Signature-Aided Air-to-Ground Video Tracking*

Pablo O. Arambel, Jeffrey Silver, and Matthew Antone
BAE Systems – Advanced Information Technologies
6 New England Executive Park,
Burlington, MA 01803, U.S.A.
pablo.arambel, jeffrey.silver, matthew.antone@baesystems.com

Thomas Strat
DARPA Information Exploitation Office
3701 N Fairfax Drive
Arlington, VA 22203
thomas.strat@darpa.mil

Abstract – Tracking ground moving objects using aerial video sensors is very challenging when the objects go through periods of occlusion caused by trees or buildings. If the occlusion interval is relatively large, there are confusing objects in the vicinity, or the object performs abrupt maneuvers while occluded, maintaining continuous tracks after the occlusion requires advanced exploitation of the imagery. This paper presents a signature-aided multiple hypothesis tracking system where signatures are extracted during periods of certainty and used after the occlusion to resolve association ambiguity. The discussion focuses on the interaction between the tracker and the signature extraction/exploitation module, as well as other tracking aspects within the signature-aided tracking paradigm.

Keywords: Tracking, video, data association, estimation, signature-aided tracking.

1 Introduction

Steerable video cameras are rapidly finding their way into most air-to-ground surveillance platforms (Figure 1). The fact that imagery can be easily interpreted by the operators makes them very popular. Tracking moving objects for extended periods of time, however, can be very demanding for the operator. The field of view (FOV) that is most appropriate for object recognition typically results in a very small coverage area, and the simple task of aiming the camera to maintain the object within the FOV requires uninterrupted attention. Thus, any means to automate the task of aiming the camera and tracking objects on the ground is usually very welcome by the operators.

Tracking ground objects automatically requires a means to detect the objects in the imagery, a means to associate these detections with established tracks, and a means to aim the camera at an appropriate point on the ground. All these tasks are very challenging when the objects go through periods of occlusion by trees or buildings and the objects are not visible. Thus, to have continuous tracks after periods of occlusion, we need to detect the object when the object reappears, and recognize those detections as corresponding to the same object that was being tracked before the occlusion. The association can be done based on kinematics if the occlusion is relatively short, in the order of a few seconds, but can be very challenging if the occlusion interval is relatively large, there are confusing objects in the vicinity, or the object performs abrupt maneuvers while occluded. Much higher performance and reliability is achieved by further exploiting the imagery to associate new detections to established tracks.

This paper presents a signature-aided multiple hypothesis tracking system to track objects through periods of occlusion and/or coverage gaps. The system architecture and main components were developed under the VIdeo Verification of IDentity (VIVID) program sponsored by the US Defense Advanced Research Projects Agency (DARPA), and continue to be extended under other programs with similar objectives.

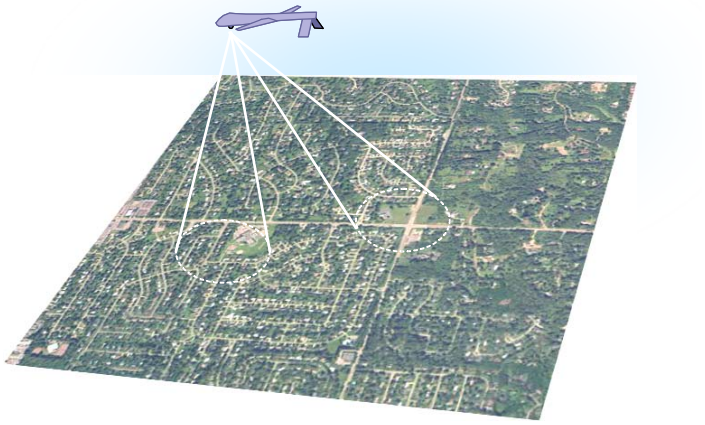


Figure 1: Steerable video cameras are becoming very popular in air-to-ground surveillance and tracking applications.

The paper is organized as follows. Section 2 describes the main functional blocks of the system. Section 3 describes the interaction between the signature extraction and exploitation module with both the multiple hypothesis tracker and the sensor resource manager that controls the camera. Section 4 presents two examples: one example in which the FOV of the camera is maintained fixed, and another example when the FOV is changed to improved

* Approved for Public Release, Distribution Unlimited.

This work was supported by DARPA under the VIVID program, contract # NBCHC030069. VIVID is focused solely on military applications involving unmanned aircraft engaging military target vehicles on the ground. At times, the VIVID Program makes use of imagery of civilian vehicles for experimental purposes when it is more expedient or cost effective to do so. ALL imagery of civilians and civilian vehicles used within the program has been collected with the consent of the individuals involved. Any representations of civilians or civilian vehicles are an artifact of the experimental process and are used purely to expedite the research.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUL 2006		**00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Signature-Aided Air-to-Ground Video Tracking	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defense Advanced Research Projects Agency (DARPA),Information Exploitation Office,3701 N Fairfax Drive,Arlington,VA,22203**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
9th International Conference on Information Fusion, 10-13 July 2006, Florence, Italy. Sponsored by the International Society of Information Fusion (ISIF), Aerospace & Electronic Systems Society (AES), IEEE, ONR, ONR Global, Selex - Sistemi Integrati, Finmeccanica, BAE Systems, TNO, AFOSR's European Office of Aerospace Research and Development, and the NATO Undersea Research Centre.

