breakage metric. The ideal number for this metric is zero.

5.4.2 Bandwidth Metrics

The bandwidth metrics we use assess to compare the different tracking architectures measure the amount of spectral information recorded over the scenario as the result of simulated sensor feedback. Specifically we use the following metrics to measure the cost of collecting spectral information in support of feature aided tracking:

- **Percentage of Full HSI Cube Interrogated.** This metric measures the percentage of the full HSI data cube interrogated at a particular scan time. In this work, a single image contains 2640 x 1680 pixel. The full HSI data cube recorded at a scan time consists then of 2640 x 1680 x 61 pixel scans. The percentage measures the percentage of samples recorded from the full cube at a particular scan time.

- **Theoretical HSI Frame Rate.** This metric assumes an adaptive HSI sensor that follows a theoretical design as described for the tunable MEMS Fabry-Perot Etalon sensor in Sec. 3.2. This sensor is in principle able to scan an individual wavelength of a pixel in approximately 15-16 ms resulting in a theoretical scan rate of approximately 66 bands per second. Thus, if a spectral approach requires recording of the full spectrum, its theoretical HSI frame rate is 1.1 Hz. On the other hand, if it requires scanning of only 3 bands, its theoretical HSI frame rate is approximately 22 Hz. The metric is computed at a scan time by averaging the number of bands scanned per track over the number of tracks and multiplying the result with 0.015 s.

6 Tracking System Components

The thrust of the present work was to develop and test the adaptive sensing technique. To properly test it, a tracking system is needed. Therefore, a tracking prototype (based on previous Numerica work) was implemented so that it could incorporate the adapative sensing technique. Note that the point of the work was not to research tracking techniques and the prototype tracking system does not represent a state-of-the-art tracking system. Future work can address incorporating the adaptive sensing techniques developed in this work with a state-of-the-art tracking system such as Numerica's Multiple Hypothesis Tracker (MHT). This section describes the different tracking components implemented in the prototype tracking system used in this work.

6.1 Detection

In this work we focus on the adaptive waveband selection of target features and not on the detection process. In the proposed architecture, target detection is performed using the panchromatic images that are assumed to be available at high update rates. Moving target detection in video is a well studied problem [25, 26]. The typical approaches use background estimation to separate moving "foreground" object pixels from the background through a background subtraction process. Given knowledge about background classes from, e.g., a GIS database or a background modeling process that uses HSI data [9], allows us to incorporate road network information to restrict detections to areas of interest. In this work we will use an idealized detection process that derives detections from available HSI truth data, as discussed in Sec. 5.2. This process generates a detection set restricted to vehicle detections and allows distinguishing between *unobstructed* and *under tree/foliage* detections. As illustrated in Fig. 14, in tracking simulations we will only use unobstructed detections as input to the tracking system. The full set of detections is used to generated truth tracks for the purpose of evaluating performance of the tracking system.

6.2 Data Association

We briefly describe the two-dimensional multi-assignment problem linking tracks with measurements. Let $\mathcal{I} = \{1, 2, \ldots, m\}$ enumerate a set of established tracks that need to be linked to a set of measurements enumerated by $\mathcal{J} = \{1, 2, \ldots, n\}$. Let the cost c_{ij} denote the cost for track $i \in \mathcal{I}$ associating with $j \in \mathcal{J}$. Let $\mathcal{A}$ denote the set of all feasible assignments between objects $i \in \mathcal{I}$ and $j \in \mathcal{J}$. Then, the two-dimensional multi-assignment problem can be expressed as [27]:

$$
\begin{aligned}
&\text{Minimize} \sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij} \\
&\text{Subject to (1)} \sum_{j\in\mathcal{A}(i,\cdot)} x_{ij} \leqslant a_i,\ (i = 1, \ldots, m),\ \mathcal{A}(i,\cdot) = \{j|(i,j) \in \mathcal{A}\}, \\
&\qquad\quad (2) \sum_{i\in\mathcal{A}(\cdot,j)} x_{ij} \leqslant b_j,\ (j = 1, \ldots, 1),\ \mathcal{A}(\cdot,j) = \{i|(i,j) \in \mathcal{A}\}, \\
&\qquad x_{ij} \in \{0, 1\}.
\end{aligned}
\tag{1}
$$

The subset $\mathcal{A}(i,\cdot) \subseteq \mathcal{A}$ enumerates all detections j that are feasible with a track i. Similar, the set $\mathcal{A}(\cdot,j) \subseteq \mathcal{A}$ enumerates all tracks i that are feasible with a detection j. The set $\mathcal{A}$ is derived through gating as discussed in Sec. 6.3. The inequality (1) limits track i to be assigned to at most a_i measurements. Using $a_i > 1$ is relevant to treat break-up of objects or resolution of objects from previously unresolved objects (i.e., unresolved convoy being resolved over time). Furthermore, it is relevant to the current work in principle since multiple detections may arise from a single object. Using $a_i \geq 1$ would allow for multiple detections to update a single track. The inequality (2) limits measurement j to be assigned to at most b_j tracks. This is relevant if measurements that were previously resolved become unresolved due to, e.g., change in sensor-to-target view. An example would be an resolved convoy of vehicles that may at times be unresolved. To limit scope in the current work, we restrict the assignment to single assignment, i.e., $a_i = b_j = 1$. The formulation for the cost coefficients is based on negative log likelihood ratios as discussed in Sec. 6.4.

This work does not use a full Multiple Hypothesis Tracking (MHT) [28] system for performance evaluation but restricts processing to single-frame tracking using the two-dimensional assignment formulation shown in Eqn. (1). Numerica has applied MHT tracking to the problem of tracking vehicles in HSI images in previous work [8, 9, 29]. This work focuses on testing the value of features for tracking and development of algorithms to manage the collection of features to support tracking. Thus, we use a Single Hypothesis Tracking (SHT) system since this provides a simpler framework for pursuing these objectives.

Uncertainty is a key concept in tracking of closely spaced objects. When using Multiple Hypothesis Tracking (MHT) one can exploit the likelihood of different hypotheses (k-best solutions) to estimate the association ambiguity and Numerica has developed this concept [30] and applied it to various tracking problems. In this work we use a simpler approach that bases ambiguity on gating and scoring: if multiple targets gate with an existing track and feasibly could update it (log-likelihood score for association is negative) then we say that the respective detections are *ambiguous* with the track. This is used in the procedure for dynamic waveband selection.

6.3 Gating

As above, let $\mathcal{I} = \{1, 2, \ldots, m\}$ enumerate a set of established tracks that need to be linked to a set of measurements enumerated by $\mathcal{J} = \{1, 2, \ldots, n\}$. Gating determines which subset of measurements $\mathcal{J}(i)$ are dynamically feasible with a track i using coarse dynamical target constraints such as maximum target velocity. Gating is essential for real-time performance of a tracking system to limit the number of correlation hypotheses that need to be scored and maintained over time. Given a track i at scan s, let $\boldsymbol{r}_i(s)$ and $\boldsymbol{v}_i(s)$ denote the estimated position and velocity of track i in the image, respectively. Further, let $\boldsymbol{r}_j(s')$ denote the (centroid) position of a measurement in scan s'. A coarse gating rule compares the track position predicted from the time $t(s)$ of scan s to the time $t(s')$ of scan s':

$$\boldsymbol{r}_i(s') = \boldsymbol{r}_i(s) + \boldsymbol{v}_i(s)\Delta t$$

against $\boldsymbol{r}_j(s')$. It is $\Delta t = t(s') - t(s)$. Let x, y denote the pixel coordinates of the position vector, i.e., $\boldsymbol{r} = [x,\ y]^T$. We say that track i gates with measurement j if both $|x_i(s') - x_j(s')| \leqslant g_x(\Delta t)$ and $|y_i(s') - y_j(s')| \leqslant g_y(\Delta t)$ are satisfied, where g_x, g_y are gating thresholds that depend on the time between measurements. In our work we use $g_x = g_y = 25 + \min(25\ \Delta t)$.

In this work we use gating to determine which targets are potentially ambiguous. In other words, the tracks that associate detections that are within the gate region of another track are considered ambiguous with this track for the purpose of waveband selection. Figure 15 shows two examples of ambiguous targets based on gating.

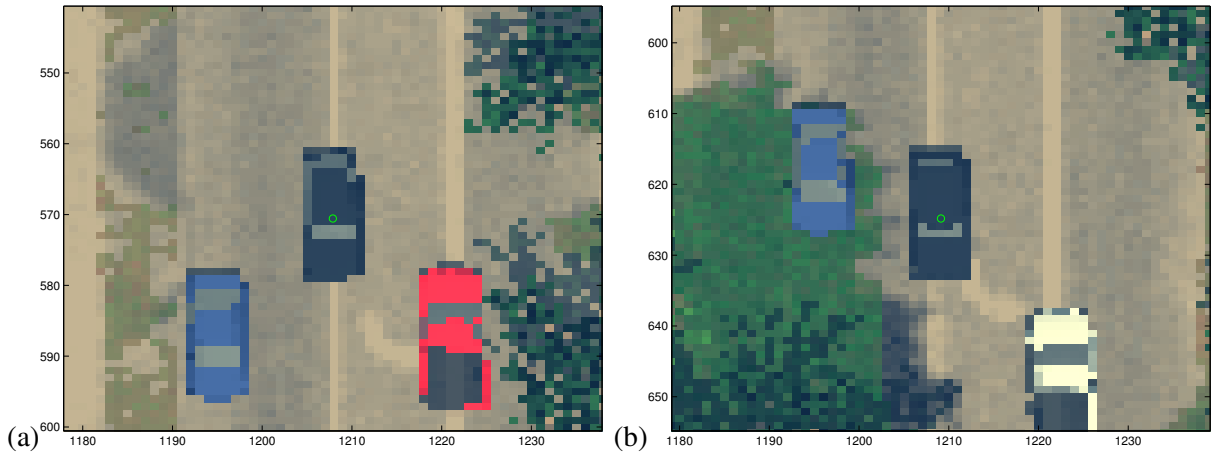


Figure 15: Gate regions for two sets of vehicles that are considered ambiguous with each other during the time periods where their tracks are within the respective gate regions.

