Fusion of Imagery Attributes with Non-Imaging Sensor Reports by Truncated Dempster-Shafer Evidential Reasoning*

ALEXANDRE JOUAN, Langis Gagnon, Elisa Shahbazian, Pierre Valin
Lockheed Martin Canada, 6111 Royalmount Ave., Montréal, Québec, H4P 1K6, Canada
tel: (514) 340-8310, extensions 8442, 8715, 8343, 8703 fax: (514) 340-8318
alexandre.jouan@lmco.com, langis.gagnon@lmco.com, elisa.shahbazian@lmco.com, pierre.valin@lmco.com

Abstract - *This paper describes an on-going effort to perform data fusion of uncertain, incomplete or poor quality information coming from imaging and non-imaging sensors present on-board the CP-140 (Aurora) Canadian maritime patrol aircraft. An Adaptable Data Fusion Testbed (ADFT) has been constructed where simulated and real data can be analyzed by modular algorithms for each of the main fusion functions and image interpretation algorithms in order to derive requirement specifications and help in the design phase towards fielding a real Data Fusion (DF) system.*

Keywords: Imagery, Identification, Attribute, Dempster-Shafer, Testbed

1. Introduction

The Adaptable Data Fusion Testbed (ADFT) is designed to accommodate modular interchangeable algorithm implementation and performance evaluation of: (1) Fusion of positional data from imaging and non-imaging sensors; (2) Fusion of attribute information obtained from imaging and non-imaging sensors and other sources such as communication systems, and (3) Object Recognition (OR) in imaging data. More precisely, the data comes from the non-imaging and imaging sensors that are typically present on airborne platforms performing maritime surveillance. The *non-imaging* sensors are a 2-D radar (AN/APS-506), an Electronic Support Measures (ESM), an Identification Friend of Foe (IFF) and a datalink (Link-11). The *imaging* sensors present on the Aurora are the Forward Looking Infra-Red (FLIR) and a Synthetic Aperture Radar (SAR) which can operate in Strip Map, Range Doppler Profiling and Spotlight modes (Adaptive or Non-Adaptive). This SAR is presently in the Advanced Development Model (ADM) phase and is currently in integration and testing phase for deployment in late 1998. Real data is also available from the previous eXperimental Development Model (XDM) phase. The attribute data that these sensors can provide is important in determining the identification of target platforms, particularly the long range features that the Spotlight SAR can furnish about navals targets in the maritime surveillance missions of the Aurora.
The ADFT is generic enough to demonstrate fusion of any sensor suite that comprises surveillance radars, communications data (datalinks) as well as imaging and non-imaging sensors. As such, it could also be adapted to other platforms, and for different purposes, without any major modifications. The same type of capabilities are also necessary for many other applications such as remote sensing, medical imaging, robotics, air traffic control, etc

* Proceedings of Conference FUSION ‘98, Las Vegas, July 1998

2. Implementation of Fusion Functions

Any generic DF application must contain the following set of sequential functions to act on simulated (where ground truth is known) or real data (where ground truth is not necessarily known):

1. *registration* to first perform spatial and temporal alignment of the simulated (or real) sensor data,

2. an *association* mechanism to then correlate the new incoming data with possible existing tracks found in the track database and send associated positional data to positional fusion and associated attribute data to attribute fusion,

3. *positional estimation* to then update the tracks in the time domain with the associated new data and write this positional information to the track database, possibly extracting attribute data such as speed, acceleration and sending to attribute fusion,

4. *identification (ID) estimation* (or *attribute fusion*) to then fuse all attribute data through evidential reasoning, whether they originate from imaging (through image understanding and feature extraction) or non-imaging sensors and consequently update the dynamic track database.

For non-imaging sensor reports, a Nearest-Neighbour *association* scheme performs well for low clutter, which is presently the case studied. Such an optimal single-scan associator is the Jonker-Volgenant-Castanon algorithm. A Multiple Hypothesis Tracker is envisioned for high clutter and dense targets. Imaging sensor acquisition (particularly the SAR) is queued by the operator on a given target of interest so the *association* mechanism is straightforward in this case.