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**8**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

the recognizability of the target. Finally, Section 5 provides a summary.

2 System Components

The multiple target video tracker is comprised of four sub-components as shown in Figure 2 below: a Video Processor (VP), a Multiple Hypothesis Tracker (MHT), a Confirmatory ID (CID), and a Sensor Resource Manager (SRM). The function of the VP is to detect moving objects and generate "micro-tracks" and high-confidence associations for the downstream tracker by processing the raw sensor streams of motion imagery and metadata. The VP makes few assumptions about the scene content, operating almost exclusively in the focal plane domain, and exploits the spatial and temporal coherence of the video data. Three main components of the VP subsystem are a Point Tracker, which detects and associates sparse interest points from frame to frame; a Motion Segmenter, which clusters interest points and classifies moving regions using several video frames; and a Template Matcher, which holds track on slow-moving and stopped targets that would otherwise be overlooked by motion-based algorithms.

The Point Tracker detects interest points in a given frame using a metric based on eigenvalues of the Hessian of the spatial image gradient at each pixel. Peaks in the metric function are detected, and a multi-level spatial bucketing technique ensures that points are both distributed across the image extent and separated sufficiently from one another to describe the primary scene content. In subsequent frames, these points are replenished as necessary to describe newly visible content. Each point is tracked from frame to frame at sub-pixel accuracy using an iterative multi-resolution gradient descent algorithm that seeks to minimize the sum of squared errors between pixel colors within a fixed patch around each point. Further details of the algorithms associated with the VP can be found in [1]. Figure 3 below shows an example with the original frame, extracted interest points, points stabilized to a common background, and micro-tracks.

As a result of the detection mechanism, the VP also generates bounding boxes and masks associated with those detections. We refer to the bounding box and mask, in conjunction with the portion of the image associated with the bounding box, as image chip. The image chips associated with the detections are sent to the Confirmatory ID (CID) module for further processing. The CID module uses these chips to extract object signatures. We use the term signature in a broad sense to indicate any set of features used by the CID module to distinguish that particular object from other objects. These features can comprise color histograms, edge maps, shape histograms, and other attributes. The CID generates signatures from image chips that correspond to established tracks, and compares them with image chips corresponding to new micro-tracks to determine whether those signatures correspond to the same object or not. This information is passed back to the tracker in the form of approximated likelihoods. Exact likelihoods are very difficult to generate because the amount of data used to

