Feature-Aided Tracking of Ground Vehicles using Passive Acoustic Sensor Arrays

VISHAL CHOLAPADI RAVINDRA
YAAKOV BAR-SHALOM
THYAGARAJU DAMARLA

Tracking of a moving ground target using acoustic signals obtained from a passive sensor network is a difficult problem as the signals are contaminated by wind noise and are hampered by road conditions, terrain and multipath, etc., and are not deterministic. Multiple target tracking becomes even more challenging, especially when some of the vehicles are light (e.g., wheeled) and some are heavy (e.g., heavy wheeled vehicles like trucks, tracked vehicles like tanks, etc.). In such cases the stronger acoustic signals from the heavy vehicles can mask those from the light vehicles, leading to poor detection of such targets. The full position estimates of emitters (targets), obtained following the association of the DoA angle estimates from multiple sensor arrays at each time scan, are used for target tracking. However, because of the particular challenges encountered in multiple ground vehicle scenarios, this association using kinematic (DoA angle) measurements only is not always reliable and can lead to lost as well as false tracks.

In this paper we propose a new feature-augmented static association algorithm where feature augmented DoA angle measurements from multiple sensors are associated to localize targets and obtain composite measurements (position estimates) using a static multidimensional assignment (MDA) framework. We present a novel DoA detection scheme followed by a feature extraction technique designed from and for real data. Dynamic S-D and feature-aided S-D (multidimensional) assignment algorithms are presented to assign composite measurements and feature-augmented composite measurements, respectively, to tracks. The techniques are developed based on real data sets and tested on real data based on a field experiment.

Manuscript received August 06, 2009; revised February 03, 2010; released for publication August 17, 2010.

Refereeing of this contribution was handled by Dr. Huimin Chen and Dr. Stefano Coraluppi.

Authors' addresses: V. C. Ravindra and Y. Bar-Shalom, Department of Electrical & Computer Engineering, University of Connecticut, Storrs, CT 06269, E-mail: (vishalcr@nal.res.in, ybs@engr.uconn.edu); T. Damarla, Army Research Laboratories, Adelphi, MD 20783, E-mail: (rdamarla@arl.army.mil).

1557-6418/10/$17.00 © 2010 JAIF

1. INTRODUCTION

Ground vehicle tracking using acoustic data obtained from passive sensor networks is a very challenging problem as the signals are contaminated by wind noise, and are hampered by road conditions, terrain, multipath, etc., and are not deterministic. Multiple vehicle tracking becomes even more challenging, especially when some of the vehicles are light vehicles, some are heavy wheeled vehicles and some are tracked like tanks, and are closely spaced. Passive acoustic sensors are gaining in popularity because of their low cost, ease of deployment, and the fact that they can be deployed on the ground. As passive sensors do not emit their own signals, unlike active sensors, there is no danger of being detected. Passive acoustic sensor networks are being used in battlefield monitoring as well as in civilian surveillance applications.

In single target scenarios with active sensors, kinematic measurements (such as range, bearing, etc.) are obtained which can be used to estimate the trajectory of targets [2]. When a network of passive sensors is used, however, the range of a target can be obtained only after associating the direction of arrival (DoA, or line of sight—LoS) angle estimates obtained by at least three sensors. This is because ghosting or false intersections occur with just two sensors with multiple targets in the same plane. Hence, in order to eliminate or reduce ghosting one has to use DoAs from at least three sensors (see [1, Sec. 8.8.2]). This makes the problem computationally expensive for a large number of measurements. There is added difficulty in data association when targets stay close together over an extended period of time (as acoustic signals from some targets can fade and then re-appear or can be masked by stronger signals from other targets), because one has to associate the DoA angle estimates from the *same target* to obtain its position.

Feature-aided tracking (FAT) is a rapidly developing research area [15, 16, 19, 33], as various FAT techniques exploit certain properties of the received signals at the sensor level to augment the kinematic measurements to alleviate the difficulties encountered in data association and tracking using kinematic measurements only. For scenarios such as the one considered in the present paper, traditional data association and target tracking algorithms can be enhanced by integrating features with the kinematic data. Typical features are estimated target dimensions, radar cross-section, signature data and received signal properties among others. Features have been typically used to enhance target identification, i.e., classification, discrimination or recognition [7] and to enhance data or track association [1, 15].

Classification aided tracking is presented in [15, 16] where it is offered as an alternative to the traditional approach (with the classifier being treated separately from the tracker). However, a unique mathematical model is needed to be constructed for each target type in order to make such an algorithm robust. In [8] a multilevel feature-based association algorithm to simultaneously track and identify targets is presented, where features extracted from high range resolution (HRR) profiles are correlated with target signatures for identification. The simultaneous use of target classification information and target kinematic measurements for target tracking using a multidimensional assignment framework is presented in [4]. Another paper [28] studies the combining of identification and tracking, as each task can be enhanced by fusing it with the other; however, these methods require the availability of target signatures beforehand. In [22] kinematic measurements from ground moving target indicator (GMTI) and local motion features from HRR are combined in a probabilistic logic based tracker for dense multitarget scenarios. Features extracted from HRR profiles, by approximating it with a Gaussian mixture density using the expectation maximization (EM, see [17]) algorithm, are combined with kinematic measurements in a probabilistic framework for multiple target tracking in [33]. Another algorithm is presented in [21] to combine GMTI measurements with features extracted in the wavelet domain from HRR profiles for multitarget tracking in a joint probabilistic data association (JPDA, see [1, 3]) context. Various feature-aided classification and FAT techniques for single target as well as multiple targets have been proposed and implemented in the recent past for ground vehicle scenarios using acoustic signals from sensor arrays. Classification and identification of multiple targets using acoustic signatures is presented in [12]. In this approach, bearing tracking and data association is first performed. Once the bearings of the targets are established, beamforming is performed in each individual direction to extract harmonic features, and a multi-variate Gaussian classifier is applied on each feature set for identification and classification. Many target classification algorithms based on acoustic signatures in the literature [20, 24, 25, 34] assume that a single target is present and use statistical parameters, namely the mean and variance of several harmonics of the fundamental engine firing rate of each target for identification and classification. When multiple targets are present within the surveillance region of acoustic sensor arrays, the measurements no longer exhibit the same statistics as they did for the individual targets when alone, due to masking and interference. Hence, the algorithms developed with a single target assumption perform poorly [12]. In [11] an algorithm to track multiple ground vehicles was presented based on a template of the DoA angles for the leading target or the target closest to the sensor array and hence the loudest with the strongest signal to noise ratio (SNR). This template was used to predict the DoA angles of all the other targets. A distributed fusion algorithm is developed in [38] that extracts features in time as well as frequency domain from the acoustic signals and integrates classification results from different sensor arrays to increase the classification accuracy. Acoustic signatures from air and ground vehicles produced from their engine or propulsion mechanisms are used by neural network based pattern recognition algorithms for classification in [35]. Target tracking is easier if the identities of signal sources (targets) are known beforehand based on classification techniques [12, 13]. However, this is an unreasonable assumption in the scenario considered in the present paper due to the various challenges discussed earlier. Moreover, the extraction of reliable signatures is challenging because of the environment, interference from other moving parts and nearby vehicles and their nonstationarity. Further, a majority of the classification techniques require building reliable mathematical models and templates as well as an extensive library of training data for each target type.

In the present paper we present a feature-aided data association technique that employs a *feature-augmented* measurement set (as compared to just kinematic measurements) in a static as well as dynamic assignment framework for target localization and tracking. The major challenge is the development of reliable models to extract and characterize features. Section 2 describes the generation of the PSD of the signals received by the passive sensors using the minimum variance distortionless response (MVDR, see [36]) spectral estimation technique, and a novel DoA angle detection scheme using the PSD. Section 3 describes a novel algorithm used to extract features from the PSD. Section 4 introduces and describes the target localization problem, i.e., obtaining their full position estimates (composite measurements) using a multidimensional assignment (MDA[1]) framework. The composite measurements are assigned to tracks using both conventional as well as feature-aided dynamic MDA algorithms in Section 5. This multisensor information processing is configuration III from [1, Sec. 8.2]. Section 6 describes the real data scenario considered in the present paper and Section 7 provides the target localization and tracking performance comparison results between the conventional and feature-aided MDA algorithms. The techniques are developed based on real data sets and tested on real data based on a field experiment.

2. PSD GENERATION AND DOA DETECTION

Circular 2-D passive sensor arrays made up of M microphones arranged equidistantly, as shown in Fig. 1, are employed in a passive sensor network on the ground to listen to multiple ground vehicles. Assuming that the vehicles are at a sufficiently large distance from the sensor array, so that the received signals from the targets can be approximated by a planar wavefront, various wideband beamforming algorithms described in

[1]MDA is also known as S-D assignment. Note the use of S for the dimension of the assignment (the number of lists).

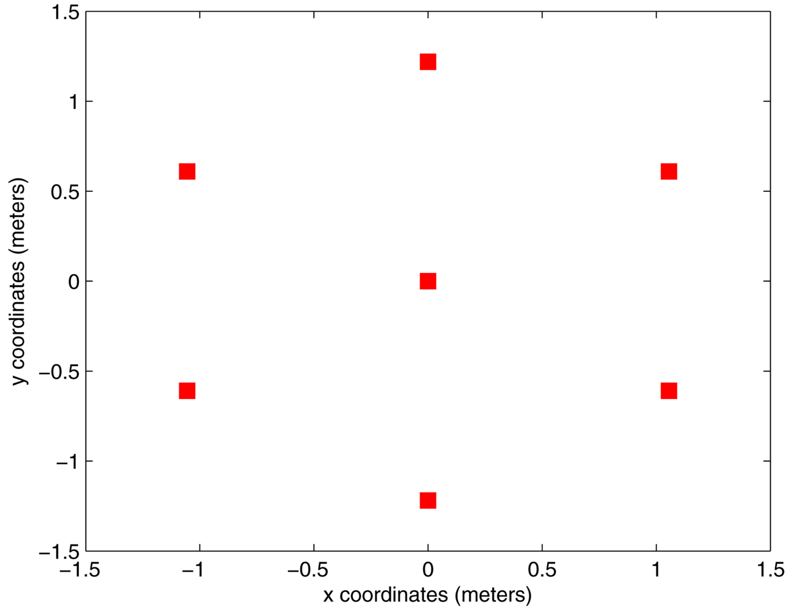


Fig. 1. Circular sensor array made up of 7 microphones.

[9, 10, 13, 23, 26, 27, 36, 37, 39, 41] can be used to detect signal sources and estimate their DoA angles with respect to the known sensor array positions.

At each sampling time, a fast Fourier transform (FFT) is performed on the raw acoustic data from each microphone in an array and a discrete frequency band from f_1 to f_{n_b}, where n_b is the number of bins with bin intervals of 1 Hz, is chosen for processing. The frequency bins are chosen to conform to the typical frequencies of the acoustic signals emitted by engines and other moving vehicle parts. As a result, we have a wideband processing algorithm that uses data from n_b frequency bins. We denote the FFT data used for processing at each sample by $X(m,f_b)$, where $m \in \{1,\ldots,M\}$ represents each microphone and $f_b \in \{f_1,\ldots,f_{n_b}\}$ indicates the frequency bins. The minimum variance distortionless response (MVDR) algorithm provides, for each sensor array, an estimate of the power spectral density (PSD) of the acoustic signals impinging the array at a particular scan

$$\hat{P}(\theta_d,f_b) = \frac{1}{\underline{a}(f_b)'\hat{R}(f_b)^{-1}\underline{a}(f_b)},$$

$$\theta_d \in \{\theta_1,\ldots,\theta_D\} \quad \text{and} \quad f_b \in \{f_1,\ldots,f_{n_b}\} \tag{1}$$

and

$$\underline{a}(f_b) = \{X\}'_{f_b} V \tag{2}$$

with $\{X\}_{f_b}$ denoting the column of the matrix X corresponding to column f_b, V denoting the steering matrix (of dimension $M \times 360$, representing all possible directions in degrees, see Ch. 2, [36]), and

$$\hat{R}(f_b) = \frac{1}{M}\sum_{m=1}^{M} X(m,f_b)^* X(m,f_b) \tag{3}$$

the estimated covariance matrix.