Positional estimation involves an adaptive Kalman filter of some type. For manoeuvering air targets, an Multiple Model with rule-based switching has been found to work well (with constant-velocity motion models differing only by their process noise matrix), while an adaptive sinlgle-motion model Kalman filter is sufficient for slow-moving naval targets which are the prime targets for the Aurora's main missions. The *positional estimation* function uses radar (range, bearing data), IFF (range, bearing and possibly altitude data), ESM (bearing only data) and Link-11 track data.

ID estimation processes IFF, ESM, Link-11, tracking information and imaging features to achieve a correlation with attributes found in platform, emitter and geo-political databases. LM Canada has developed a truncated Dempster-Shafer (DS) algorithm for the *ID estimation* function which avoids the NP-hard aspect of DS while possessing robustness under countermeasures and affording a proper handling of data with poor or uncertain quality or uncertain interpretation. *ID estimation* proceeds by reasoning over attributes provided by the sensors and also by positional estimation and will thus be referred to as *attribute fusion* from now on. Each of these attributes must be listed in data bases such as the Platform Data Base (PDB), the Emitter Name List (ENL) containing all the emitters that can be identified by the ESM, and the Geo-Political List (GPL) containing the allegiance of each country.

The rest of the paper is thus organized as follows: first a description of the attribute contents of the databases, next a discussion of the fuzzification of the measured attribute "values" to render them uncertain, and finally the reasoning scheme utilized to obtain a list of possible ID with associated belief and plausibilities, eventually culminating in a single ID (a singleton of the PDB) with a high belief and a plausibility near 1.

3. Databases And Their Contents

The most important of the databases is the PDB. It must contain first a comprehensive list of all the physical parameters that can be measured by the sensors [1]. These are *real-valued* variables that can, in principle, be measured (directly or indirectly) by the sensors to an accuracy given by the sensors or by the method used to process the sensor's measurements. An example of a direct measurement is the length of a ship as ascertained from an image interpretation algorithm of known performance, thus giving an error estimate for the measurement. An indirect measurement would be a value for the speed as reported by *positional estimation* with an estimate of the error being provided in the track's covariance matrix. Other direct measurements are Radar Cross-Section (RCS), height, width, and another indirect measurement is acceleration.

The other type of information that must be present in the database concerns more directly the platform ID, namely its category, type, class, a list of its known emitters and the like. These are *not real-valued* variables and no error can be assigned to a report. However some confidence can be associated to the sensor report in the form of a belief that the proper ID was achieved (e.g. for the ESM) or in the expected algorithmic performance of an image interpretation algorithm of the Image Support Module (ISM) by looking at its image training set. This assumes some access to the ground truth scenario and appropriate imagery and is most easily appraised by using simulated data and images.

The complete list of the attributes present for each platform entry (labelled both by a character string and a number to be used in proposition formation) in the PDB follows below, with the sensors likely to provide information being listed in square brackets (Tracking refers to parameters associated to the stable track in the PDB):