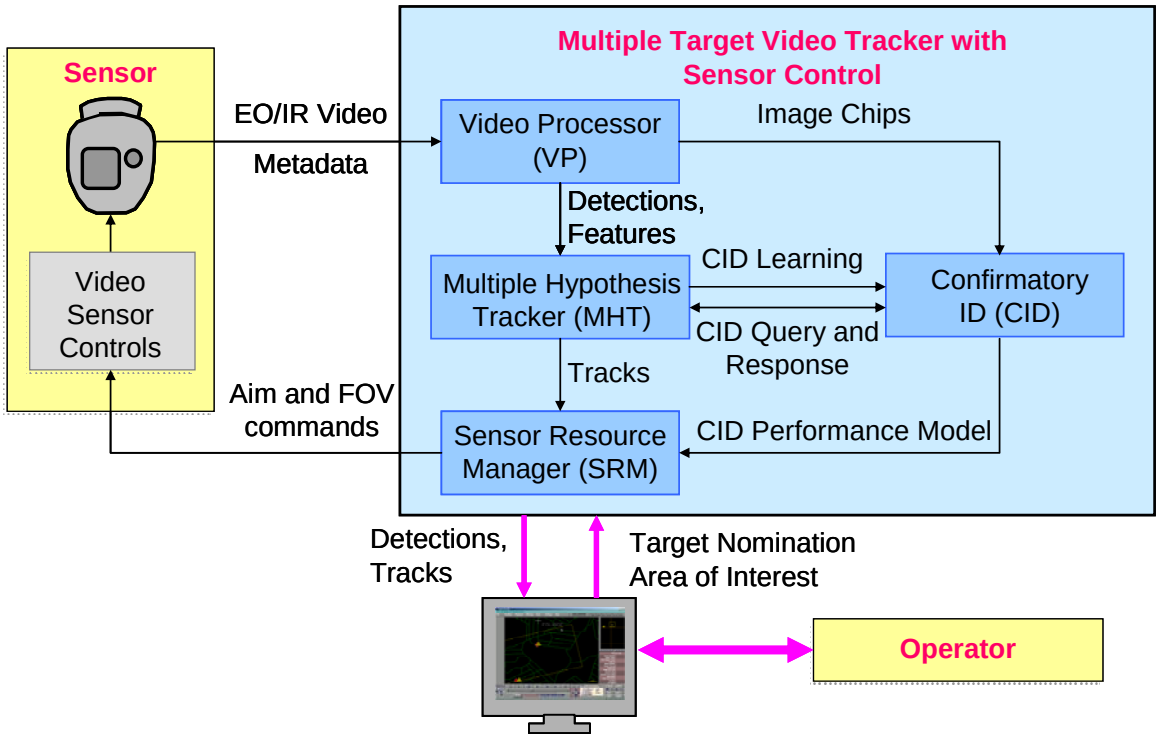


Figure 2: The video tracker automatically detects moving objects and extracts signatures for each object that is being tracked. After occlusion and other periods of uncertainty, it exploits the signatures to resolve the ambiguity and maintain track continuity. The sensor resource manager commands the sensor to support the entire signature extraction and exploitation process.

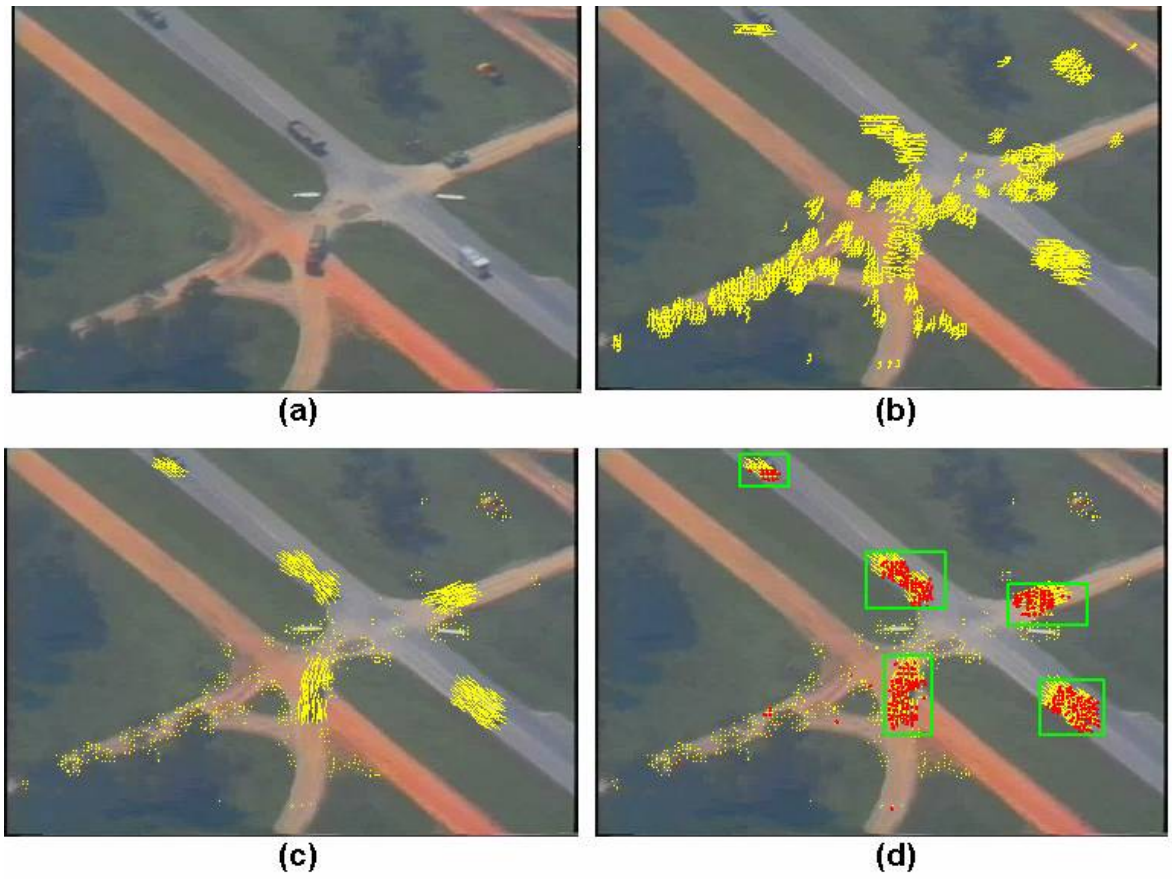


Figure 3: An example of motion segmentation: interest point tracks (b) from a frame (a) are stabilized to a common background (c). Outliers are detected and clustered according to consistency of motion and proximity (d).

generate the signatures is relatively small, but these likelihoods are needed by the tracker to combine CID-extracted information with kinematic likelihoods.

