Efficient Integration of Spectral Features for Vehicle Tracking Utilizing an Adaptive Sensor

Burak Uzkent[a], Matthew J. Hoffman[b], and Anthony Vodacek[a]

[a]Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY, USA

[b]School of Mathematical Sciences Department, Rochester Institute of Technology, 85 Lomb Memorial Drive, Rochester, NY, USA

ABSTRACT

Object tracking in urban environments is an important and challenging problem that is traditionally tackled using visible and near infrared wavelengths. By inserting extended data such as spectral features of the objects one can improve the reliability of the identification process. However, huge increase in data created by hyperspectral imaging is usually prohibitive. To overcome the complexity problem, we propose a persistent air-to-ground target tracking system inspired by a state-of-the-art, adaptive, multi-modal sensor. The adaptive sensor is capable of providing panchromatic images as well as the spectra of desired pixels. This addresses the data challenge of hyperspectral tracking by only recording spectral data as needed. Spectral likelihoods are integrated into a data association algorithm in a Bayesian fashion to minimze the likelihood of misidentification. A framework for controlling spectral data collection is developed by incorporating motion segmentation information and prior information from a Gaussian Sum filter (GSF) movement predictions from a multi-model forecasting set. An intersection mask of the surveillance area is extracted from OpenStreetMap source and incorporated into the tracking algorithm to perform online refinement of multiple model set. The proposed system is tested using challenging and realistic scenarios generated in an adverse environment.

Keywords: spectral likelihoods, persistent vehicle tracking, performance-driven sensor, adaptive forecasting

1. INTRODUCTION

Aerial surveillance in urban environments has been extensively studied in the literature. Most of them try to tackle tough tracking problems with limited information. A large volume of studies focus on using Ground Moving Target Indicator radars, which measure the locations of moving ground objects as well as their doppler velocities, to tackle problems in ground target tracking.[1–4] With radar, however, we are limited to kinematic information to achieve persistent tracking, and kinematic data alone is very likely to fail in challenging situations due to occlusions, parallax, clutter, etc. To achieve a more robust system, one can consider utilizing unique fingerprints of objects in addition to the kinematic data.

Most of the narrow-area surveillance studies extract local features from the scene to identify their targets. These features can be texture, color histogram, SIFT, optical flow, edges, corners, etc.[5–7] However, in a wide-area ground-based tracking system, one can not rely on these features due to low resolution. Instead, the hyperspectral data sampled from a single pixel of a target of interest (TOI) can be beneficial since it provides unique fingerprint. Hyperspectral imaging involves of acquiring data in hundreds of narrow adjacent spectral bands. Thus, by inserting extended data, classification of objects can be further improved. It is impossible to transmit the volume of data from a conventional full hyperspectral sensor in real time tracking. However, with the recent advancements in the sensor technology, it becomes possible to quickly collect only desired spectral data.[8] As a result, such a sensor in fact can be employed in a real time tracking system.

In this paper, a persistent target tracking system utilizing an adaptive optical multi-modal sensor is proposed. Since some portion of spectral data to be recorded is determined by the performance of the filter, we incorporate

Further author information: (Send correspondence to Burak Uzkent)
Burak Uzkent: E-mail: bxu2522@rit.edu, Telephone:+1 585-301-0989

Video Surveillance and Transportation Imaging Applications 2015, edited by Robert P. Loce, Eli Saber,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 9407, 940707 · © 2015 SPIE-IS&T
CCC code: 0277-786X/15/$18 · doi: 10.1117/12.2082266

an intersection mask of the scene where the true target probability density can be multi-modal. This way we can gain prior knowledge of a possible maneuver and adaptively tune the transition models. However, abrupt motion changes might still yield to highly nonlinear true target probability densities. In these cases, it is extremely challenging to collect spectra at the TOI pixels by relying on prediction-based spatial sampling. Hence, we perform measurement-based sampling in addition to the prediction-based sampling to maximize the likelihood of collecting data from the TOI.

2. ADAPTIVE MULTI-MODAL SENSOR AND ITS MANAGEMENT

A sensor capable of collecting spatial and spectral data is required to perfom high rate tracking of objects of interest. For this reason, Rochester Institute of Technology Multi-object Spectrometer (RITMOS) is considered as an adaptive performance-driven sensor together with a performance-driven algorithm to detect, identify and track targets in highly cluttered scenes.[8] It utilizes a micromirror array to reflect the light to the one of two sensors; spectrograph and panchromatic channel. The switch from panchromatic to spectral data mode for a pixel can happen very fast due to the compactness and speed of micromirror arrays. To capture the panchromatic image of the scene, an array of micromirrors reflect the light to a panchromatic imaging array. Individual micromirrors imaging the object are then tilted to reflect the light to collect the full spectrum of a specified pixel. All these configurations can be performed via a user controlled or automated computer interface.

