SCC-R-236-2

SPACE COMPUTER CORPORATION

DoE Phase II SBIR: Spectrally-Assisted Vehicle Tracking

Final Report

Date Prepared: **February 28, 2013**

Submitted to:

U.S. Department of Energy, NA-22

Attn: Victoria Franques

1000 Independence Ave, S.W.

Washington, DC 20585

Prepared Under: **DoE Grant No. DE-SC0000689**

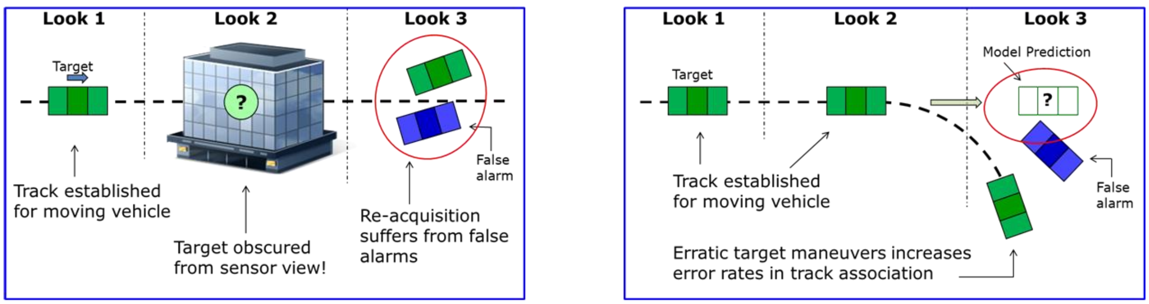
Space Computer Corporation
Los Angeles, California
www.spacecomputer.com

Contents

Space Computer Corporation
Los Angeles, California
www.spacecomputer.com

1. Introduction

This document is a Final Report for our Phase II SBIR for work on the topic of Spectrally-Assisted Vehicle Tracking. All program goals have been met, particularly in advancing the state-of-the art in hyperspectral-aided vehicle tracking.



a) Obscured target b) Erratic motion

Figure 1. Tracking performance suffers when target is obscured from view or undergoes erratic motion.

Figure 2 illustrates the basic concept of leveraging target spectral information in complicated scenarios to help resolve tracking ambiguities between targets and false alarms. This simple example illustrates a scenario where a green target is initially exhibiting motion that is well characterized by the tracker's kinematic model (e.g. piece-wise linear velocity). The green target is thus easily tracked from the first to the second observation. Complication arises in the third observation as two complications occur: 1) the green target moved erratically, and 2) a false alarm red vehicle entered the scene.

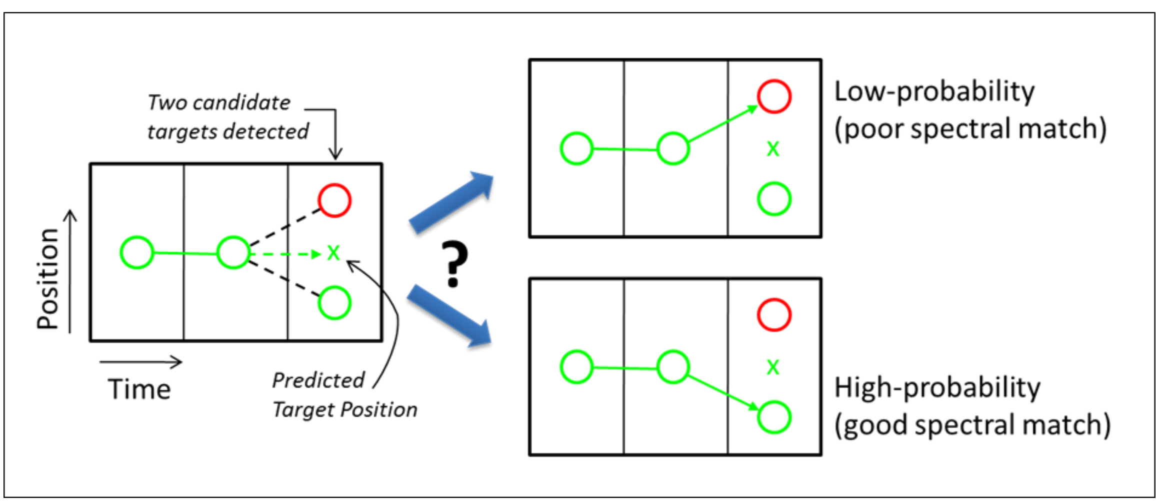


Figure 2. Target's spectral signature used to disambiguate confusing tracking scenarios.

1.1 Program Goals

The goal of this Phase II SBIR is to develop a prototype software package to demonstrate spectrally-aided vehicle tracking performance. The primary application is to demonstrate improved target vehicle tracking performance in complex environments where traditional spatial tracker systems may show reduced performance. Example scenarios in Figure 1 include a) the target vehicle obscured by a large structure for an extended period of time, or b), the target engaging in extreme maneuvers amongst other civilian vehicles.

The target information derived from spatial processing is unable to differentiate between the green versus the red vehicle. Spectral signature exploitation enables comparison of new candidate targets with existing track signatures. The ambiguity in this confusing scenario is resolved by folding spectral analysis results into each target nomination and association processes. Figure 3 shows a number of example spectral signatures from a variety of natural and man-made materials.

The work performed over the two-year effort was divided into three general areas: algorithm refinement, software prototype development, and prototype performance demonstration. The tasks performed under this Phase II to accomplish the program goals were as follows:

1. Acquire relevant vehicle target datasets to support prototype.
2. Refine algorithms for target spectral feature exploitation.
3. Implement a prototype multi-hypothesis target tracking software package.
4. Demonstrate and quantify tracking performance using relevant data.

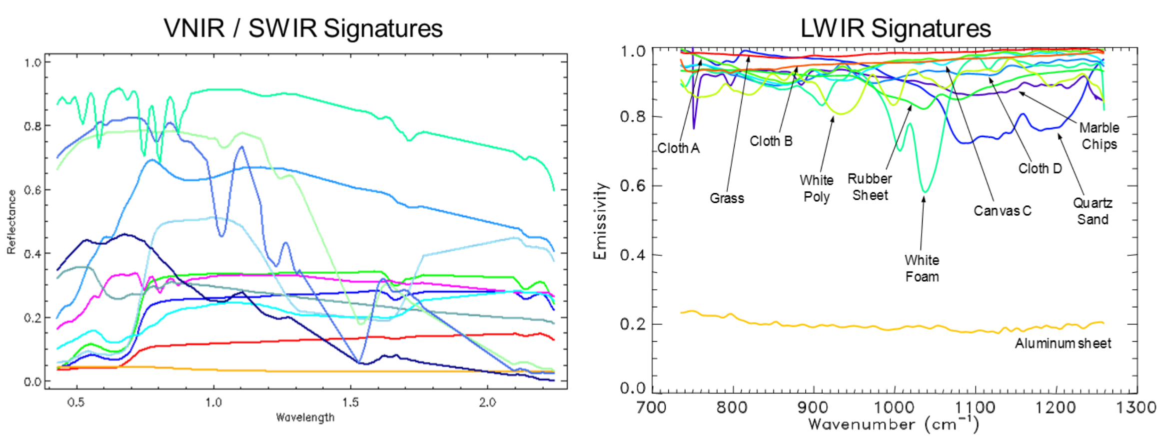


Figure 3. Spectral signatures from a multitude of natural and man-made materials in both the solar and thermal regions.

1.2 Schedule

The schedule of program tasks are illustrated below in Figure 4. Details for each of the work tasks are elaborated upon in later sections of this report.

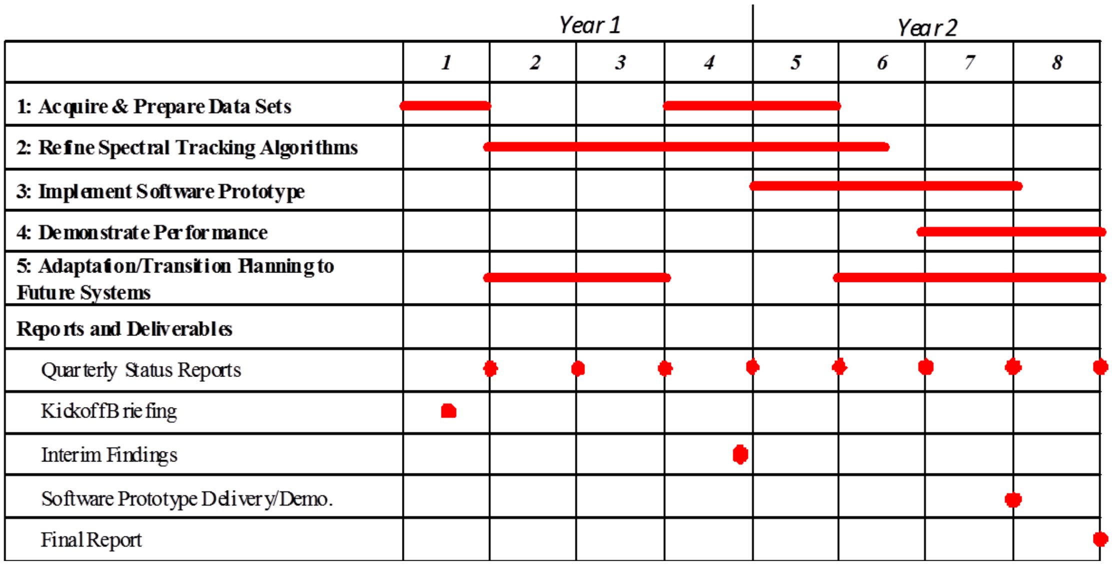


Figure 4. Two-year Phase II SBIR project schedule.

1.3 Kickoff meeting

A kickoff meeting was held at Los Alamos National Laboratory on February 10, 2011. Meeting attendees from DOE included Victoria Franques, Herb Fry, Kevin Mitchel, Brad Henderson, James Theiler, and Bernie Foy. Representatives from SCC included Pierre Villeneuve, Jeffrey Weinheimer, and Scott Beaven.

The presented briefings included:

1. Target tracking overview and problem description
2. Prior SCC efforts and results in hyperspectral-aided tracking
3. Requirements and goals for representative data for algorithm development
4. Review of potential system concepts for future prototype software
5. Preliminary results for an innovative joint spatial-spectral target feature descriptor

Follow-up discussions focused on available test data and focus of algorithm refinement effort. Herb Fry offered to make available sample data from the DoE's Blue-Grey experiment from 2005 involving both LWIR and VNIR/SWIR HSI sensors over an industrial complex. While this data does not support the rapid repeat coverage rates required for tracking moving targets, it does offer multiple observations of static targets over a two-day period.

Kevin Mitchel and Herb Fry expressed an interest in investigating cross-band target phenomenology for advanced target exploitation. For example, consider simultaneous LWIR and VNIR/SWIR hyperspectral observation of a target of interest, along with sufficient metadata to properly geo-register imagery to a common spatial grid. Given target signatures for both VNIR/SWIR and LWIR (reflectance and emissivity spectra), one would expect highly-correlated detector response for those pixels over the target. Background clutter in the LWIR versus the VNIR/SWIR is not expected to be highly correlated since they derive from different phenomenology sources. Thus combined VNIR/SWIR and LWIR target detection performance is expected to be greater than VNIR/SWIR or LWIR alone. Efforts in this direction will support target-tracking algorithm development as well as longer-term DOE mission interests.

2. Sensor Data Acquisition

2.1 Acquired DOE Blue-Grey AL-7 Sample Data

Sample data from the Blue-Grey collection experiment consists of imagery recorded over the AL-7 target site. Sensor data comprised GLASS LWIR HSI, COMPASS VNIR/SWIR HSI, and a Nikon RGB high-resolution context camera. The primary objective of the AL-7 Blue-Grey collection experiment had little connection to the problem of spectrally-aided target tracking. It is well-characterized data of opportunity that is expected to provide additional test case scenarios for developing and demonstrating performance of tracking system concepts. An example of imagery from the RGB context camera is shown in Figure 5. The highlighted red inset shows an example of the diversity of static targets in a parking lot.

Figure 5. Blue-Grey AL-7 example RGB context camera imagery over parking lot area.

Figure 6 shows an example of COMPASS imagery covering the same parking lot area. The COMPASS sensor system collects imagery in whiskbroom fashion. Each scan is collected while the scan mirror projects the line of pixels from one side to the other. The red parallelogram highlights the same area as the inset in Figure 5 above. Sample target and background spectra from the COMPASS sensor are shown in Figure 7 from selected vehicle targets and the ground in the parking lot.

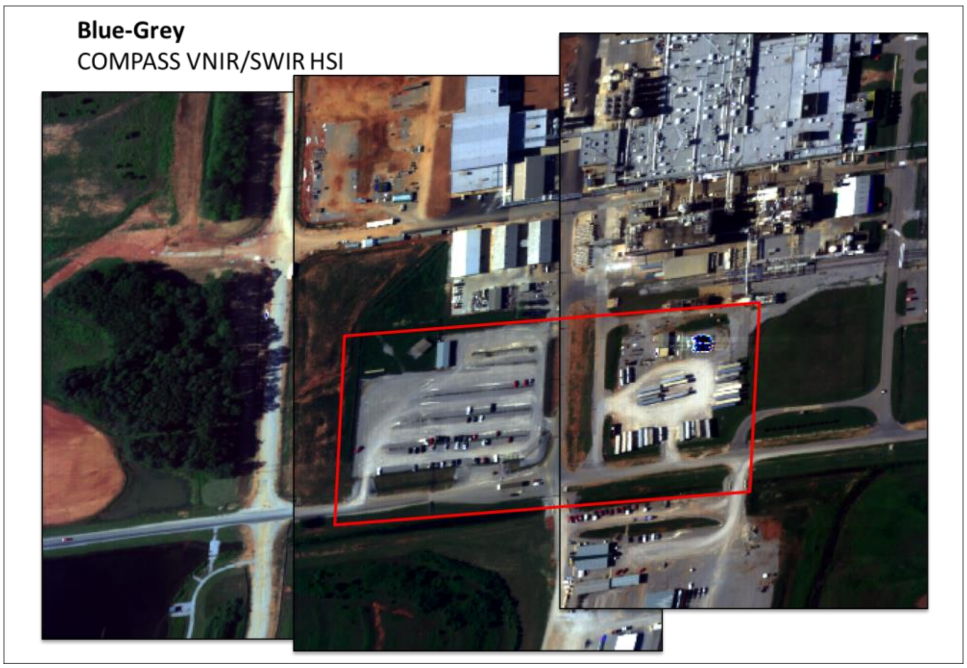


Figure 6. Blue-Grey AL-7 example COMPASS imagery.