The MHT processes the LOS and feature reports from the VP to create and update moving object tracks. Tracks comprise position and velocity estimates, feature estimates, error covariances, and hypothesis likelihoods. The tracker also maintains a probabilistic estimate of the occlusion status of each target. This is required so that the SRM can correctly evaluate the probability of detecting a target when determining whether resources should be expended to observe that target. The tracker understands each target to have a binary occlusion state, and updates an associated probability of occlusion for that target forward in time according to a Poisson-Bernoulli model of occlusion state evolution in the absence of any external information from the video processor. (This will occur, for example, if the camera is otherwise tasked and is pointed away from where the tracker predicts the target should be located.) When the camera does attempt to view the target at a location predicted by the tracker, the video processor then confirms that the target either was seen or was not seen during that attempt. This information is then used to update the probability of occlusion based on the computed probability that the target is within the camera footprint, and VP performance parameters indicating the probability the target would be detected given it is both unoccluded and within the camera footprint, and the probability of false alarm generation. The tracker also interacts with the CID module as described in Section 3.1.

The SRM manages the information collection necessary to support kinematic tracking, multi-target track association, as well as the acquisition of high resolution imagery to support the CID functions. Typically, the image resolution required by the CID module to generate reliable signatures is higher than the resolution required to detect and track targets; therefore, the SRM selects the appropriate FOV depending on the current tasks. The SRM balances the expected payoff of alternative viewing options against the costs due to sensor slewing, settling and collection time and tasks the sensor to optimize tracking performance. Section 3.2 describes the CID performance modeling used by the SRM to determine when to collect new imagery to support the CID functions.

3 Signature-aided Tracking

3.1 MHT and CID Interaction

This section describes the MHT internal structure and its interaction with the CID module. An MHT functional diagram is shown in Figure 4. The MHT receives detections and features that the VP has extracted from the imagery, and associates them to predicted track hypotheses to form multiple track hypotheses. These

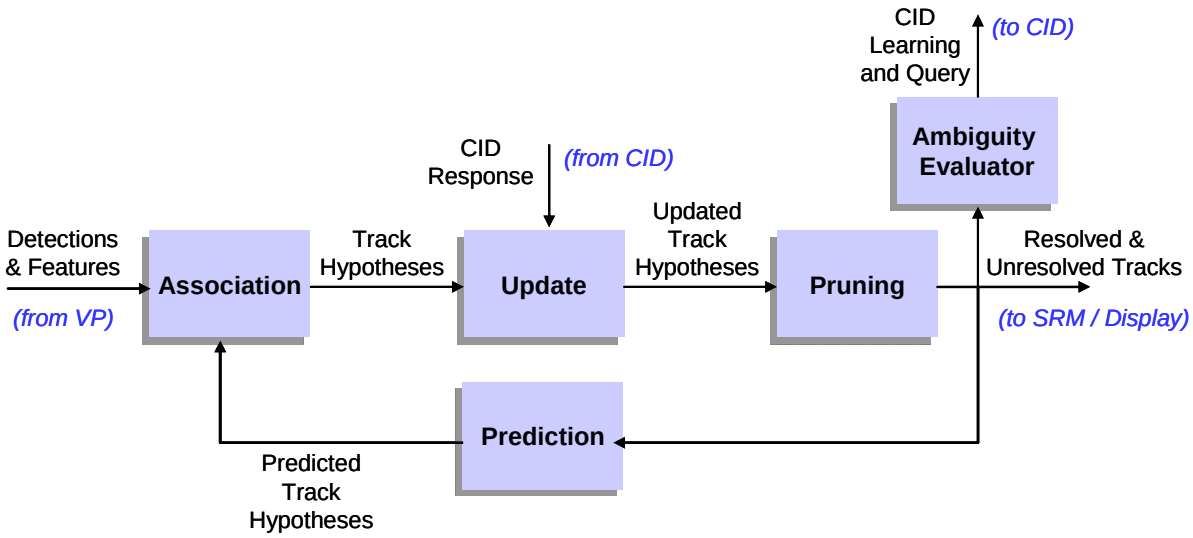


Figure 4: The MHT processes VP detections, creates and maintains tracks, and interacts with the CID module.

hypotheses represent different possibilities regarding detection-to-track associations that for example include the association of the detection to an existing track (hypothesis 1), the association of the detection to another nearby track (hypothesis 2), the association of a track to a missed detection (hypothesis 3), and other similar hypotheses. The tracker then updates the state estimates for each one of these track hypotheses, and updates the hypothesis likelihoods as well. A Kalman Filter is used to estimate the state of each track hypothesis. The likelihood of each alternative combines both the kinematic information as well as the information provided by the CID module via the CID response message, as indicated in the figure. A highly confident CID response that favors one hypothesis immediately triggers the pruning of the alternative hypotheses. CID confidence is derived from the ratio of the likelihoods of the different alternatives in the CID response message.

The MHT then prunes the hypothesis tree, where the hypotheses with the lowest likelihood are removed from the system to avoid an explosion of hypotheses. The pruning process also removes the ambiguity in some assignments, since some of the track hypotheses that contain a given detection are removed, possibly leaving a single track with the associated detection. When that happens, that detection is unambiguously assigned to a particular target. This is desirable because that detection becomes a candidate for a learning message and can be sent to the CID module to improve the signatures for that particular target. The output of the pruning module consists of a set of alternative hypotheses that explains the sequence of detections that have been received from the VP since the beginning of the process.