The performance-driven tracking algorithm needs to be designed in a way that it meets the specifications of RITMOS. It requires about $0.1 - 0.125$ s to obtain panchromatic image of a scene by RITMOS. On the other hand, the full spectrum of a single pixel in the visible to NIR wavelength takes 1 ms. Spatial and spectral data can not be collected simultaneosly. In addition, it does not allow selection of conflicting pixels. In other words, a pixel, that is in the same column with an already selected pixel, can not be fed to RITMOS. The physical models outperforming the others in terms of target probability density approximation can be prioritized for spectra acquisition. The frame rate 1 s is used in this study, meaning that sufficient track estimates are updated at every second. Panchromatic images are used to accomplish motion segmentation. While the sensor collects panchromatic image, the tracking algorithm outputs the prior results that will be fed to the sensor resource management system for spectral data acquisition. Full spectrum data for 100 pixels can be collected in about 0.1 s. The workflow of the tracking algorithm matching these sensor specifications is demonstrated in fig. 1.

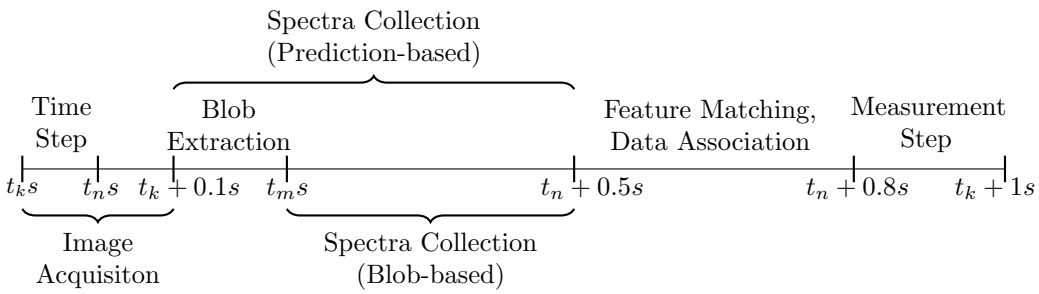


Figure 1: The workflow of the proposed tracking system using a performance-driven sensor.

3. SCENARIOS

To develop and test the system in a controlled environment that allows us a knowable ground truth, we use synthetic imagery generated by the Digital Imaging and Remote Sensing Image Generation (DIRSIG) model.[9] In addition, the Simulation of Urban Mobility (SUMO) traffic simulator has been integrated with DIRSIG to produce dynamic imagery for tracking scenarios.[10]

The motivation for using synthetic data is that, since we know the true positions and characteristics in a synthetic image, we can accurately compute performance metrics for the tracking system. Furthermore,

multiple scenarios and sampling strategies within those scenarios can also be carried out without running multiple experiments. The scenario used in this paper comes from DIRSIG Megascene I, which is built to resemble part of Rochester, NY, USA. The simulation uses hyperspectral imaging from a fixed aerial platform assuming a static sensor mount. The spectral range is 400 to 1000 nm with a spectral resolution of 10 nm. Thus, generated synthetic images have 61 wavelength bands. Eighty six vehicles are placed in the 130 seconds long simulation. Among them, we will focus on tracking four vehicles at separate runs. The first three vehicles are obscured by the tree canopies and shadows about 20% of their travel time whereas the last one is obscured in five frames. Twenty four paint models are assigned to the vehicles with equal probabilities. Overall, the scenarios are generated in a crowded scene with severe occlusions and vehicles with similar paint models.

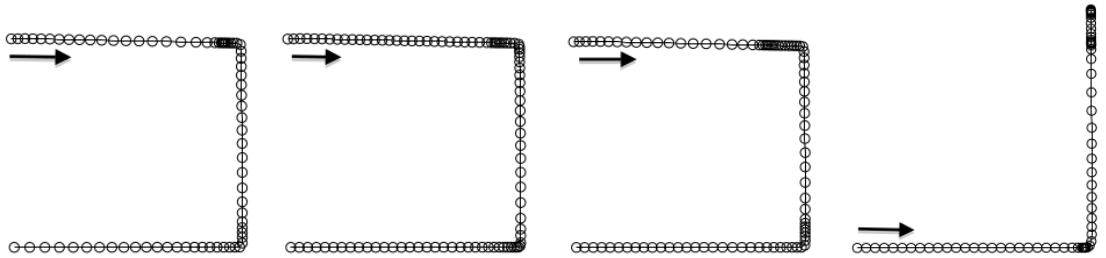


Figure 2: Trajectories of the vehicles to be tracked.