6.4 Computation of Assignment Score

The assignment score for track formation is typically based on the definition of a Likelihood Ratio (LR). Given a dataset D this is defined as [28]:

$$\text{LR} = \frac{p(D|H_1)P_0(H_1)}{p(D|H_1)P_0(H_0)} := \frac{P_T}{P_F},$$

with true target hypothesis H_1 of probability P_T and false alarm hypothesis H_0 of probability P_F. Further, $p(D|H)$ denotes the probability density function evaluated with the received data under the assumption that H is correct and $P(H)$ denotes the a priori probability of H.

As discussed in [28], assuming the accuracy of the measurement process is independent of the target dynamics, the likelihood ratio can be partitioned into a product of two terms LR_{meas} and LR_{sig} that represent measurement and signal-related contributions, respectively. The measurement term can be based on a Mahalanobis distance between time-aligned track state estimate and measurement vector. The signal related term incorporates information about the sensor such as a detection probability and a false alarm probability. It is specific for the sensor making the measurement. Given S scans of data, and assuming scan-to-scan independence of the measurement errors, LR can be expressed as a product over the individual scan LRs:

$$\text{LR}(S) = \frac{P_0(H_1)}{P_0(H_0)} \prod_{s=1}^{S} \text{LR}_{\text{meas}}(s)\text{LR}_{\text{sig}}(s). \tag{2}$$

It is convenient to use a Negative Log Likelihood Ratio (NLLR) as a track score:

$$c(S) = -\ln(\text{LR}(S)) = -\sum_{s=1}^{S} \left(\text{NLLR}_{\text{meas}}(s) + \text{NLLR}_{\text{sig}}(s)\right) - \ln\left(P_0(H_1)/P_0(H_0)\right). \tag{3}$$

Now assume given a (kinematic) detection $\boldsymbol{z}_j$ that is obtained from a panchromatic image with corresponding spectral feature vector $\boldsymbol{f}_j$ that was obtained through tasking of an adaptive HSI sensor. One can form a *feature-augmented measurement vector* [31]:

$$\hat{\boldsymbol{z}}_j = [\boldsymbol{z}_j,\ \boldsymbol{f}_j]^T.$$

Further assume given a track i with associated kinematic target state estimate $\boldsymbol{x}_i$ and estimated spectral feature vector $\boldsymbol{\phi}_i$. Assuming that the kinematic and spectral measurement errors are independent, the generalized likelihood function that the augmented measurement $\hat{\boldsymbol{z}}_j$ is from target consistent with the estimated

of track i is given by [31]:

$$p(\boldsymbol{z}_j, \boldsymbol{f}_j|\boldsymbol{x}_i, \boldsymbol{\phi}_i) = p(\boldsymbol{z}_j|\boldsymbol{x}_i)p(\boldsymbol{f}_j|\boldsymbol{\phi}_i).$$

Thus, interpreting the measurement of Eqn. (3) as a joint kinematic and feature measurement, the NLLR track score Eqn. (3) can be written in the form

$$c(S) = -\sum_{s=1}^{S} \left(\text{NLLR}_{\text{kin}}(s) + \text{NLLR}_{\text{sig}}^{\text{kin}}(s)\right) + \left(\text{NLLR}_{\text{feat}}(s) + \text{NLLR}_{\text{sig}}^{\text{feat}}(s)\right) - \ln\left(P_0(H_1)/P_0(H_0)\right), \quad (4)$$

where now $\text{NLLR}_{\text{kin}}(s)$ is the NLLR for the pure kinematic measurement and $\text{NLLR}_{\text{feat}}(s)$ is the NLLR for the feature part associated with the measurement. We also broke the signal related part into two components since one may want to model the kinematic and feature sensing processes differently. The incremental track score (assignment score) for assigning track i to detection j at frame s can then be written as:

$$c_{ij} := \left(\text{NLLR}_{\text{kin}}(s) + \text{NLLR}_{\text{sig}}^{\text{kin}}(s)\right) + \left(\text{NLLR}_{\text{feat}}(s) + \text{NLLR}_{\text{sig}}^{\text{feat}}(s)\right).$$

In particular, the following form of the score can be derived [28, 31]:

$$c_{ij} = -\ln V_{\text{kin}} - \ln P_d^{\text{kin}} + \ln P_{fa}^{\text{kin}} - \ln p(\boldsymbol{z}_j|\boldsymbol{x}_i) - \ln V_{\text{feat}} - \ln P_d^{\text{feat}} + \ln P_{fa}^{\text{feat}} - \ln p(\boldsymbol{f}_j|\boldsymbol{\phi}_i), \quad (5)$$

where P_d, P_{fa} denote probabilities of detection and false alarm, respectively, and V_{kin} and V_{feat} are the surveillance volumes in kinematic and spectral measurement space, respectively. The above score function Eqn. (5) requires detailed modeling of all parameters and in practice one often uses a simplified form where all tuning parameters are lumped into a single parameter γ that then represents a tuning threshold determined from analyzing the scores observed for correct associations. We follow a similar approach and use the following score function in this work:

$$c_{ij} = w_{\text{kin}} L_{ij}^{\text{kin}} + w_{\text{feat}} \left(\frac{\chi_{\text{kin}}(n_{\text{kin}})}{\chi_{\text{feat}}(n_{\text{feat}})}\right) L_{ij}^{\text{feat}} - \gamma_1 \left(n_{\text{feat}} - n_{\text{kin}}\right) - \gamma_2, \quad (6)$$

where the χ represent the dimension dependent tail probabilities of a chi-square distribution for normal densities. We use the 99.9% confidence values, e.g., $\chi(2) = 13.82$ and $\chi(61) = 100.88$ [32]. The number of dimensions of the kinematic and feature spaces are denoted by n_{kin}, n_{feat}, respectively. Further, we used the notation $L_{ij}^{\text{kin}} = -\ln p(\boldsymbol{z}_j|\boldsymbol{x}_i)$ and $L_{ij}^{\text{feat}} = -\ln p(\boldsymbol{f}_j|\boldsymbol{\phi}_i)$. The w_{kin}, w_{feat} denote weight values providing different weighting of kinematic versus feature scores. We use $w_{\text{kin}} = 1, w_{\text{feat}} = 0$ for pure kinematic tracking, $w_{\text{kin}} = 0, w_{\text{feat}} = 1$ for pure feature tracking and $w_{\text{kin}} = 0.5, w_{\text{feat}} = 0.5$ for combined kinematic/feature tracking.

The kinematic NLLR score is given as [28]

$$L_{ij}^{\text{kin}} = 0.5 \left[\left(\boldsymbol{y}_{ij}^{\text{kin}}\right)^T \left(\mathbf{S}_{ij}^{\text{kin}}\right)^{-1} \boldsymbol{y}_{ij}^{\text{kin}} + n_{\text{kin}} \ln(2\pi) + \ln|\mathbf{S}_{ij}^{\text{kin}}|\right],$$

with residual vector $\boldsymbol{y}_{ij}^{\text{kin}} = \boldsymbol{x}_i - \boldsymbol{z}_j$ and measurement residual covariance $\mathbf{H}_{\text{kin}}\mathbf{P}_i^{\text{kin}}\mathbf{H}_{\text{kin}}^T + \mathbf{R}_j^{\text{kin}}$, where $\mathbf{H}_{\text{kin}}$ represents the transformation matrix from kinematic track state to kinematic measurement space. We use a simplified approach in this work with constant covariances $\mathbf{P}_i^{\text{kin}} = \mathbf{R}_j^{\text{kin}} = \mathbf{R}_{\text{kin}}$, where $\mathbf{R}^{\text{kin}}$ is modeled as a diagonal covariance matrix with standard deviation $\sigma_{\text{kin}} = 10$ pixel.

The feature NLLR score L_{ij}^{feat} is similar to L_{ij}^{kin}. As constant feature covariance matrix $\mathbf{R}^{\text{feat}}$ we use a fixed covariance matrix estimated from the truth tracking process with standard deviations for the diagonal elements of the matrix for the different bands as shown in Fig. 16. This figure shows $\sigma^{\text{feat}}(b)$, the square-root of the diagonal elements of the covariance $\mathbf{R}^{\text{feat}}$ as a function of band number b.

The above formulation Eqn. (6) is motivated by the fact that the kinematic and feature likelihoods are derived from different dimensional and metric spaces. Thus, it is necessary to transform the likelihoods to within similar mean values and confidence ranges before combining them into a single score function.

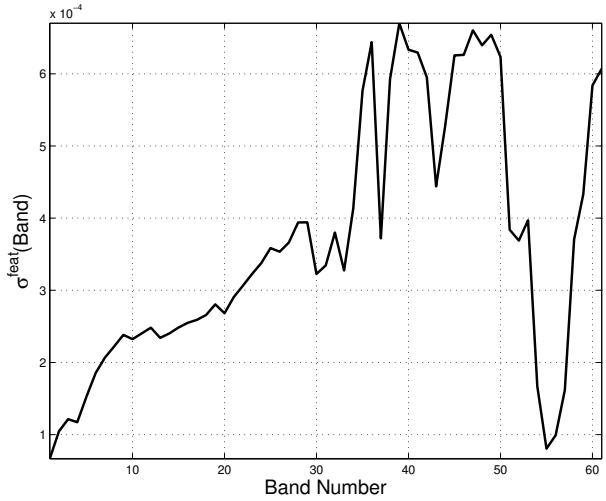


Figure 16: Square root of the diagonal elements of the feature covariance matrix estimated from truth tracking data.

The tuning value γ_1 is derived from simulation runs by comparing the kinematic scores of correct associations with the corresponding feature scores. In our work, we only use the position information as a kinematic feature, i.e., $n_{\text{kin}} = 2$. Figure 17(a) shows the distributions of (top) the kinematic scores obtained from a 30 frame scenario and (bottom) the feature scores using the full set of bands, i.e., $n_{\text{feat}} = 61$. The shifted feature scores are obtained using $\chi_{\text{kin}} = 13.82$, $\chi_{\text{feat}} = 100.88$, and $\gamma_1 = 0.22$. As can be seen from Fig. 17(b), the kinematic and shifted feature scores are within similar range.