The global hypothesis with the highest likelihood is sent to the display. The best global hypothesis represents the best description of the tracking status that the video tracker can provide at a given time. A more complete status of the MHT that includes all the relevant track hypotheses and their likelihoods is sent to the SRM module. The SRM uses that information to select the best aimpoint and FOV to track and to remove ambiguities. The hypothesis tree is also reviewed by the ambiguity evaluator to determine the detection ambiguity status. If a detection has been unambiguously associated to a target and no other detections have been associated to that target at that particular time, the detection becomes a CID learning candidate. If the detection is ambiguously associated to more than one target, or there are multiple detections associated to a target at a given time, that detection becomes a CID query candidate. To maintain the object signatures as pure as possible, not all of the candidates become part of learning or query messages. For example, if the length and width measurements in a stream of detections are unstable, this may indicate that the target is going behind trees or other occlusion, and the detections are deemed not suitable for CID purposes. Checking that the detections are stable prevents the corruption of the object signatures.

The loop is closed by the prediction module that takes the existing track hypotheses and predicts the motion of the vehicles using a Kalman Filter. The prediction includes both the state estimate (position and velocity) as well as the covariance matrices that indicate the prediction uncertainty. This is passed to the association step that combines it with the measurement errors to create gates that reject very unlikely associations. The predictions and corresponding uncertainties are also used to compute the likelihood ratios.

3.2 SRM and CID Interaction

We now discuss the SRM interaction with CID capability. The SRM manages two distinct decision tasks associated with the CID module. First, the SRM has to schedule learning sequences of video to be captured during times when the MHT can clearly disambiguate relevant targets. In order to do this, the SRM has to assess both the value of the targets of interest (which may be derived from either their nomination status or their proximity to other nominated targets because of anticipated confusion zones) as well as the current requirements of the CID module to perform efficiently when an identification query is submitted. Second, the SRM has to schedule CID queries of relevant targets when appropriate. In order to do this, the SRM has to assess both the scene confusion

content (as calculated within the MHT) as well as the anticipated performance of the CID module before the query is submitted.

Both the assessments of target value and scene confusion are functions that the SRM can perform with only direct contact with MHT. However, for both the assessment of CID requirements in the scheduling of learning sequences, and for the prediction of anticipated benefit in determining if images should be collected for CID query submission, the SRM requires an explicit quantitative estimate of the impact of additional camera video sequences on expected CID performance. In support of this, the tracker maintains a probabilistic estimate of binary target CID recognizability status, similarly as it does with target occlusion status, according to a Poisson-Bernoulli model of recognizability devolution in the absence of processed learning sequences. (This decay modeling is necessary to capture real effects of recognizability degradation because of temporal and environmental variations in visibility conditions due to fog, solar ephemeris, etc.) Additionally, when learning images are acquired they impact the CID recognition probability according to a precisely defined probability model of affirmatively transitioning CID recognizability status at a given pose (aspect and grazing angle) and resolution. The associated CID recognition probability maintained for each target inside the tracker is therefore a dynamically estimated probability of the CID module's ability to perform successfully as a function of the target pose and target resolution with respect to the airborne platform.

Letting $P_t^{(\mathrm{T})}(\theta)$ denote the probability of CID recognizability for target T at pose θ and time t (and suppressing dependence on resolution for notational simplicity), the above motion modeling assumptions imply the temporal decay update equation over a time interval Δt:

$$\frac{\Pr\left(\text{target T is CID recognizable } @t+\Delta t\right)}{\Pr\left(\text{target T is CID recognizable } @t\right)} = \frac{P_{t+\Delta t}^{(\mathrm{T})}(\theta)}{P_t^{(\mathrm{T})}(\theta)} = e^{-\lambda \Delta t}$$

where λ denotes the Poisson rate parameter in the recognizability devolution model. Similarly, letting $P_{t^-}^{(\mathrm{T})}(\theta)$ and $P_{t^+}^{(\mathrm{T})}(\theta)$ denote the probabilities of CID recognizability for target T at pose θ and time t before and after processing of an acquired learning image, respectively, the above impact modeling assumptions imply the instantaneous impact update equation at time t:

$$\frac{\Pr\left(\begin{array}{l}\text{target T is } \textit{not} \text{ CID recognizable } @t \\ \textit{after} \text{ learning image acquisition}\end{array}\right)}{\Pr\left(\begin{array}{l}\text{target T is } \textit{not} \text{ CID recognizable } @t \\ \textit{before} \text{ learning image acquisition}\end{array}\right)} = \frac{1-P_{t^+}^{(\mathrm{T})}(\theta)}{1-P_{t^-}^{(\mathrm{T})}(\theta)} = 1-p_{\mathrm{CID}}(\Delta\theta)$$

where $\Delta\theta$ denotes the pose deviation from learned image and $p_{\mathrm{CID}}(\Delta\theta)$ denotes the CID impact model for generic targets; i.e. $p_{\mathrm{CID}}(\Delta\theta)$ is the probability that a target becomes CID recognizable at a given pose after a learning image is acquired at pose deviation $\Delta\theta$, given that it was not recognizable before the learning image was acquired.