Synthetic spectral data is extremely helpful for the reasons mentioned before. However, it is crucial to simulate data matching real-world phenomenology. DIRSIG output has to be further processed to meet this need. DIRSIG yields sensor reaching radiance output relative to the sensor specifications such as aperture width. The ultimate goal is to process sampled data obtained by a representative of a real world imaging scpectrometer. This process converts the sensor reaching radiance values into the digitally sampled values. In this process, we account for factors that exists in the real-world phenomenology such as filter effects, shot noise, readout noise, integration time, detector elements, and analog to digital converters. An extensive treatment on this radiometric sampling process is given in.[11]

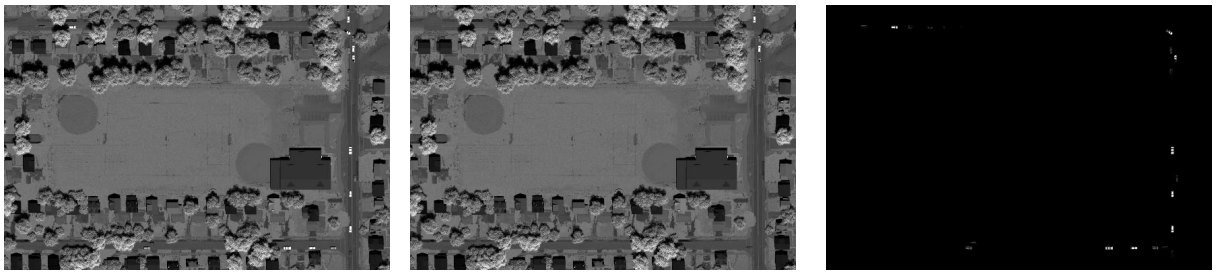
4. MOTION SEGMENTATION

The first fundamental step in most the tracking studies is motion segmentation. The goal is to detect the moving objects in the surveillance scene. This step provides us candidate measurements likely to be associated to the tracks to correct prior estimations. In this study, we use the background subtraction technique on the panchromatic images from RITMOS. Background subtraction is a simple technique that subtracts two frames to detect moving objects. These two frames are the current frame $I(x, y, k)$ at time k and the background frame $I_B(x, y)$ for the given x and y coordinates of pixels. The output $I_D(x, y, k) = abs(I(x, y, k) - I_B(x, y))$ is generaly applied an empirical threshold. We apply the Median filter to the previous $n - 1$ number of frames in addition to the current frame to generate a reference frame I_B as

$$I_B(x, y) = median\{I(x, y, k - i)\}, \quad i = 0, 1, ..., n - 1. \tag{1}$$

The Median filter is a computationally cheap method which makes it eligible for our tracking system. It is especially applicable for background dominated scenes and it perfectly fits in our system since we only perform vehicular simulation. In this study, we consider 9 previous scans in addition to the current scan to generate a background mask. In fig. 3, you can see the estimated background and foreground mask results at a time step in the generated scenario.

The difference image (fig. 3c) is applied a threshold. Next, the resultant binary image is applied the morphological closing operation to remove the noise due to lighting changes, tiny nonstationary objects, and other sources and fill in the gaps. Finally, the connected component analysis is applied to uniquely label extracted blobs.



(a) (b) (c)

Figure 3: (a) Panchromatic image of the scene, I, at k (b) background mask generated by Median filtering multiple frames ($n = 10$), I_B, (c) foreground mask after background subtraction, I_D.

5. GAUSSIAN SUM KALMAN FILTER

The GSF represents a non-Gaussian distribution by a finite mixture of Gaussian distrubutions. It allows us to implement a multiple model set strategy that can be crucial in multi-modal true target probability density cases. The GSF together with intersection network concept might lead to an adaptive multiple model set strategy in the junctions. This way, nonlinear true target probability density in the intersections can be better approximated which in turn results in directing the sensor to collect useful spectral features. The state-space and covariance matrix of each density kernel is updated by a linear or nonlinear recursive estimator such as the Kalman filter (KF), Extended Kalman filter (EKF), or Unscented Kalman filter. We have both linear and nonlinear models, thus, the KF will be used to transit Gaussians with linear models whereas the EKF is used for nonlinear ones. The EKF represents a nonlinear model by linearizing it based on the first two terms of the Taylor expansion series. Large linearization error might occur in highly nonlinear systems. However, the EKF fits well in this study since we use mildly nonlinear motion models.

Assume we have a process and observations, X_{k-1} and $Y_{k-1} = \{y_{k-1}|i = 1, 2, ..., k-1\}$, the uncertainty $P_{k|k-1}$ associated with the state vector $X_k = [c_x, c_y, w, h, V_x, V_y]$ is determined by $p(X_k|Y_{k-1})$. In the state vector, c_x and c_y represent the coordinates of a target, whereas V_x and V_y denote the horizontal and vertical velocities and w and h denote the width and height of a target. The GSF estimates the prior mean and uncertainty of the mixture as

$$x_{k|k-1} = \Sigma_{n=1}^{M} w_{k-1}^{n}(X_{k-1|k-1}^{n}), \tag{2}$$

$$p_{k|k-1} = \Sigma_{n=1}^{M} w_{k-1}^{n}([P_{k-1|k-1}^{n}(X_{k-1|k-1}^{n} - x_{k|k-1})(x_{k|k-1} - X_{k-1|k-1}^{n})^{T}], \tag{3}$$

where $x_{k|k-1}$ and $p_{k|k-1}$ are the expected mean and covariance of the prior mixture and w_{k-1}^{n} and M represent the weight for the nth Gaussian kernel and number of kernels. The expected mean and covariance of the posterior density $p(X_k|Y_k)$ is estimated in the same way.