The DoA angles are estimated in the present paper by a novel DoA angle detection scheme from the PSD (illustrated in Fig. 2). DoA detection, at each angle θ, can be performed by applying a thresholding algorithm on the frequency averaged estimated power spectrum corresponding to direction θ

$$\hat{P}(\theta) = \frac{1}{n_b}\sum_{b=1}^{n_b} \hat{P}(\theta,f_b). \tag{4}$$

Fig. 2 shows a snapshot of the estimated PSD at a particular scan for a sensor array. Along each angle θ the frequency averaged PSD denoted by $\hat{P}(\theta)$ in (4) is illustrated in Fig. 3(a). In [13], thresholding on $\hat{P}(\theta)$ was used to detect DoA angles, i.e., a DoA angle detection was declared at θ_d if a peak was observed in the spectrum $\hat{P}(\theta)$ at θ_d. The first derivative $d\hat{P}(\theta)/d\theta$ is shown in Fig. 3(b). A DoA detection is made at angle θ_d if a positive peak is detected at θ_d in $-d^2\hat{P}(\theta)/d\theta^2$ shown in Fig. 3(c). This is because a thresholding algorithm based just on the frequency averaged power spectrum is likely to miss peaks which can be detected by a thresholding algorithm that is based on the second derivative of the spectrum. In the example shown in the figure, the DoA angles detected are 239°, 123°, 219° and 189°, arranged in decreasing order of their corresponding amplitudes.

In a typical ground vehicle tracking scenario where the targets are moving on road or off-road conditions, there are a variety of extraneous factors which affect the acoustic signal, such as road conditions, sound generated by moving parts, wind, etc. This results in numerous false alarms, i.e., power detections in frequency-angle bins not directly due to the engine or moving parts of the targets. There are also missing DoA angle estimates (missed detections) due to signal attenuation or the possible masking of signals from lighter vehicles

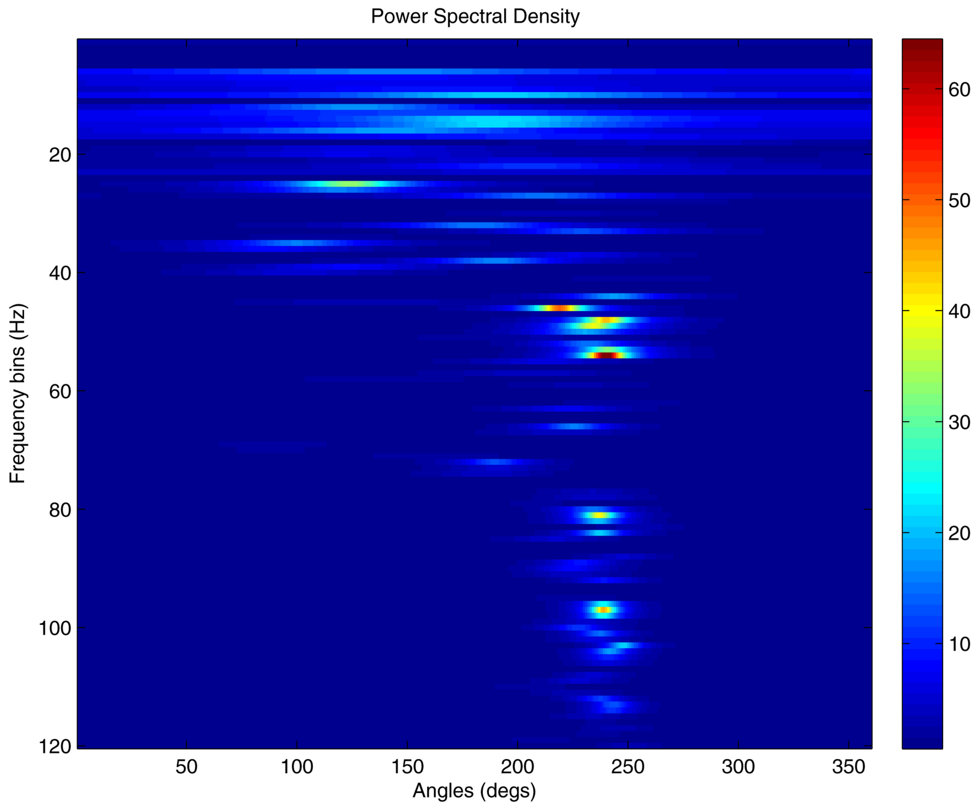


Fig. 2. Power spectral density as a function of DoA angles and frequency bins obtained from the MVDR algorithm for one sensor array at a particular scan (data set [12]).

by those from the heavier vehicles, especially when the targets are closely spaced. As a result, the quality of the data association is low if just the DoA angle estimates are used as measurements. Therefore, the power spectrum will be exploited to generate features which could enhance the accuracy of data association. Motivated by this, a new feature extraction procedure is described in the next section.

3. FEATURE EXTRACTION

In feature-aided ground vehicle tracking applications the peak amplitudes of the spectrum have been used as features [12, 35] in the past; however, due to signal attenuation and masking in the case of multiple vehicle road convoys, they are not reliable as features. The location and the spread of the peaks carry more useful information about the signal source, compared to the amplitude, because they are not as affected by signal attenuation and because a vehicle causes frequency peaks in similar locations of the spectrum across all the sensor arrays at any given time.[2] In this paper we propose a statistical modeling of the power distribution in the frequency bins along each detected DoA and use a Gaussian mixture model (GMM) to extract feature vectors instead of just scalar features. For the problem considered only scalar features were used in [12, 13].

[2]Doppler has been neglected as it amounts to less than 1–3(assuming 5–10 m/s vehicle speed).

3.1. Fitting of a Gaussian Mixture Model (GMM)

The observed data $d(x)$, i.e., the estimated power spectrum in a particular direction x, is modeled as

$$d(x) = y(x;\beta) + \epsilon(x) \tag{5}$$

where $y(x;\beta)$ is the fitted parametric model, β is the parameter vector and $\epsilon(x)$ is the fitting error. The objective is to estimate the parameters of the model such that the error (noise) is minimized in a statistical sense.

The nonlinear least squares (NLLS) method is used to estimate the parameters of a nonlinear model—a GMM—used to fit the data. The Gaussian mixture modeling, useful for peak finding applications, is given by

$$y(x;\beta) = \sum_{l=1}^{n} \alpha^l \exp\left[-\frac{1}{2}\left(\frac{x-\mu^l}{\sigma^l}\right)^2\right] \tag{6}$$

where $\beta = [(\alpha^1,\mu^1,\sigma^1),\ldots,(\alpha^n,\mu^n,\sigma^n)]'$ is the parameter vector of the GMM such that α^l is the weight (amplitude), μ^l is the location and σ^l is the width of the peak of component l, and n is the number of components of the GMM. The NLLS method can be used to estimate the parameters of the GMM which best fit the observed data in (5). Fig. 4 illustrates the GMM fitting results when applied to power spectrum data for certain DoA detections obtained by three sensor arrays at a particular time. A GMM with $n = 4$ components was used to ex-

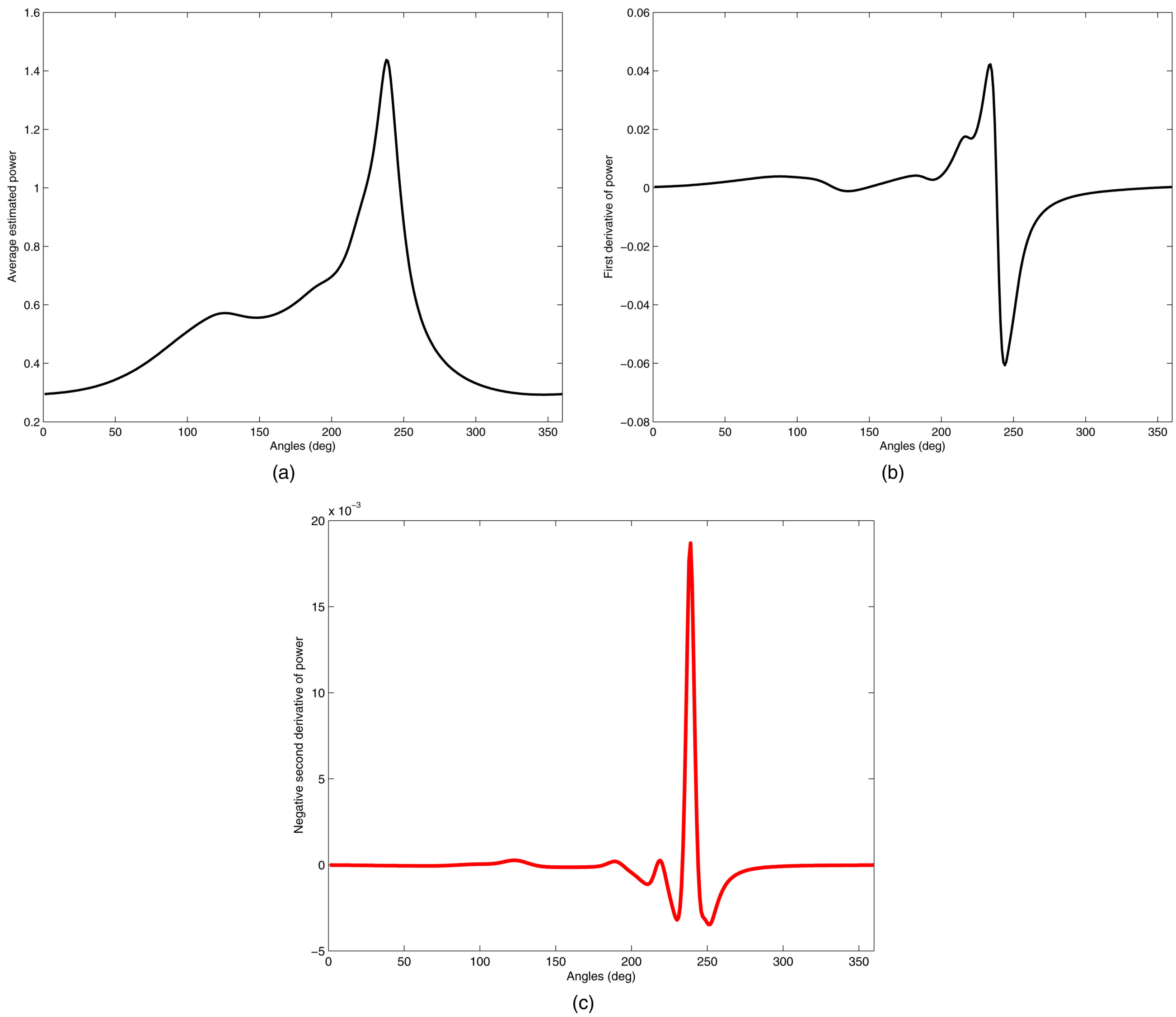


Fig. 3. The thresholding used to detect DoAs from a particular sensor array at a certain scan: (a) frequency averaged power spectrum as a function of angles only, (b) the first derivative of the function in (a), and (c) the negative of the second derivative of the function in (a).

tract the location, width and the amplitude of the peaks from the power spectrum data.[3]

From the example illustrated in Fig. 4, one can see that there are at most $n = 4$ frequency peaks from each of the $S = 3$ sensors, corresponding to three detected DoA angles θ_1, θ_2 and θ_3 (subscripted by the sensor index). These peak locations form a matrix of dimension $n \times S$ (with the column elements listed in decreasing amplitude order)

$$L = \begin{pmatrix} 40 & 78 & 40 \\ 15 & 40 & 18 \\ 36 & 117 & 37 \\ 49 & 15 & 14 \end{pmatrix}. \qquad (7)$$

[3]The selection of n is taken as a design parameter, a larger n was found to lead to excessive uncertainty in the feature model for the application considered.