The impact model $p_{\mathrm{CID}}(\Delta\theta)$ is ideally based upon the empirically measured sensitivity of the CID algorithms to deviations from learning pose. For example, if the CID module has a learning image of a vehicle acquired at a 45 degree aspect angle, $p_{\mathrm{CID}}(\Delta\theta)$ encodes how CID recognition performance degrades for comparative images obtained at 50 degrees, 55 degrees, etc. Figure 6 illustrates the model update process described for both a simple "band" defined image-based model and a more intricate "decay" model about three captured poses, assuming for simplicity that pose is defined by aspect only. Note that the performance models (i.e., maintained CID recognition probabilities) are displayed on different scales in the figure for clarification only, and that they actually both peak at probability 1. We may even incorporate information that learned images at certain poses provide evidence for CID performance at entirely different orientations, due to anticipated symmetries in target appearance from opposite directions. Thus, a learned image at 90 degrees aspect can potentially impact CID performance at 270 degrees aspect as well. Higher impact models of varying complexity may also be pursued. We remark here that what is relevant to the SRM is not the precise form of the image-specific modeling that is used, but merely that some quantitative impact model of this type exists so that expected CID performance can be correctly imputed and CID scheduling can be balanced by the SRM against competing tracker goals.

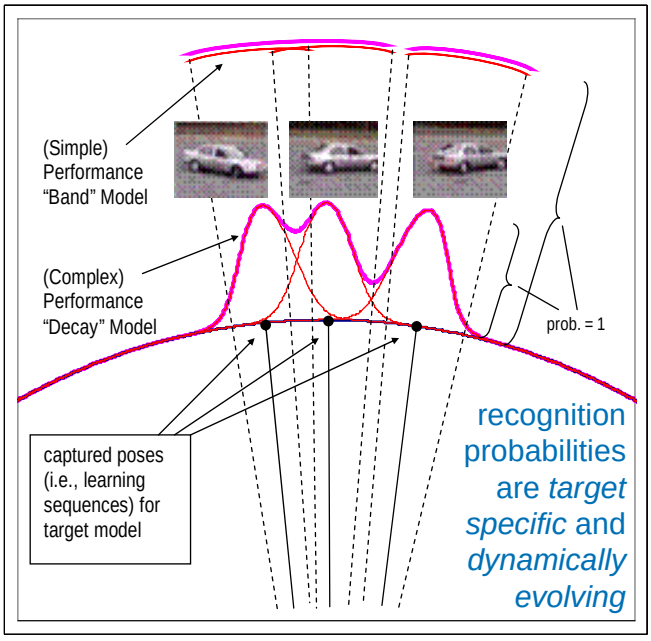


Figure 5: Image-based CID impact models support CID (recognition probability) performance modeling.

Given the performance modeling notation above, we now address how the SRM utilizes the CID recognition probabilities and MHT input in order to learn on known targets and to disambiguate target confusions. In order to schedule a learning sequence for a *known* target T at a known pose θ (as determined by the MHT and available or estimated metadata), the SRM refers to the maintained recognition probability $P_t^{(\mathrm{T})}(\theta)$ directly to determine the added benefit of collecting additional video images. Furthermore, in order to determine whether a query should be submitted for an *unknown* target at a known pose (as determined by the MHT and available or estimated metadata), the SRM determines the probability that the CID module can recognize the queried image among all suggested candidate targets by evaluating the convex combination of recognition probabilities $P_t^{(\mathrm{T})}(\theta)$ with respect to the MHT prior distribution on the candidate target model space:

$$P_t(\theta) = \sum_{\text{targets T}} P_t^{(\mathrm{T})}(\theta)\Pr(\mathrm{T})$$

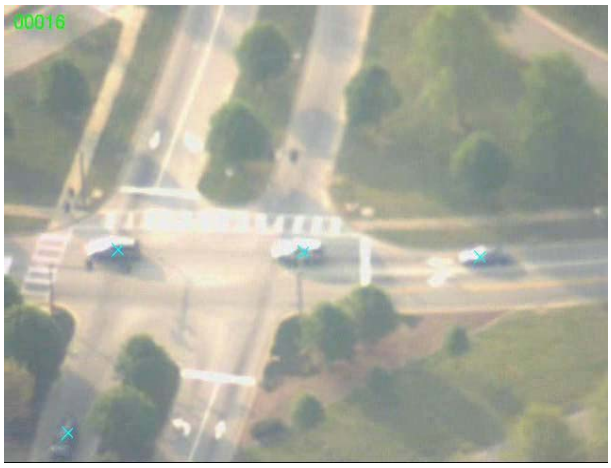
Finally, although not highlighted above for clarification purposes, we remark that resolution dependence of the maintained CID recognition probabilities is important and affects the SRM calculation by allowing it to anticipate expected CID performance differences for video acquired at distinct zoom levels, as specified by ground sample distance. Thus, for example, the SRM can determine if currently available zoom levels are compatible for CID querying against earlier learned images. The SRM can also balance the benefit of selecting wider fields of view to potentially allow the MHT to process more kinematic data for other nearby targets, against the resolution requirements that ensure good CID learning performance for the current target of interest.