We need to use an appropriate number of Gaussians to ensure robust tracking in challenging cases. A larger number of components can improve tracking, however, it also presents undesired complexity. Therefore, we want to use an optimum number of Gaussians in terms of target motion density approximation and complexity burden. A more detailed documentation of the GSF in a similar adaptive multi-modal sensor-inspired tracking system is explained in.[12,13]

6. SPECTRAL FEATURE-AIDED IDENTIFICATION

Data association is required to successfully correct prior target probability density $p(X_k|Y_{k-1})$ with observations Y_k. In this study, we propose a novel two step data association method. In the first step, we rely on the strength of spectral features to eliminate spectrally unlikely validated measurements* within a gate drawn by the GSF. In

*A measurement is considered as the pyhsical representative of a foreground blob. It is assigned the central coordinates and dimensions of a blob. For this reason, we use a linear measurement model in the filtering section unlike GMTI-based tracking studies where measurements are provided in polar coordinates. Conversion from polar to cartesian coordinates is achived by a nonlinear measurement model in the filter.

other words, we only consider spectrally similar objects to the TOI as a candidate true measurement. By doing so, we can both decrease complexity of the cost minimization problem and increase identification accuracy.

6.1 Feature Matching and Measurement Validation

Two kinds of spatial sampling are performed to feed the sensor for spectral data acquisition. The first is prediction-based spatial sampling. In this case, the sensor is tasked to collect spectral data without any prior knowledge of where measurements or segmented blobs are located. Therefore, sampling is performed by taking a subset of the pixels from where the filter predicts the target will be. This process is repeated for each Gaussian in the GSF. Assume that m pixels in the vicinity† of the centroid of each prior Gaussian are sampled.

$$D_k^n = \begin{bmatrix} d_{11} & d_{21} & d_{31} & \dots & d_{m1} \\ d_{12} & d_{22} & d_{32} & \dots & d_{m2} \\ \vdots & \vdots & \vdots & & \vdots \\ d_{1b} & d_{2b} & d_{3b} & \dots & d_{mb} \end{bmatrix}, \tag{4}$$

where b is the number of bands of and $n = 1, 2, .., M(k)$ with M representing the number of Gaussians in the GSF. Comparison is performed considering the features acquired in the first and last scan as

$$f_k^n = mean\{min\{SAM(S_k^n, D_{k-1}^{match})\}, min\{SAM(D_k^n, D_1^{user})\}\}, \tag{5}$$

where D_{k-1}^{match} represents the Gaussian yielding the best the match in the last scan and D_1^{user} denotes the features collected in the vicinity of the user selection. SAM stands for the Spectral Angle Mapper distance metric which measures the similarity of two spectral vectors. If the target is lost at $k-1$, then the match at $k-2$ is used. Finally, the Gaussian that match the TOI is used to get the prediction-based validated blob. If multiple Gaussians match the TOI, their average is computed to get the prediction-based validated blob.

The second one is measurement-based sampling. This sampling process is implemented after the motion segmentation algorithm. The sensor is tasked to collect full spectra at the pixels in the vicinity of the centroids of the segmented blobs. This step can be beneficial in non-Gaussian target probability density case since the prediction-based sampling might fail in collecting target data. Assume that m pixels for each segmented blob are sampled as in eq. 4.

$$D_k^B = \begin{bmatrix} d_{11} & d_{21} & d_{31} & \dots & d_{m1} \\ d_{12} & d_{22} & d_{32} & \dots & d_{m2} \\ \vdots & \vdots & \vdots & & \vdots \\ d_{1b} & d_{2b} & d_{3b} & \dots & d_{mb} \end{bmatrix}, \tag{6}$$

where B denotes the list of blobs validated by the filter. Comparison is performed as

$$f_k^B = mean\{min\{(SAM(D_k^B, D_{k-1}^{match})\}, min\{SAM(D_k^B, D_1^{user})\}\}, \tag{7}$$

where D_{k-1}^{match} denotes the blob yielding the best match at $k-1$ and D_1^{user} is the features collected around the user selection. The blobs matching with the TOI ($f_k^B < Threshold$) together with the prediction-based validated blob ($f_k^n < Threshold$) are then considered as the candidate blobs for data association.