1. The maximum acceleration ACC, the maximum platform speed V_MAXI and the minimum platform speed V_MINI all serve as bounds to discriminate between possible air target identifications. The acceleration and speed are assessed by *positional estimation* and can be used to infer identity proposition by the *attribute fusion* function [Tracking, Link-11].
2. ACRO is the acronym of the country name indicated in the GPL and used also to refer to the country that owns the platform in the PDB. In the PDB, ACRO is used by the attribute fusion function to link the PDB platform with the country allegiance indicated in the GPL.
3. ALT_MAXIM is the maximum altitude that a platform may reach [bound for altitude reported by IFF].
4. The variable EMITTER_LIST is an exhaustive list of the INEML number which correspond to the emitters that are carried by the platform [ESM].
5. Three geometrical dimensions which should be interpreted as fuzzy logic variables by the proposition interpreter of an attribute fusion function [SAR, FLIR]:
6. HEI, the physical height of the platform in its natural vertical axis (for a surface target this correspond to the average height between bridge and water line),
7. WID the physical width of the platform in its natural transversal axis,
8. LEN the physical length of the platform in its natural longitudinal axis.
9. IR is a field associated to the infrared signature of the platform as seen by an infrared sensor system and its operator [FLIR].
10. PLATYPE forms the first level of platform classification used in this PDB. This variable is closely related to the category descriptor given by the ISM and reflects its platform military utilization [SAR classifier, FLIR classifier].
11. RCS_FOR, RCS_SID, RCS_TOP correspond respectively to radar cross-section of the platform seen from the front, the side and the top view respectively. These fields represent variables which should be interpreted as fuzzy logic variables. The values are empirically much larger than the geometrical cross-section obtained by the product of the two relevant dimensions (HEI, WID, LEN) since metallic objects offer strong radar backscatter when compared to the geometrical cross-section, i.e. the one that can be evaluated from the outlines of the target by the ISM.
12. SUBTYPE provides a sub-classification of the platform type [SAR, FLIR].

There are two Name Lists that are linked in functionality to the PDB. The ENL links each of the INEML emitter numbers appearing in the EMITTER_LIST of the PDB entries to a character string representing that particular emitter, following standard conventional naming scheme, e.g. the one appearing in Jane's Books. The GPL links the country's ACRO variable of the PDB to the country allegiance for proper interpretation of IFF reports.

In order to give an estimate of the size of the problem at hand, the current PDB contains 142 distinct platforms, the ENL identifies 179 emitters and the GPL identifies 24 countries and specifies their allegiance, which can vary from mission to mission independently and without interfering with the PDB nor the ENL.

4. Attribute Measurements

Some sensors measure attributes quite directly. For example the ESM will provide an emitter list with some confidence level about the accuracy of the list that reflects the confidence in its electromagnetic spectral fit. The Link-11 data has to be taken at face value unless the reporting platform is known a priori to provide poor track quality data.

The IFF can however lead to some complications. An IFF response does lead to an identification of a friendly or commercial target but the lack of a response does not necessarily imply that the interrogated platform is hostile. One has to distribute the lack of a response between at least two declarations: the most probable foe declaration and a less probable friendly or neutral declaration that allows for an IFF equipment that is not working or absent.

Similar complications arise when dealing with kinematic parameters reported occasionally by the tracker in *positional estimation.* Firstly, each physical quantity has a different dimension (speed, acceleration) and an accurate determination is not necessarily needed for fusion. Indeed it is convenient to bin the attribute "speed" into fuzzy classes from "very fast" to "very slow" (separately for air and surface targets). Similar binning for acceleration could range from "very large g" to "very small g". Membership in each class is a measure of how well the measured value fits into the descriptor as described below.

Further, speed or acceleration reports must be fused only if they involve a significant change from past historical behaviour in that track. The reason is two-fold: firstly no single sensor must attempt to repeatedly fuse identical ID declarations otherwise the hypothesis that sensor reports are statistically independent is violated, and secondly the benefits of the fusion of multiple sensors is lost when one sensor dominates the reports. Furthermore, a measured value of speed (or acceleration) only indicates that the target is capable of that speed, not that is corresponds to either V_MAXI (or V_MINI) of the PDB (nor the maximal ACC of the PDB). It is a reasonable working hypothesis to fuzzify the value reported by the tracker into adjacent "bins" to account for the target being at, say only 80% of its optimal speed (a "very fast" target can occasionally travel "fast"), or travelling with a strong tailwind (a "fast" target can occasionally appear as "very fast"). Finally the concept of binning can be generalized to continuous membership functions of a fuzzy set.