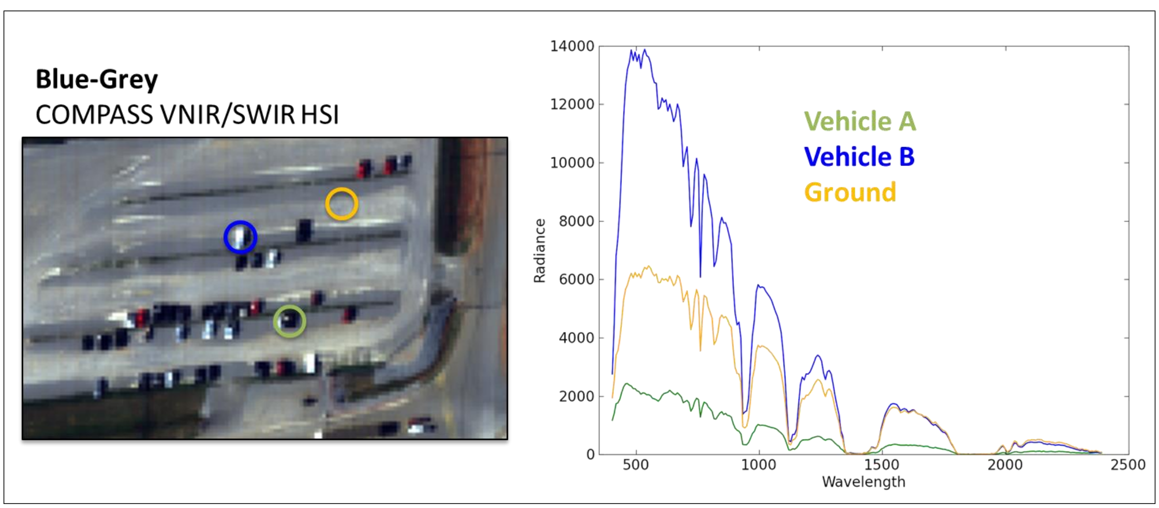


Figure 7. Blue-Grey AL-7 example VNIR/SWIR vehicle spectra from COMPASS sensor.

Figure 8 shows an example of GLASS imagery collected over the same parking lot area mentioned above. The GLASS sensor recorded HSI data in a push-broom mode where the line of focal plane pixels are projected perpendicular to the direction of flight. This data shows evidence of severe platform vibration. An initial analysis of GPS/INS metadata shows that much of this high-frequency motion was not measured by the INS system. This may introduce problems in the near future as we attempt to geo-rectify all image datasets to a common spatial grid. Sample target and background spectra from the GLASS sensor are shown in Figure 9 for two vehicles as well as the parking lot ground surface.

Figure 8. Blue-Grey AL-7 example GLASS imagery.

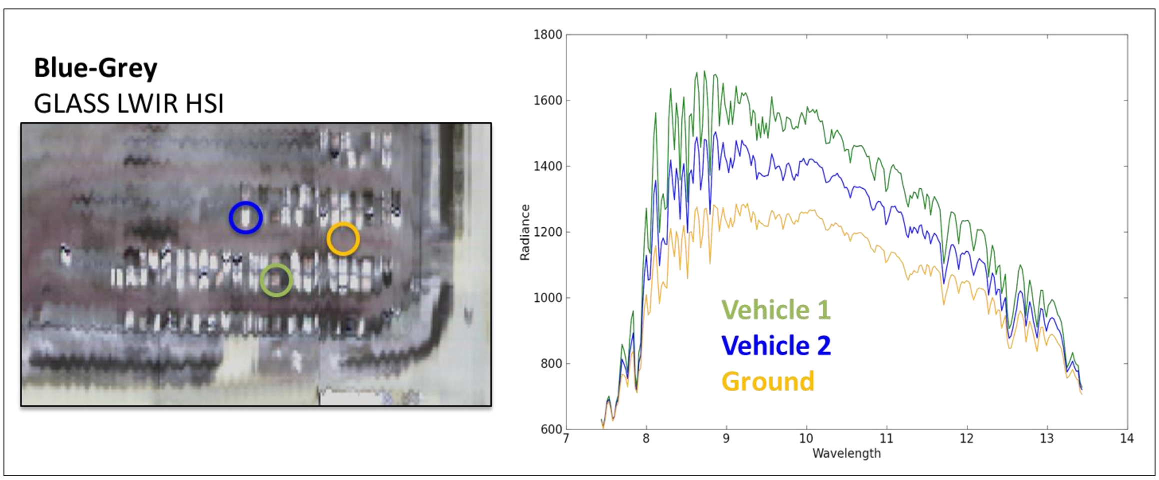


Figure 9. Blue-Grey AL-7 example LWIR vehicle spectra from GLASS sensor.

2.2 Acquired Air Force LWIR Polarimetric Data

SCC acquired an additional dataset of opportunity from the Air Force Research Laboratory in Dayton, OH. This data was collected via a LWIR polarimetric microgrid imaging sensor named PIRATE (Polarized InfraRed Advanced Tactical Experiment), [1]. The data collected by PIRATE enables estimation of the S0, S1, and S2 Stokes vector parameters at each image pixel. The S0 parameter corresponds to the LWIR broadband radiance, while S1 and S2 correspond to two orthogonal linear polarizations. Circular polarization was not measured by this sensor. Most manmade objects having smooth metal or plastic surfaces show a distinctive polarization signature that enables increased signal-to-clutter measurements.

The majority of this data consists of small remotely controlled aircraft flying in a wooded area. Some imagery shows the aircraft easily visible against the cold sky background. Other examples show the aircraft nearly invisible against the forest background. This data is expected to provide test scenarios that highlight two work areas: 1) small moving target that is difficult to track against certain backgrounds; and 2) demonstrate potential benefits in tracking performance when polarimetric phenomenology is leveraged to increase target-to-background contrast.

2.3 Collected Hyperspectral Video Test Data

We performed a small data collection experiment directly on the SCC premises using two independent sensors: and wide field-of-view high-definition RGB camcorder and a 20-band VNIR video HSI sensor on loan from Bodkin Design & Engineering (BD&E). The camcorder recorded imagery (3 bands, 1920 x 1080 pixels) at 60 frames per second, while the BD&E sensor recorded HSI video data (20 bands, 180 x 180 pixels) at 30 frames per second. The camcorder provided RGB context imagery over a wide field of view, while the video HSI was operated with a zoom lens and was manually steered to follow target objects of interest.

The experiment scenario involved civilian vehicles and civilians in a busy commercial parking lot. The camcorder was mounted static on a tripod for the entire experiment. The video HSI sensor was steered to follow a single red and white vehicle as it drove throughout the parking lot. The driver was instructed to exit the car, walk about the parking lot, and then return to the vehicle. This collection of data directly supports algorithm development and concept demonstration. Figure 10 shows sample imagery from both sensors at approximately the same time. The tracked civilian vehicle is clearly identified as a red and white Mini Cooper with a Canadian flag on the roof.

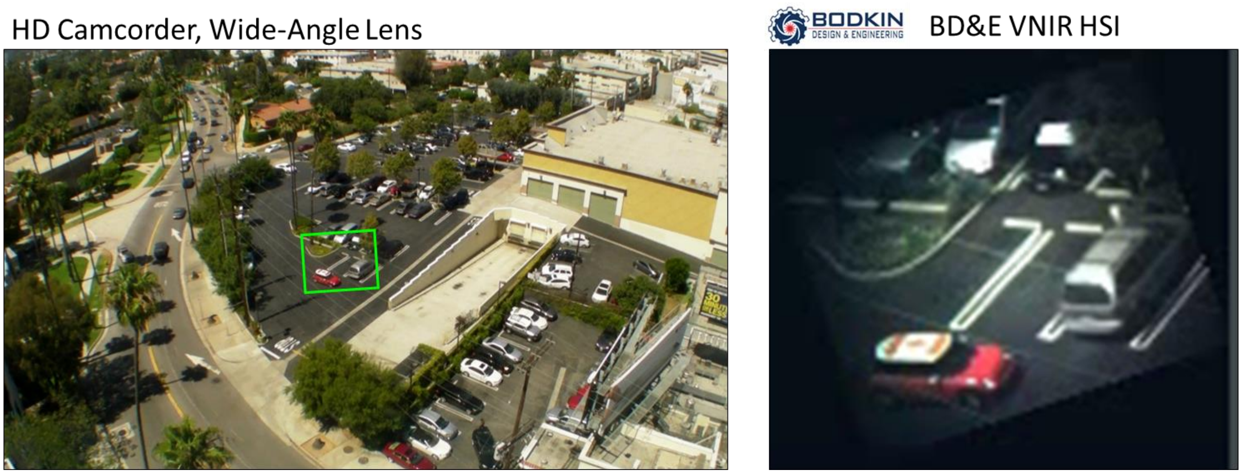


Figure 10. Sample imagery from in-house hyperspectral video data collection experiment.

Figure 11. Sample spectra from civilian vehicle and adjacent background.

2.4 Investigate Additional In-House Data Sets

SCC has access to large volumes of HSI data from a number of VNIR/SWIR and LWIR sensors. While only a fraction of this data supports quantifiable analysis of moving targets via repeat sensor coverage of a scene, it is adequate to supplement the recently acquired Blue-Grey multi-sensor data and AFRL polarimetric LWIR imagery.

The HYCAS VNIR/SWIR sensor collected over a mall parking lot with a large number of civilian vehicles (Figure 12). This example consists of four subsequent scene observations separated by approximately four minutes. A total of 180 stationary vehicles were manually masked in order to evaluate nomination performance. This data has been exploited previously to investigate a number of target nomination and association phenomenology problems. The primary restriction with this data is that the repeat observation rate spans multiple minutes, and thus it is not particularly suited to analyses that depend upon measuring a target's kinematic motion.

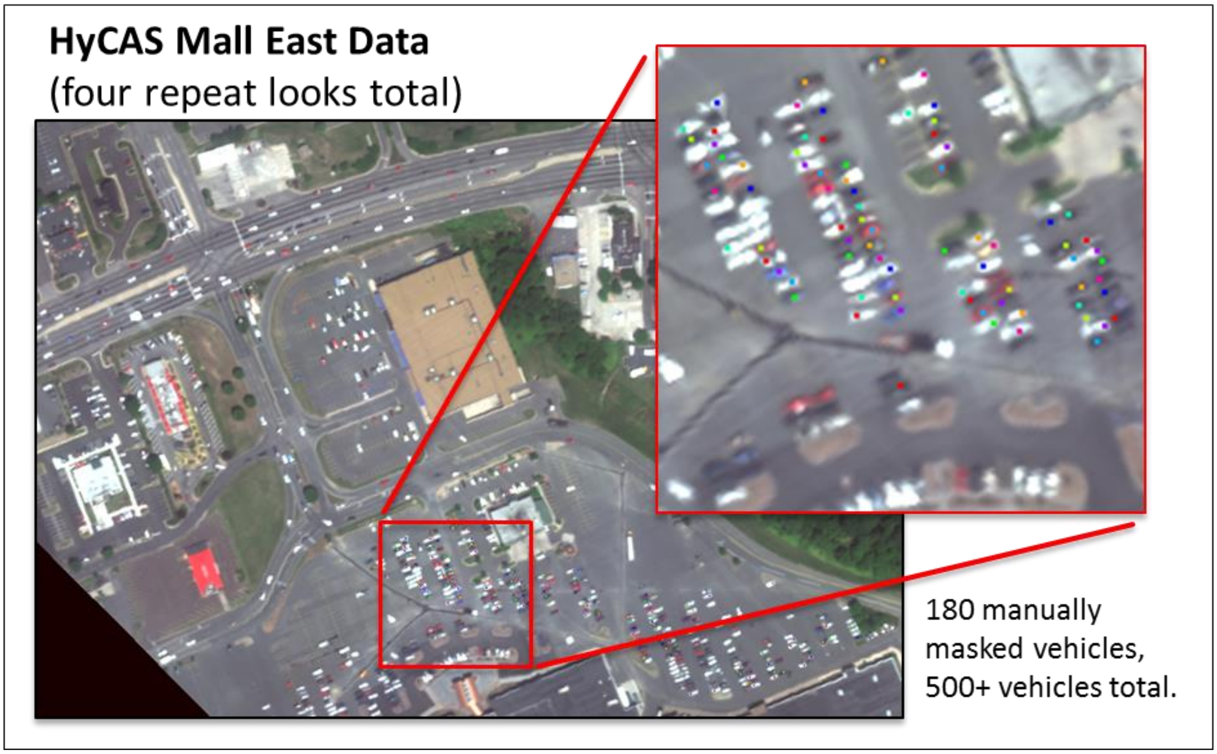
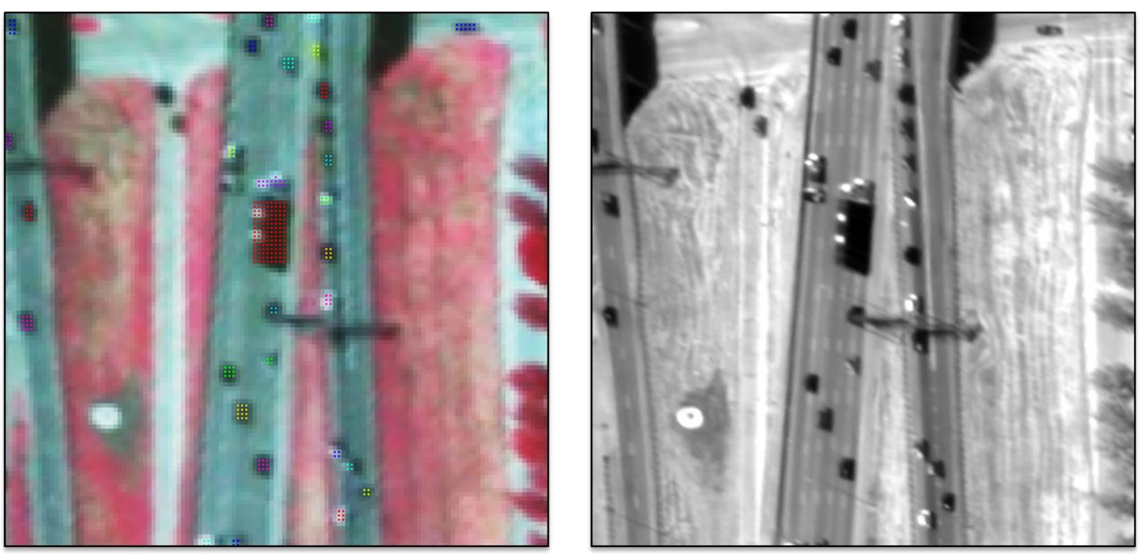