The peak location matrix given in (7) for the three sensor arrays in Fig. 4 shows that the peak locations are not necessarily matched across the S sensors, i.e., a peak location is not necessarily matched with the peak locations shown in the same row of the other $S-1$ columns (lists). For example, in (7), the first row is [40 78 40]. The location of the peak in the second column (at 78 Hz) is not close to the location of the peaks in the first and third columns (at 40 Hz). Hence, in order to properly order peak locations across lists (sensors), each peak location should be matched to other peak locations (in the other lists) in such a way that each matched S-tuple of peak locations consists of peak locations which are close to each other. If a peak is located in such a way that it cannot be matched to any other peaks in the remaining lists, it is matched to a dummy element which indicates a missing peak detection. Each peak location is matched to at most one corresponding peak location from another list. This matching of peak locations is

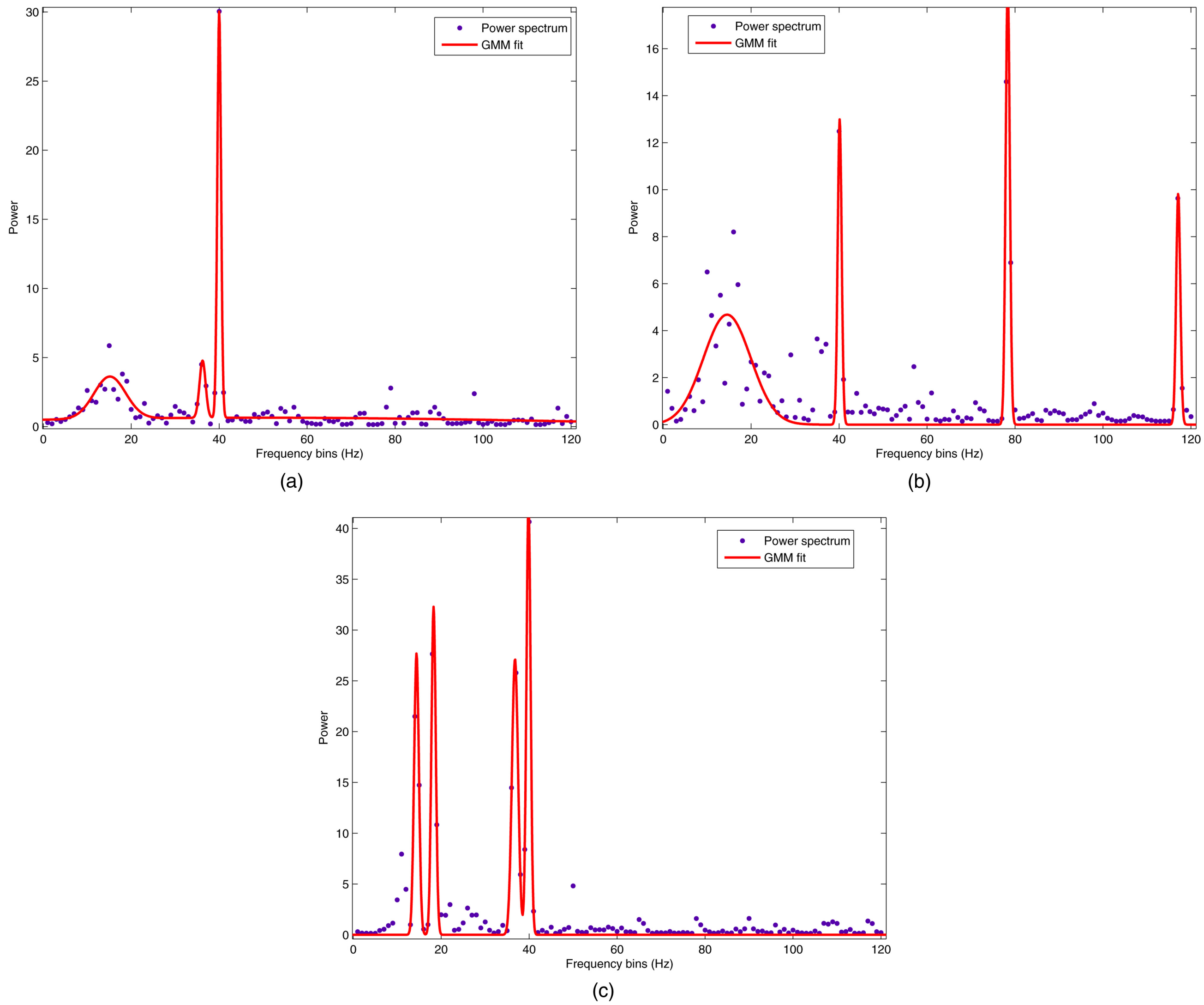


Fig. 4. Gaussian mixture model fitting to extract features: (a) feature data from sensor array 1 (DoA angle θ_1, peak locations: [40, 15, 36, 49]), (b) feature data from array 2 (DoA angle θ_2, peak locations: [78,40,117,15]), and (c) feature data from array 3 (DoA angle θ_3, peak locations: [40, 18, 37, 14]).

done by solving a generalized multidimensional assignment (MDA)[4] algorithm [1, 14, 29, 30, 31, 32] that is described in Sections 4 and 5.

3.2. Frequency Peak Location Matching to Obtain Feature Vectors

The peak location vectors (e.g., columns of the matrix in (7)), estimated by the GMM algorithm, corresponding to each of the detected DoA angles are $\Phi^p_{1i_1,\ldots,Si_S} = [\phi^p_{1i_1}, \phi^p_{2i_2}, \ldots, \phi^p_{Si_S}]$ where each peak location vector $\phi^p_{si_s}$ corresponding to DoA angle θ_{si_s}, $s = 1,\ldots,S$, is of the form

$$\phi^p_{si_s} = [\oslash \quad \mu^1_{si_s} \quad \mu^2_{si_s} \cdots \mu^n_{si_s}]' \tag{8}$$

where $\mu^l_{si_s}$ is the location of the lth component (peak) of the GMM (6) used to fit $\hat{P}(\theta_{si_s})$ and $\oslash$ indicates the dummy element[5] that signifies the missed detection of a peak.

The MDA algorithm is solved, for the purpose of matching of peak locations across lists, to obtain the following S-tuple of ***feature vectors*** corresponding to the S-tuple of detected DoA angle estimates $\theta_{1i_1}, \theta_{2i_2}, \ldots, \theta_{Si_S}$

$$\Phi_{1i_1,\ldots,Si_S} = [\phi_{1i_1}, \phi_{2i_2}, \ldots, \phi_{Si_S}] \tag{9}$$

where

$$\phi_{si_s} = [\oslash \quad \mu^{j_{s1}}_{si_s} \quad \mu^{j_{s2}}_{si_s} \cdots \mu^{j_{sn_m}}_{si_s}]' \tag{10}$$

where $\mu^{j_{sq}}_{si_s} \in \{\oslash, \mu^l_{si_s}\}$ and $q = 1,\ldots,n_m$. Each element $\mu^l_{si_s}$ from (8) appears exactly once in ϕ_{si_s} while $\oslash$ appears $n_m - n$ times. It has to be noted that n_m could vary with each S-tuple of DoA angle estimates being considered, while n remains the same as it is a GMM fitting design parameter and is fixed. This leads to "matched"

[4]The assignment problem is called generalized assignment if dummies are used.

[5]Indexed by zero, and shown as the top row.

feature vectors which is shown using the illustrative example in (7) and (11). The example presented corresponds to the same triplet ($S = 3$) of DoA angle estimates $\theta_1, \theta_2, \theta_3$, which gives rise to the peaks illustrated in Fig. 4. The triplet of ***peak location vectors*** of length $n = 4$ are as shown in (7). The triplet of feature vectors, each of length $n_m = 7$, obtained after performing the matching of peak locations as described above is

$$L^{\text{matched}} = \begin{pmatrix} 40 & 40 & 40 \\ 15 & 15 & 14 \\ 36 & \oslash & 37 \\ 49 & \oslash & \oslash \\ \oslash & 78 & \oslash \\ \oslash & 117 & \oslash \\ \oslash & \oslash & 18 \end{pmatrix}. \tag{11}$$

4. STATIC MDA PROBLEM: GENERATION OF FULL POSITION ESTIMATES

One of the most important issues in multisensor-multitarget tracking is data association [1]. Static data association, i.e., measurement to measurement association at each time scan is especially important in the case of passive sensor networks, where each sensor array obtains only DoA angle measurements at each scan. At each scan, the DoA angle measurements from the same target, from at least three sensor arrays (in order to reduce ghosting, see [1, Sec. 8.8.2]), have to be associated in order to obtain the full position estimate of its location. Recently, a class of algorithms called multidimensional assignment (MDA) algorithms have been developed to solve the data association problem using an assignment approach [6, 14, 29, 30, 31, 32]. The present paper uses the MDA approach to solving the data association problem. This approach, designated as Multisensor Information Configuration III in [1], requires first a **static association** of DoA angle measurements to DoA angle measurements across sensor arrays at each time k, resulting in full position estimates called "composite measurements" (CM). If the DoA angle measurements are augmented by features, the result will be feature-augmented composite measurements. This is to be followed by **dynamic association** of the CMs to tracks and filtering which then yields target tracks.

4.1. Conventional-Cost Based MDA

For the static problem considered, at any given time k, we are given S scans (lists) of measurements from S passive sensors gathering data in a surveillance region. Each list contains a certain number of detections, not necessarily equal to the number of targets. The objective is to obtain the full position estimates of an unknown number of targets using the S lists of DoA angle measurements. The sensors obtain measurements at discrete time samples,[6] $k = 1, \ldots, K$, with a time period of T s.

Each list consists of DoA angle measurements θ_{si_s}, where $i_s = 1, 2, \ldots, n_s$. Each measurement either originated from a true target t or from some spurious source of clutter ($t = 0$). If the measurement θ_{si_s} is originated from target t, it is modeled as

$$z_{si_s} = h(\mathbf{x}_t, \mathbf{x}_s) + w_{si_s} \tag{12}$$

where $\mathbf{x}_t$ is the true target position, $\mathbf{x}_s$ is the (fixed) sensor array position, h is the measurement function, $w_{si_s} \sim \mathcal{N}(0, \sigma^2_{si_s})$ is the measurement noise and σ_{si_s} represents the sensor error. If the measurement θ_{si_s} is originated from clutter, it is modeled as uniformly distributed within the field of view of the sensor array s, i.e.,

$$p(\theta_{si_s} \mid t = 0) = \frac{1}{V_s} \tag{13}$$

where V_s is the volume of the field of view of the sensor s. In addition, the probability of detection of a target is P_D.

The goal is to localize the targets by estimating their positions at time k. A generalized likelihood ratio[7] which uses estimated target positions instead of true target positions (which are unavailable) for candidate associations, is used to attach costs to each feasible S-tuple of measurements (or candidate associations) [1]. To account for missed detections that give rise to incomplete S-tuples a *dummy* measurement is added to each list which simplifies the notation for these incomplete measurement-to-target associations caused by missed detections. The MDA algorithm is then used to *globally minimize* the cost in order to obtain feasible associations.

The likelihood that an S-tuple of measurements Z_r (where r stands for the S-tuple index), originated from target t, with position $\mathbf{x}_t$ at some instant k is

$$\Lambda(t) = p(Z_r \mid t) = \prod_{s=1}^{S} [1 - P_{D_s}]^{1-\delta_{i_s}} [P_{D_s} p(\theta_{si_s} \mid \mathbf{x}_t)]^{\delta_{i_s}}, \quad i_s \in \{0, 1, \ldots, n_s\} \tag{14}$$

where δ_{i_s} is the measurement detection indicator function. The likelihood that the measurements Z_r are all spurious or *unrelated* to target t, i.e., ($t = 0$) is

$$\Lambda(t = 0) = p(Z_r \mid t = 0) = \prod_{s=1}^{S} \left[\frac{1}{V_s}\right]^{\delta_{i_s}}. \tag{15}$$

The cost of associating the S-tuple Z_r to target t is given by the negative log-likelihood ratio (NLLR), where the likelihoods that make up the numerator and the denominator are given in (14) and (15), respectively. However, since $\mathbf{x}_t$ in (14) is unknown, it will be replaced

[6] All the sensors are assumed synchronized with respect to time.