6.2 Data Association

The data association is a key step for a consistent tracking system since it is required to update $p(X_k|Y_{k-1})$. Inserting features from a single modality in data association is not a complete solution for consistent target tracking. Therefore, we simplify the data association problem by reducing spectrally unlikely vehicles to the TOI using the discriminative spectral features in addition to the kinematic features (see 6.1). It should be noted that, only the filter validated measurements are considered and checked if they match to the TOI spectrally. In the next step, the spectral features of both the spectrally and kinematically validated measurements ($y_k = \{y_k^i | i = 1, .., N(k)\}$) are integrated in a bayesian fashion into the data association to further improve identification rates.

† Spectral data is sampled at the central pixel of a blob and horizontal and vertical neighbors of that pixel.

In this study, we consider the Probabilistic Data Association Filter (PDAF) for the integration of kinematic and spectral likelihoods.[14] PDAF considers the validated measurements in the latest scan to perform associations. It avoids immediate hard decisions by estimating the joint likelihood of validated measurements to update $p(X_k|Y_{k-1})$. The association probability for y_k^i in nondummy measurement assignment ($i > 0$) case is estimated as

$$\beta_k^i = \frac{\theta_k^i \, \tau_k^i}{1 - P_D P_G + \sum_{j=1}^{N(k)} \theta_k^j \, \tau_k^j}, \quad i = 1, ..., N(k). \tag{8}$$

In dummy measurement assignment ($i = 0$) case it is defined as

$$\beta_k^0 = \frac{1 - P_D P_G}{1 - P_D P_G + \sum_{j=1}^{N(k)} \theta_k^j \, \tau_k^j} \tau_k^0, \tag{9}$$

where P_D and P_G are the probability of detection and detecting the target in the gate. θ_k estimates the likelihood ratio of the measurement y_k^i originating from the target rather than the clutter based on the predicted measurement $h(x_{k|k-1})$ and innovation uncertainty S.

$$\theta_k^i = \frac{N[y_k^i; h(x_{k|k-1}), S(k)] P_D}{\lambda}, \tag{10}$$

where λ is the density of the Poisson process that models the clutter. Finally, the spectral likelihood of y_k^i (τ_k^i) is estimated as a function of the similar vehicles to the TOI at time $k-1$ and k. In the case of dummy measurement ($i = 0$) it is formulated as

$$\tau_k^0 = \begin{cases} \frac{1}{\zeta_k + 1} & \zeta_k \geq \zeta_{k-1} \\ \frac{\zeta_{k-1} - \zeta_k}{\zeta_{k-1} + 1} & \zeta_k < \zeta_{k-1} \\ 1 & \zeta_k = 0 \end{cases} . \tag{11}$$

In the case of nondummy measurement assignment ($i > 0$) it is defined as

$$\tau_k^i = \begin{cases} (1 - \frac{1}{\zeta_k + 1}) F_k^i & \zeta_k \geq \zeta_{k-1} \\ (1 - \frac{\zeta_{k-1} - \zeta_k}{\zeta_{k-1} + 1}) F_k^i & \zeta_k < \zeta_{k-1} \\ 0 & \zeta_k = 0 \end{cases} , \tag{12}$$

where ζ_k is the number of spectrally similar vehicles ($N(k)$) to the TOI. F denotes the normalized spectral scores and is estimated as shown below.

$$F_k^i = \frac{Th - f_k^i}{\Sigma_{j=1}^{N(k)} Th - f_k^j}, \tag{13}$$

where Th is the spectral threshold which is determined offline. By incorporating spectrally similar vehicles information into τ, we can have a better knowledge of a target loss. For instance, if the number of similar vehicles goes down significantly from $k-1$ to k, it is very likely that the target is lost at k. In this case, τ for dummy measurement is higher than the nondummy measurement assignment.

7. ROAD NETWORK CONSTRAINED FILTERING

In a given scenario, we aim to cover all possible paths a target can follow in a particular time step so that we can collect target spectral data in prediction-based sampling. On a straight road there is no need for assigning a turn model since the likelihood of a manuever is extremely low and noise can account of non-straight paths. On the other hand, on an intersection a target is very likely to change its direction. For this reason, in this paper, specific motion models are adaptively removed/inserted in the GSF based on the external intersection map. Details of OpenStreetMap data-based intersection extraction can be found in.[12,15] On a straight road,

the Stop model[‡] is employed in about 10% of the Gaussians whereas the remaning ones are given the Constant Velocity (CV) model with low or high amount of noise to account for rapid or slow accelarations.[3] Meanwhile, in the intersections, around 60% of the Gaussians are applied left or right turn models with low process noise whereas the other 30% and 10% are applied the CV and Stop models.[16]

Intersection maps provide us intersection geometry which can be used to assign proper turn rates. This prior information is used to model turn rate values with a normal distribution. Once the target is estimated to be in an intersection, turn rate values for the turn models are randomly drawn from the normal distribution of that intersection. In the following frames, turn rates are updated by the filter.