Figure 12. HyCAS Mall East data with 180 manually masked target vehicles over four repeat looks of same area.



a) SPIRITT HSI false-color RGB imagery with masked vehicle targets.

b) SPIRITT high-resolution context imagery.

Figure 13. SPIRITT imagery with manually defined vehicle target masks.

SCC has processed several examples of SPIRITT VNIR/SWIR "spot mode" data (Figure 13) from several large cities in the US. The SPIRITT system is an embedded suite of multiple sensors viewing a scene through a common optical path. This system generally operates at high altitude (50 kft) and observes targets at long slant range (25 km). The current configuration includes four channels: VNIR hyperspectral, SWIR hyperspectral, VIS panchromatic, and MWIR panchromatic. A small number of datasets are available where SPIRITT operated in "spot mode" and flew large circles (25 km radius) while observing target areas for extended periods. These datasets have proven useful while investigating use of a single sensor to perform both spatial and spectral target tracking tasks.

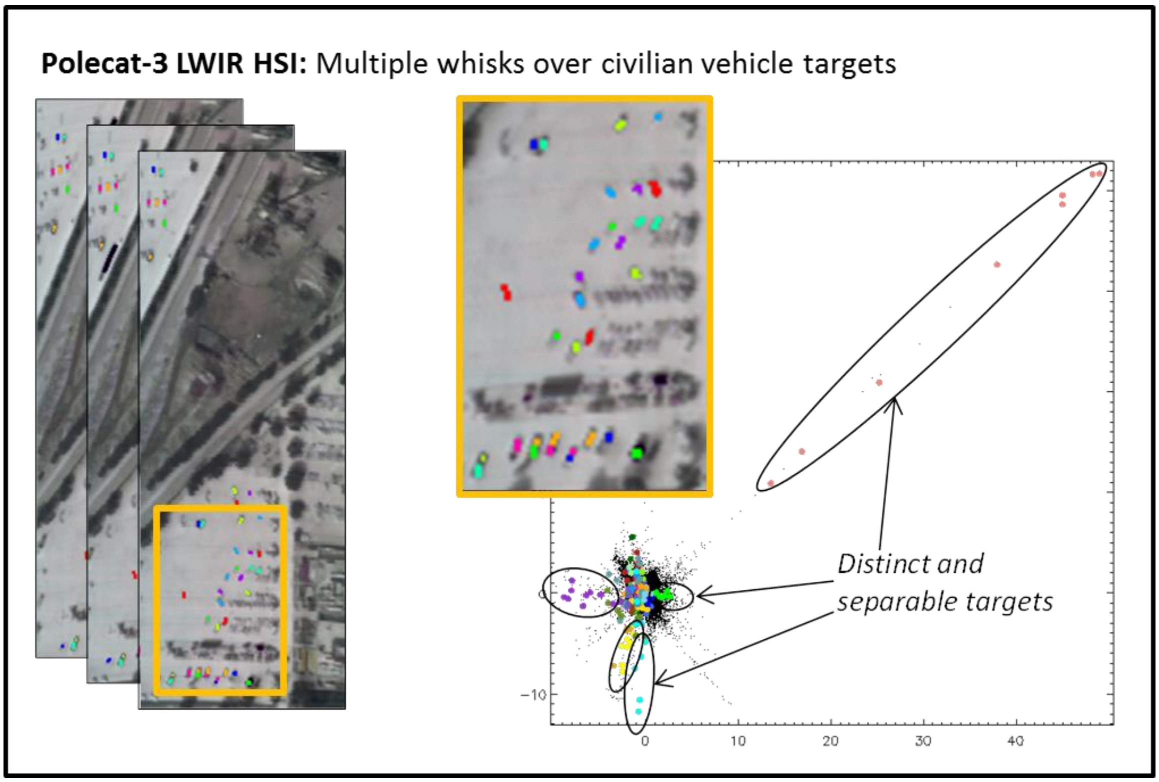


Figure 14. Polecat-3 LWIR HSI with multiple repeat observations of civilian vehicle targets.

Example LWIR vehicle target data exists in the form of Polecat-3 Red-Snapper three-whisk sets collected over an El-Segundo, CA parking lot populated with a large number of civilian vehicles. The site was observed multiple times over a period of many hours, thus allowing for controlled testing of repeated target observations. Vehicle target spectral phenomenology was investigated by manually selecting samples from numerous on-target pixels.

3. Algorithm Refinement and Data Analysis

3.1 Scene-Based Image Registration

The LWIR polarimetric data collected by PIRATE [1], mentioned earlier in Section 2.2, enables investigation of the potential improvement that polarimetric sensors can bring to tracking small hard-to-see targets. Most manmade objects having smooth metal or plastic surfaces show a distinctive polarization signature that enables increased signal-to-clutter measurements.

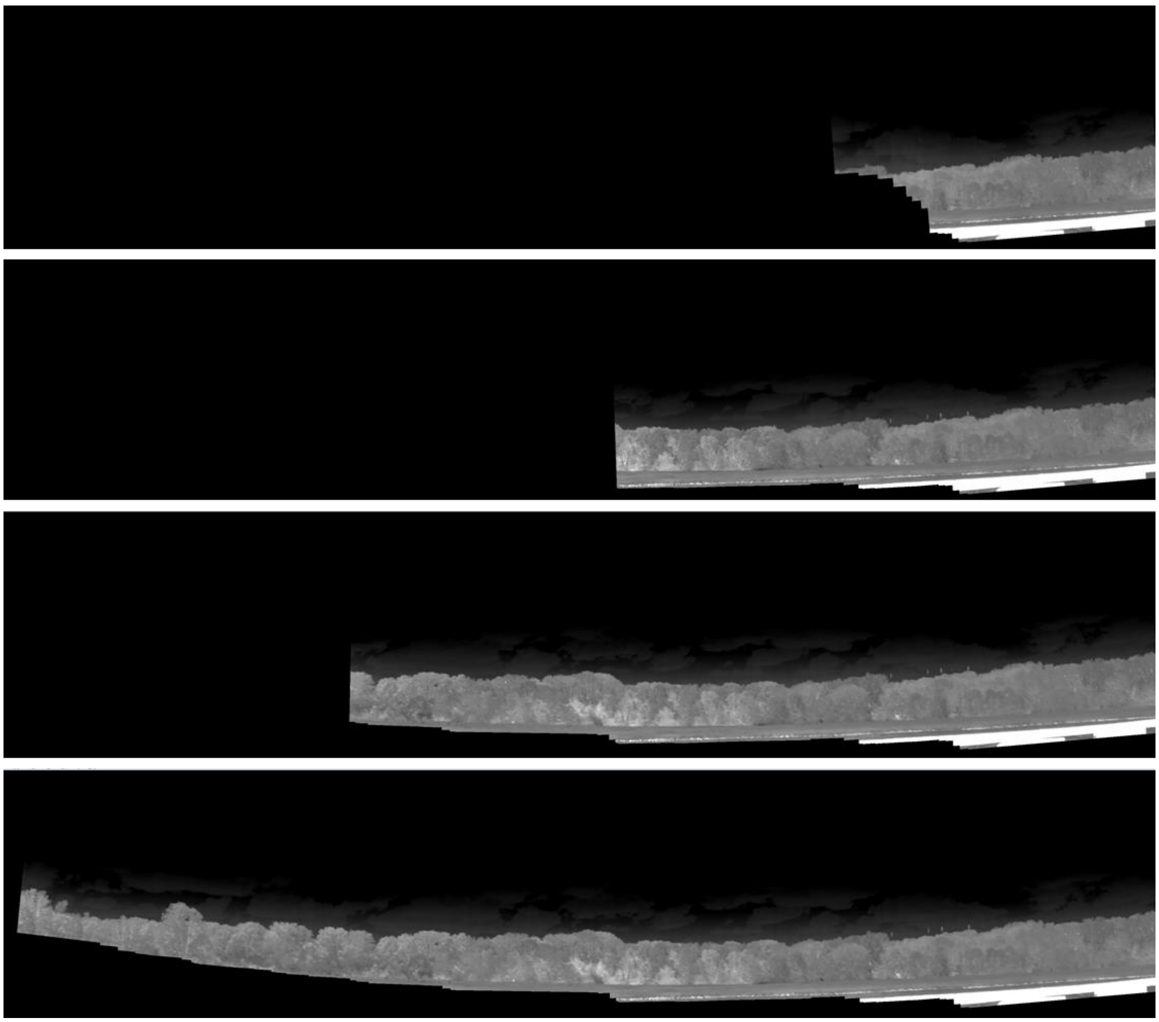


Figure 15. Scene-based pair-wise image warping applied to PIRATE LWIR sensor data.

The majority of this data consists of small remotely controlled aircraft flying in a wooded area. Some imagery shows the aircraft easily visible against the cold sky background. Other examples in this data show the aircraft nearly invisible against the forest background. This data is expected to provide test scenarios that highlight two work areas: 1) small moving target that is difficult to track against certain backgrounds; and 2) demonstrate potential benefits in tracking performance when polarimetric phenomenology is leveraged to increase target-to-background contrast.

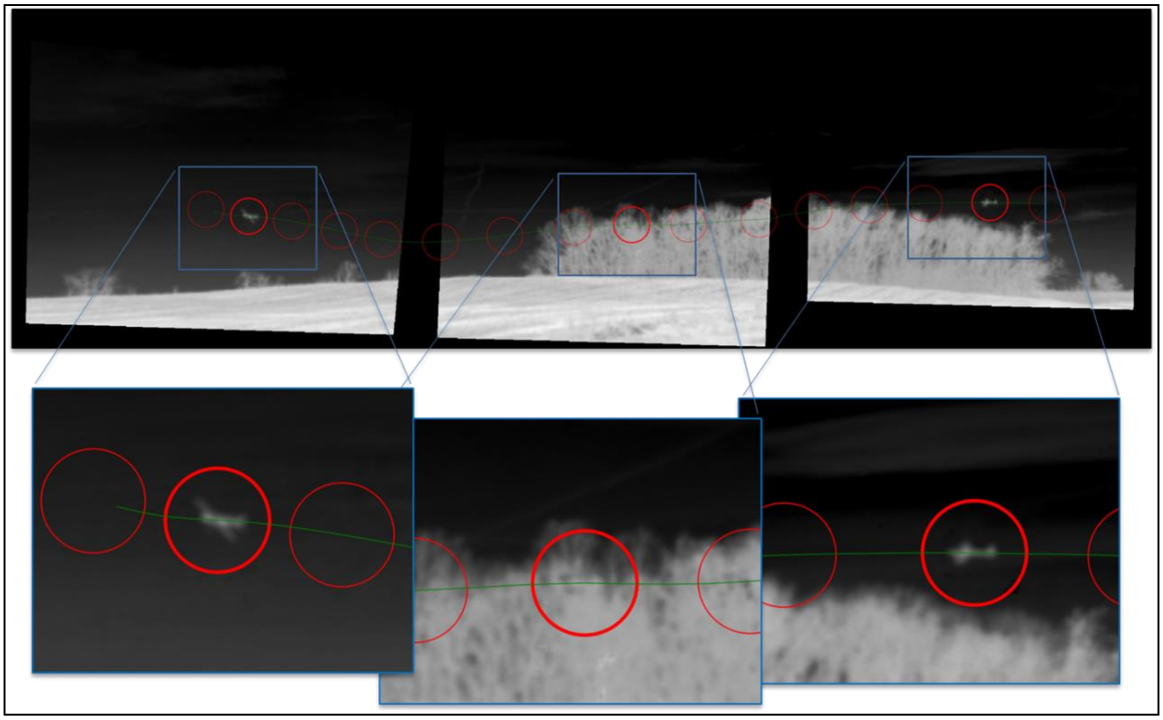


Figure 16. PIRATE S0 imagery analogous to LWIR broadband. Tracking airplane through vegetation clutter is difficult dues to reduced thermal contrast.