Similarly, the image interpretation module for either the SAR or the FLIR can generate a nearly infinite set of declarations from a single given image. Care must be taken to preserve as much independence between the declarations and certainly prevent any conflict. Such an independence can be achieved to a reasonable extent if different features are extracted from the image in different steps or if totally different mathematical algorithms are used in each step. The so-called Image Support Module (ISM) which LM Canada is designing for image interpretation of SAR data (presently XDM, but later this year ADM) is the 2-D equivalent of the ESM's 1-D signal interpretation. The present design involves the four steps described in Figure 1 below of which the first three have been implemented and tested [2]. The design shown for ship features extraction and ship classification involves a hierarchical decision tree whose logic is displayed in Figure 1.

Given the image acquisition parameters and the navigation data from the Aurora's DF tracker, Step 1 checks if proper ship orientation is achieved (e.g. the image is sufficiently elongated), and, if so, an image segmentation process detects a target whose image is simply connected. In Step 2, a Hough transform then permits an estimation of the ship length, which is immediately sent to DF for the ID estimation process. The length reported has an intrinsic uncertainty which is fuzzified into length bins whose width is further discussed below. The reported length is then fused with the length attribute listed for all the ships in the PDB.

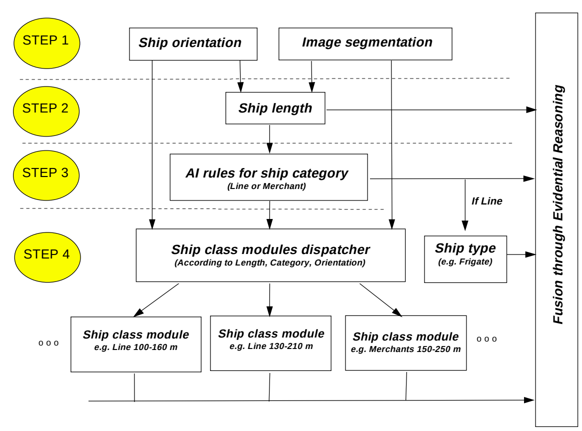


Figure 1. Hierarchical ship features extraction and ship classification

In Step 3, Artificial Intelligence rules based on the relative position and number of main scatterers (as identified by pixel intensities being above a certain threshold) allow the determination of ship category into "line" or "merchant" categories by locating its superstructure. The presently implemented method is a Neural Net (NN) trained on 37 production rules based on the location of the main radar scatterers in 9 different regions along the length of the ship. The possible outputs of the NN are "line", "merchant" or "unrecognized". It should be noted that these categories are only a subset of the NATO STANAG where "line" is only a subset of combatant ships (other combatants include Amphibious Warfare, Mine Warfare, and Patrol ships) and "merchant" is a subset of so-called "non-naval" entities (which include also Fishing, Leisure, and Law Enforcement ships). They are however the main categories relevant for the Aurora missions mentioned earlier. An "unrecognized" declaration from the NN indicates that it could reach an ID and consequently that declaration is assigned to the ignorance in the DS algorithm for evidential reasoning.

Step 3 also performs an attempts at identifying ship class if the NN declaration for "line" is sufficiently large (say >50%). This is due to the observed correlation between ship length and ship class observed from a survey of about 100 classes of ships in Jane's Fighting Ships, as shown in Figure 2 below. Note that this survey is arbitrarily normalized to its maximum value to better show the relative spread in length values for each class. For an actual mission, some knowledge about the relative population of each class would renormalize the data. The smallest width of these distributions can serve as an indication of the binning size needed for the fuzzification of the length attribute in Step 2. The line types which are generated in this fashion can discriminate between frigates, destroyers, cruisers, battleships and carriers (as identified in the PDB). An indication of the fuzziness of the declaration is given by the relative overlap between classes for a given measured length. From Figure 2, it can be seen that a length measurement around 250 meters generates propositions for battleships and carriers with roughly the same confidence levels, but a measurement of 180 meters is almost exclusively assigned to cruisers.