[7] An extensive treatment on the choice of a likelihood ratio as a cost function can be found in [5].

by its maximum likelihood estimate (MLE)

$$\hat{\mathbf{x}}_t = \arg\max_{\mathbf{x}_t} p(Z_r \mid t) \tag{16}$$

and the cost is

$$c_r = -\ln \frac{\hat{\Lambda}(t)}{\Lambda(t=0)} \tag{17}$$

where

$$\hat{\Lambda}(t) = \prod_{s=1}^{S} [1 - P_{D_s}]^{1-\delta_{i_s}} [P_{D_s} p(\theta_{si_s} \mid \hat{\mathbf{x}}_t)]^{\delta_{i_s}}. \tag{18}$$

Using (12), (15) and (18) in (17), the cost of the candidate association of the S-tuple of measurements Z_r to a target t is

$$c_r = -\sum_{s=1}^{S} [1 - \delta_{i_s}] \ln(1 - P_{D_s}) + \delta_{i_s} \ln[P_{D_s} V_s p(\theta_{si_s} \mid \hat{\mathbf{x}}_t)]. \tag{19}$$

4.2. Feature-Aided MDA

In the feature-aided MDA problem, at time k, we are given S lists of feature-augmented measurement vectors instead of S lists consisting only of DoA angle measurements. An S-tuple of DoA angle measurements Z_r can be augmented with an S-tuple of feature vectors $\Phi_r = \{\phi_{1i_1}, \ldots, \phi_{Si_S}\}$, where ϕ_{si_s} denotes the feature vector corresponding to a DoA angle measurement θ_{si_s}, to form an S-tuple of feature-augmented measurement vectors (denoted by boldface)

$$\mathbf{Z}_r = [Z_r, \Phi_r]'. \tag{20}$$

The generalized likelihood that the feature-augmented S-tuple $\mathbf{Z}_r$ consists of feature-augmented measurement vectors from a target t is[8]

$$\hat{\mathbf{\Lambda}}(t) = p(\mathbf{Z}_r \mid t) = p([Z_r, \Phi_r] \mid (\hat{\mathbf{x}}_t, \hat{\varphi}_t)) \tag{21}$$

where $\hat{\mathbf{x}}_t$ is the MLE of the target position given in (16) and $\hat{\varphi}_t$ is the estimated feature vector of target t. Assuming that the DoA angle measurement errors and the feature vectors are independent, we have

$$\hat{\mathbf{\Lambda}}(t) = p(Z_r \mid \hat{\mathbf{x}}_t) p(\Phi_r \mid \hat{\varphi}_t) = \hat{\Lambda}_K(t) \hat{\Lambda}_\phi(t) \tag{22}$$

where $\hat{\Lambda}_K(t)$ and $\hat{\Lambda}_\phi(t)$ represent the kinematic and feature generalized likelihood functions, respectively. Note that $\hat{\Lambda}_K(t)$ is the same as the generalized likelihood function given in (18).

The generalized likelihood that an S-tuple of feature-vector measurements, Φ_r, corresponding to an S-tuple of angle measurements Z_r are from a target t with feature estimate $\hat{\varphi}_t$ is given by

$$\hat{\Lambda}_\phi(t) = p(\phi_{1i_1}, \phi_{2i_2}, \ldots, \phi_{Si_S} \mid \hat{\varphi}_t) \tag{23}$$

where ϕ_{si_s} is the feature vector corresponding to the DoA angle measurement θ_{si_s} from the sth list

$$\phi_{si_s} = [\mu^1_{si_s}, \mu^2_{si_s}, \ldots, \mu^{n_r}_{si_s}]' \tag{24}$$

where $\mu^l_{si_s}$ ($l = 1, \ldots, n_r$) represents either a feature (detected matched peak location) or a dummy (missed detection of the peak in list s), as described in Section 3.2.[9]

Assuming independence between the measurement errors across lists, (23) can be simplified

$$\hat{\Lambda}_\phi(t) = \prod_{s=1}^{S} \{p(\phi_{si_s} \mid \hat{\varphi}_t)\}^{\delta_{i_s}}. \tag{25}$$

Substituting (24) in (25), and assuming that the components ϕ_{si_s} of the feature vector are uncorrelated, we have

$$\hat{\Lambda}_\phi(t) = \prod_{s=1}^{S} \left\{ \prod_{l=1}^{n_r} [1 - P_{D^l_s}]^{1-\delta_{s_l}} [P_{D^l_s} p(\mu^l_{si_s} \mid \hat{\varphi}_t)]^{\delta_{s_l}} \right\}^{\delta_{i_s}} \tag{26}$$

where $P_{D^l_s}$ is the (nonunity) probability of detection of the features in list s, while δ_{s_l} is the detection indicator function for the feature $\mu^l_{si_s}$. The feature $\mu^l_{si_s}$ is assumed to be distributed as follows

$$p(\mu^l_{si_s} \mid \hat{\varphi}_t) = \mathcal{N}(\mu^l_{si_s}; \hat{\mu}^l(t), (\sigma^l_{si_s})^2), \qquad i_s \in \{1, \ldots, n_s\} \tag{27}$$

where

$$\hat{\mu}^l(t) = \frac{\sum_{s=1}^{S} \mu^l_{si_s} \delta_{s_l}}{\sum_{s=1}^{S} \delta_{s_l}}. \tag{28}$$

The standard deviation $\sigma^l_{si_s}$ of the feature $\mu^l_{si_s}$ is obtained from the GMM fitting in (6).

The likelihood that the S-tuple $\mathbf{Z}_r$ consists of feature-augmented measurement vectors which are all spurious or are unrelated to a target t is

$$\begin{aligned} \mathbf{\Lambda}(t=0) &= p(\mathbf{Z}_r \mid t=0) = \prod_{s=1}^{S} \{p(\theta_{si_s} \mid t=0) p(\phi_{si_s} \mid \varphi(t)=0)\}^{\delta_{i_s}} \\ &= \prod_{s=1}^{S} \left\{ p(\theta_{si_s} \mid t=0) \cdot \left[\prod_{l=1}^{n_r} p(\mu^l_{si_s} \mid \varphi(t)=0) \right]^{\delta_{s_l}} \right\}^{\delta_{i_s}}. \end{aligned} \tag{29}$$

Assuming that the angle and feature clutter measurements are uniformly distributed, we have

$$\mathbf{\Lambda}(t=0) = \prod_{s=1}^{S} \left\{ \frac{1}{V_s} \left[\prod_{l=1}^{n_r} \frac{1}{V^f_s} \right]^{\delta^l_s} \right\}^{\delta_{i_s}} \tag{30}$$

where V_s and V^f_s are the surveillance volumes in angle and frequency, respectively, of array s.

[8]The likelihood based on the augmented (kinematic and feature) measurements is denoted by boldface ($\mathbf{\Lambda}$).

[9]The length of the feature vector n_r varies for each candidate S-tuple of feature-augmented measurement vectors, as explained in Section 3.2. Additional notation is omitted for simplicity.

The cost of assigning the S-tuple $\mathbf{Z}_r$ is given by the NLLR obtained from the likelihood functions (22) and (30), and substitution from (18) and (26)

$$\mathbf{c}_r = -\ln\frac{\hat{\Lambda}(t)}{\Lambda(t=0)} = -\ln\frac{\hat{\Lambda}_K(t)\hat{\Lambda}_\phi(t)}{\Lambda(t=0)}. \tag{31}$$

This cost function can be simplified to the following form

$$\mathbf{c}_r = -\sum_{s=1}^{S}[1-\delta_{i_s}]\ln(1-P_{D_s}) + \delta_{i_s}\ln[P_{D_s}V_s p(\theta_{si_s}\mid\hat{\mathbf{x}}_t)]$$
$$+\,\delta_{i_s}\left\{\sum_{l=1}^{n_r}(1-\delta_s^l)\ln(1-P_{D_s^l}) + \delta_s^l\ln[P_{D_s^l}V_s^f p(\mu_{si_s}^l\mid\hat{\varphi}(t))]\right\}. \tag{32}$$

The most likely set of S-tuples such that each feature-augmented measurement vector in a list is assigned to either other measurement vectors, or declared false, with the constraint that each assigned S-tuple receives at most only one measurement vector from each list, is obtained by solving a global MDA optimization problem using the cost function (32). The assignment problem formulation can be found in [1, 14, 29, 30, 32]. Various Lagrangian relaxation techniques have been developed [14, 30, 31] to solve the MDA problem and the details are discussed in the following section. The solution to the MDA problem at each scan k results in a set of assigned feature-augmented measurement vectors (with at least 3 non-dummy DoA angles), and using triangulation (see Ch. 8, [1]) on the assigned S-tuples of DoA angle measurements the full position estimates (the composite measurements) are obtained.

5. TARGET TRACKING WITH DYNAMIC ASSIGNMENT

The static multidimensional assignment algorithm described in Section 4 yields composite measurements (CM) at each scan. The classical dynamic assignment problem [1, 29], i.e., the assignment of CMs from the current scan k or a window of scans up to scan $k-S+1$ to established tracks from the previous scan $k-1$, can be solved using an S-D[10] dynamic assignment framework. When the list of CMs at the current scan only is considered, the 2-D[11] dynamic assignment framework is used. Solving a 2-D problem is computationally less demanding than the case where $S>2$, as it can be solved optimally in polynomial time using a generalized auction algorithm [6, 18]. In a 2-D dynamic assignment framework, each CM at scan k is to be assigned to either an established track $T_u(k-1)$ from scan $k-1$, where $u\in\{1,\ldots,U(k-1)\}$, or is used to start a new track if assigned to $u=0$ (a dummy), based on the global solution of a 2-D generalized assignment algorithm using a likelihood ratio-based cost function. However, the information about track evolution which can be gained by using multiple scans of CMs, is lost. By using multiple scans it is also possible to modify previous assignments (excluding assignments from the scan at the tail of the sliding window, which cannot be modified in the next scan) which is not possible in 2-D dynamic assignment. Using a multidimensional algorithm (S-D) rather than a 2-D algorithm may yield better tracking results. In S-D dynamic assignment a sliding window of the latest $S-1$ scans of CMs is to be assigned to a list of established tracks. At scan k, a list of pre-existing tracks from scan $k-S+1$, is to be assigned to CMs from the window of scans with depth $S-1$, i.e., measurements from scan $k-S+2$ to the current scan k, using a likelihood ratio based cost function. It is well known that when $S\geq 3$, the MDA is an NP-hard problem, i.e., the complexity of an optimal algorithm increases exponentially with the size (number of dimensions) of the problem, as a result suboptimal algorithms with acceptable accuracy have to be solved that have polynomial complexity. In [30] a 3-D assignment problem was solved using a Lagrangian relaxation technique that successively solves a series of generalized 2-D assignment problems with the worst case complexity of $\mathcal{O}(3in^3)$ [30], where i is the number of relaxation iterations and n is the number of reports in each scan (or from each sensor in the case of static MDA discussed in Section 4). However, when $S\geq 4$ multiple sets of constraints have to be relaxed. In [31] the constraints are relaxed one set at a time corresponding to each list and the resulting $S-1$-D assignment problem is solved iteratively with a feasible solution to the original S-D problem being reconstructed subsequently. In the present paper the technique developed in [14] is used where all the S-2 lists are relaxed simultaneously. The problem then is a 2-D assignment problem and can be solved optimally in polynomial time. The Lagrangian multipliers associated with multiple constraint sets are updated simultaneously for faster convergence [14]. The computational complexity is $\mathcal{O}(i(S-1)Cn^3)$ where i is the maximum number of iterations, C is the range of the cost coefficient and n is the number of reports at each scan.