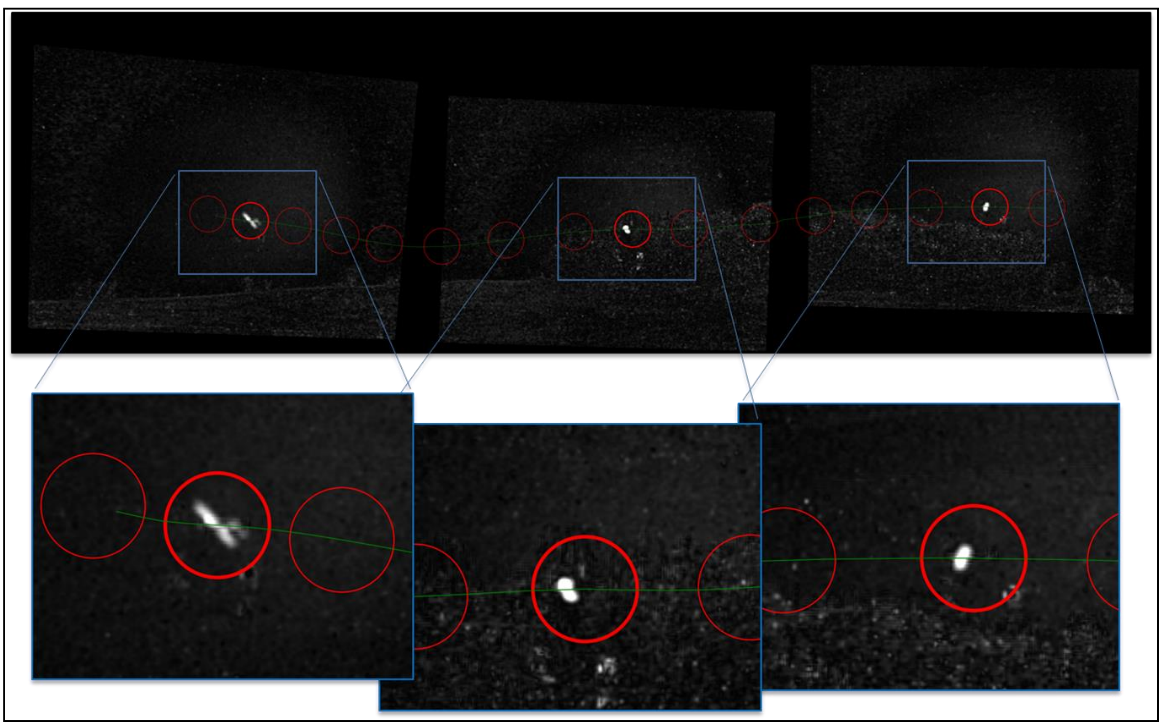


Figure 17. PIRATE Degree of Linear Polarization (DoLP) processing greatly enhances moving target tracking performance.

This particular dataset involved observations of a small aircraft flying in front of a cluttered vegetation background. This data offered an opportunity to investigate the potential gain in target-tracking performance obtained from processing the aircraft's polarization signature. Camera motion was accounted for by computing the homographic image transform between all adjacent image pairs. These transforms were chained back-to-back in order to derive the global transform required to warp each image to the coordinate system of a user-selected reference image.

Pairwise image warping homographies were derived from analysis of in-scene feature point coordinates determined using the SURF algorithm [2]. The SURF image feature is a 64-dimensional vector derived from the local area histogram of image gradients. The homography transform between image pairs is computed using a fast implementation of the RANSAC algorithm [3]. This solution method is very efficient at rejecting feature correspondence pairs that do not agree with the expected image-to-image projection transform. Results of computing and applying image-to-image homography transforms are show above in Figure 15.

3.2 LWIR Polarimetric Target Tracking

A baseline tracker algorithm was applied to the airplane target viewed in the thermal broadband imagery and with the DoLP polarimetric image data product. Preliminary results show qualitative improvements in tracking performance as shown below in both Figure 16 and Figure 17. These figures show three individual frames selected from the original video sequence of several hundred frames. Notice that the airplane is virtually undetectable in the S0 broadband data as shown in the center inset of Figure 16. The corresponding image frame in Figure 17 shows an easily detectable target.

We applied an alpha-beta tracker [4] [5] to a sequence of 122 corresponding S0 and DoLP image frames. The alpha-beta tracker is based upon a first-order kinematic prediction model, i.e., it assumes that objects of interest travel at a constant velocity. Target nominations are performed through sequential image differencing after scene-based registration. It is worth noting that we applied the tracker to single-channel DoLP imagery, rather than the full linear Stokes vector data, due to time constraints, although we have plans to pursue such a modified tracker in the future. Alpha-beta tracker performance is directly tied to the quality of the frame-to-frame image registration. All registration was performed on a frame-to-frame basis using the following procedure:

1. **Identify Distinct Image Features**: The SURF feature descriptor algorithm19was applied to two sequential image frames to nominate distinctive spatial regions within each image.
2. **Associate Common Image Features**: Features are compared across the image pairs and similar ones are matched. Image pairs are rejected that violate the expected image-to-image projection transform.
3. **Compute Projection Transform**: Matched features are used to compute the projective image transform.
4. **Project Images to a Common Coordinate Space**: Each image is warped onto a common coordinate space according to the estimated projection transform.

After registration, the initial position of the target was selected by hand and the alpha-beta tracker was applied independently to the S0 and DoLP image frames. Figure 16 and Figure 17 show a timeline of image frames extracted from each tracker-applied video sequence. In both cases the tracker is performing well, indicated by the target being within the yellow tracker prediction circle. For these aircraft, the strong thermal signature from the aircraft engine appears to provide enough contrast for the tracker to remain locked onto the target in both the S0 and DoLP sequences. However, there is a case where the aircraft turns to fly towards the trees and the thermal S0 signature is not observable by the camera for several frames. Once the aircraft turns and the thermal engine signature is again visible, the tracker reacquires the aircraft and successful tracking resumes [1].

3.3 Joint Spatial-Spectral Target Feature Descriptor

Prior efforts at applying spectral processing methods in support of vehicle tracking have primarily focused on exploiting a single spectrum extracted from on-target HSI data. This tied in very easily with existing algorithms and tools for spectral matched signature processing. The drawback to this simple approach is that the spatial variability of the vehicle's signature was ignored. We have investigated this issue by leveraging state-of-the-art machine vision pattern-matching algorithms. The DAISY spatial feature descriptor [6] in particular showed promise as it enables processing of local image gradient information in a manner similar to hyperspectral imagery. The DAISY algorithm computes an N-dimensional feature vector for each image pixel which is then exploited very much like a traditional hyperspectral data cube. Our effort has focused on a joint spatial-spectral exploitation algorithm by applying DAISY to each band of a HSI cube.

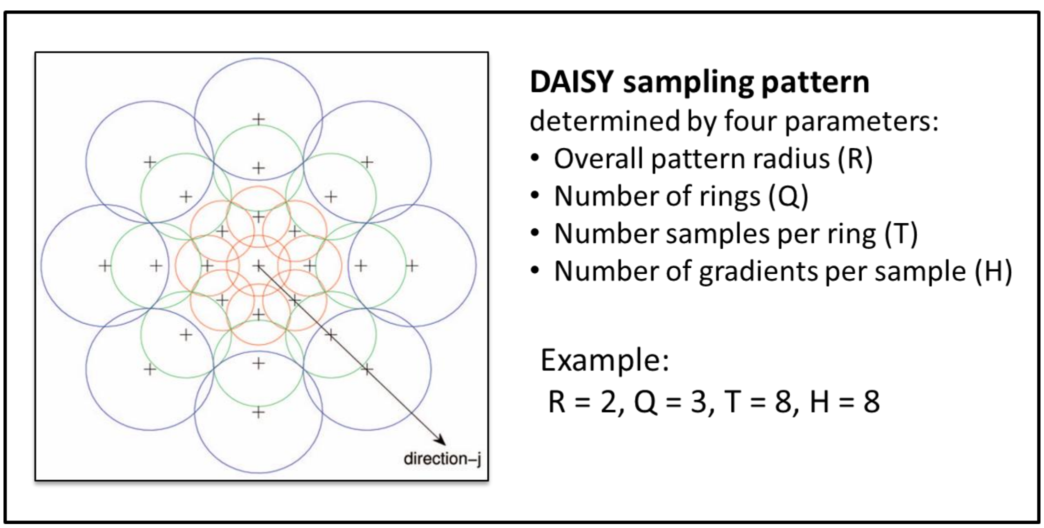


Figure 18. DAISY feature-local sampling pattern.

An example application of DAISY to support spectrally-aided tracking is shown in Figure 19 for HyCAS HSI data collected over a civilian mall parking lot. This figure shows a single target highlighted in the black square on the left panel of Figure 19. The three inset panels show three types of target signature extraction: a) single pixel spectrum per target with 95 spectral bands, b) grey-scale image with DAISY (17 petals) computed at each pixel, and c) joint DAISY and hyperspectral feature with 17 DAISY petals computed for each pixel and each spectral band.

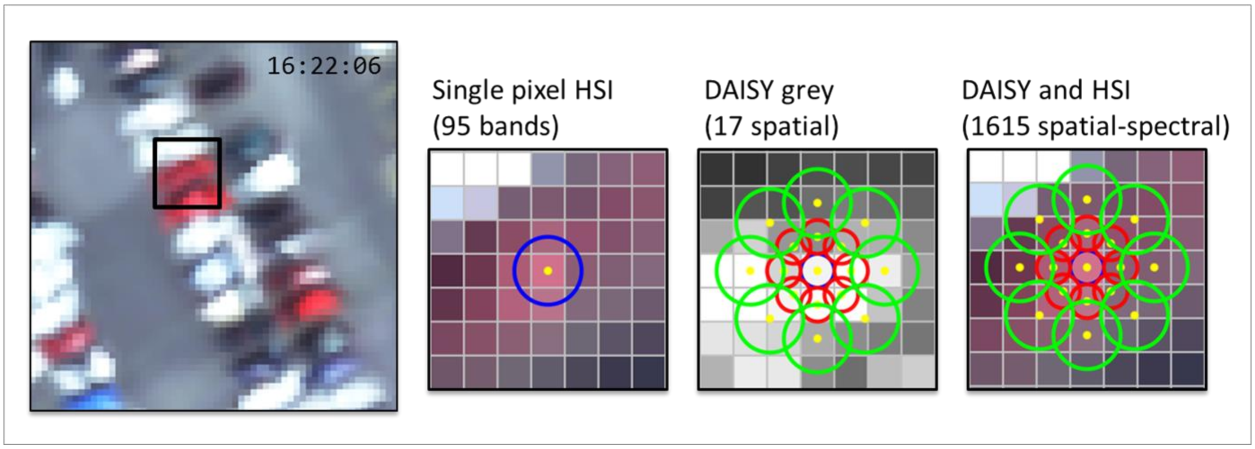


Figure 19. Joint spatial-spectral target feature descriptor using DAISY for each band.

This test data was evaluate for target association performance using a user-defined target truth mask indicating the positions of 180 individual vehicles over a total of four repeat HyCAS flight passes. Track association performance was evaluated by computing the frequency by which each target properly associated with itself in the three alternate looks while using ACE as a spectral matching metric. In this case the term “spectral” is used loosely to include the three feature signature modes defined in Figure 19.

Quantitative performance results [7] [8] are shown below in Figure 20 in the form of probability of detection (or association) versus average number of false alarms. DAISY applied to the grey-scale imagery yielded higher performance than traditional hyperspectral processing, while joint spatial-spectral processing yielded significantly higher performance than either spatial alone or spectral alone.

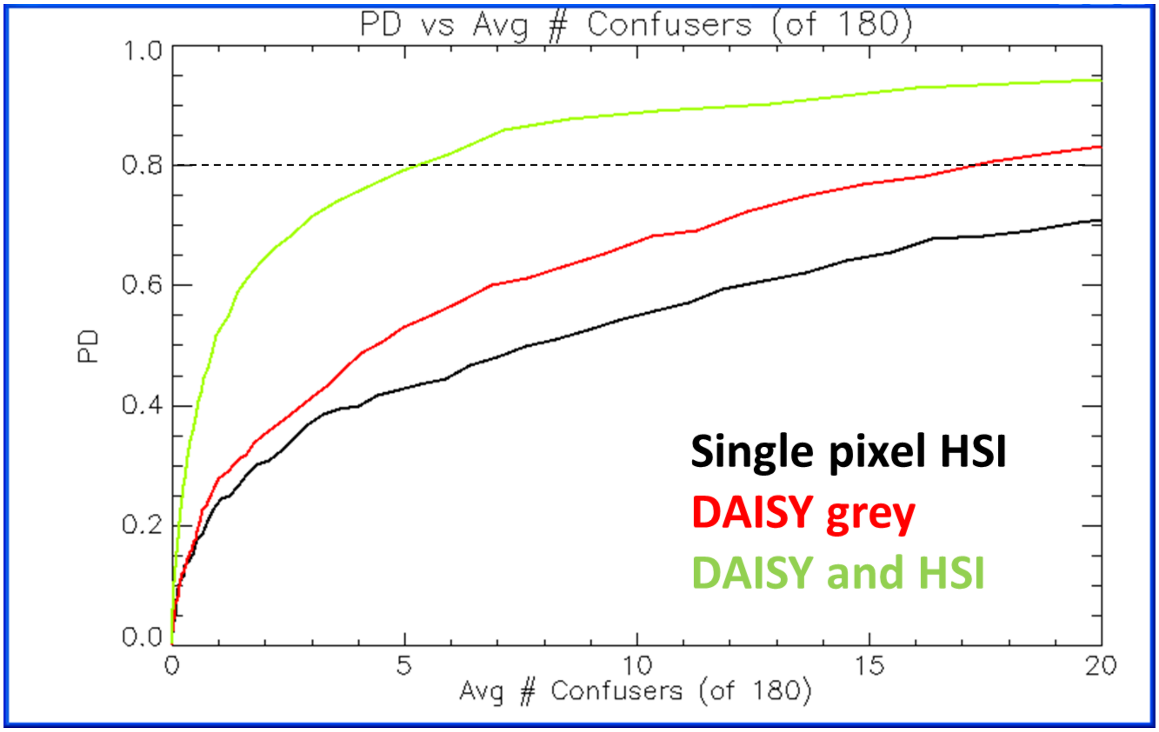


Figure 20. Performance curve comparing spatial versus spectral information.