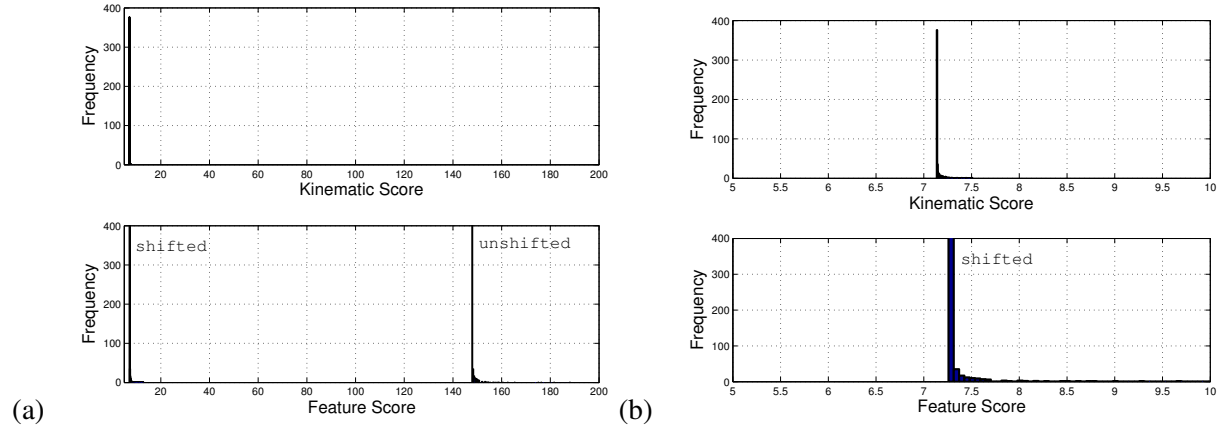


Figure 17: Distribution of kinematic, unshifted and shifted feature scores for correct associations in a 30 frame scenario using $n_{\text{feat}} = 61$. The shifted feature scores are obtained using $\chi_{\text{kin}} = 13.82$, $\chi_{\text{feat}} = 100.88$, and $\gamma_1 = 0.22$.

The tuning value γ_2 is derived from simulation runs by comparing the scores of correct associations with the scores of incorrect associations. Figure 18(c) shows the distributions of (top) the correct correlation scores obtained from a 30 frame scenario using $w_{\text{kin}} = w_{\text{feat}} = 0.5$ and (bottom) the scores for incorrect associations (for targets in a gate). Similar, Fig. 18(a),(b) shows the distributions for pure kinematic and pure feature scores. We note that $\gamma_2 > 10$ appears to be a good scoring value that will ensure negative scores for correct associations and positive values for incorrect associations. We say that a correlation ij with a score $c_{ij} < 0$ is feasible. In our simulations we use $\gamma_2 = 15$. The figures show that the distributions of incorrect

and correct associations in the gate overlap slightly. The reason for formulating the tracking problem as an assignment problem is to ensure that the correct associations are selected if multiple detections are feasible with a given track in a scan s.

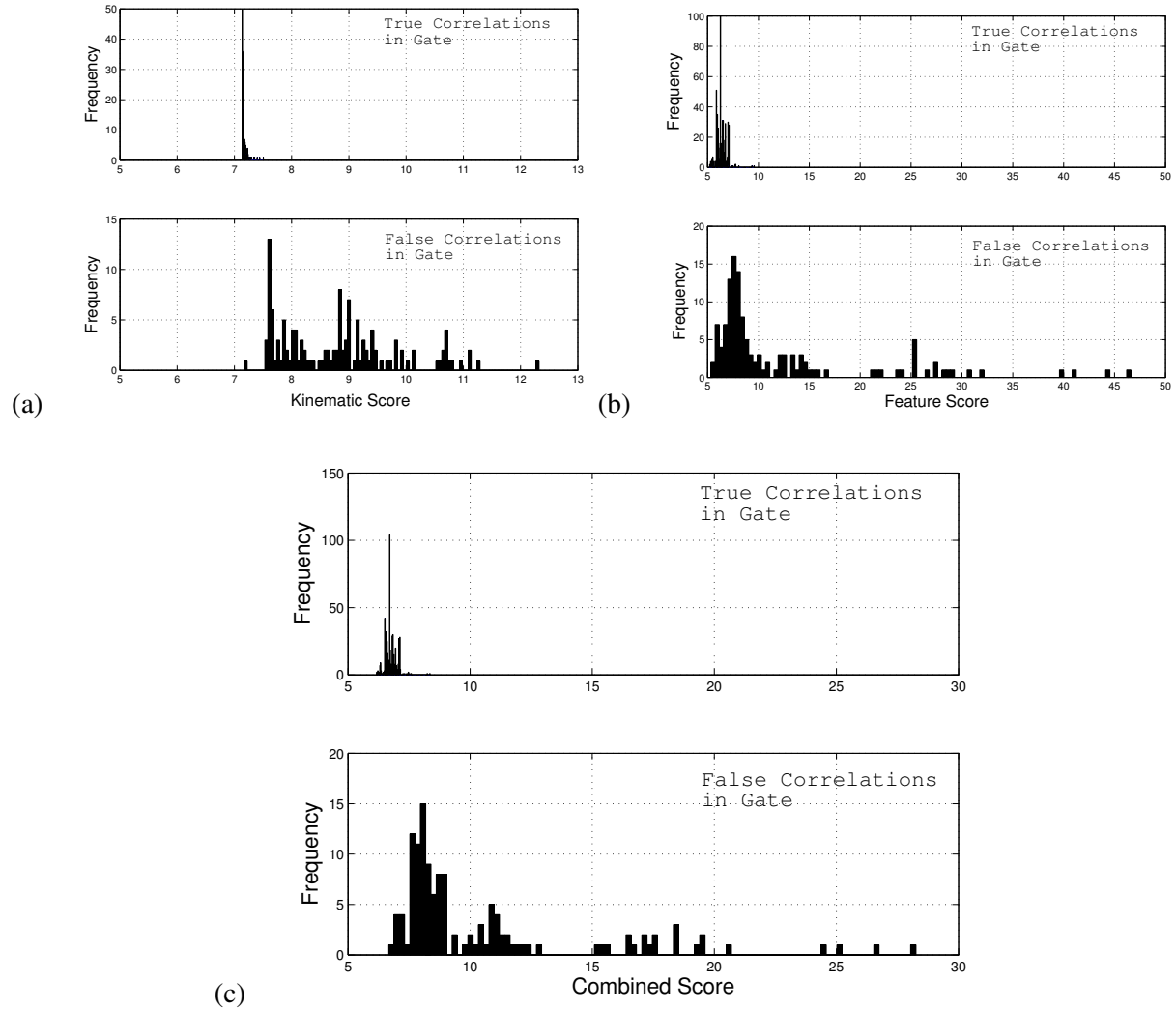


Figure 18: Distribution of (a) kinematic, (b) feature, and (c) combined scores for (top) correct associations and (bottom) incorrect associations in a gate in a 30 frame scenario.

6.4.1 Spectral Distance Measure

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ denote the spectral signature vectors for two different classes, e.g., truth object and detection. These signature vectors are computed by averaging the class signatures over all pixel in that class at each wavelength. We define the distance between the two vectors using the mixed measure of Spectral Angle Mapper (SAM) and Spectral Information Divergence (SID), as defined by Du *et al.* [33]:

$$d_{\text{feat}}(\boldsymbol{x}, \boldsymbol{y}) = \text{SID}(\boldsymbol{x}, \boldsymbol{y}) \cdot \tan\left(\text{SAM}(\boldsymbol{x}, \boldsymbol{y})\right)$$

with

$$\text{SAM}(\boldsymbol{x}, \boldsymbol{y}) = \text{acos}\frac{< \boldsymbol{x}, \boldsymbol{y} >}{||\boldsymbol{x}|| \cdot ||\boldsymbol{y}||} \text{ and } \text{SID}(\boldsymbol{x}, \boldsymbol{y}) = D(\boldsymbol{x}|\boldsymbol{y}) + D(\boldsymbol{y}|\boldsymbol{x}).$$

SID is defined through the cross entropies, or Kullback-Leibler divergences, $D(\boldsymbol{x}|\boldsymbol{y})$ and $D(\boldsymbol{y}|\boldsymbol{x})$ defined through

$$D(\boldsymbol{x}|\boldsymbol{y}) = \sum_{i=1}^{n_{\text{bands}}} p_i \log\left(\frac{p_i}{q_i}\right), \text{ with } p_i = \frac{x_i}{\sum_{j=1}^{n_{\text{bands}}} x_j}, \; q_i = \frac{y_i}{\sum_{j=1}^{n_{\text{bands}}} y_j},$$

where n_{bands} is the number of bands in the spectrum.

As noted in [11], the above distance measure is not defined for the case of $n_{\text{bands}} = 1$ and the Fisher Discriminant Ratio (FDR) [34] is proposed to be used in this case. The FDR for a single band is defined as

$$\text{FDR} = \frac{\mu_1 - \mu_2}{\sigma_1^2 + \sigma_2^2},$$

where μ_1, μ_2, σ_1^2, and σ_2^2 are the mean and variances of pixel values from corresponding classes in the respective band.

The distance measures need to be extended if more than two classes are considered. In this work, this case arises in waveband selection, where we will need a distance measure that measures the distance of a target spectrum $\boldsymbol{x}$ to multiple spectra classes $\boldsymbol{y}_s$, where the first spectrum $\boldsymbol{y}_1$ corresponds to the local background surrounding the targets and the remaining spectra correspond to other objects that are potentially ambiguous with the target detection. A multi-class distance measure is defined in this work as [11]:

$$d_{\text{avg}} = \left(\frac{1}{M}\sum_{j=1}^{M} d_j^{-1}\right)^{-1},$$

where M is the number of class pairs that can be constructed between the target spectrum $\boldsymbol{x}$ and the spectra $\boldsymbol{y}_s$. Thus, for waveband selection we form only pairs which contain the spectrum $\boldsymbol{x}$ and one of the spectra $\boldsymbol{y}_s$.

6.5 Track Extension

If the indicator assignment $x_{ij} = 1$ is in the solution of the data association problem Eqn. (1) for a track i and a detection j, then the track i is "extended" with the detection j. That is, the track state of track i is updated with the measurement corresponding to detection j. In this work we use an α, β filter for state estimation [35]. This filter is a simplified filter for estimation that is related to Kalman filters and does not require a detailed system model.