5.1. Conventional-Cost Based S-D Dynamic Assignment

The likelihood that the $S-1$-tuple of CMs $z_{m_{k-S+2}}(k-S+2),\ldots,z_{m_k}(k)$ is from track $T_u(k-S+1)$, where $m\in\{0,1,\ldots,M\}$ and $u\in\{1,\ldots,U(k-S+1)\}$ is given by

$$\Lambda_{m_{k-S+2},\ldots,m_k,u}(k) = p[z_{m_{k-S+2}}(k-S+2),\ldots,z_{m_k}(k)\mid T_u(k-S+1)]$$
$$= \prod_{i=k-S+2}^{k}(1-P_D)^{1-\delta_i}\cdot[P_D\Lambda_{m_i,u}(i)]^{\delta_i} \tag{33}$$

[10] S denotes the number of scans in dynamic assignment, it has no relation to the number of sensor arrays used in the static case.

[11] The 2-D dynamic assignment is a special case of the S-D sliding window dynamic assignment algorithm where the size of the window is 1, i.e., $S=2$.

where

$$\Lambda_{m_i,u}(i) = p[z_{m_i}(i) \mid T_u(k-S+1), \{z_{m_j}(j)\}_{j=k-S+2}^{i-1}] \tag{34}$$

is the likelihood if the CM $z_{m_i}(i)$ is associated with track $T_u(k-S+1)$ continued with measurements $\{z_{m_j}(j)\}_{j=k-S+2}^{i-1}$ [4]. The kinematic CM prediction errors are assumed to be independent across lists, and $z_{m_i}(i)$ is the m_ith CM at scan i. The probability of detection and the indicator function of measurement $z_{m_i}(i)$ are P_D and δ_i, respectively, with the latter defined as

$$\delta_i = \begin{cases} 1 & \text{if} \quad m_i > 0 \\ 0 & \text{if} \quad m_i = 0 \end{cases}. \tag{35}$$

The likelihood that the measurements forming the $S-1$-tuple $z_{m_{k-S+2}}(k-S+2),\ldots,z_{m_k}(k)$ are from none of the established tracks (i.e., they are from false alarms)[12] is

$$\begin{aligned} &\Lambda_{m_{k-S+2},\ldots,m_k,0}(k) \\ &\quad = p[z_{m_{k-S+2}}(k-S+2),\ldots,z_{m_k}(k) \mid T_0(k-S+1)] \\ &\quad = \prod_{i=k-S+2}^{k} \left[\frac{1}{V}\right]^{\delta_i} \end{aligned} \tag{36}$$

where a uniform distribution is assumed, with V being the volume of the surveillance region.

The cost of assigning an $S-1$-tuple, $z_{m_{k-S+2}}(k-S+2),\ldots,z_{m_k}(k)$, of CMs to a track $T_u(k-S+1)$ where $u \in \{0,1,\ldots,N(k-S+1)\}$ is

$$c_{m,u,S} = \begin{cases} 0 & \text{if} \quad u = 0 \\ -\ln\left[\dfrac{\Lambda_{m_{k-S+2},\ldots,m_k,u}(k)}{\Lambda_{m_{k-S+2},\ldots,m_k,0}(k)}\right] & \text{if} \quad u > 0 \end{cases} \tag{37}$$

where the two likelihood functions are defined in (33) and (36).

5.2. Feature-Aided S-D Dynamic Assignment

In the conventional-cost based S-D dynamic assignment approach, only the kinematic CMs obtained by the sensor network are used for target tracking and the information contained in the feature vectors is lost. The solution of the static MDA algorithm described in Section 4 results in $M(i)$ CMs, at scan i, that include the feature vectors corresponding to the assigned DoA angle measurements that give rise to each CM. A composite feature vector $\Omega_{m_i}(i)$, where i is the scan index $i \in \{k-S+2,\ldots,k\}$, can be obtained from the S-tuple of feature vectors corresponding to the S-tuple of assigned DoA angle measurements that give rise to $z_{m_i}(i)$. Appendix A describes the procedure to obtain a composite feature vector.

[12] All the measurements from time $k-S+2$ deemed false by the assignment are used to initialize new tracks.

In feature-aided S-D dynamic assignment, a sliding window of the latest $S-1$ scans (lists) of *feature-augmented CM vectors*, rather than just kinematic CMs, is assigned to an established track list. At scan k, one has a list of existing tracks at scan $k-S+1$. To this list one assigns the feature-augmented CM vectors from a window of scans with a window depth of $S-1$, i.e., measurement vectors from scan $k-S+2$ to current scan k, using a likelihood ratio based cost function. Each CM $z_{m_i}(i)$, at scan $i \in \{k-S+2,\ldots,k\}$, is augmented by its corresponding composite feature vector $\Omega_{m_i}(i)$ to form a feature-augmented CM vector (using bold face notation)

$$\mathbf{z}_{m_i}(i) = [z_{m_i}(i)' \quad \Omega_{m_i}(i)']'. \tag{38}$$

The likelihood that the $S-1$-tuple of feature-augmented CM vectors $\mathbf{z}_{m_{k-S+2}}(k-S+2),\ldots,\mathbf{z}_{m_k}(k)$ is from track $T_u(k-S+1)$, where $m \in \{0,1,\ldots,M\}$ and $u \in \{1,\ldots,U(k-S+1)\}$ is given by (again using bold face notations)

$$\begin{aligned} &\boldsymbol{\Lambda}_{m_{k-S+2},\ldots,m_k,u}(k) \\ &\quad = p[\mathbf{z}_{m_{k-S+2}}(k-S+2),\ldots,\mathbf{z}_{m_k}(k) \mid T_u(k-S+1)] \\ &\quad = \prod_{i=k-S+2}^{k} (1-P_D)^{1-\delta_i} \cdot [P_D \boldsymbol{\Lambda}_{m_i,u}(i)]^{\delta_i} \end{aligned} \tag{39}$$

where

$$\boldsymbol{\Lambda}_{m_i,u}(i) = p[\mathbf{z}_{m_i}(i) \mid T_u(k-S+1), \{\mathbf{z}_{m_j}(j)\}_{j=k-S+2}^{i-1}] \tag{40}$$

where P_D and δ_i are the probability of detection and the indicator function, respectively, of $z_{m_i}(i)$. Assuming independence between the kinematic CM prediction errors and their corresponding composite feature vectors, we have using (38)

$$\begin{aligned} &\boldsymbol{\Lambda}_{m_i,u}(i) \\ &\quad = p[z_{m_i}(i) \mid T_u(k-S+1), \{z_{m_j}(j)\}_{j=k-S+2}^{i-1}] p[\Omega_{m_i,u}(i)] \\ &\quad = \Lambda_{m_i,u}(i) p[\Omega_{m_i,u}(i)] \end{aligned} \tag{41}$$

where $\Lambda_{m_i,u}(i)$ is the (kinematic) likelihood defined in (34) and $\Omega_{m_i,u}(i)$ is the composite feature vector at scan i matched to the features from track u whose distribution is derived in Appendix A.

The likelihood that the feature-augmented CM vectors forming the $S-1$-tuple $\mathbf{z}_{m_{k-S+2}}(k-S+2),\ldots,\mathbf{z}_{m_k}(k)$ are false alarms, assuming independence between the kinematic measurement errors and the features as before, is

$$\begin{aligned} &\boldsymbol{\Lambda}_{m_{k-S+2},\ldots,m_k,0}(k) \\ &\quad = p[\mathbf{z}_{m_{k-S+2}}(k-S+2),\ldots,\mathbf{z}_{m_k}(k) \mid T_0(k-S+1)] \\ &\quad = \prod_{i=k-S+2}^{k} \{p[z_{m_i}(i) \mid T_0(k-S+1)] p[\Omega_{m_i}(i) \mid T_0(k-S+1)]\}^{\delta_i}. \end{aligned} \tag{42}$$

Assuming uniform distributions like before, we have

$$\boldsymbol{\Lambda}_{m_{k-S+2},\ldots,m_k,0}(k) = \prod_{i=k-S+2}^{k} \left\{ \frac{1}{V} p[\Omega_{m_i,0}(i)] \right\}^{\delta_i} \tag{43}$$

where V is the volume of the surveillance region, $\Omega_{m_i,0}(i)$ represents the matched feature vector at scan i, if its elements are from clutter, and its distribution is given in Appendix A.

The cost of assigning an $S-1$-tuple of feature-augmented CM vectors $\mathbf{z}_{m_{k-S+2}}(k-S+2),\ldots,\mathbf{z}_{m_k}(k)$ to a track $T_u(k-S+1)$ where $u \in \{0,1,\ldots,N(k-S+1)\}$ is

$$\mathbf{c}_{m,u,S} = \begin{cases} 0 & \text{if} \quad u = 0 \\ -\ln\left[\dfrac{\boldsymbol{\Lambda}_{m_{k-S+2},\ldots,m_k,u}(k)}{\boldsymbol{\Lambda}_{m_{k-S+2},\ldots,m_k,0}(k)}\right] & \text{if} \quad u > 0 \end{cases} \tag{44}$$

where the two likelihood functions are defined in (39) and (43).

5.3. Track Initiation

A one-point initialization technique [40] is used to initialize tracks at the beginning of the scenario (i.e., $k = 0$),[13] where $M(0)$ tracks are initialized at scan 0 from $\{z_m(0)\}_{m=1}^{M(0)}$ measurements. The initial state estimate of each track is of the form $\hat{x}_u(0 \mid 0) = [\xi_m \ \ \eta_m \ \ \dot{\xi}_m \ \ \dot{\eta}_m]$ where u is the track index. The velocity state components $\dot{\xi}_m$ and $\dot{\eta}_m$ are initialized at 0 m/s, as the motion direction of the target is assumed unknown. The initial state estimate covariance of each track is of the form

$$P_u(0 \mid 0) = \begin{bmatrix} R_m & \mathbf{0}_{2\times 2} \\ \mathbf{0}_{2\times 2} & \left(\dfrac{v_{\max}}{2}\right)^2 \mathbf{I}_{2\times 2} \end{bmatrix} \tag{45}$$

where $v_{\max}$ is the assumed maximum target speed and R_m is the measurement covariance matrix of the 2-dimensional Cartesian CM z_m. As it is not readily available, it is approximated by the Cramer-Rao lower bound (CRLB) [2, 29, 32] corresponding to the likelihood function of the S-tuple of DoA angle measurements that yielded it, and $\mathbf{0}_{n\times m}$ and $\mathbf{I}_{n\times m}$ are the all-zero matrix and the identity matrix of dimension $n \times m$, respectively. The maximum target speed $v_{\max}$ is assumed to be 9 m/s in the present paper.

An interacting multiple model (IMM) estimator with two second-order linear kinematic models (white noise acceleration, WNA) with two process noise levels is used (see [2], Ch. 11). The one with the lower noise level (with standard deviation 0.25 m/s^2) is used to model the uniform motion and the other one with standard deviation 5 m/s^2 for the maneuvers. The mode transition probability matrix

$$\pi_{ij} = \begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix} \tag{46}$$

is used. A 5-sigma ($\gamma = 25$) validation region (see [1], Ch. 3) is used for gating, and in both 2-D and S-D dynamic assignment, only measurements that fall within the validation regions of tracks are candidates for assignment. A probability of detection of 0.85 is assumed for the CMs and 0.8 for the features used in feature-aided dynamic assignment. The volume of the entire surveillance region of the sensor network is assumed to be 100,000 m^2.

5.4. Track Confirmation and Deletion

After solving the S-D dynamic assignment problem (Sections 5.1 and 5.2) at each scan, each assigned measurement from scan $k-S+2$, i.e., the scan at the tail of the sliding window, is used to update the existing tracks, and the unassigned measurements are allowed to form new candidate tracks. The unassigned measurements from scans $k-S+3$ to k are retained as is in the sliding window for the S-D assignment at the next scan. Hence, they are given again a chance to be assigned to tracks. Candidate tracks are formed solely for track initiation and they can be thought of as "second class" members of the track list while the established tracks can be thought of as "first class" members. A candidate track is either confirmed as an updated track or rejected, after a maximum of 5 scans, using a Markov chain cascaded logic (2/2&2/3) (see [1, Sec. 2.6.3]). A confirmed track is deleted if it is not assigned any measurements in 3 consecutive scans. Fig. 5 is a flowchart describing the S-D dynamic assignment based tracker.