3.4 Spatial-Spectral Signature versus Spectral Window

The results of the previous section show that joint exploitation of a target's spatial-spectral features has potential to significantly improve vehicle tracking performance. This work leads naturally to pose questions about potential sensor system configurations where practical trade-offs must be made between spatial resolution versus field-of-view versus spectral resolution versus spectral bandwidth. This is tremendously large parameter space, especially given that most fielded systems will likely involve two or more sensors operating collaboratively.

Our initial investigation of this trade space has focused on the relative impact of spectral sampling for a given configuration of the HyCAS sensor, using the same data as in the previous section [9]. A simple scalar measure of performance is derived from the ROC-type curves shown in Figure 20 by computing the area under each respective performance curve. This metric is directly related to overall performance of a particular sensor and/or algorithm configuration. An observed increase in the ROC area metric may be directly interpreted as increased performance for target track association as an average over all vehicles measured in the study.

The spatial-spectral performance analysis of the previous section were repeated for a single band at a time, thereby eliminating spectral correlation as an exploitable feature and focusing on effective of the DAISY feature at each band. This yields the results shown in Figure 21 where the target track association integrated performance is plotted as an independent quantity at each of the 95 HyCAS sensor bands. These results yield insight into the performance trade space as a function of band selection for a given sensor and for a given local spatial feature sampling template. Several interesting observations may be made from this result:

1. The blue and green portions of the visible spectrum yield higher performance versus the red and NIR portions of the spectrum.
2. The SWIR bands between 2.1 μm and 2.4 μm show similar high performance levels, even though these send bands have lower overall SNR levels due to reduced solar illumination levels.
3. Overall performance increases as more bands are added to the process up until a plateau is reached (green curve in Figure 22), beyond which additional bands yields no improvement.
4. The final set of bands which yielded no improvement corresponds to the red and NIR portions of the spectrum.

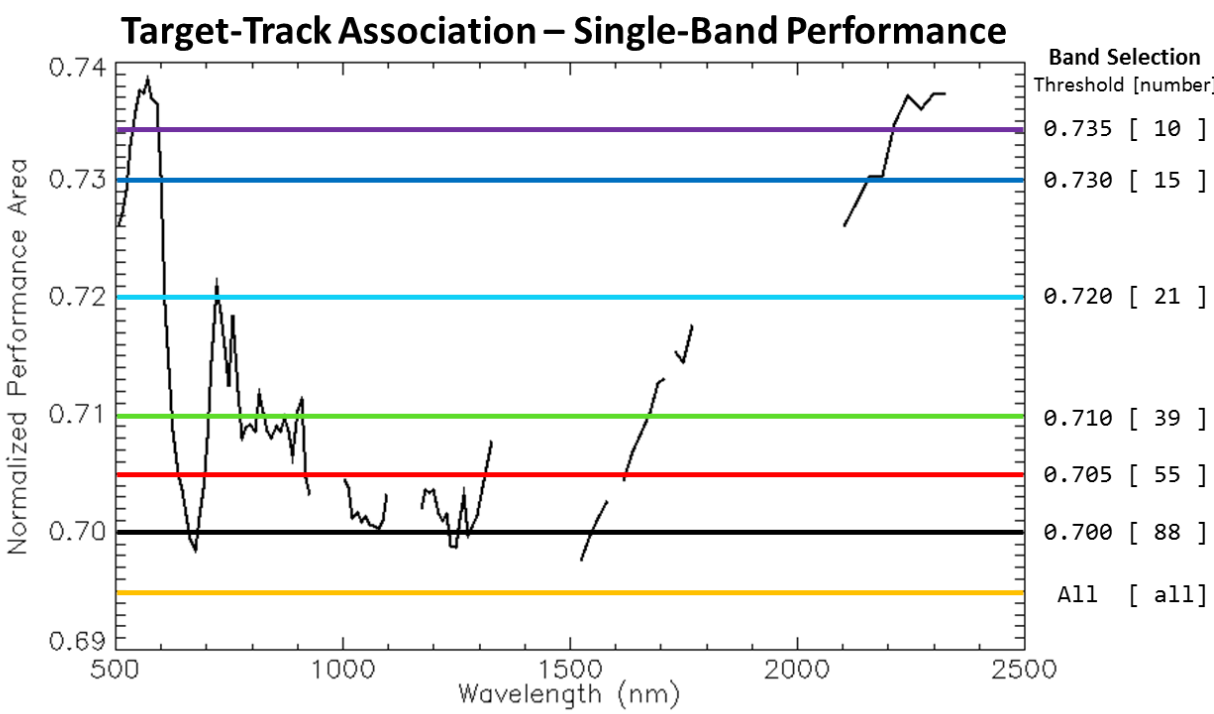


Figure 21. Single-band performance versus multi-band selection threshold.

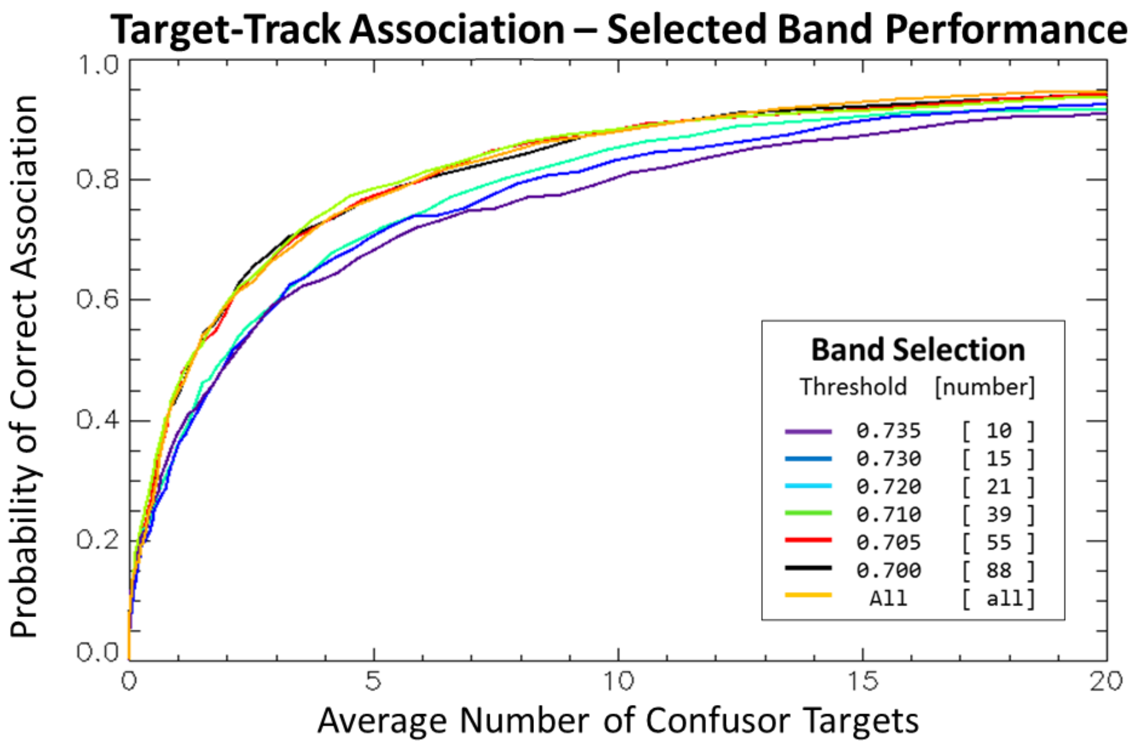


Figure 22. Target association performance versus selected band subset.

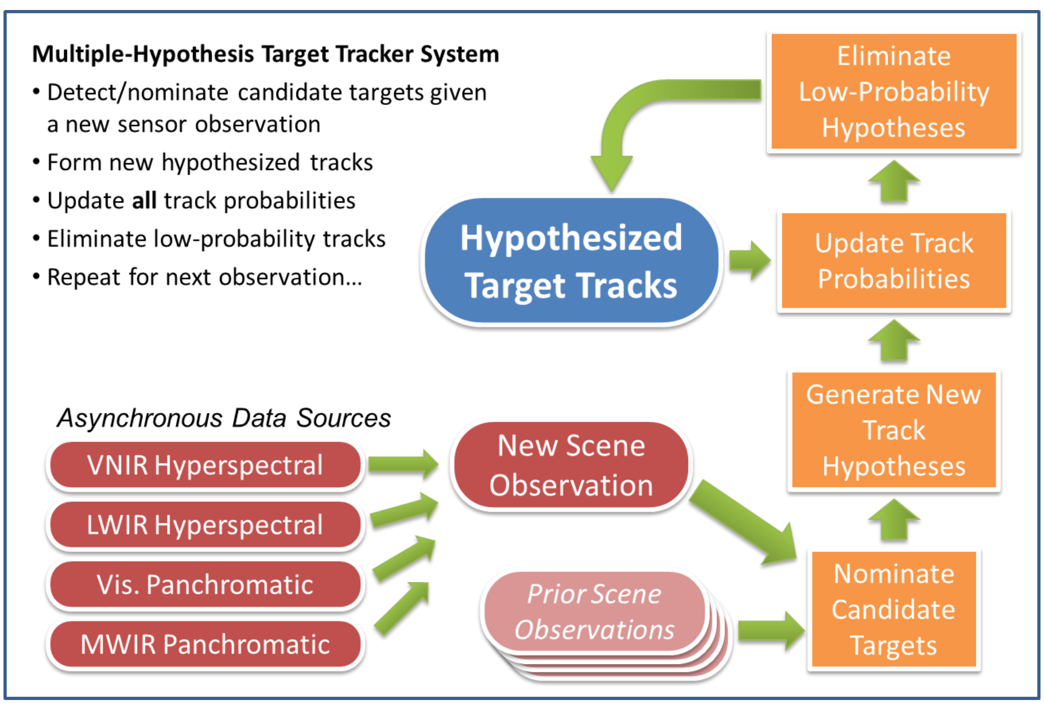


Figure 23. Conceptual framework for a multiple-hypothesis tracker (MHT) system.

3.5 Multiple-Hypothesis Tracker

In the terminology of a Multiple-Hypothesis Tracker (MHT) [10], a hypothesis is the state of motion and color models of all the current tracks along with the history of the tracks gathered so far. When a new measurement is associated to one of the tracks in a hypothesis, the hypothesis is updated. In a standard tracker, a motion model and gating criteria are used to associate a new measurement to one of the existing tracks. If a measurement could credibly be associated to more than one track, the most likely track is chosen.

In a multi-hypothesis tracker, if a measurement could reasonably be associated to more than one track, the hypothesis is forked. Multiple child hypotheses are generated. Each hypothesis is assigned a scalar likelihood that all its tracks are appropriately associated. This overall likelihood number is based on how well the new measurement fits the prediction of the motion and color models for the track to which it was assigned. With each new set of observations, the child hypotheses are ranked in order of probability so that unlikely hypotheses can be removed. To compute the likelihood that a hypothesis could occur requires a model. The model used computes the product of the probabilities that each new observation has been appropriately assigned to a false alarm, a new track, or to a continuation of any of the previous tracks.

3.6 Particle-Filter Tracker

While the MHT framework described above is convenient for managing numerous targets over extended periods of time, a tracker based upon the particle filter (PF) concept has recently proven successful as a powerful tool for numerically propagating a target's model state vector and making comparisons with newly-nominated candidate targets. We have performed preliminary evaluation of several methods for implementing particle filters. The basic concept is illustrated in Figure 24 where at each time step two basic calculations are performed:

1. Monte-Carlo forward propagate large number of random realizations (particles) of current target (dark red points in Figure 24).
2. Weight each particle by the likelihood of that model instance versus the new data observation. Reject those particles with low likelihood (light-grey points in Figure 24) and accept only those particles with high likelihood (bright red particles in Figure 24).

The mean value of the final accepted particles represented the expected state of the target given the new data observation. The Monte-Carlo approach behind particle filters enables complex model not easily adapted to analytical probability distributions. In other words, the particle filter enables joint combination of numerous target parameters and descriptive features. Finally, the numerical values output by each particle filter are readily ingested by the MHT system described above.