Given a track i at scan s, let $\boldsymbol{r}_i(s)$ and $\boldsymbol{v}_i(s)$ denote the estimated position and velocity of track i in the image, respectively. Further, let $\boldsymbol{r}_j(s')$ denote the (centroid) position of a measurement in scan s'. As for gating, we first need to compute the predicted track position at time of scan s':

$$\hat{\boldsymbol{r}}_i(s') = \boldsymbol{r}_i(s) + \boldsymbol{v}_i(s)\Delta t,$$

with $\Delta t = t(s') - t(s)$. The prediction error, or residual, is then given by $\boldsymbol{r}_j(s') - \hat{\boldsymbol{r}}_i(s')$. This is the residual used to compute the kinematic part of the association score. The updated track position becomes:

$$\boldsymbol{r}_i(s') = \hat{\boldsymbol{r}}_i(s') + \alpha_r\left(\boldsymbol{r}_j(s') - \hat{\boldsymbol{r}}_i(s')\right).$$

An analogous update is used for the velocity component where we assume that the predicted velocity remains constant: $\hat{\boldsymbol{v}}_i(s') = \boldsymbol{v}_i(s)$. A similar procedure is used to update the spectral feature state $\boldsymbol{f}_i$, where we also assume that the predicted feature state remains constant: $\hat{\boldsymbol{f}}_i(s') = \boldsymbol{f}_i(s)$. The measured spectral feature state is the feature state that corresponds to the detection. In our configuration we use $\alpha_r = 0.5$, $\alpha_v = 0.2$, $\alpha_f = 0.2$.

6.6 Track Stitching

If $x_{ij} = 0$ is in the solution of the data association problem Eqn. (1) for a track i and a detection j, then the detection j is not associating to any existing active track. Before initiating a new track with this detection, we check this track initiation candidate against inactive tracks in the neighborhood. To this end, we gate the detection j with the set of inactive tracks using a slightly larger gate region than for a regular gate. If inactive tracks are found in the vicinity of the detection j, we compute assignment scores between the detection j and these tracks. If a feasible solution exists to this assignment problem, than the respective inactive track is updated, i.e., *stitched*, with the detection j and set from inactive to active. As the performance results demonstrate, track stitching is an essential component in an urban tracking system to address frequent target occlusions.

6.7 Track Initiation

Detections that do not extend an active track or stitch to an inactive tracks initiate new tracks. When initiating a new track, the initial position and feature state is derived from the detection. The initial velocity is set to zero.

6.8 Waveband Selection and Dynamic Target Feature Update

When initiating a new track, a track feature is initially computed through a waveband selection algorithm, as discussed in Sec. 6.8.1. The purpose of waveband selection is to reduce bandwidth and data collection constraints by maintaining a minimal feature set sufficient to improve tracking performance. In a dynamic environment, it is necessary to update the feature set over time as the local environment changes. Figure 19 shows the dynamic target feature updating process. Given an active track, the target feature is updated for new tracks and for tracks with a degrading track score, once the track score exceeds a threshold. In this case, a full HSI subimage is recorded in the track neighborhood and in the neighborhood of any ambiguous tracks that are in the vicinity of the respective track. From the subimages, a target signature vector, local background signature vector and ambiguous target signature vectors are computed as input to the waveband selection algorithm. The waveband selection algorithm, as discussed in Sec. 6.8.1, computes a small set of m_{max} bands from which the updated target feature model signature vector is computed.

Instead of using the full set of wavebands, one can use a smaller set of wavebands over which to perform waveband selection. In previous work, we used a set of *candidate wavebands* [14] that is estimated for a track based in local target and background data. This set of candidate wavebands is selected as local maximum between target and background signature as shown in Fig. 20 and estimates for a track once at track initiation. In feature update computations, instead of collecting the full HSI spectrum, the spectrum is only collected over the candidate wavebands.

6.8.1 Waveband Selection Algorithm

In this section we present the waveband selection process that was implemented for this program. The waveband selection algorithm attempts to find a subset of bands that provides best separability between the target signature and the n other signature vectors (local background plus ambiguous tracks). The objective is to find a minimal spectral feature set that minimizes the amount of data to be collected and transmitted while allowing the tracking system to maintain track on the target. Assume given a target signature vector $\boldsymbol{x}$, derived from the detected target pixels, and a local background signature vector $\boldsymbol{y}_1 := \boldsymbol{y}_{\text{bgr}}$, derived from a spectrum recorded for local background pixels of the target. In addition, assume given a set of $n-1$ signature vectors $\boldsymbol{y}_2$ through $\boldsymbol{y}_n$ that are derived from detections associated to potentially ambiguous tracks in the vicinity of the target of interest.

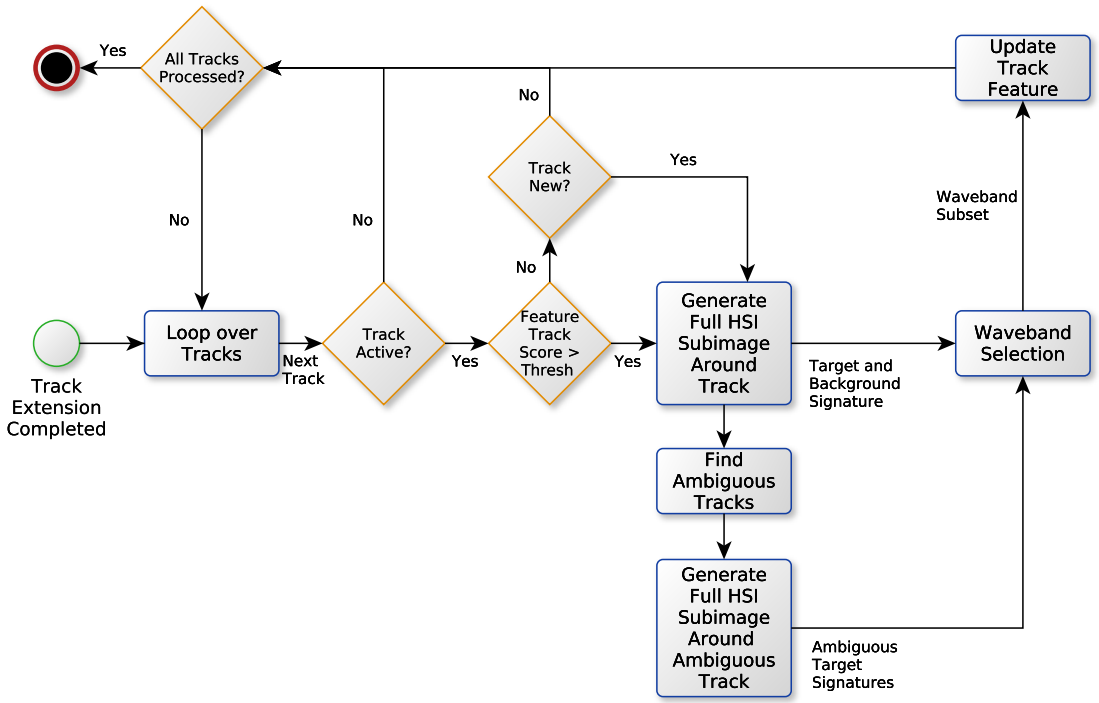


Figure 19: Illustration of target feature update process.

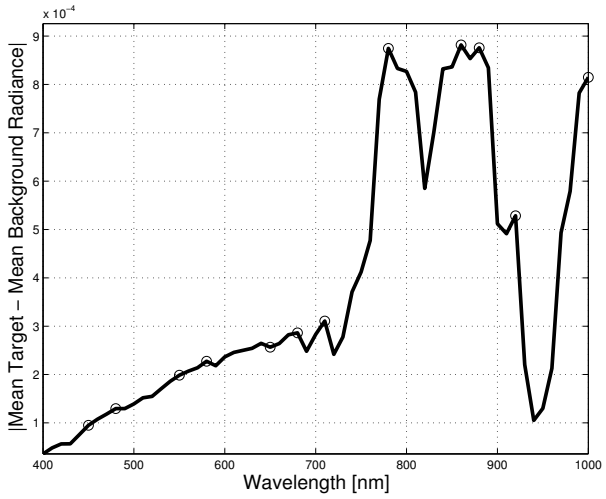


Figure 20: Candidate wavebands selected as local maximum between target and background signature.

As described by Lu [11], we use Sequential Forward Floating Search (SFFS) for waveband selection. The SFFS algorithm [36] is a specific type of feature selection algorithm that is an optimization technique that, given a set of m features, attempts to select a subset of size n that leads to the maximization of some criterion function. For the search algorithm optimization measure we use the multi-class distance measure d_{avg} defined in Eqn. (6.4.1). To compute this average distance we consider only pairs that contain the elements of the target signature vector $\boldsymbol{x}$ as one component in the pair. The other component is derived from the elements of a signature vector $\boldsymbol{y}_i \in \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$. Thus, the average distance measure is computed over n pairs. If no ambiguous track is found with a target track, then the waveband selection considers only a

single pair consisting of the target signature and the target local background signature.

The following describes the steps of the waveband selection algorithm, where $m_{\max}$ denotes the maximum number of desired wavebands. This number is a pre-defined parameter.

1. Find waveband b_1 that maximizes the multi-class FDR.
2. Find best two-band combination $\{b_1, b_2\}$, that includes b_1 from the previous step. The best two-band combination maximizes the multi-class d_{feat} between the target signature and the n other signature vectors.
3. Given the best $(m_{\max} - 1)$-band combination that includes the bands $\{b_1, b_2, \ldots, b_{m_{\max}-1}\}$ found in previous steps, find the best $m_{\max}$-band combination that maximizes the multi-class d_{feat}.
4. Repeat the last step by finding the best $m_{\max}$-band combination that includes the $m_{\max} - 1$ bands $\{b_2, \ldots, b_{m_{\max}}\}$ from previous steps. This step sometimes can find a better alternative to b_1.

Note in the above that the algorithm is finding bands by *maximizing* the average multi-class distance metric. This is because we are attempting to find a band combination that enhances the "contrast" or separability between target and background/ambiguous objects.

7 Performance Results

In this section we show performance results with the adaptive target dependent feature tracking algorithm over a 100 frame scenario (collected at a 10 Hz rate). This scenario represents a 10 s surveillance of the simulated scene. This short scenario is representative for the general dynamics that occur in the simulated scene with vehicles entering and leaving the scene throughout the scenario. In addition, this scenario length contains vehicles that traverse almost half of the image in some regions of the image while passing through temporary occlusions throughout the scenario. While the required track lifetime in EO/IR urban surveillance applications must be longer than 10 s, the chosen scenario shows the typical scenario characteristics of the available image data.