6. REAL DATA SCENARIO

The algorithm developed above was exercised on real data obtained from a field experiment conducted by the U.S. Army Research Laboratory [12] at Aberdeen Proving Ground, Maryland. A passive acoustic sensor network was placed within a path that was traveled by two targets, a heavy vehicle and a light vehicle. Fig. 6 shows the actual trajectories based on the instrumentation that was installed in the two vehicles. However, due to some alignment errors these could not be used as ground truth for our evaluation. They will, however, be used to compare the life of the tracks with the duration of these trajectories.

All the acoustic sensor arrays used in the experiment were circular arrays of microphones. The sampling rate for each array was set at 1024 Hz, with a cutoff frequency of 312 Hz and a gain of 100 selected for all microphones in the array. The acoustic sensor arrays that make up the passive sensor network were located at known fixed positions $\mathbf{x}_s$, $s = 1,\ldots,S$, where $S = 4$ is

[13] This is preferable to two-point initialization when, due to the small sampling interval (1 s) the variance of the two-point velocity estimate [1] is significantly larger than the maximum speed.

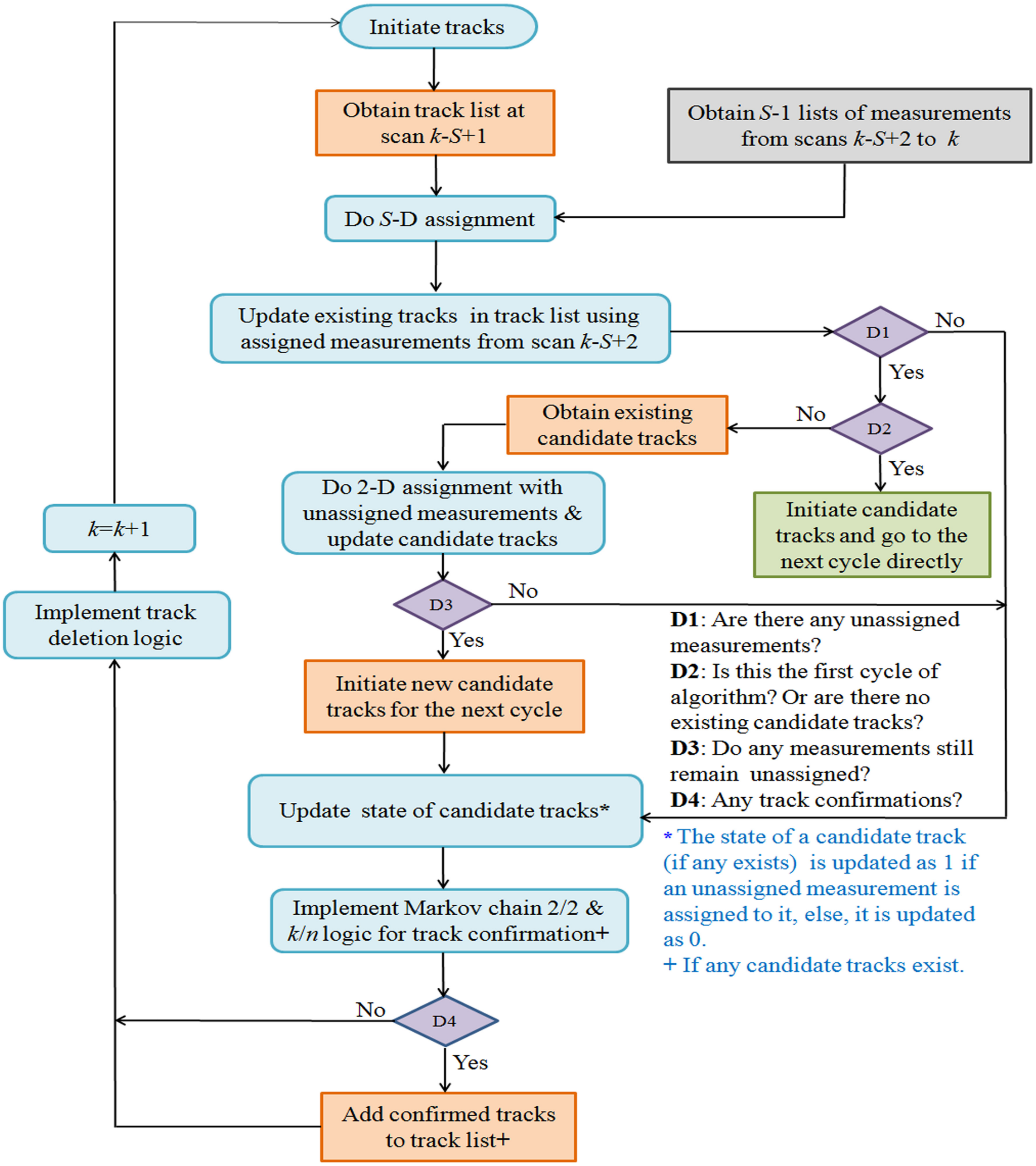


Fig. 5. Flowchart describing the dynamic S-D assignment-based tracker.

the number of sensor arrays. Each sensor array has a field of view of 360° and is made up of 7 equi-distant microphones as shown in Fig. 1. DoA angle measurements are estimated at each sampling time (1 s), from the PSD of the acoustic signals received by each sensor array, with an angular measurement error standard deviation of 1°, as described in Section 2. A low-pass spectrum of 1–120 Hz, divided into 120 bins of 1 Hz each, is used by the MVDR algorithm to estimate the PSD. The probability of detection of a DoA angle is assumed to be 0.8, the probability of detection of the features is assumed to be 0.8 and the surveillance region volume in frequency is assumed to be $V_s^f = 120$ Hz (the width of the frequency spectrum).

7. RESULTS

Fig. 7 shows the tracking results obtained using the 2-D dynamic assignment algorithm for a 2 target scenario: one target (a heavy vehicle) traveling counter-clockwise on an oval gravel path and the other target (a light vehicle) on a southeast-northwest asphalt path. The scenario starts at time 50 s and ends at time 185 s.[14]

Fig. 7(a) shows the tracks obtained when the 2-D dynamic assignment tracker uses the CMs obtained from the conventional-cost based static (without fea-

[14]This is the time interval when there is a significant ambiguity due to the interference between the two vehicles—the second target appears shortly after time 50 s.

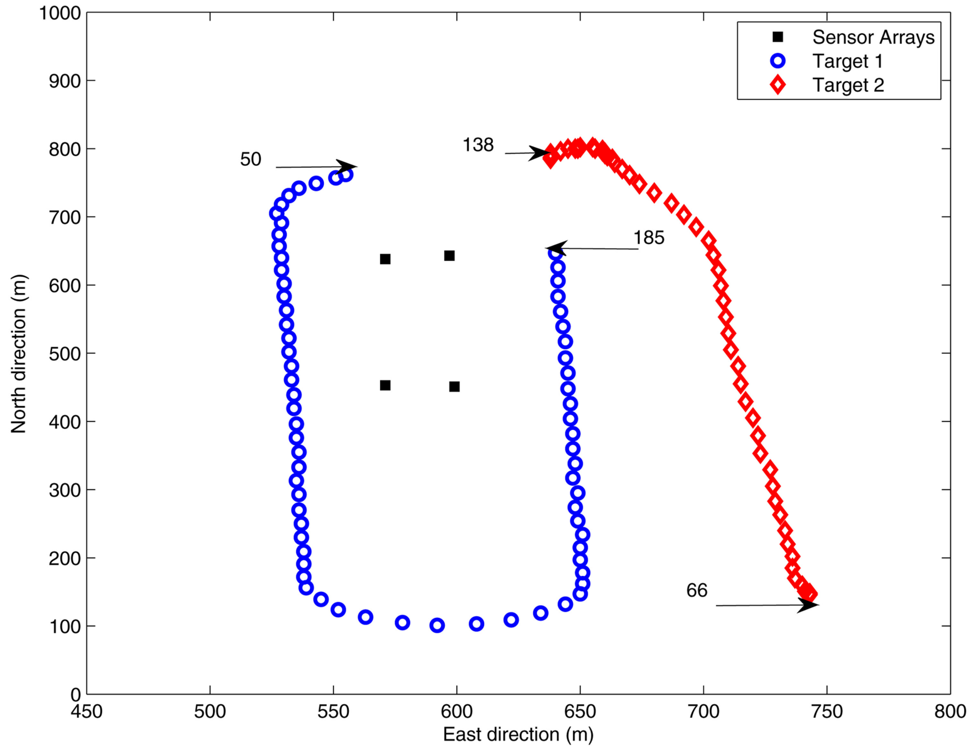


Fig. 6. True trajectories along with the start and end times (in seconds) for each target.

tures) MDA algorithm described in Section 4.1. One can see from comparing Fig. 7(a) with Fig. 6 that there are two false tracks at the bottom of the oval path when target 1 is turning. Track 2 is confirmed at scan 96 and continues towards the north, incorrectly, between scans 120 and 126 after which it gets back on track (compare with Fig. 6). In the case where the tracker uses CMs obtained from the feature-aided static MDA algorithm (Section 4.2), shown in Fig. 7(b), the tracker is able to follow the maneuver of target 1 at the bottom of the oval path better than the one in Fig. 7(a). Track 2 is confirmed at scan 89 (i.e., 7 s earlier) and is not lost until scan 117, after which it is reacquired at scan 125. It can also be seen from both Figs. 7(a) and 7(b) that track continuity is not as good at the bottom of the oval track as when the target is closer to the sensors, as the CMs obtained in that region are ill-conditioned because the DoA angle measurements from all 4 arrays are very closely spaced.

Fig. 8 shows the results obtained when the S-D assignment tracker with $S = 4$, i.e., when 2 additional lists of measurements are added. This improves over the performance when $S = 2$, as well as when $S = 3$ (the latter not shown). For $S > 4$, there is no further perceptible performance improvement. In the result shown in Fig. 8(a), when the CMs obtained from the conventional-cost based (without features) static MDA algorithm are used, target 1 is lost at scan 119 and there is a false track between scans 130–135. Target 1 is reacquired at scan 136 and continues till the end of the scenario (scan 185), along the oval path. The tracker has a false track between scans 83–89 and reacquires target 2 at scan 96 and keeps it until scan 127. In the result shown in Fig. 8(b), one can see that while track 1 is lost between scans 117–130, there are no false tracks like in Fig. 8(a). Track 2 is confirmed at scan 86 (10 s sooner than without features) and is not lost till it stops moving (130 s). The 4-D dynamic assignment based tracker performs much better at the bottom of the oval track when features are used to obtain CMs, when compared to the case where features are not used to obtain CMs.

Fig. 9 shows the results obtained when the feature-aided dynamic 2-D and feature-aided dynamic 4-D assignment based trackers are used. In the scenario shown in Fig. 9(a), when the feature-aided dynamic 2-D tracker is used, target 1 is lost at scan 120 (instead of 117 in the dynamic assignment without features) and is reacquired at scan 128 (vs. 130). Track 2 is confirmed at scan 86 and is not lost till target 2 stops moving. In the results shown in Fig. 9(b), target 1 is lost between scans 120–128. Track 2 is confirmed at scan 86 and is not lost till target 2 stops moving. From Fig. 9(a) it can be seen that when the 2-D feature-aided tracker is used, target 1 is lost between scans 65–69; this loss is eliminated when the 4-D feature aided tracker is used, as shown in Fig. 9(b). While there is a clear benefit from using features in the static association, the results in the dynamic part indicate that the benefits are that target 1 is lost at the bottom of the oval path for a shorter duration when compared to the dynamic assignment scenarios where features are not used.