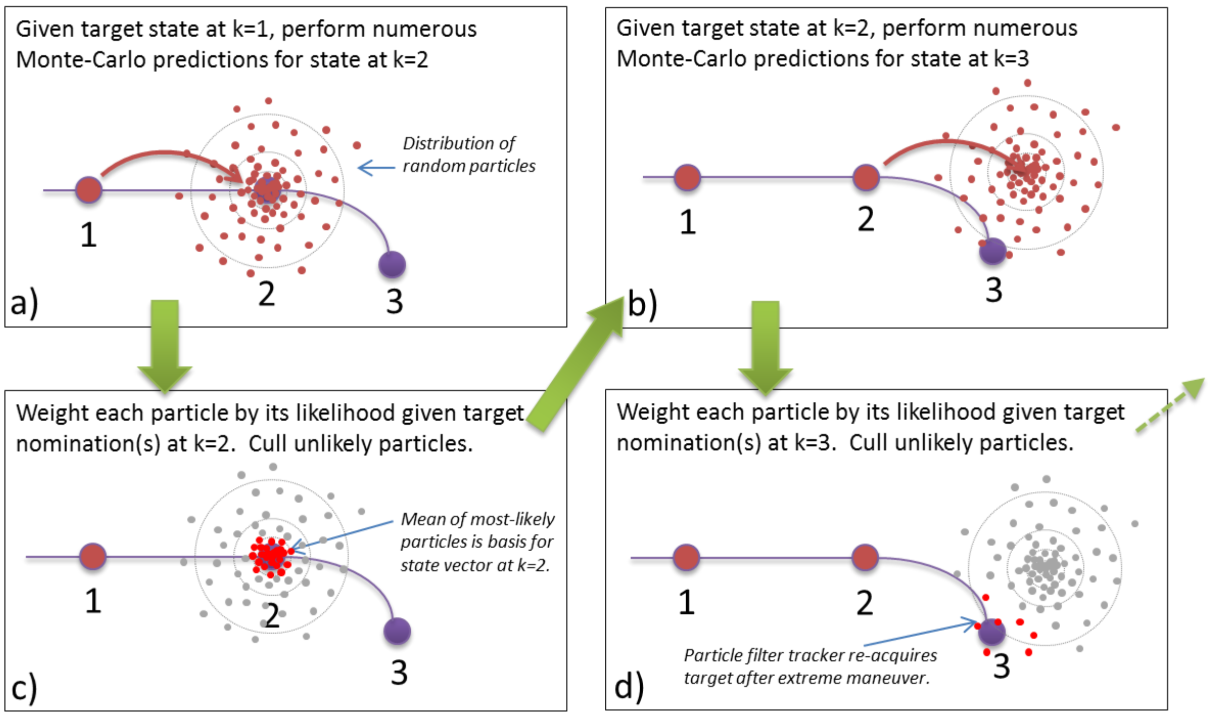


Figure 24. Example scenario illustrating a particle filter-based tracker system.

4. Software Tracker Framework

This section describes the functionality and usage of the Hyperspectral Assisted Tracking (HAT) Suite for investigating spectral-aided tracking methodologies and providing a real-time interface for a tracking display. This foundation architecture supports adaptation to a wide scope of future sensor conops configuration. Figure 25 illustrates the four initial functional requirements as a data-flow diagram:

1. Data source:
   - Ingest sensor image data from one or more sources
   - Trigger main event loop
   - Data preprocessing (e.g. geo-registration)
   - Deliver scene image update to tracker module
2. Target tracker:
   - Accept new images from data source
   - Read prior results from storage
   - Application logic and tracker algorithms
   - Target nomination and track association
   - Maintain target(s) state information vs. time
   - Write new results to storage

3. Data storage:
    - Practical archive for imagery and track results
    - E.g. File system structure, or SQL database
    - Receives data from tracker
    - Sends data to tracker and user control
    - Contains sufficient information to "replay" tracking scenario
4. Use interface:
    - Primary display of tracking results
    - Information update triggered by data source event loop
    - Primary interface for algorithm & sensor control parameters

This portion of our Phase II SBIR effort has focused on the clear definition of software interfaces for the three large green arrows shown in Figure 25 where data is passing from one module to another. The software functionality within each of these components is extended from previous government-funded efforts. The notable exception in this case is the implementation of our most recent advances in joint spatial-spectral target feature exploitation.

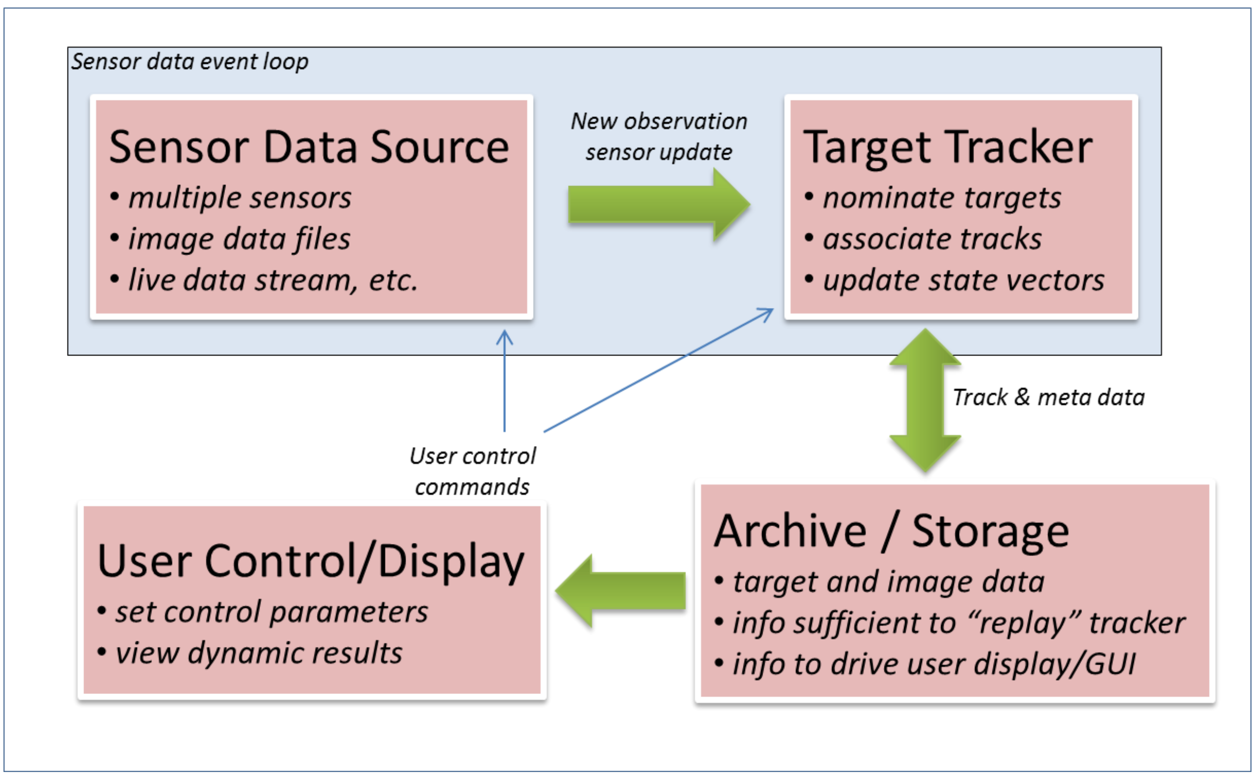


Figure 25. Prototype tracker software architecture with four independent components: 1) Data source, 2) Tracker, 3) Data storage, and 4) User control.

A significant amount of our task for software prototype development has focused on the target tracker module. This involves adapting the spatial-spectral tracker to the multiple-hypothesis tracker (MHT) framework, which was described earlier in concept form in Figure 23. The MHT framework is important in order to manage multiple targets observed over extended periods of time. MHT utilizes a soft-decision approach based on probabilities, where multiple possibilities for the target are maintained for some period of time and then discarded when they are determined to have low likelihood of being correct. The decision is delayed so that more information can be collected to make a better decision. In this framework, a potential track consists of a history of nominated objects across the looks up to the current time. Tracking involves properly assigning nominated objects to the correct tracks.

For this application, a target truth mask is used to initiate target tracks. The first two observations of a target are used to compute the targets initial state. The process of tracking proceeds as follows. The filtered position and velocity of the track from the previous look are used to predict the position of the target in the current look. A nominated object is associated with the track. The filtered position and velocity are then computed to be used in the next look. The final implemented data flow software architecture is shown below in Figure 26.

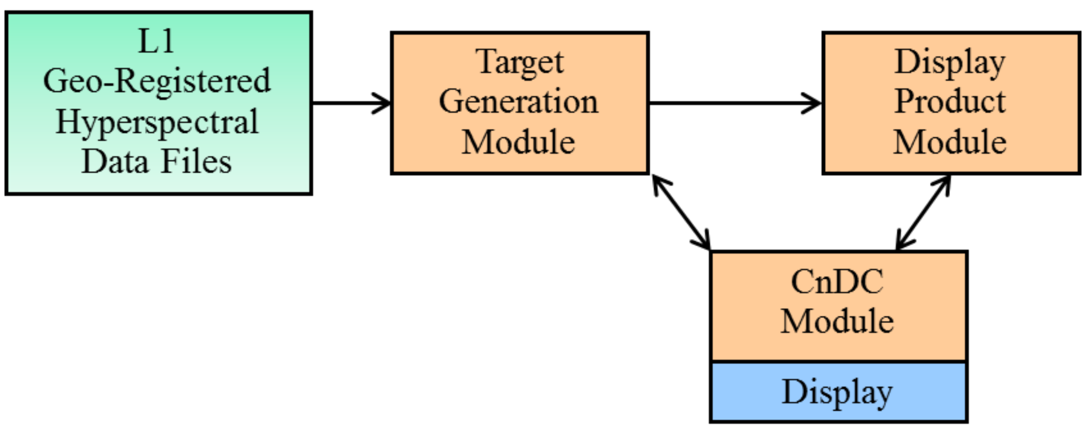


Figure 26. Software data flow for the Hyperspectral Assisted Tracking module.

The software foundation framework consists of three application modules that communicate over TCP/IP for data transfer and control purposes. The Target Generation module reads sensor data and generates target tracks based on currently selected settings. The tracking results along with imagery are then sent to the Product Display module that constructs the RGB images of the tracking results. These results are then passed to the Command and Control Computer (CnDC) for display to the user. The CnDC also allows the user to change the tracking parameters

The modules are set up to operate in the manner in which many real-time systems have already been successfully implemented. The suite operates on a single look of sensor imagery at a time. In principle the target generation module could receive L1 and Geo-Registered data from a real-time system and perform tracking and display in the same manner as from data files with little modification.

One goal of this suite was to make it easier to test new tracking modules. In previous tracking tools, one application would generate the tracking results and another application would display the results. This meant running the target generation application many times in order to establish the best tracking parameters. In the HAT Suite, changes to the parameters are immediately applied and displayed for quick feedback.

The use of TCP/IP networking messages makes solution highly flexible. The modules can all run on separate computers which can be useful if processing power is an issue without resulting in too much latency.

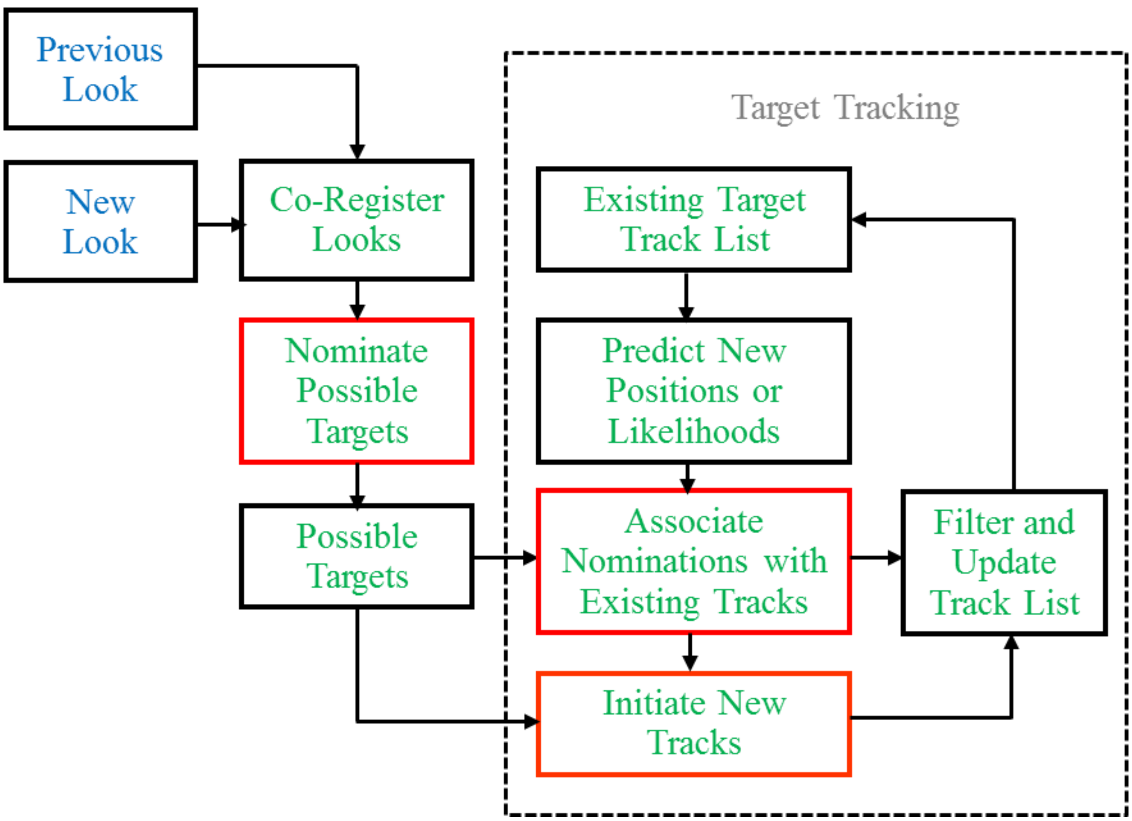


Figure 27. Target tracking algorithm data flowchart.

When a new look of imagery is available it is co-registered with the previous look to put the new look in the same frame of reference as the previous looks. Possible targets are then nominated for tracking. The possible targets are then associated with existing tracks or used to create a new track if no suitable track already exists. The tracks are then updated with the new nomination measurements. A filtered position is calculated which is used to predict the location of the target in the next look. Targets are tracked by maintaining a state of each target in every look. The kinematic model consists of a constant velocity model which is valid if the time between looks is sufficiently low.