During the simulation, the tracking system is fed with "unobstructed" detections, i.e., no detections are received for vehicles in the case that a vehicle pixel is under a tree. As we had discussed in Sec. 5, this represents approximately half of the full set of detections.

7.1 Tested Algorithm Configurations

Simulation results are provided for five different configurations of the tracking system:

1. **Kinematic Only Tracking (KOT)**: Use $w_{\text{kin}} = 1.0$ and $w_{\text{feat}} = 0.0$ for both track extension and track stitching. In this case, spectral feature data is not used in the tracking system.
2. **Track Derived Feature Only Tracking (TDFOT)**: Use $w_{\text{kin}} = 0.0$ and $w_{\text{feat}} = 1.0$ for both track extension and track stitching. In this case, only spectral feature data is used in the tracking system for computation of assignment scores used in track extension and track stitching. Spectral features are dynamically selected through waveband selection by selecting three bands ($m_{\text{max}} = 3$) using the waveband selection algorithm described in Sec. 6.8.
3. **Track Derived Feature Aided Tracking (TDFAT)**: Use $w_{\text{kin}} = 0.5$ and $w_{\text{feat}} = 0.5$ for both track extension and track stitching. In this case, both kinematic and spectral feature data is used in the tracking system for computation of assignment scores used in track extension and track stitching. Spectral features are dynamically selected through waveband selection by selecting three bands using the waveband selection algorithm described in Sec. 6.8.

4. **Detection Derived Feature Aided Tracking (DDFAT)**: Use $w_{\text{kin}} = 0.5$ and $w_{\text{feat}} = 0.5$ for both track extension and track stitching. In this case, both kinematic and spectral feature data is used in the tracking system for computation of assignment scores used in track extension and track stitching. As spectral features, the full set of wavebands is used for each target, as shown in Fig. 5.

5. **Track Derived Feature Aided Tracking and No Track Stitching (TDFAT_NTS)**: Similar to TDFAT but not performing any track stitching.

We show track picture and bandwidth metrics for these different algorithm configurations in the following two subsections.

7.2 Track Picture Metrics

Table 1 shows the track picture metrics for the different algorithm configurations discussed in Sec. 7.1. In the scenario, detections on a total of 47 truth targets were received. In all cases, more tracks were initiated than truth targets exist, indicating that tracks were broken due to obstruction of view since the simulation provides only detections on unobstructed targets. However, comparing the results of tracking without track stitching (TDFAT_NTS) with the results when using track stitching shows that the track stitching greatly improves results by reducing the number of initiated tracks by almost a factor of two. Comparing the results of using kinematic only tracking (KOT) with the results when using dynamic feature updating with track derived feature aided tracking (TDFOT, TDFAT) we note that the feature tracking provides slightly better performance. In particular, best performance is obtained when using the combined approach that uses both kinematic and feature data (TDFAT). Comparing the results obtained when using the track derived feature data (TDFAT) with the approach that uses the full set of wavebands (DDFAT), we observe identical results. Thus, the dynamic approach does not degrade performance while, as we will see in the next subsection, greatly reducing the required bandwidth and improving the theoretical feature update rate.

Table 1: Track picture metrics over 100 frame scenario for different algorithm configurations.

Metric	**KOT**	**TDFOT**	**TDFAT**	**DDFAT**	**TDFAT_NTS**
# Truth Tracks	47	47	47	47	47
# Track Initiated	59	56	56	56	95
Max. # Redundant Tracks	1	1	1	1	1
Max. # Missing Tracks	4	4	3	3	13
Total # Swaps	6	7	4	4	3
Mean # Tracks per Truth Object	1.38	1.32	1.28	1.28	2.09
# Truth Objects Tracked by Multiple Tracks	16	13	12	12	28

Figure 21 shows the tracks for three different tracked vehicles. All shown tracks represent pure tracks. The figure shows as green dots the positions of detections that updated the respective track. Stitched lines connect consecutive updates of a track. In all three cases we see that tracks are maintained through significant periods of obstruction.

Figure 22 shows the tracks for two different tracked vehicles where dynamic feature updating is performed. Green dots in the figure indicate track updates and dashed lines connect consecutive updates. A red star indicates a time when a dynamic feature update is performed. In Fig. 22(a) we notice that the dynamic feature updating is performed only in the area of the image where the local background changes its characteristic. For the track in Fig. 22(b) dynamic feature updating is performed only three times, once shortly after initiation and two times after track stitching when local background characteristics have changed. Throughout the scenario, dynamic feature updating was only performed a total of 27 times for the complete set of tracks.

Figure 21: Tracks for three different tracked vehicles. Green dots indicate track updates. Dashed lines connect consecutive updates.

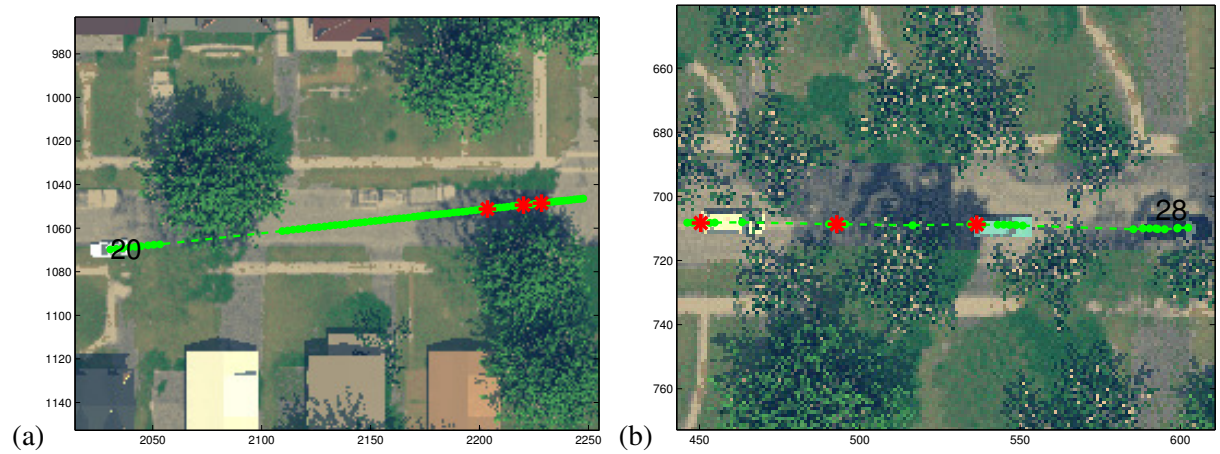


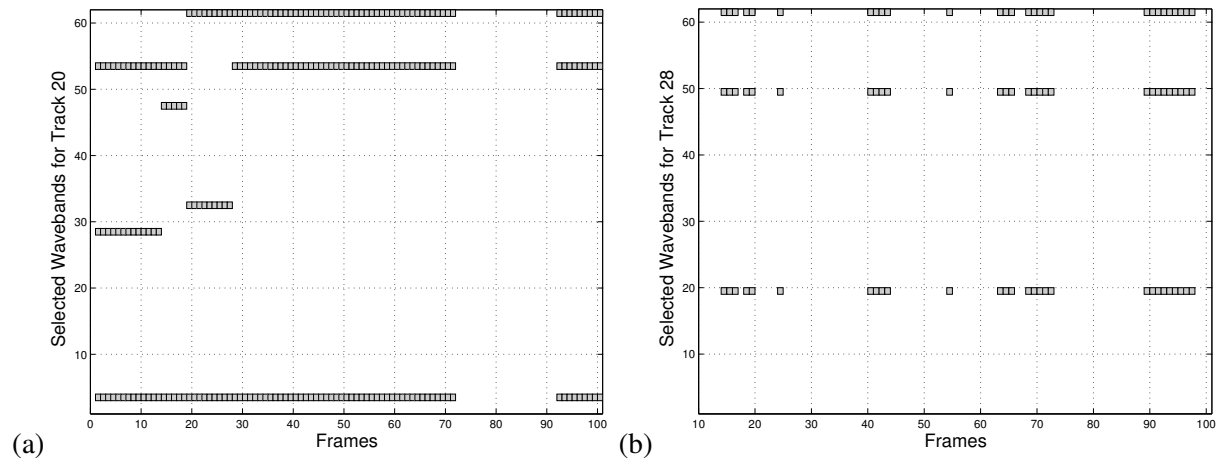
Figure 22: Tracks for two different tracked vehicles where dynamic feature updating is performed. Green dots indicate track updates. Dashed lines connect consecutive updates. A red star indicates a time when a dynamic feature update is performed.

Figure 23 shows the wavebands used for feature aided tracking for the two tracks shown in Fig. 22. In the first case, new wavebands are selected for track 20 when dynamic feature updating is performed. In the second case, although the algorithm is run to check if a better set of wavebands exists, no new wavebands are selected throughout the scenario for track 28. Note that a different set of wavebands is maintained for each track.

7.3 Bandwidth Metrics

This section presents bandwidth metrics that compare the amount of recorded HSI data in support of feature aided tracking. We compare the two tracking system configurations DDFAT and TDFAT. In the DDFAT architecture, no adaptive track-based management of the target features is performed. Instead, the full HSI cube is collected in a foveal vision type manner by only collecting the data in those areas where it is needed for target tracking. The DDFAT architecture itself thus requires interaction with a controllable HSI sensor but it does not maintain a dynamic subset of bands that form a target-dependent spectral feature set. In the TDFAT architecture on the other hand, when initiating a new track, a target dependent feature set is determined using a waveband selection algorithm. This spectral subset is updated over time as local conditions change in a way that may degrade the tracking performance.