4 Examples

4.1 Single FOV Example

Figure 6 illustrates the behavior of the MHT and the interaction with the CID module in a typical situation. It is assumed that the image resolution is good enough for the CID module to operate, so that the SRM does not need to zoom the camera in to acquire higher resolution imagery.

- In Figure 6A, the VP detects objects moving on the ground and sends detections and temporal coherence information to the MHT.
- In Figure 6B, the MHT creates tracks after a number of associated detections are consistent with the motion of a ground target. If the temporal coherence information is high, the MHT only generates one track hypothesis. The figure only shows the track associated with the first vehicle.
- In Figure 6C, the target goes behind some trees and is not detected by the VP. The MHT maintains a track and predicts the motion of the vehicle.
- In Figure 6D, the target reappears and is again detected by the VP. Since the detections are consistent with the motion of the first target, a new hypothesis is created that postulates that the new detections belong to the target that is being tracked. At the same time, another hypothesis that postulates that the first target is still occluded is maintained.
- In Figure 6E, the two hypotheses are maintained until enough evidence that favors one of the competing hypotheses can be accumulated. Evidence can be derived from kinematic consistency, but can also be derived from the CID response message. In this example, a number of detections have been linked to each other and are sent to query the CID module. In plain words, the query message posses the following question: "do these reports belong to this particular target?"
- In Figure 6F, the CID response contains likelihoods that correspond to the alternative hypothesis. In this example, the likelihood that "the detections are generated by this particular target" is much higher than the likelihood that "the detections are NOT generated by this particular target." Since the confidence is high, the MHT removes the alternative hypothesis and maintains only that one that contains the new detections. These detections are now unambiguously associated to that target, and can therefore be sent to the CID module for learning. In the hypothetical case that the CID likelihoods are not decisive enough, they are combined with the total likelihood for that track. This typically triggers a subsequent query unless one of the competing hypotheses gets removed for another reason.



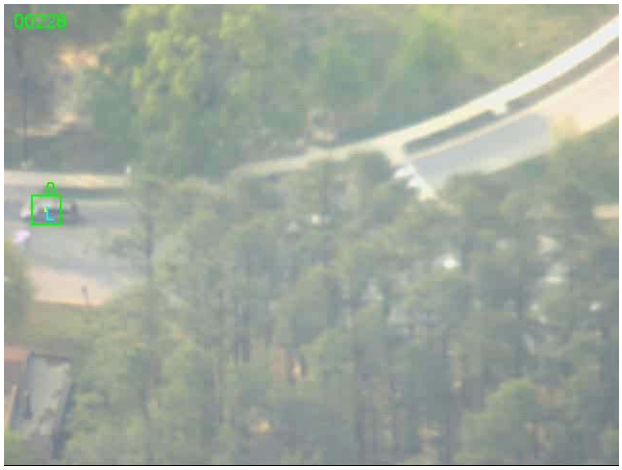
A: The VP detects objects moving on the ground.

B: A track is created if a number of associated detections are consistent with the motion of a vehicle. A CID Learning message is generated if the detections are unambiguously associated to a single target. The picture shows the only track hypothesis that is created for the first target. Other tracks are not displayed.

C: An occluded target can not be detected, but the motion is predicted by the MHT.

D: The reappearing target is detected by the VP and the MHT creates two hypotheses that represent two possible alternatives: the new detection belongs to the target, or the new detection belongs to another vehicle.

E: The CID module is queried if a number of consistent detections are deemed to be appropriate for the CID.



F: A highly confident CID response induces the pruning of the least likely hypothesis by MHT. In this case, the hypothesis that claims that “this is the same target” is maintained while the alternative hypothesis that claims that “this is another target” is removed.

Figure 6: Example of interaction between MHT and CID to create and exploit object signatures. A blue cross indicates detections by the VP; a blue ‘L’ indicates that the detections are used in the learning message and therefore used to create object signatures by the CID; a blue ‘Q’ indicates that the detection is used in the query message; green boxes indicates different track hypotheses for the leading vehicle.

4.2 Multiple FOV Example

The example above was generated under the assumption that the CID module can operate with Ground Sample Distances (GSD) in the order of 20cm per pixel. However, the performance of the CID module is typically higher for higher resolution. Thus, the camera needs to be zoomed in for the learning and query cycles to be most useful. The sequence in Figure 7 illustrates a typical MHT/CID interaction when the sensor is zoomed in and out. Changes of FOV and pointing angles are commanded by the SRM. The image on the upper left corner shows the target that is being tracked. When the target is nominated, the SRM zooms in on the target to acquire high resolution images, as shown in the upper right corner (zoomed images in the figure are simulated). As the target enters the occlusion and no more detections are generated, the target location is predicted by using the last estimated ground position and velocity. Naturally, the uncertainty of the estimated location grows with time. When the target emerges after the occlusion (bottom left of Figure 7), the tracker creates two hypotheses: one that postulates that the current detections correspond to the nominated target, and another one that postulates that the nominated target is still occluded and therefore the current detections correspond to another target or are false alarms. To resolve this ambiguity the SRM sends a command to aim the camera at the detections using a narrow FOV to acquire high resolution images (bottom right of Figure 7). The MHT system then sends a query to the CID module to confirm the identity of the target. If the answer is affirmative—as in this example—the MHT removes alternative hypotheses. If the answer is negative, the MHT removes the hypothesis that postulates that these detections correspond to the nominated target and continues to coast the track associated with the nominated target.