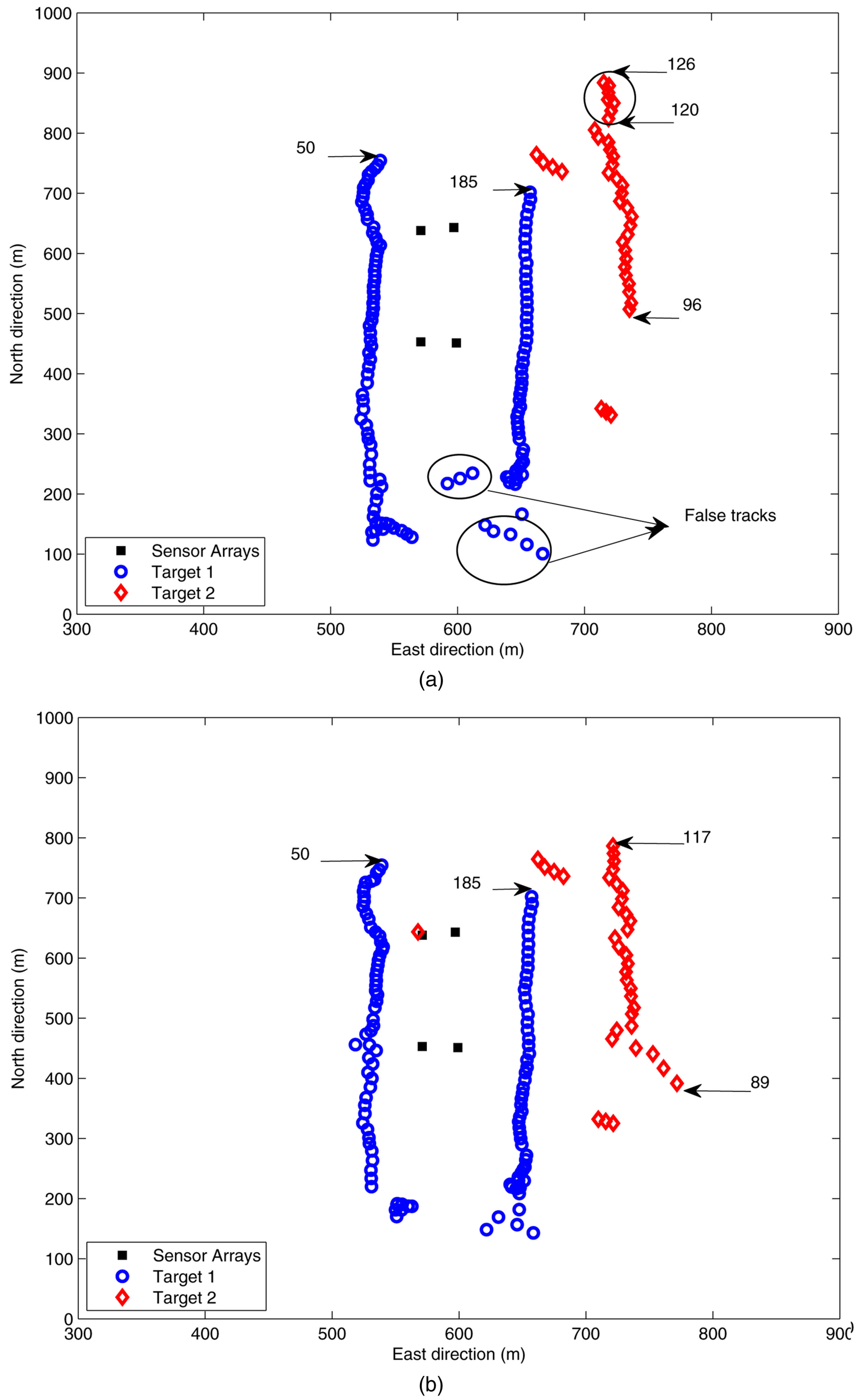


Fig. 7. Target tracking using conventional-cost based 2-D dynamic assignment: (a) using CMs obtained from ***conventional-cost based*** static MDA, and (b) using CMs obtained from ***feature-aided*** static MDA.

8. CONCLUSIONS

In this paper a static data association algorithm to obtain full position estimates (composite measurements —CMs) of multiple ground targets using real acoustic signal data, obtained by a passive sensor network, has been presented using a novel feature-aided static MDA (multidimensional assignment) framework. While the methodology developed is general, it is illustrated on real data collected by a sensor network comprised of four sensor arrays which listen to two vehicles, a heavy vehicle and a lighter vehicle. The CMs are assigned to tracks using dynamic 2-D and S-D (with $S = 4$) assignment.

A novel detection scheme has been presented to detect DoAs from real acoustic data and a new feature extraction technique based on GMM (Gaussian mixture model) based fitting and multidimensional matching has been presented to extract feature vectors which aug-

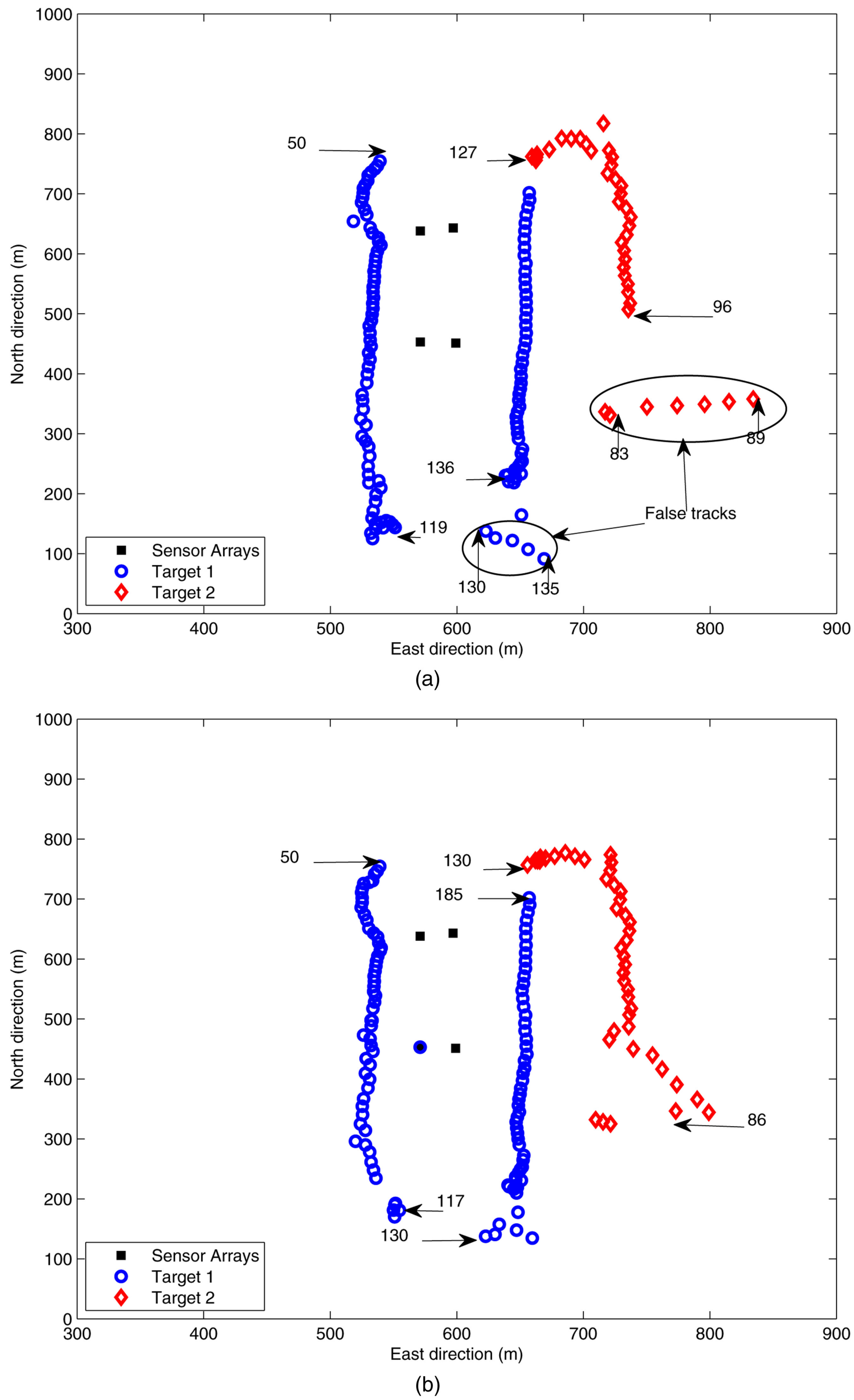


Fig. 8. Target tracking using conventional-cost based 4-D dynamic assignment: (a) using CMs obtained from ***conventional-cost based*** static MDA, and (b) using CMs obtained from ***feature-aided*** static MDA.

ment the corresponding DoA angle measurements. An MDA algorithm is solved at each scan, using feature-augmented likelihood ratio based cost functions, to obtain composite measurements that are the full position estimates of targets. Composite measurements are assigned to tracks using both conventional cost-based as well as feature-aided S-D dynamic assignment.

It is observed from running the dynamic 2-D and 4-D assignment algorithms on real data that the tracking performance is significantly better when features are used to generate the CMs. One track is initiated earlier, and the breakages in the other are reduced. It is also observed that the tracks using the feature-aided tracker have shorter breakages. Its performance could be improved if the shifts in frequency due to the motion of vehicles are accounted for. This is a topic for further research.

Modeling of the variation in time of the features and track segment association can be further future topics of investigation.

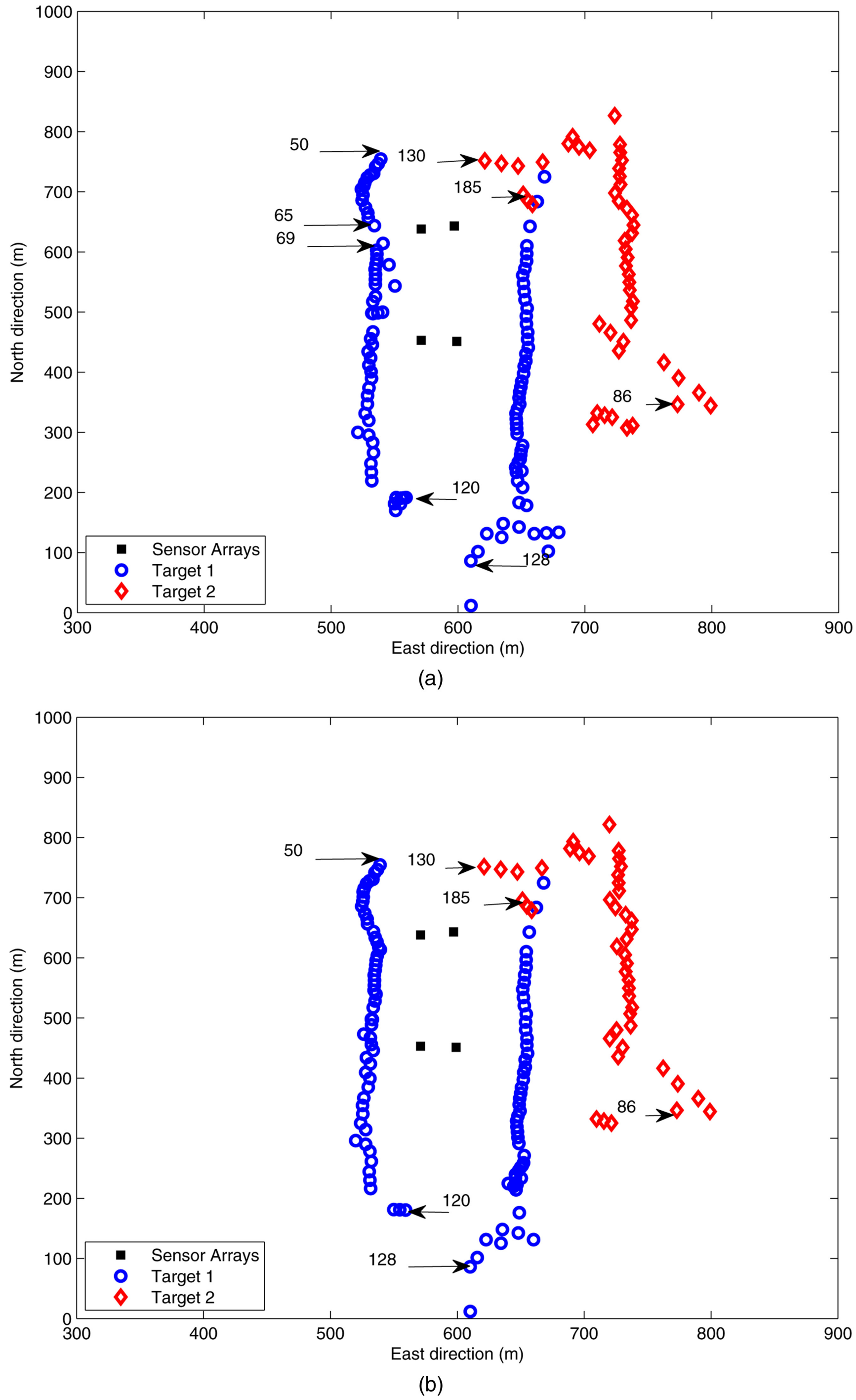


Fig. 9. Target tracking using feature-augmented cost based 2-D and 4-D dynamic assignment: (a) feature-aided 2-D dynamic assignment, and (b) feature-aided 4-D dynamic assignment.