The time interval between observations is given by T and the state vector at time nT is given by the expression $\vec{x}(n) = [x(n), \dot{x}(n)]^T$. The constant velocity model is given by $\vec{x}(n) = \Phi\vec{x}(n-1)$, and with $\Phi = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$. An alpha-beta tracker is used to produce a filtered estimate for position based on a new measurements $z(n)$ for each look: $\vec{z}(n) = H\vec{x}(n) + v(n)$, with $H = [1, 0]$ and $v(n)$ is a Gaussian noise process. The filtered estimate based on new measurement and previously filtered estimates is given by $\vec{x}'(n) = [i - K(n)]\Phi\vec{x}'(n-1) + K\vec{z}(n)$. The vector K is defined as $K = \begin{bmatrix} \alpha \\ \beta/T \end{bmatrix}$, where α controls position smoothing, and β control velocity smoothing. Values for α and β each range between 0 and 1.

4.1 Target Nomination by Frame Differences

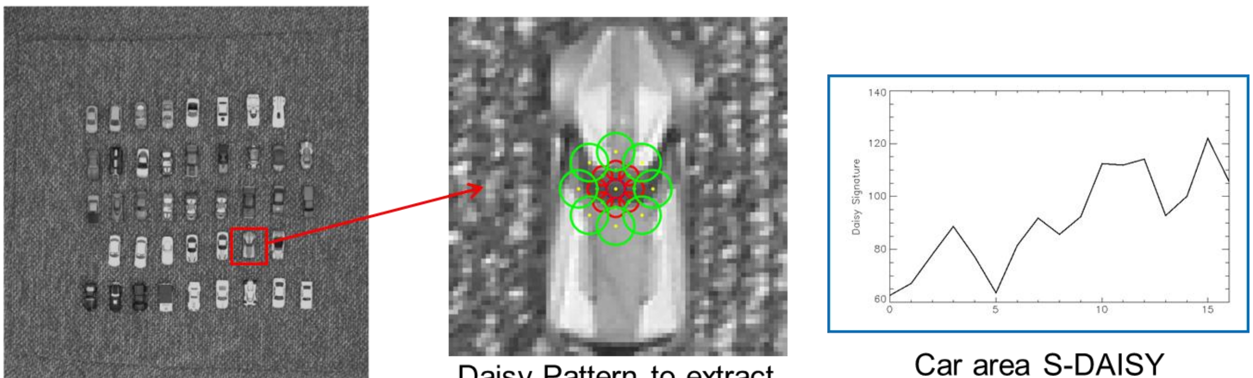
Nominating targets for tracking can be achieved by several methods. This application has the option to utilize frame differencing to identify moving objects and a single band DAISY descriptor which is roughly equivalent to spatial template matching. A standard technique used to identify targets for tracking is by looking at the difference between subsequent frames of imagery to identify the movers as illustrated in Figure 28.

Figure 28. Target nomination via frame difference calculation.

Computing the difference between two looks of imagery in a single band identifies moving objects that can be tracked. The two images on the left are images of a freeway off-ramp taken ~2s apart. The image on the right shows the difference between the two images. The white and black spots show the locations of vehicles in the two looks. A Threshold is applied to provide a list of nominated targets. This method works poorly if the targets are stationary.

4.2 Target Nomination by Single-Band S-DAISY

When there is sufficient spatial resolution of the imagery, the spatial features of the targets can be used to locate the same target in the next look by searching in the vicinity of the predicted position. S-DAISY [7] is a spatial descriptor that can be treated as a signature in signature matched detection as illustrated in Figure 29 and Figure 30.



A single broadband image generated from an RGB image of several toy cars.

Daisy Pattern to extract descriptor for a portion of the car.

Car area S-DAISY signature.

Figure 29. Single Band S-DAISY descriptor is used as a spatial feature signature to find the target in the next look. (Left) A broadband image of a set of toy cars with different spatial features. (Middle) An S-DAISY descriptor centered on one of the cars. (Right) The resulting spatial descriptor signature for the car.

Application of the Single Band DAISY descriptor to detect the same car in a subsequent look using a matched filter. The Matched Filter was computed for every pixel in the scene using the spatial signature of the car from the previous look (Figure 30). The background statistics were computed on spatial cube generated by applying the same S-DAISY pattern to every pixel in the scene. The matched filter shows the strongest detection for the target where the signature was extracted. More spatial pixels on the cars will generally lead to better detection performance.

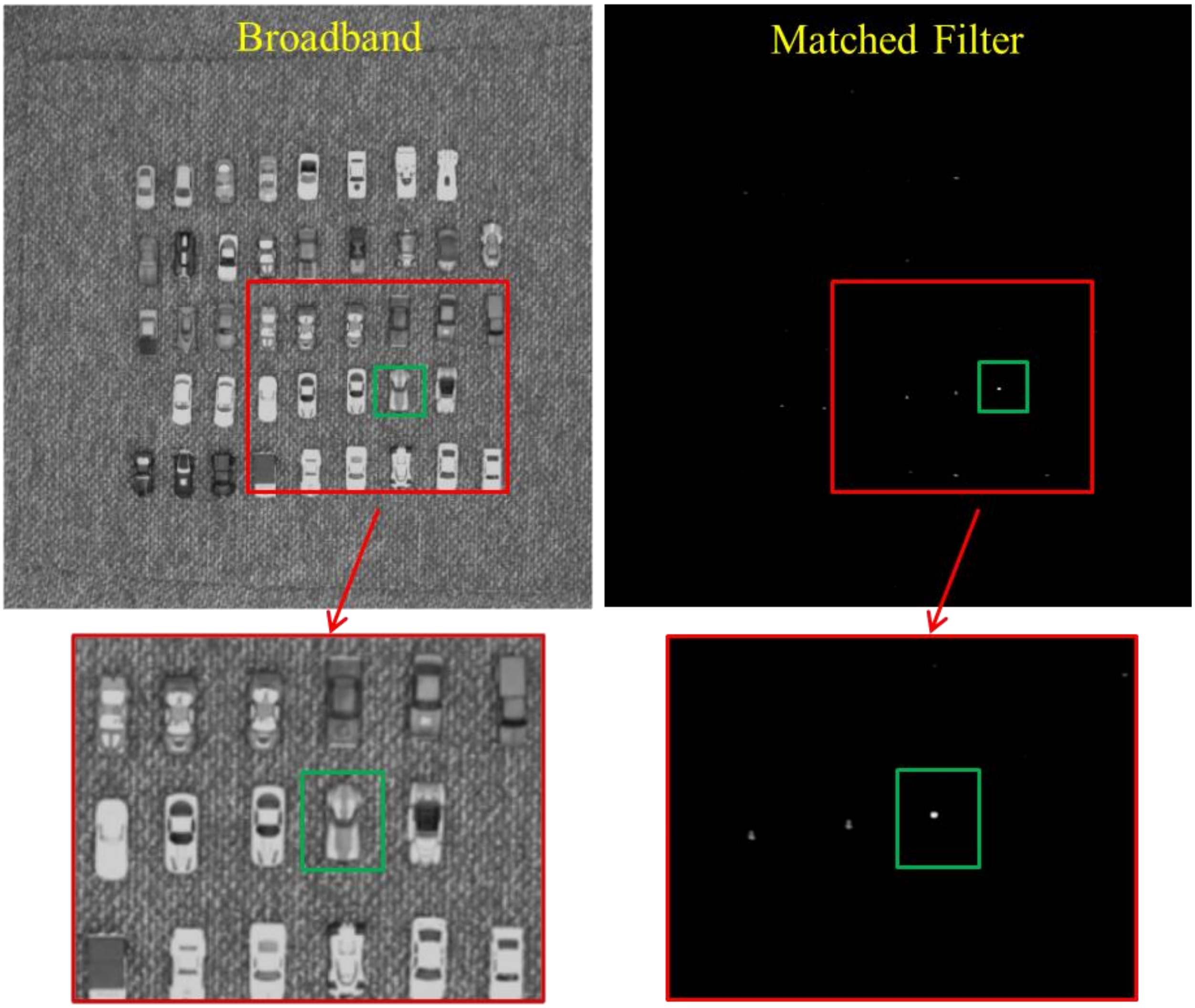


Figure 30. S-DAISY descriptor implemented as detector for nomination.

4.3 Target Association by Nearest-Neighbor

Association attempts to associate the correct nominated target out of several targets with a track. In this application we use the closest target to the predicted position and the most similar target based on the S-DAISY spectral-spatial feature. One of the simplest ways to associate nomination with tracks is to choose the one closest to the predicated track position. This method works well if the time between looks is sufficiently low and the vehicle remains unobstructed from view.

4.4 Target Association by Full DAISY Comparison

A spectral-spatial feature can be used to associate the nomination that has the most similar features to the target in track. For this purpose we use the S-DAISY signatures applied to several bands constructed as shown in Figure 31.

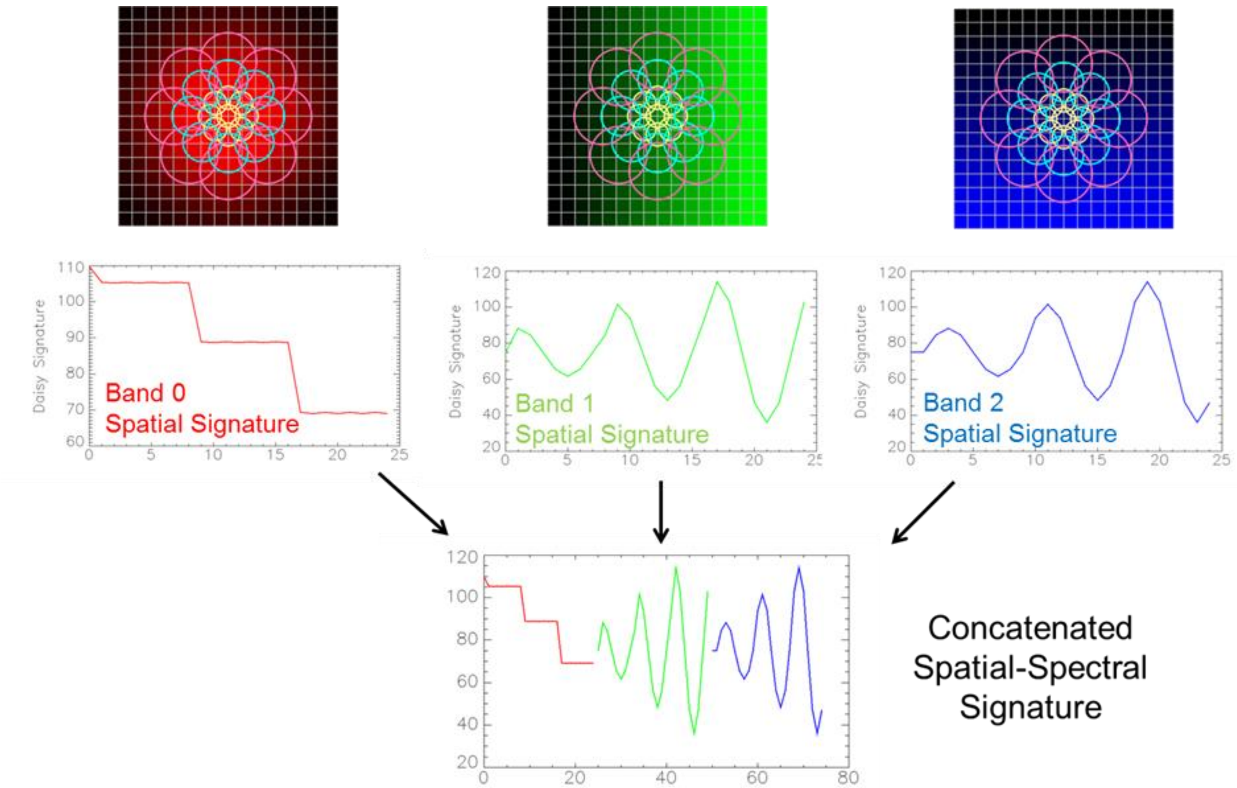


Figure 31. Example construction of the DAISY spatial-spectral signature.

A notional example for producing a spectral-spatial feature for a 3-band image is shown in Figure 31. The target consists of a simple Gaussian in the red band, a gradient in the x-direction in the green band and a gradient in the y-direction in the blue band. The concatenated signatures encode both the spatial and spectral information of object.

Since the signatures contain both the spatial and spectral features of the target we can compare the signatures of targets in one look to targets in another look to find the most similar target across looks. Ideally the targets should have the highest similarity metric with itself. Figure 32 illustrates how the similarity matrix is computed for a group of targets across two looks for re-association.