8. SIMULATION RESULTS

We focus on three metrics to evalute tracking and identification performance in the given scenarios. First, the Root Mean Square Error (RMSE) metric is considered to measure positioning accuracy in East-North-Up (ENU) real world coordinates since we utilize this coordinate system in DIRSIG simulation. The second performance metric is track purity (TP). It measures how many frames a tracker maintains a correct track identity within an estimated gate of the actual target position during the track life. Finally, the Current Assignment Ratio (CAR) metric is utilized to measure the ratio of maximum number of times a ground truth is associated to a track to the duration of ground truth. The TP metric does not take track termination into account. It only considers the track life, so, it might be misleading in cases where track terminations occur frequently. On the other hand, the CAR metric considers the life of ground truth so that potential misleading information due to TP is prevented. Extensive treatment on these metrics can be found in.[17]

The proposed spectral feature-aided tracking (SFAT) system is compared to a (1) kinematic only tracker (KT) and a (2) spectral only tracker (ST). This way, the benefit of further reduction of validated measurements and integrating spectral likelihoods in target identification is evaluated. In the KT method, spectral features are ignored whereas in the ST, only SAM scores in the last scan, k, are considered. In the KT case, τ in eq. 8 and 9 is neglected and no spectral elimination of the validated measurements is performed. The number of Gaussians in the GSF is 33 in all cases. Obviously, more Gaussians can be employed to better approximate true target probability density at a higher computational cost (fig. 5). Six number of pixels ($m = 6$) in the vicinity of the validated measurements' centroid are considered for spectra acquisition. A hundred Monte Carlo runs were carried out for each TOI to minimize the randomness effect on the results. A track is abandoned when it has not been associated with any measurement for more than 7 s. Table 1 displays TP and CAR scores for the TOIs.

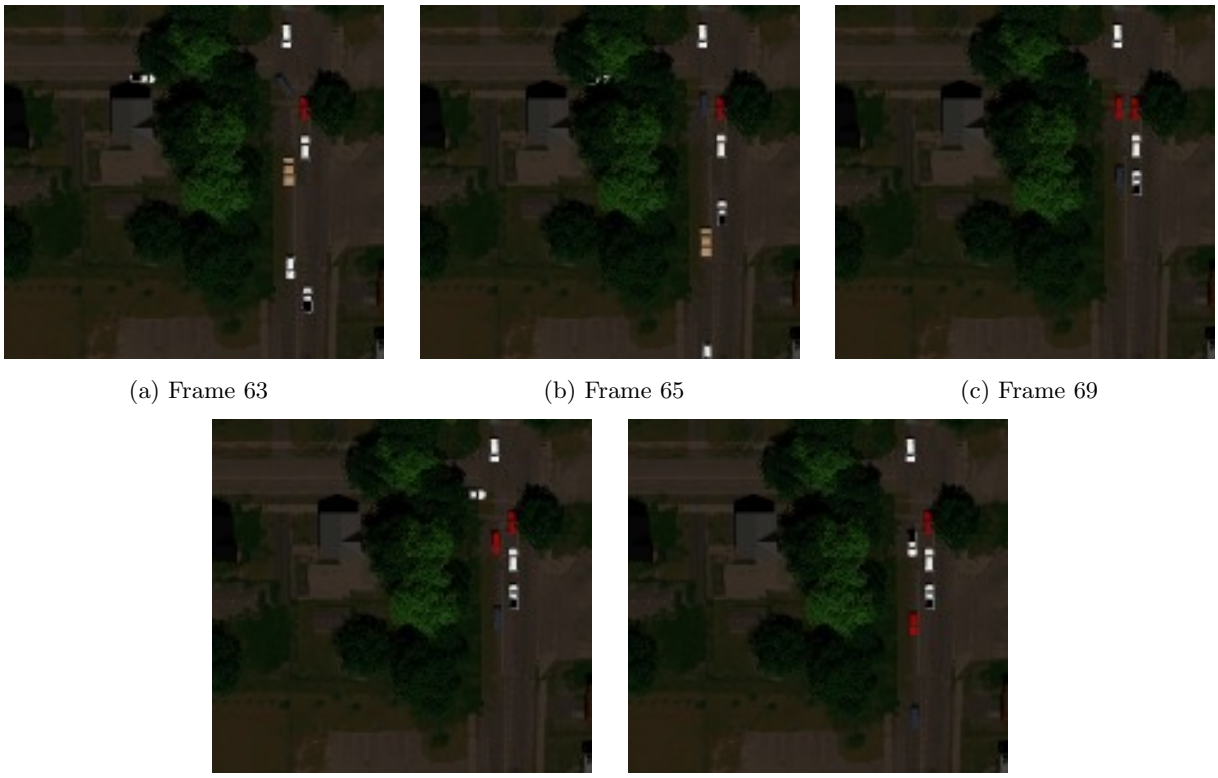
Table 1: TP and CAR scores for TOIs with different tracker types.