Figure 24 shows bandwidth metrics for the two different tracking configurations. In Fig. 24(a) we show the percentage of the full HSI cube interrogated metric (see Sec. 5.4.2) as a function of frame number. We note that the TDFAT algorithm, using $m_{\max} = 3$, records significantly fewer points of the HSI cube. The

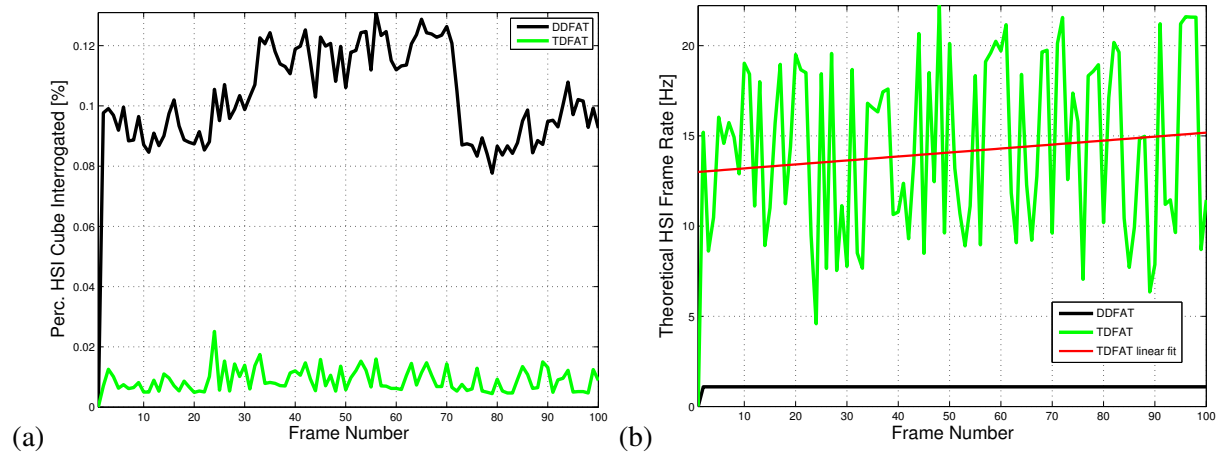


(a) (b)

Figure 23: Selected wavebands for two different tracked vehicles where dynamic feature updating is performed.

TDFAT algorithm requires recording a significant portion of the spectrum (the candidate wavebands) only at track initiation and when updating the feature data set. In Fig. 24(b) we show the theoretical HSI frame rate metric. As discussed in Sec. 5.4.2, this metric measures the average frame rate of recorded spectral data assuming a sensor that can record a single band per pixel in approximately 15 ms. Due to new track initiations, this metric fluctuates. Thus, we also show a linear fit to the data to identify the trend. For the DDFAT configuration, this frame rate is approximately 1 Hz since a total of 61 bands needs to be recorded per target. For the TDFAT configuration, the frame rate increases significantly. The theoretical frame rate for the TDFAT configuration fluctuates since more spectral data needs to be recorded when new tracks are initiated.

Note that a higher theoretical frame rate is better since it allows more frequent collection of feature data to aid the tracking and ID processes. This will lead to improved track continuity in challenging surveillance scenarios and reduced times needed for target identification.



(a) (b)

Figure 24: Bandwidth metrics: (a) Percentage of full HSI cube interrogated, and (b) HSI frame rate per target under TSP assumptions.

In summary, the performance results shown in this section demonstrated (1) the value of feature aided tracking on a relevant scenario, and (2) the benefit of performance driven adaptive sensing to reduce data collection requirements and to improve theoretical feature update rates. In particular it was shown that the adaptive track-based feature update algorithm produces the same tracking performance compared to using

the full spectral feature set in FAT, while significantly reducing bandwidth requirements and increasing theoretical feature update rate. Furthermore, results demonstrated that FAT improves the performance over regular kinematic tracking.

8 Compressed Sensing

While the presented Track Derived FAT (TDFAT) architecture estimates and maintains a minimal set of spectral features in support of target tracking over time, it is still necessary to record a significant HSI data cube in the vicinity of the target at track initiation, to compute the initial feature estimate. Let

$$Z \in \mathbb{R}^{m \times n \times b} \tag{7}$$

denote the subimage HSI data cube that needs to be recorded for the purpose of initiating a spectral feature data set. In Eqn. (7), m and n denote the size of the subimage (m and n are on the order of 20 in this work when tracking vehicles) and b denotes the number of bands to be recorded (61 in this work). Depending on the application, the area of interest for which HSI data is to be collected may be large, and recording and communicating the raw HSI data is a computationally challenging task. Thus, even in the context of performance-driven sensing, there may be a benefit in various compression methods for the HSI data. The following reviews work related to HSI compression as well as the more recent, and quite intriguing, field of compressed sensing.

8.1 Literature Review Related to HSI Compression and Compressed Sensing

Recently, the Multispectral Hyperspectral Data Compression Working Group (MHDC) of the Consultative Committee for Space Data Systems (CCSDS) has released a recommendation for a standard for lossless compression of multispectral and hyperspectral image data and a format for storing the compressed data [37]. Unfortunately, lossless compression typically does not offer very large compression ratios required for some real world applications. In fact, merely loading the uncompressed lossless HSI image into memory may require large computational resources. Thus, lossy compression schemes are of particular interest for HSI images. Vílchez et al. [38] analyze the impact of lossy compression on HSI classification and unmixing. Patrick et al. [39] investigate lossy compression for HSI and propose that a sensor manager can dictate the compression related to the needed classification accuracy. Chen et al. [40] explore the use of linear dimensionality reduction techniques and investigate their effect on the performance of classical target detection and classification techniques for HSI images. Note, they also discuss various compressed sensing and non-linear dimensionality reduction approaches (such as ISOMAP [41] and LLE [42]), but they instead focus on more classic techniques.

Chen et al. [40] choose not to pursue non-linear dimension reduction because of the purported computational expense of such methods. While the methods they discuss certainly suffer from the issues they describe, there are also many methods in this domain that could be quite applicable to the problems of interest. For example, a Kernel PCA method [43], especially one using a non-linear Mercer kernel [44], can be quite revealing as to the underlying non-linear structure of the data.

Chen et al. [40], as well as many others, also point out the recent flurry of activity in the field of Compressed Sensing (CS) that has attracted considerable research interest. The application of CS concepts to HSI data has been studied by Lv et al. [45] and Huo et al. [46]. Pfeffer and Zibulevsky [47] have studied the design of a micro-mirror array based system for CS of hyperspectral data. An HSI camera was also described by Duarte [17]. The RITMOS sensor, discussed in Sec. 3, also uses a micro-mirror array and can support compressed sensing of a HSI subimage data cube. The more advanced tunable MEMS Fabry-Perot Etalon HSI sensor design discussed in Sec. 3 allows simultaneous recording of hyperspectral imaging information across a Focal-Plane Array (FPA). Given such a sensor with these capabilities, it is interesting to

consider the applicability of a CS mechanism. In particular, if the regions of interest are all known *a priori* then it is not clear that compression or CS has much to offer except to *transmit* the data. On the other hand, if global context is important then a CS mechanism may still have a role to play.

As a parting shot let us also mention that, while seemingly dissimilar, non-linear dimension reduction and compressed sensing are actually interconnected. For example, see very recent work by Gao et al. [48]. In particular, Numerica has a small measure of expertise in such methods [49–51] and one could imagine applying such methods to various problems in HSI processing, generalizing the approach in [40].

8.2 Introduction to Basis Pursuit and Compressive Sampling

Space does not permit a full review of the vast literature of compressed sensing, but a small measure of background material will produce dividends in analyzing the applicability of compressed sensing to problems in hyperspectral imaging. In particular, there are the two closely related topics of basis pursuit [52] and compressive sampling [53] that have been active areas of research over the past several years and bear directly to the problems at hand.

To wit, let us begin by introducing a small measure of notation by considering a signal vector $\mathbf{y} \in \mathbb{R}^m$. One can think of $\mathbf{y}$ as either a one dimensional signal, a vectorization of an image $\mathbf{M} \in \mathbb{R}^{m \times n}$, or a vectorization of a hyperspectral data cube $\mathcal{Z} \in \mathbb{R}^{m \times n \times b}$ as above.

8.2.1 Compressive Sampling

Compressive sampling revolves around the reconstruction of $\mathbf{y} \in \mathbb{R}^l$ given a small set of linear samples. Specifically, given a matrix $\mathbf{A} \in \mathbb{R}^{k \times l}$ with $k \ll l$ and provided $\mathbf{b} \in \mathbb{R}^k$ with $\mathbf{A}\mathbf{y} = \mathbf{b}$, the reconstruction of $\mathbf{y}$ from $\mathbf{A}$ and $\mathbf{b}$ is desired. Of course, there are infinitely many solutions to $\mathbf{A}\mathbf{y} = \mathbf{b}$, so some principle must be chosen to select the $\mathbf{y}$ of interest. In compressed sensing that principle is *sparsity* (i.e. the solution with the smallest number of non-zero entries). The interested reader may look at [54] (and references therein) for details, but the intuition of such methods is that a sparse $\mathbf{y}$ is recoverable if $\mathbf{A}$ satisfies a *restricted isometry principle*. In other words, one requires that $\mathbf{A}$ is a unitary transform (e.g. a rotation) on all sufficiently sparse vectors $\mathbf{y}$ to allow recovery of a sparse $\mathbf{y}$. Surprisingly, such matrices are actually quite common and certain classes of random matrices satisfy this property with overwhelming probability. Compressive sampling in the underlying principle for the famed "single pixel camera" [55, 56].

The advantage of compressed sampling is that one can reconstruct the original data $\mathbf{y}$ from the *much smaller* vector $\mathbf{b}$. We also observe that the compression part of the algorithm (i.e. computing $\mathbf{A}\mathbf{y} = \mathbf{b}$) is just a matrix-vector multiply, so it can be *implemented very efficiently in hardware*. In fact, this is precisely the principle used in Pfeffer and Zibulevsky [47] for addressing HSI problems.

At this juncture it is important to spend a few words emphasizing the role that randomness plays in compressive sampling. Random matrices, with high probability, satisfy a restricted isometry principle. Of course, there are certainly non-random matrices that satisfy the same condition (a fact that will become important in our discussion of basis pursuit), but having random matrices at our disposal provides several interesting advantages for HSI problems. First, $\mathbf{A}$ can be very compactly communicated. For example, $\mathbf{A}$ can be represented as a seed value to a pseudo-random number generator and an agreed upon ordering for producing the elements of $\mathbf{A}$. Second, and perhaps more importantly, using a random matrix provides a "universal encoding" scheme that is not dependent on the particular form of the HSI signal. As is beautifully stated in [57]:

> This encoding/decoding scheme is of course very different from those commonly discussed in the literature of information theory. In this scheme, the encoder would not try to know anything about the signal, nor would exploit any special structure of the signal; it would blindly correlate the signal with noise and quantize the output - effectively doing very little work. In

other words, the encoder would treat each signal in exactly the same way, hence the name "universal encoding." – Candes and Tao [57]

8.2.2 Basis Pursuit

Basis pursuit, on the other hand, revolves around the sparse representation of a vector $\mathbf{y}$ using a dictionary of basis vectors $\mathbf{D}$. Specifically, given a matrix $\mathbf{D} \in \mathbb{R}^{m \times p}$ with $p > m$, it is of interest to find an $\mathbf{x} \in \mathbb{R}^p$ such that $\mathbf{Dx} = \mathbf{y}$. Again, there are many such solutions $\mathbf{x}$ and the principle is again to choose the solution which is the sparsest. Note that while $\mathbf{x}$ is high dimensional, $\mathbf{D}$ can be chosen so that $\mathbf{x}$ has very few non-zero entries and therefore $\mathbf{x}$ can be transmitted efficiently. In effect, the difference between the methods is that compressed sensing attempts to reconstruct $\mathbf{y}$ after it has been down-sampled by $\mathbf{A}$, while basis pursuit attempts to find a sparse representation of $\mathbf{y}$ in the basis $\mathbf{D}$. Both methods have applicability to the problem of compressing data for HSI.