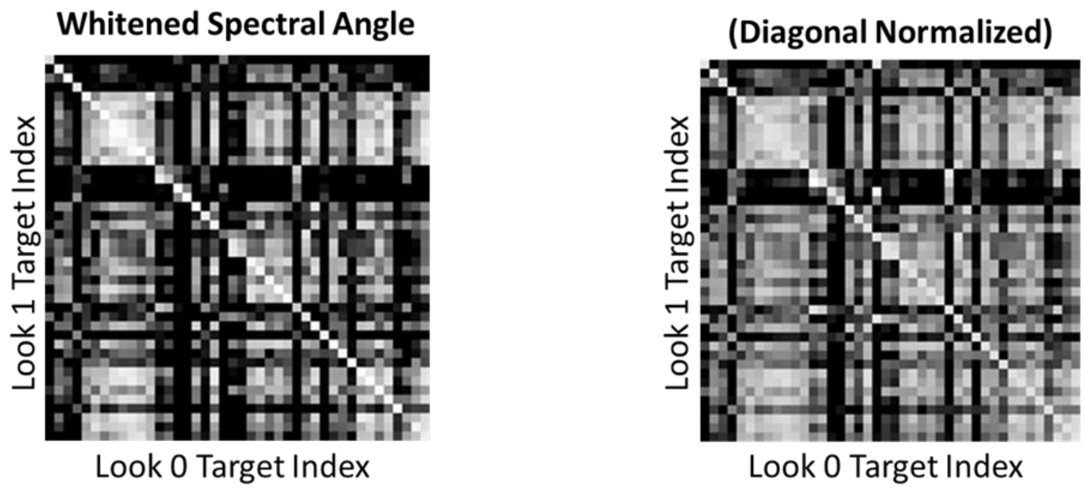


Figure 32. Target similarity metric computed from extracted spatial-spectral feature vectors.

An example target similarity matrix (Figure 32, left) constructed by computing a whitened spectral angle similarity metric between every target in one look with all of the targets in another look using the statistics of the background descriptor cube. The diagonal represents the similarity of the targets with themselves and the off-diagonal represent the similarity of targets with the other targets. Perfect performance occurs when the maximum similarity metric occurs for each target with itself so that no other targets have a higher similarity. By dividing each column of the matrix with the value of the corresponding diagonal the association performance becomes more apparent (Figure 32, right).

In previous work [8] [5] [4] we have shown how this is an effective measure of similarity. The association performance improves with more target spatial resolution and spectral information. In the target tracking application, the similarity metric is computed between the target in track and the nominated targets. The target nomination with the highest similarity metric is assigned to the track.

4.5 Input Data Formats

The modules operate on data that stored on a one look per file basis in a common data directory.

- **Scene Spectral cubes:**

  The scene files consist of a series of ENVI cubes (samples, lines and bands) stored one look per file with the following naming format:

  <SceneFname>XXXX.ext

  The file can start with any valid filename characters. An N-digit number indicating the look index follow the filename start. This ensures that the files are properly sorted based on their look index. The extension does not matter as long the file is ENVI data. An ENVI header is required with each file in with either of the following names to match the scene filename.

  <SceneFname>XXXX.hdr

  or

  <SceneFname>XXXX.ext.hdr

- **Validity Masks:**

  These files contain a mask of the valid imagery pixels with the associated scene spectral cube. These files must contain only one band with a mask of the valid pixels in the scene data. A 0 indicates invalid data, while a 1 indicates a valid pixel.

  The filenames based on the scene must match:

  <SceneFname>XXXX_validity.ext

  An appropriate header with each file is required.

- **Truth Masks:**

  These files contain the known locations of the targets for the first two observations and are used to initiate the tracks tracking. The files are integer format with 0 for the location of no targets and nonzero target IDs for each target. The average position of pixels for each target is used for the position measurements in track initiation. The target IDs must be consistent across looks so the targets can be associated properly across the first two observations. The files can contain truth masks for all looks, but only the first two observations will be used in the tool.

  The filenames based on the scene must match:

  <SceneFname>XXXX_truth.ext

  An appropriate header with each file is required.

- **Time Maps:**

  The time at each pixel may not be the same depending on how the spectral data cubes were acquired. For a line scanning system, the time at each position in the cube will vary since the image is built up by scanning. This difference in time must be appropriately accounted for or tracking performance will suffer. If the time map files are not present, the time of the look will be assumed to correspond to the look index. This will be appropriate for data collected with a framing system.

  The filenames based on the scene must match:

  <SceneFname>XXXX_time.ext

  An appropriate header with each file is required.

Example:

A set a data files might look like the following as an example:
scene_data_0000.dat
scene_data_0000.dat.hdr
scene_data_0000_time.dat
scene_data_0000_time.dat.hdr
scene_data_0000_truth.dat
scene_data_0000_truth.dat.hdr
scene_data_0000_validity.dat
scene_data_0000_validity.dat.hdr
scene_data_0001.dat
scene_data_0001.dat.hdr
scene_data_0001_time.dat
scene_data_0001_time.dat.hdr
scene_data_0001_truth.dat
scene_data_0001_truth.dat.hdr
scene_data_0001_validity.dat
scene_data_0001_validity.dat.hdr

4.6 Application User Interface

The main application control panel is shown below in Figure 33. The file menu is used to load the data for processing and exiting the program. The selection **Open** opens a file dialog for selection of the spectral imagery for processing. The user chooses a single spectral cube file and the application determines all of the associated files. The selection **Exit** closes the application.

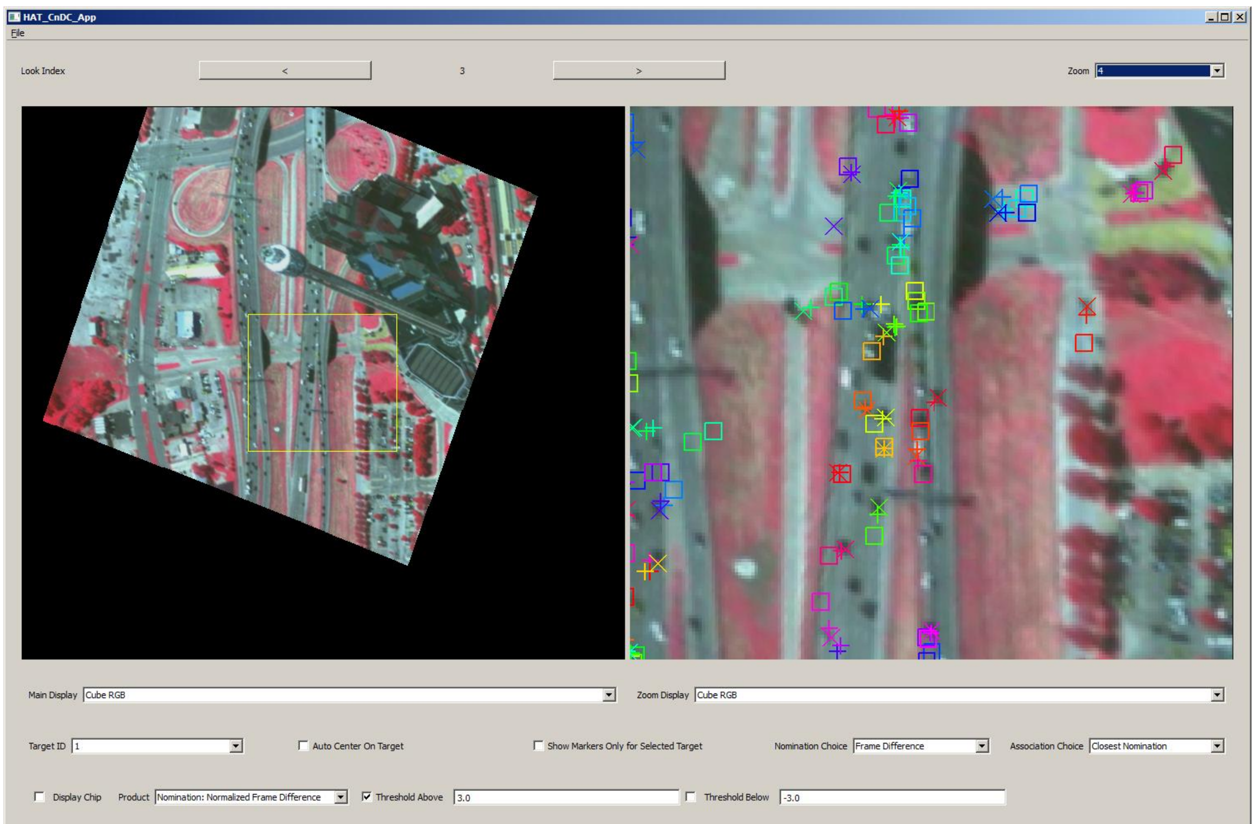


Figure 33. Application primary control panel.

The left window of the control panel above shows the full data imagery, while the right window shows a zoomed in section with overlaid tracking results. The arrow keys control cycling through the looks via the look index control panel. When the look index is incremented, tracking is performed with the currently selected options. When the look index is decremented, the previous track states are loaded and no tracking is performed. The current look index is displayed between the arrows that control the look index. The yellow box in the left window represents the currently selected region of interest. The target markers in the right window consist of a measured position (X), a predicted position (box), and a filtered position (plus). The color of the markers is consistent for each target from look to look. The shown displays are false-color RGB images of the input spectral data. The size of the zoom window is controlled by the dropdown list at the top-right of the control panel.

An optional display of the normalized difference image that can be displayed in the right control panel window is shown below in Figure 34. This shows the difference image computed between the two looks along with overlaid target markers. The bright and dark blobs are the basis for target nomination via frame difference (Section 4.1).

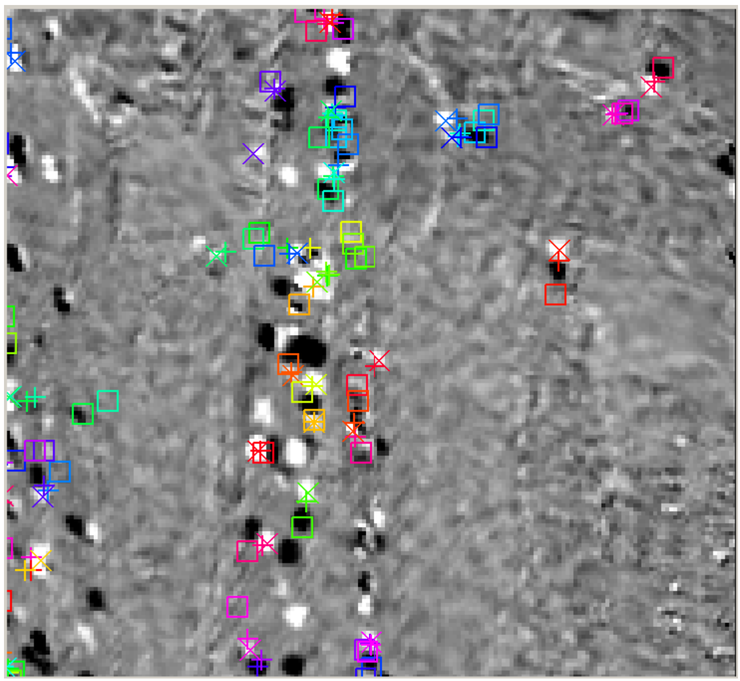


Figure 34. The zoomed image showing the normalized difference data product along with target markers.

The nomination metrics are threshold to identify objects for tracking. The parameters set the threshold values. The options allow for a two sided threshold where pixel above a certain value are kept and below another value are kept. This option is useful for the frame difference nomination. If no thresholding is chosen then the tracking results will result in the targets remaining in their constant velocity predicted states without updating with new measurements. These thresholded values affect the tracking in addition to the various chip display options.

5. Conclusions

Long-term continuous moving-target surveillance from airborne electro-optical sensors provides critical information for tactical awareness situations. Tracking civilian vehicles in urban environments is a challenging problem for existing systems, which generally rely on high-resolution video imagery to identify targets by their spatial characteristics. It is difficult for current spatial-based trackers to re-acquire target lock once the subject has been obscured from view for even moderate lengths of time.

The goal of this effort has been to demonstrate the exploitation of a vehicle target's unique spectral signature as a means to improve overall tracker performance. An example solution is to leverage the combined strengths of both spatial and spectral feature matching capabilities into a generalized multiple-hypothesis tracker system. This effort has demonstrated effective joint exploitation of spatial-spectral target features via the S-DSAISY algorithm. A foundational software framework was developed that will enable future exploration of additional target tracking algorithm concepts.

6. References

[1] B. M. Ratliff, D. A. LeMaster, R. T. Mack, P. V. Villeneuve, J. J. Weinheimer and J. R. Middendorf, "Detection and Tracking of RC Model Aircraft in LWIR Microgrid Polarimeter Data," in *Proc. SPIE Remote Sensing*, San Diego, 2011.

[2] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, "SURF: Speeded Up Robust Features," *Computer Vision and Image Understanding (CVIU),* pp. 346-359, 2008.

[3] M. Fischler and R. Bolles, "Random Sample Consensus (RANSAC): A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Comm. of the ACM 24,* pp. 381-395, 1981.

[4] J. J. Weinheimer and S. G. Beaven, "Hyperspectral Assisted Target Tracking," in *Proc. Military Sensing Symposia on Camouflage, Concealment and Deception*, Orlando, 2007.

[5] J. J. Weinheimer and S. G. Beaven, "Development and Evaluation of Hyperspectral Assisted Target Tracking," in *Proc. Military Sensing Symposia on Camouflage, Concealment and Deception*, Las Vegas, 2008.

[6] Tola, Lepetit and Fua, "A fast local descriptor for dense matching," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2008.

[7] J. J. Weinheimer, P. V. Villeneuve and S. G. Beaven, "Spectral DAISY: A Combined target Spatial-Spectral Dense Feature Descriptor for Improved Tracking Performance," in *Proc. SPIE, Signal and Data Processing of Small Targets*, 2011.

[8] J. J. Weinheimer and S. G. Beaven, "Spectral Feature-Aided Multi-Hypothesis Target Tracking for Airborne HSI Data," in *Proc. Military Sensing Symposia on Passive Sensors*, Orlando, 2010.

[9] J. J. Weinheimer, S. G. Beaven and P. V. Villeneuve, "Optimal Band Selection for Target Association in HSI Data," in *Proc. Military Sensing Symposia on Passive Sensors*, Pasadena, 2012.

[10] M. L. Gran, "Color-augmented target tracking using a liquid crystal tunable filter camera," in *Proc. SPIE*, Orlando, 2009.