	Track Purity (%)					Current Assignment Ratio (%)				
Type/ID	1st	2nd	3th	4th	Overall	1st	2nd	3th	4th	Overall
KT	39.35	26.91	43.04	37.93	36.81	38.96	25.98	18.02	37.92	30.22
ST	20.01	87.48	65.04	50.24	55.69	4.82	65.05	87.48	43.63	50.25
SFAT	53.84	80.83	79.33	86.47	75.12	47.26	80.83	79.22	86.47	73.45

The KT method suffers from the vehicles with similar trajectories in highly cluttered scenes. In addition, obscurations lead to a series of wrong assigmnents in the KT experiments. The accuracy rate of KT can be increased by increasing the frame rate. This way target probability density approximation problem is simplified. On the other hand, the SFAT and ST handle these challenges by eliminating vehicles with similar trajectories utilizing the spectral signatures. Furthermore, if we can capture spectral features from the obscured target once it becomes visible, then we can re-associate the track with the true measurements. The ST method is confused in obscured scenes with many vehicles having similar paint models regardless of their kinematic features and performs poorly. On the other hand, the SFAT method outperforms the ST since it utilizes kinematic features

[‡]Stop model is not applied any Gaussian process noise due to the fact that any additional noise would yield a nonstationary track.

in addition to the spectral features. It can not perform well in the first TOI case since the target travels nearby many similarly painted vehicles before, during, and after occlusions as seen in fig. 4. The PDAF algorithm is confused once the TOI is occluded due to three other spectrally similar vehicles within the gate (fig. 4c). When the TOI becomes visible again, the PDAF faces measurements with similar spectral and kinematic features and weigh them similarly to estimate the joint likelihood. In the following frames, the TOI travels down whereas the others stop at the traffic lights for some more time and this leads the PDAF algorithm to assign higher weights to the stationary ones and wrong measurements for the rest of the scenario.



(a) Frame 63 (b) Frame 65 (c) Frame 69

(d) Frame 70 (e) Frame 73

Figure 4: The first TOI (white) is approaching a tree canopy occlusion (a), and partially occluded in the frame 65 (b). Later, it is fully occluded (c), and becomes partially visible again in frame 70 (d). Finally, it travels nearby other white vehicles in (e).

In fig. 5, we demonstrate the overall tracking and average run time performances in terms of the number of Gaussians in the GSF. Optimum number of Gaussians considering the tracking accuracy and average run time is 33. However, one can also decrease or increase the number of Gaussians based on the priorities. For instance, one can go with less number of Gaussians to track multiple targets. In this case, less spectral data is acquired for each target since allocated time for spectra acquisiton does not change. Less number of Gaussians corresponds to more erroneous target pdf approximation which yields more redundant data and degraded identification accuracy. However, we can tune the number of Gaussians online utilizing the intersection networks since nonlinearities mostly occur in the intersections.

9. CONCLUSIONS

Inserting spectral features in a wide area air-to-ground surveillance system can increase the persistency of tracking in cases where a single modality struggles. For this reason, we considered an adaptive multi-modal performance-driven sensor capable of adaptive spectral data acquisition. Discriminative spectral features are used to eliminate some of the filter validated measurements in the gate to compensate for the spectral data complexity and increase

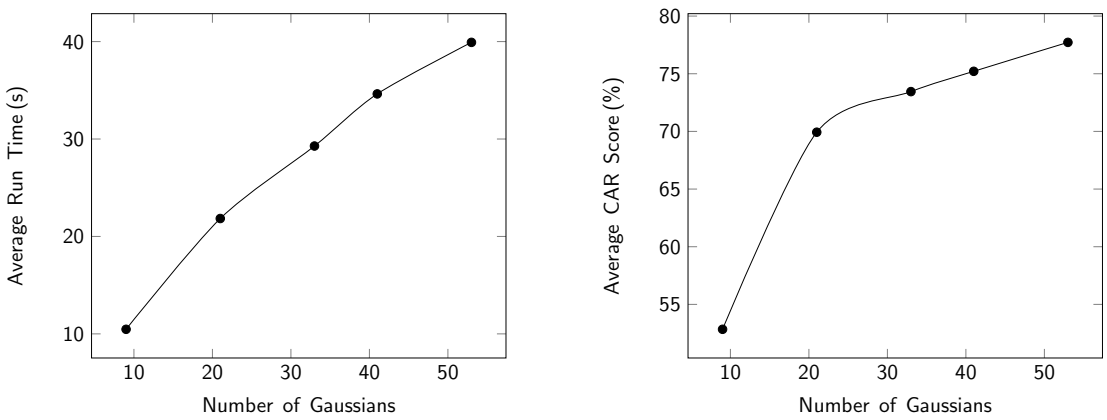


Figure 5: Average run time and average CAR scores for the SFAT algorithm in terms of the number of Gaussians in the GSF.

tracking accuracy. Then, spectral likelihoods are estimated and fused into the data association algorithm. The proposed SFAT method outperforms both KT and ST in the given scenarios. In addition, only around 1.5% of the spectral data of the full scene is collected in each revisit. One drawback of the proposed SFAT method is its performance may degrade in a cluttered scene with severe obscurations as in the first TOI experiments. In the future, we plan on addressing this issue by integrating spectral likelihoods into a multi-dimensional assignment algorithm to capture the evolution of the target trajectory which is not possible with the PDAF. However, in this case we have to allocate higher amount of time for data association step which results in less allocated time for spectra acquisiton.