The advantage of basis pursuit is that $\mathbf{D}$ can be chosen judiciously to minimize the amount of information that needs to be sent to represent the data. For example, one can choose a collection of basis functions, such as Fourier or wavelet, with desirable properties.

There are other important implications of such methods. First, point-wise error constraints are straightforward to guarantee. The point-wise error constraint is imposed on the compression side so that the compressor can use the freedom they afford for constructing a low bit-count representation of the data set. Second, de-noising (and perhaps thresholding) can be done automatically as part of the error bounds. Of course we are not the first to observe this connection and other similar ideas — sometimes called "Basis Pursuit Denoising" — which can be found in [52]. Finally, given large data sets where segmentation [58, 59] is desirable, we note that there are several ideas based upon basis pursuit, such as Morphological Component Analysis [60, 61], in this domain.

Of course, basis pursuit methods also require modification for the current context. In particular, a classic basis pursuit method requires more computation on the compression side than does a classic compressed sensing algorithm. Fortunately, Numerica, as well as many others [62–69], have performed substantial work in the area of efficient basis pursuit algorithms, and one of the capabilities that Numerica brings to the table for compressed sensing HSI processing is a novel scheme (with a pending patent) using the Bregman iteration for point-wise constrained basis pursuit problems. For example, consider solving a problem with 131,072 unknowns. Solving such a problem using a linear programming approach with an interior point technique [70] would take an (estimated) 4.4×10^7 seconds while using our techniques one can reduce that time to approximately 49 seconds! To make the above comparison fair we have used the scripting language Python [71] for both the standard solver (since that is its native language) and the Numerica developed solver. We note that the Numerica solver is also available in a C++ version where the difference is even more pronounced. The Numerica solver can solve a problem of the same size in 5.8 seconds using only 1.1 megabytes of memory.

8.2.3 Other Applications for Compressed Sensing

While compression, either by way of compressed sampling or basis pursuit, are certainly important for HSI applications, there are other interesting ways in which compressed sensing can also be applied for *HSI analysis*.

For example, compressed sensing is used in [72, 73] for *unmixing*. The idea in this domain is that each pixel of a particular HSI image is a mixture of some collection of surface properties. Specifically, [72] uses a compressed sampling idea to represent a given HSI spectra in terms of a spectra library without ever generating the full HSI data cube, while [73] uses ideas from basis pursuit and Total Variation minimization to indicate subpixel locations of features.

As another example, Golbabee et al. [55, 74] use compressed sensing to detect and analyze correlated structures in HSI data. The work in [55] revolves around the assumption that the HSI data is highly correlated and that modeling as a linear combination of a small number of independent sources leads to efficient algorithms requiring only a small number of measurements. [74] builds upon that work to produce efficient algorithms using joint-sparse representations.

9 Conclusion

This work presented tracking architectures in support of the goals of performance-driven sensing or dynamic data-driven application systems. We discussed two sensors, capable of collecting hyperspectral data in a commanded manner on a per-pixel and per-band level, that support hyperspectral data collection in these type of application systems. By using a set of DIRSIG simulated hyperspectral images we implemented tracking simulations of three different tracking configurations: kinematic only, detection derived feature aided, and target derived feature aided. The latter configuration represents a fully adaptive tracking configuration where spectral feature data is managed and optimized on a per track level using local information.

We described the different tracking architectures and the integrated algorithm components and presented simulation results that compare the different configurations on a simulated scenario that represent a typical urban scene. The track picture metrics shows that the combination of kinematic and HSI spectral data can improve tracking performance over pure kinematic or pure spectral feature tracking. Furthermore, the results demonstrated that the adaptive track-dependent feature tracking configuration (TDFAT) produces the same track picture performance results as when using the detection derived feature aided tracking configuration (DDFAT) that collects significantly more spectral data.

The results demonstrate that the TDFAT configuration realizes the goals of the performance-driven sensing paradigm: By only collecting a subset of features as needed, the data collection and bandwidth requirements are significantly reduced. The dynamic management of the feature data is designed to produce a feature set that supports the specific signal processing task (target tracking in our work) and to collect spectral data only where needed. By managing a small set of features, the update rate for feature data collected on individual targets is greatly increased which in turn should offer improved tracking and Object ID performance. Further improvements can be expected by applying principles from the field of Compressed Sensing motivated in Sec. 8.

10 FutureWork

The following provides some recommendations for future work to extend the developed algorithms.

- We tested the algorithms on a limited set of simulated data and future work could perform more extensive testing over longer and more widely varying scenarios. In particular, we could test feature based track stitching algorithms that rejoin broken tracks over longer time periods and test the algorithms at varying frame rates.

- The present work used a specific feature selection algorithm based on SFFS. Future work could integrate more advanced feature selection algorithms within an adaptive sensing concept.

- The current work integrated algorithms to support adaptive sensing with a prototype tracking system. In future work, the algorithms could be integrated with an operational state-of-the-art MHT tracking system and demonstrated on recorded data. In particular, it would be desired to demonstrate the algorithms in-the-loop with an operational HSI sensor.

- Given an integration with an MHT tracker would also allow us to investigate further the aspects of ambiguity management, where some use of the HSI sensor is controlled by an algorithm that attempts to minimize association uncertainty across all targets.

- The present work has only considered hyperspectral data. As has been shown in the literature, polarimetric data is also useful in target surveillance applications. Thus, future work could extend the present work to include more general feature data.

- The present algorithms rely on the collection of a full set of spectral data in a subregion of a surveillance region to initialize the feature data on a newly observed target or to update the feature data over time. This is a domain in which compressed sensing can pay substantial dividends in making HSI sensors cheaper, increasing computational efficiency, and providing more information to the user. *Accordingly, two directions immediately present themselves for possible future work*. First, and perhaps most obvious, is the issue of efficient calculations. Compressive sensing is, perhaps almost by definition, efficient on the compression side. In fact, as we discuss here, one is even able implement such methods in hardware so that *the full data cube Z never need be computed*. Unfortunately, for the large data sets arising from modern HSI sensors, the decompression task is non-trivial. Especially for real time applications, efficient decompression is an open problem. Thankfully, algorithms based upon iterative approaches like Bregman iterations [65], hold promise for progress. In fact, in other problem domains, Numerica has already submitted two patents on precisely this type of idea that could be leveraged to accelerate HSI decompression. Second, beyond merely reconstructing the original HSI data, compressive sensing algorithms can be used for *novel analysis*. The above discussion of spectral unmixing is a prime example. While there may be many different collections of surface properties that give rise to a particular measured spectra, under mild assumptions, the decomposition into a *sparse set of surface properties* is unique. Again, in other problems domains, Numerica is working on precisely such signal separation problems, with a special focus on anomaly detection. Such features promise to provide a resource for downstream processing, such as feature aided tracking.

References

[1] T. Wang and Z. Zhu, "Intelligent multimodal and hyperspectral sensing for real-time moving target tracking," in *Applied Imagery Pattern Recognition Workshop, 2008. AIPR '08. 37th IEEE*, 2008.

[2] J. Kerekes, M. Muldowney, K. Strackerjan, L. Smith, and B. Leahy, "Vehicle tracking with multitemporal hyperspectral imagery," in *Proc. of SPIE, Vol. 6233*, 2006.

[3] D. Rosario, "Hyperspectral target tracking," in *Aerospace Conference, 2011 IEEE*, 2011.

[4] J. P. Kerekes, M. D. Presnar, K. D. Fourspring, Z. Ninkov, D. R. Pogorzala, A. D. Raisanen, A. C. Rice, J. R. Vasquez, J. P. Patel, R. T. MacIntyre, and S. D. Brown, "Sensor modeling and demonstration of a multi-object spectrometer for performance-driven sensing," in *Proc. of SPIE, Vol. 7334*, 2009.

[5] "AFOSR: Adaptive multimodal sensing and THZ-speed electronics." http://www.wpafb.af.mil/library/factsheets/factsheet.asp?id=12580.

[6] R. Meyer, K. Kearney, Z. Ninkov, C. Cotton, P. Hammond, and B. Statt, "RITMOS: a micromirror-based multi-object spectrometer," in *Proc. of SPIE, Vol. 5492*, 2004.

[7] A. Rivas, "Tunable micro-electro mechanical fabry perot etalon," Master's Thesis, Chester F. Carlson Center for Imaging Science at Rochester Institute of Technology, Rochester, NY, August 2011.

[8] A. Rice, J. Vasquez, J. Kerekes, and M. Mendenhall, "Persistent hyperspectral adaptive multi-modal feature-aided tracking," in *Proc. of SPIE, Vol. 7334*, 2009.

[9] A. Rice and J. Vasquez, "Context-aided tracking with an adaptive hyperspectral sensor," in *2011 Proceedings of the 14th International Conference on Information Fusion (FUSION)*, 2011.

[10] H. Nguyen, A. Banerjee, and R. Chellappa, "Tracking via object reflectance using a hyperspectral video camera," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2010.

[11] Z. Lu, A. Rice, J. Vasquez, and J. Kerekes, "Target discrimination via optimal wavelength band selection with synthetic hyperspectral imagery," in *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2010.

[12] S. Nakariyakul, "A new feature selection algorithm for multispectral and polarimetric vehicle images," in *2009 16th IEEE International Conference on Image Processing (ICIP)*, 2009.