5 Conclusions

This paper described the interactions of a signature extraction and exploitation module with the multiple hypothesis tracker and the sensor resource manager. These modules and a front-end video processor are the main components of a signature-aided system that exploits airborne imagery to track multiple ground targets. The signature-aided system enables the maintenance of the tracks after relatively large periods of occlusion and coverage gaps. This system was developed under the VIdeo Verification of IDentity (VIVID) program sponsored by the US Defense Advanced Research Projects Agency (DARPA) and is currently being transitioned to an operational platform.

References

[1] P.O. Arambel, M. Antone, M. Bosse, J. Silver, J. Krant, and T. Strat, *Performance Assessment of a Video-based Air-to-ground Multiple Target Tracker with Dynamic Sensor Control*, SPIE Defense and Security Symposium, Signal Processing, Sensor Fusion, and Target Recognition XIV, Orlando, FL, March 28, 2005.\

[2] C. Stauffer and W.E.L. Grimson. *Adaptive Background Mixture Models for Real-Time Tracking*. In Proceedings of CVPR, 1999, pp. 246-252.

[3] A. Chao and S. Berning, *Precise Image Calibration and Alignment*, in Proceedings of the 55th Annual Meeting of the Institute of Navigation, June 1999.

[4] R. Hartley and A. Zisserman. Multiple View Geometry. Cambridge University Press, 2000.

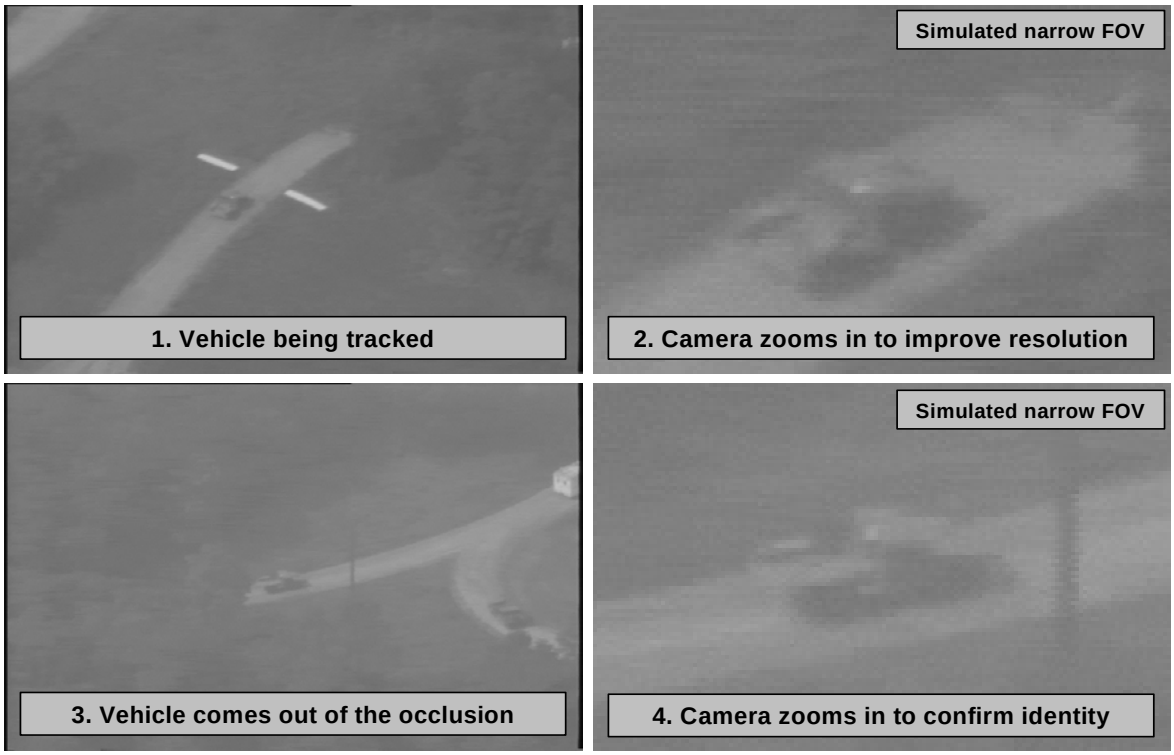


Figure 7: Tracking through extended occlusions requires the use of the Confirmatory ID module to confirm the identity of the target when this is reacquired.