ACKNOWLEDGMENTS

The authors would like to thank Bin Chen for his assistance on intersection mask incorporation into the tracking algorithm and Adam Goodenough and Scott D. Adam for their contributions on converting SUMO output into DIRSIG compatible input. This work is supported by the Dynamic Data Driven Applications Systems Program, Air Force Office of Scientific Research, under Grant FA9550-11-1-0348.

REFERENCES

[1] Arulampalam, M. S., Gordon, N., Orton, M., and Ristic, B., "A variable structure multiple model particle filter for GMTI tracking," in [*Information Fusion, 2002. Proceedings of the Fifth International Conference on*], **2**, 927–934, IEEE (2002).

[2] Payne, O. and Marrs, A., "An unscented particle filter for GMTI tracking," in [*Aerospace Conference, 2004. Proceedings. 2004 IEEE*], **3**, IEEE (2004).

[3] Kirubarajan, T. and Bar-Shalom, Y., "Tracking evasive move-stop-move targets with a GMTI radar using a VS-IMM estimator," *Aerospace and Electronic Systems, IEEE Transactions on* **39**(3), 1098–1103 (2003).

[4] Sinha, A., Kirubarajan, T., and Bar-Shalom, Y., "Autonomous ground target tracking by multiple cooperative UAVs," in [*Aerospace Conference, 2005 IEEE*], 1–9, IEEE (2005).

[5] Wang, J. and Yagi, Y., "Integrating color and shape-texture features for adaptive real-time object tracking," *IEEE Transactions on Image Processing* **17**(2), 235–240 (2008).

[6] Shin, J., Kim, S., Kang, S., Lee, S.-W., Paik, J., Abidi, B., and Abidi, M., "Optical flow-based real-time object tracking using non-prior training active feature model," *Real-Time Imaging* **11**(3), 204–218 (2005).

[7] Zhou, H., Yuan, Y., and Shi, C., "Object tracking using SIFT features and mean shift," *Computer vision and image understanding* **113**(3), 345–352 (2009).

[8] Meyer, R. D., Kearney, K. J., Ninkov, Z., Cotton, C. T., Hammond, P., and Statt, B. D., "RITMOS: a micromirror-based multi-object spectrometer," in [*Astronomical Telescopes and Instrumentation*], 200–219, International Society for Optics and Photonics (2004).

[9] Ientilucci, E. J. and Brown, S. D., "Advances in wide-area hyperspectral image simulation," in [*AeroSense 2003*], 110–121, International Society for Optics and Photonics (2003).

[10] Krajzewicz, D., Hertkorn, G., Rössel, C., and Wagner, P., "Sumo (simulation of urban mobility)," in [*Proc. of the 4th middle east symposium on simulation and modelling*], 183–187 (2002).

[11] Rice, A., Vasquez, J., Mendenhall, M., and Kerekes, J., "Feature-aided tracking via synthetic hyperspectral imagery," in [*Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2009. WHISPERS'09. First Workshop on*], 1–4, IEEE (2009).

[12] Uzkent, B., Hoffman, M., Vodacek, A., and Chen, B., "Feature Matching with an Adaptive Optical Sensor in a Ground Target Tracking System," *IEEE Sensors Journal* **15**(1), 510–519 (2015).

[13] Uzkent, B., Hoffman, M. J., Vodacek, A., Kerekes, J. P., and Chen, B., "Feature Matching and Adaptive Prediction Models in an Object Tracking DDDAS," *Procedia Computer Science* **18**, 1939–1948 (2013).

[14] Bar-Shalom, Y., Daum, F., and Huang, J., "The probabilistic data association filter," *Control Systems, IEEE* **29**(6), 82–100 (2009).

[15] Chen, B., Sun, W., and Vodacek, A., "Improving image-based characterization of road junctions, widths, and connectivity by leveraging OpenStreetMap vector map," in [*Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*], 4958–4961, IEEE (2014).

[16] Li, X. R. and Jilkov, V. P., "Survey of maneuvering target tracking. Part I. Dynamic models," *Aerospace and Electronic Systems, IEEE Transactions on* **39**(4), 1333–1364 (2003).

[17] Blasch, E. P. and Valin, P., "Track purity and current assignment ratio for target tracking and identification evaluation," in [*Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*], 1–8, IEEE (2011).