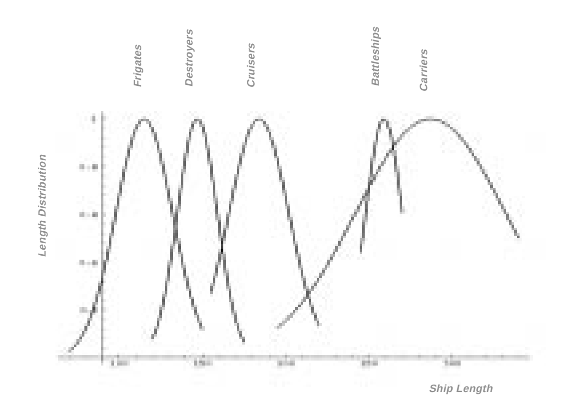


Figure 2. Length Distribution for Line ships

Finally in Step 4, specialized NNs trained on subsets of the database of ship images (artificially created from a simulator for various aspect and depression angles), that span a given length interval, refine the ID declaration to ship class (e. frigate of Halifax class, destroyer of Spruance class). The outputs of the neural net for each possible class are again numbers between 0 and 1 which are interpreted as the level of confidence in obtaining the correct class ID. The neural net also provides an "unrecognized" class which again reflects its inability to reach a conclusion about ship class. This is then attributed to the ignorance in the DS sense, as in Step 3.

5. Attribute Fusion By Truncated Dempster-Shafer

The Dempster-Shafer (DS) theory of evidence offers a powerful approach to manage the uncertainties within the problem of target identity. Every sensor declaration about the M possible "values" of an attribute assigns a Basic Probability Mass (BPM) value m_i (i=1...M) to that attribute (present in the database) and generates M propositions which are just the numerical list of platforms in the PDB that can attain the said value for the attribute. For a PDB containing N platforms, the numerical list of platforms which forms a proposition is represented in the current implementation by a string of zeroes and ones in the location of a string of N bits. This is done to speed up calculations by bit manipulations for ensemble operations such as union and intersection, which are needed in DS theory. For physical quantities like speed, length, RCS and image classification attributes like category or class, M is usually greater than 1. This is due either to the fuzzification of the physical quantity or because of the inherently complex nature of the algorithmic determination of the attribute (e.g. by NN outputs). DS theory is particularly suited for our application because it requires no a priori information, can resolve conflicts (present in hostile environments due to countermeasures), and can assign a mathematical meaning to ignorance (which is the result of some of the chosen algorithms).

However, traditional DS has the major inconvenience of being an NP-hard problem. As various evidences are combined over time, DS combination rules will have a tendency to generate more and more propositions which in turn will have to be combined with new input evidences. Since this problem increases exponentially, the number of retained solutions must be limited. Our truncated version of DS theory of evidence performs the conventional combination rules of DS theory but retains the final solution proposition according to the following criteria [3]:

1. All combined propositions which have a BPM higher than MaxB are retained.

2. All combined propositions which have a BPM lower than MinB are eliminated.

3. If the number of retained propositions in step 1 is smaller than MaxNum, the subroutine will retain, by decreasing BPM, the propositions consisting of one element (singleton) until MaxNum is reached. If MaxNum is not reached, one retains, by decreasing BPM, the propositions consisting of two elements. The process is repeated until MaxNum is reached.

The values for MaxB, MinB and MaxNum presently used are 0.05, 0.001 and 8 respectively. Step 3 takes into consideration that the platforms commanding officer favours the singleton propositions. With this truncation, the previously mentioned commutativity still applies but the associativity sometimes does not apply. When the BPMs are small and many propositions are fused, the process loses the associative property and the results (identification) become dependent of the combination order. From our investigation and experience, in the worst case, this violation has a small effect and changes the calculated BPM by no more than 5 % compared to the standard theory. The worst case is the one where the MaxNum selected propositions have all a BPM of the order of MaxB. In such cases there is a large portion of BPM that goes to the dropped propositions. Fortunately, in practical applications sensors provide attribute with enough evidence (more than 30%) to avoid this situation.