[13] A. Vodacek, J. Kerekes, and M. Hoffman, "Adaptive optical sensing in an object tracking DDDAS," in *International Conference on Computational Science, ICCS 2012, Procedia Computer Science 9 (2012) 1159-1166*, 2012.

[14] S. Gadaleta, J. Kerekes, and K. Tarplee, "Autonomous target dependent waveband selection for tracking in performance-driven hyperspectral sensing," in *Proc. SPIE 8390, 839023*, 2012.

[15] "DDDAS Dynamic Data-Driven Application Systems." http://www.dddas.org/.

[16] A. Rivas, J. Kerekes, and A. Raisanen, "Tunable single pixel MEMS Fabry-Perot interferometer," in *Computational Optical Sensing and Imaging (COSI), Toronto, Canada*, 2011.

[17] "Hyperspectral imaging." http://cnx.org/content/m37370/latest/?collection=col11133/1.5.

[18] S. Pierce and J. Vasquez, "Context aided tracking in aerial video surveillance," Proceedings of the SPIE, Conference on Signal and Data Processing of Small Targets, 2008.

[19] J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw, "Review of passive imaging polarimetry for remote sensing applications," *Appl. Opt.*, vol. 45, pp. 5453–5469, Aug 2006.

[20] J. Schott, ed. Bellingham, WA: SPIE Press, 2009.

[21] Digital Imaging and Remote Sensing Laboratory at Rochester Institute of Technology, "Digital Imaging and Remote Sensing Image Generation (DIRSIG)." http://dirsig.cis.rit.edu/.

[22] E. Ientilucci and S. Brown, "Advances in wide area hyperspectral image simulation," in *Proc. of SPIE, Vol. 5075*, 2003.

[23] "ENVI File Formats." http://geol.hu/data/online_help/ENVI_File_Formats.html.

[24] R. Rothrock and O. Drummond, "Performance metrics for multiple-sensor multiple-target tracking," in *Proc. SPIE 4048*, 2000.

[25] V. Reilly, H. Idrees, and M. Shah, "Detection and tracking of large number of targets in wide area surveillance," in *Proceeding ECCV'10 Proceedings of the 11th European conference on computer vision conference on Computer vision: Part III*, pp. 186–199, 2010.

[26] A. Czyzewski, G. Szwoch, P. Dalka, and P. Szczuko, "Multi-stage video analysis framework," in *Video Surveillance* (W. Lin, ed.), InTech, 2011.

[27] A. B. Poore and S. M. Gadaleta, "The group assignment problem arising in multiple target tracking," in *Theory and Algorithms for Cooperative Systems*, pp. 423–449, 2004.

[28] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999.

[29] A. Rice, J. Vasquez, M. Mendenhall, and J. Kerekes, "Feature-aided tracking via synthetic hyperspectral imagery," in *Proceedings of the First IEEE Workshop on Hyperspectral Image and Signal*

Processing: Evolution in Remote Sensing (WHISPERS), 2009.

[30] S. Gadaleta, B. Slocumb, S. Herman, B. Brewington, F. Obermeyer, M. Levedahl, and A. Poore, "Short-term ambiguity assessment to augment tracking data association information," in *Eighth International Conference on Information Fusion, Philadelphia, PA*, 2005.

[31] V. Ravindra, Y. Bar-Shalom, and T. Damarla, "Feature-aided tracking of ground vehicles using passive acoustic sensor arrays," *Journal of Advances in Information Fusion*, vol. 5, no. 2, pp. 88–107, 2010.

[32] "Table of Chi-square statistics." http://home.comcast.net/ sharov/PopEcol/tables/chisq.html.

[33] Y. Du, C.-I. Chang, H. Ren, F. D'Amico, and J. Jensen, "A new hyperspectral discrimination measure for spectral characterization," *Optical Engineering*, vol. 43, no. 8, pp. 1777–1786, 2004.

[34] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. San Diego, CA: Academic Press, 2003.

[35] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory, Algorithms, and Software*. J. Wiley, 2001.

[36] A. Jain and D. Zongker, "Feature-selection: Evaluation, application, and small sample performance," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 153–158, 1997.

[37] J. Sánchez, E. Auge, J. Santaló, I. Blanes, J. Serra-Sagristá, and A. Kiely, "Review and implementation of the emerging CCSDS recommended standard for multispectral and hyperspectral lossless image coding," in *First International Conference on Data Compression, Communications and Processing*, 2011.

[38] F. García-Vílchez, J. M. noz Marí, M. Zortea, I. Blanes, V. González-Ruiz, G. Camps-Valls, A. Plaza, and J. Serra-Sagristá, "The impact of lossy compression on hyperspectral image classification and unmixing," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, 2011.

[39] J. Patrick, R. Brant, and E. Blasch, "Hyperspectral imagery throughput and fusion evaluation over compression and interpolation," in *11th International Conference on Information Fusion*, 2008.

[40] Y. Chen, N. Nasrabadi, and T. Tran, "Effects of linear projections on the performance of target detection and classification in hyperspectral imagery," *Journal of Applied Remote Sensing*, vol. 5, 2011.

[41] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–23, 2000.

[42] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323–2326, Dec. 2000.

[43] C. Bishop, *Pattern recognition and machine learning*. Springer New York., 2006.

[44] A. B. Lee and L. Wasserman, "Spectral Connectivity Analysis," *Journal of the American Statistical Association*, pp. 1241–1255, 2009.

[45] J. Lv, Y. Li, B. Huang, and C. Wu, "Hyperspectral compressive sensing," in *Proc. of SPIE Vol. 7810*, 2010.

[46] C. Huo, R. Zhang, and D. Yin, "Compression technique for compressed sensing hyperspectral images," *International Journal of Remote Sensing*, 2011.

[47] Y. Pfeffer and M. Zibulevsky, "A micro-mirror array based system for compressive sensing of hyperspectral data." Technical Report, Dept. of Electrical Engineering, Technion - Israel Institute of Technology, 2010.

[48] J. Gao, Q. Shi, and T. Caetano, "Dimensionality reduction via compressive sensing," *Pattern Recognition Letters*, vol. 33, no. 9, pp. 1163–1170, 2012.

[49] R. Paffenroth, P. D. Toit, R. Nong, L. L. Scharf, A. Jayasumana, and V. Bandara, "Space-time signal

processing for distributed pattern detection in sensor networks," *Journal of Selected Topics in Signal Processing (submitted)*, vol. 6, no. 1, 2013.

[50] R. Paffenroth, P. D. Toit, L. Scharf, A. P. Jayasumana, V. Bandara, and R. Nong, "Distributed pattern detection in cyber networks," in *Cyber Sensing*, 2012.

[51] R. Paffenroth, P. D. Toit, L. Scharf, A. P. Jayasumana, V. Bandara, and R. Nong, "Space-time signal processing for distributed pattern detection in sensor networks," in *Signal and Data Processing of Small Targets*, vol. 8393, 2012.

[52] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1999.

[53] D. L. Donoho, "Compressed Sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.

[54] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math*, vol. 59, pp. 1207–1223, 2005.

[55] M. Golbabaee, S. Arberet, and P. Vandergheynst, "Multichannel Compressed Sensing via Source Separation for Hyperspectral Images," in *Proceedings of Eusipco 2010*, 2010.

[56] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Proc. Mag.*, vol. 25, no. 2, pp. 83–91, 2008.

[57] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: {U}niversal encoding strategies?," *IEEE Trans. Inform. Theory*, vol. 52, pp. 5245–5406, Dec. 2006.

[58] S. Safavian, H. Rabiee, and M. Fardanesh, "Projection pursuit image compression with variable block size segmentation," *IEEE Signal Processing Letters*, vol. 4, pp. 117–120, May 1997.

[59] J. Vaisey and A. Gersho, "Image compression with variable block size segmentation," *IEEE Transactions on Signal Processing*, vol. 40, no. 8, pp. 2040–2060, 1992.

[60] S. Cea, J. l. Starck, F. Gif, Y. Cedex, F. Murtagh, and M. Louys, "Astronomical Image Compression," vol. 136, pp. 579–590, 1999.

[61] J. Bobin, J.-L. Starck, J. M. Fadili, Y. Moudden, and D. L. Donoho, "Morphological component analysis: an adaptive thresholding strategy.," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 16, pp. 2675–81, Nov. 2007.

[62] T. Goldstein, X. Bresson, and S. Osher, "Geometric Applications of the Split Bregman Method: Segmentation and Surface Reconstruction." http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.154.3959, 2009.

[63] J.-f. Cai, S. Osher, and Z. Shen, "Fast Linearized Bregman Iteration for Compressed Sensing." http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.139.7640, 2008.

[64] J.-f. Cai, S. Osher, and Z. Shen, "Convergence of the Linearized Bregman Iteration for L1-norm Minimization." http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.139.8779, 2008.

[65] S. Osher, Y. Mao, B. Dong, and W. Yin, "Fast linearized Bregman iteration for compressive sensing and sparse denoising," *Commun. Math. Sci.*, vol. 8, pp. 93–111, 2010.

[66] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for 1-minimization with applications to compressed sensing," *SIAM J. Imaging Sci*, pp. 143–168, 2008.

[67] J.-f. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration, submitted." http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.157.9714, 2009.

[68] W. Yin, S. Osher, J. Darbon, and D. Goldfarb, "Bregman iterative algorithms for compressed sensing and related problems," *SIAM J. Imaging Sciences*, vol. 1, no. 1, pp. 143–168, 2008.

[69] J. Xu and S. Osher, "Iterative Regularization and Nonlinear Inverse Scale Space Applied to Wavelet-Based Denoising," *IEEE Transactions on Image Processing*, vol. 16, pp. 534–544, Feb. 2007.

[70] J. Dahl and L. Vandenberghe, "CVXOPT." http://abel.ee.ucla.edu/cvxopt/index.html.

[71] M. Lutz, *Programming Python*. O'Reilly and Associates, 1996.

[72] C. Li, T. Sun, K. F. Kelly, and Y. Zhang, "A compressive sensing and unmixing scheme for hyperspectral data processing," *IEEE Transactions on Image Processing*, vol. 21, pp. 1200–10, Mar. 2012.

[73] Z. Guo, T. Wittman, and S. Osher, "L1 unmixing and its application to hyperspectral image enhancement," vol. 7334, pp. 73341M–73341M–9, Spie, 2009.

[74] M. Golbabaee and P. Vandergheynst, "Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery," *The 37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012)*, 2012.