6. Maritime Air Area Operations

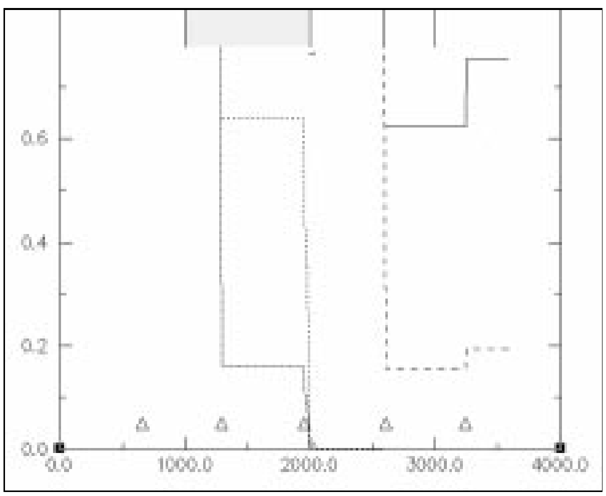
The truncated DS scheme for fusion of imaging features with non-imaging sensor reports is being tested on four scenarios that are mission-critical for the CP-140 aircraft: Maritime Air Area Operations, Direct Fleet Support, Counter Drug Operations and Maritime Sovereignty Patrols. Results are only presented here for the first one.

Although the Maritime Air Area Operations scenario contains various air targets for tracking and various merchant ships, the most relevant targets for evidential reasoning are a flotilla of 3 ships from the former USSR, one destroyer of

Udaloy class, one cruiser of Kara class, and one frigate of Mirka class, with common heading and common speed of 30 knots, a value close to their maximum attainable speed (according to the values in the PDB). These ships are imaged by the SAR when the Aurora is flying by at right angles and the relative heading is 45 degrees (in order to obtain a good simulated image [2]) and the long range features of the first 3 steps of the classifier are fused with (mainly) ESM reports from the same platforms. In addition, ESM reports from the Udaloy are intentionally chosen to be occasionally incorrect, in order to test the robustness of the DS algorithm under countermeasures.

6.1 ID Estimation of the Mirka-II

Figure 3 shows the temporal evolution of four propositions having, during a certain amount of time, the highest mass. Five triangles at the bottom of the figure represent the time at which an ESM report has been fused.



After the first 10 minutes (t=656 s), the Kara-Azov and the Mirka are not properly resolved (within an angle of 1º). The emitter #92 belonging to the Kara-Azov and other platforms is detected, provoking the initiation of proposition 1 in which the Mirka-II is absent. Then, at t=1293 s, the emitter #103 is detected which belong to the Mirka-II and to the Kara-Azov. As a result, proposition 2 emerges. The ground-truth shows that, this time, it is emitted by the Mirka-II.

Prop.	Platforms
1	{Kirov-Ushakov/Lazarev/Velikyi, Kara-Azov}
2	{Kara-Azov}
3	{Mirka-I, Mirka-II, Sam-Kotlin}
4	{Mirka-II}

At t=1950 s, the emitter #56 is detected which only belong to the Mirka-II. The mass associated to proposition 2 decreases. A SAR image is acquired an analysed at time t=1980 s. The fusion of the Ship-Length attribute confirms the elimination of proposition 3 (the Kara is a cruiser two times longer than the Mirka-II) and proposition 3 becomes preeminent. The fusion of the Ship Type attribute at time t=2040 s increases the mass of the proposition 2. Then, at time t=2606 s and 3243 s, two emitters (#44, #55) belonging only to the Mirka-II provokes the mergence and confirmation of the proposition 4.

6.2 ID Estimation of the Kara-Azov

Figure 4 shows, for the Kara-Azov, the same type of information shown in Figure 3.

At time t=1275 s, the emitter #45 is detected which belong to a large list of platforms creating proposition 1.Then, at time t=1950 s, the emitter