9. ACKNOWLEDGMENT

This material is based upon work supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office under Grant 51018-CI. The authors thank Dr. Javier Areta for his help in implementing the Lagrangian relaxation based MDA algorithm. The authors are also grateful to the Editor and the reviewers for their comments and suggestions to improve the paper.

APPENDIX A. THE COMPOSITE FEATURE VECTOR

The composite feature vector $\Omega_{m_i}(i)$ corresponding to the CM $z_{m_i}(i)$ is given by

$$\Omega_{m_i}(i) = [\omega^1_{m_i}(i), \omega^2_{m_i}(i), \ldots, \omega^{n_{m_i}}_{m_i}(i)]'. \tag{47}$$

Each element $\omega^l_{m_i}(i)$ in (47) is the weighted average of the lth elements of the feature vectors corresponding to the assigned S-tuple of DoA angle measurements that give rise to $z_{m_i}(i)$ and is obtained using (28).

The composite feature vector corresponding to track T_u is the composite feature vector corresponding to the CM that has been assigned and used to update the track u and is given by

$$\Omega_u = [\omega_u^1, \omega_u^2, \ldots, \omega_u^{n_u}]'. \qquad (48)$$

The composite feature vector in (47) is matched to the composite feature vector in (48) using the same matching technique described in Section 3.2. The matched composite feature vector at scan i, i.e., $\Omega_{m_i,u}(i)$, is given by

$$\Omega_{m_i,u}(i) = [\omega_{m_i}^{j_1}(i), \omega_{m_i}^{j_2}(i), \ldots, \omega_{m_i}^{j_{n_{m_i,u}}}(i)]' \qquad (49)$$

where $\{\omega_{m_i}^{j_q}\}_{q=1}^{n_{m_i,u}} \in \{\oslash, \omega_{m_i}^l(k)\}$; the resulting dummy elements (missed detections) after matching are represented by $\oslash$ and $\omega_{m_i}^l(i)$ is an element of the composite feature vector in (47), with $l \in \{1,\ldots,n_{m_i}\}$ being the index of its elements.[15] Similarly, the matched composite feature vector of track T_u is given by

$$\Omega_u^\dagger(k-1) = [\omega_u^{j_1}, \omega_u^{j_2}, \ldots, \omega_u^{j_{n_{m_i,u}}}]' \qquad (50)$$

where $\{\omega_u^{j_q}\}_{q=1}^{n_{m_i,u}} \in \{\oslash, \omega_u^l\}$, with $l \in \{1,\ldots,n_u\}$ being the index of the elements of the composite feature vector (48).

The likelihood that the matched composite feature vector $\Omega_{m_i,u}(i)$ at scan k is from track T_u is given by

$$p[\Omega_{m_i,u}(i)] = \prod_{q=1}^{n_{m_i,u}} (1 - P_{D_q})^{1-\delta(m_{iq})} \cdot [P_{D_q} p(\omega_{m_i}^{j_q}(i))]^{\delta(m_{iq})} \qquad (51)$$

assuming that the individual matched composite feature vector elements are uncorrelated and where P_{D_q} and $\delta(m_{iq})$ represent the probability of detection and the indicator function for the elements of the matched composite feature vector (49). The elements of the matched composite feature vector are assumed to be distributed as

$$p(\omega_{m_i}^{j_q}(i)) = \mathcal{N}(\omega_{m_i}^{j_q}(i); \hat{\omega}^{j_q}(i), \sigma_{j_q}^2) \qquad (52)$$

where

$$\hat{\omega}^{j_q}(i) = \frac{\omega_{m_i}^{j_q}(i)\delta(m_{iq}) + \omega_u^{j_q}(i)\delta(u_q)}{\delta(m_{iq}) + \delta(u_q)} \qquad (53)$$

$\delta(u_q)$ is the indicator function of the elements of the matched composite feature vector (48), and σ_{j_q} is the standard deviation of the distribution of $\omega_{m_i}^{j_q}(k)$ and is assumed to be the minimum of all the standard deviations $\sigma_{si_s}^l$ (see Section 4.2) of the features that give rise to the composite feature vector.

The likelihood that the matched composite feature vector $\Omega_{m_i,u}(i)$ at scan i is from a source of clutter, assuming a uniform distribution with V^f being the volume of the surveillance region of the sensor network in frequency, is given by

$$p[\Omega_{m_i,0}(i)] = \prod_{q=1}^{n_{m_i,u}} \left[\frac{1}{V^f}\right]^{\delta(m_{iq})}. \qquad (54)$$

[15] Each element of (47) appears only once in (49).

REFERENCES

[1] Y. Bar-Shalom and X. Li
Multitarget-Multisensor Tracking: Principles and Techniques.
Storrs, CT, YBS Publishing, 1995.

[2] Y. Bar-Shalom, X. Li, and T. Kirubarajan
Estimation with Applications to Tracking and Navigation.
Wiley, 2001.

[3] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe
Sonar tracking of multiple targets using joint probabilistic data association.
IEEE Journal of Oceanic Engineering, **8**, 3 (July 1983), 173–184.

[4] Y. Bar-Shalom, T. Kirubarajan, and C. Gokberk
Tracking with classification-aided multiframe data association.
IEEE Transactions on Aerospace and Electronic Systems, **41**, 3 (July 2005), 868–879.

[5] Y. Bar-Shalom, S. S. Blackman, and R. J. Fitzgerald
Dimensionless score function for multiple hypothesis tracking.
IEEE Transactions on Aerospace and Electronic Systems, **43**, 1 (Jan. 2007), 392–400.

[6] D. Bertsekas
The auction algorithm: A distributed relaxation method for the assignment problem.
Annals of Operations Research: Special Issue on Parallel Optimization, **14** (1988), 105–123.

[7] B. Bhanu
Automatic target recognition: State of the art survey.
IEEE Transactions on Aerospace and Electronic Systems, **22**, 4 (July 1986), 364–379.

[8] E. Blasch and L. Hong
Simultaneous feature-based identification and track fusion.
In *Proceedings of IEEE International Conference on Decision and Control*, Tampa, FL, Dec. 1998, 239–244.

[9] J. C. Chen, K. Yao, and R. E. Hudson
Source localization and beamforming.
IEEE Signal Processing Magazine, (Mar. 2002), 30–39.

[10] H-W. Chen and J-W. Zhao
Wideband MVDR beamforming for acoustic vector sensor linear array.
In *IEE Proceedings of Radar, Sonar and Navigation*, **151**, 3 (June 2004), 158–162.

[11] T. R. Damarla
Tracking a convoy of multiple targets using acoustic sensor data.
In M. K. Masten and L. A. Stockum (Eds.), *Proceedings of SPIE Acquisition, Tracking and Pointing XVII*, vol. 5082, Orlando, FL, Apr. 2003, 37–42.

[12] T. R. Damarla, T. Pham, and D. Lake
An algorithm for classifying multiple targets using acoustic signatures.
In *Proceedings of SPIE Sig. Proc., Sensor Fusion and Targ. Recognition*, 2004, 421–427.

[13] T. R. Damarla and G. Whipps
Multiple target tracking and classification improvement using data fusion at node level using acoustic signals.
In *Proceedings of SPIE Unattended Ground Sensor Technologies and Applications VII*, vol. 5796, Orlando, FL, Apr. 2005, 19–27.

[14] S. Deb, M. Yeddanapudi, K. Pattipati, and Y. Bar-Shalom
A generalized S-D assignment algorithm for multisensor-multitarget state estimation.
IEEE Transactions on Aerospace and Electronic Systems, **33**, 2 (Apr. 1997), 523–538.

[15] O. E. Drummond
Feature, attribute and classification aided target tracking.
In *Proceedings of SPIE Signal and Data Processing of Small Targets*, vol. 4473, San Diego, CA, Aug. 2001, 610–621.

[16] O. E. Drummond
On categorical feature aided tracking.
In *Proceedings of SPIE Signal and Data Processing of Small Targets*, vol. 5204, San Diego, CA, Aug. 2003, 544–558.

[17] N. M. Laird, A. P. Dempster, and D. B. Rubin
Maximum likelihood from incomplete data via the EM algorithm.
J. Roy. Soc. Statist., Series B, **39**, 1 (1977), 138.

[18] A. V. Goldberg, S. A. Plotkin, and E. Tardos
Combinatorial algorithms for the generalized auction problem.
Mathematics in Operations Research, **16** (1991), 351–381.

[19] B. Gu and L. Hong
Tracking 2-D rigid targets with invariant constraints.
Information Science, **138** (2001), 79–97.

[20] M. E. Holil, A. Rotolo, and J. Chang
A cylinder classification and target ID algorithm based on a reduced feature space representation for ground vehicles.
In *Proceedings of 2003 Meeting of the IRIS Speciality Group on Acoustic and Seismic Sensing*, Sept. 2003.

[21] L. Hong, S. Cong, M. T. Pronobis, and S. Scott
Wavelets feature aided tracking (WFAT) using GMTI/HRR data.
Signal Processing, **86** (2003), 2683–2690.

[22] L. Hong, N. Cui, M. Pronobis, and S. Scott
Local motion feature aided ground moving target tracking with GMTI and HRR measurements.
IEEE Transactions on Automatic Control, **50**, 1 (Jan. 2005).

[23] B-C. Kim and I-T. Lu
High resolution broadband beamforming based on the MVDR method.
OCEANS 2000 MTS/IEEE Conference and Exhibition, vol. 3, 2000, 1673–1676.

[24] D. Lake
Harmonic phase coupling for battlefield acoustic target identification.
In *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, vol. 4, Seattle, WA, May 1998, 2049–2052.

[25] D. Lake
Tracking fundamental frequency for synchronous mechanical diagonostic signal processing.
In *Proceedings of 9th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, Portland, OR, Sept. 1998, 200–203.

[26] W. Liu, S. Weiss, J. G. McWhirter, and I. K. Proudler
Frequency invariant beamforming for two-dimensional and three-dimensional arrays.
Signal Processing (Elsevier), **87**, 11 (Nov. 2007), 2535–2543.

[27] W. Liu and S. Ding
An efficient method to determine the diagonal loading factor using the constant modulus feature.
IEEE Transactions on Signal Processing, **56**, 12 (Dec. 2008), 6102–6106.

[28] D. H. Nguyen, J. H. Kay, B. J. Orchard, and R. H. Whiting
Classification and tracking of moving ground vehicles.
Lincoln Lab. Journal, **13**, 2 (2002), 275–308.

[29] K. Pattipati, R. Popp, and T. Kirubarajan
Survey of assignment techniques for multitarget tracking.
In Y. Bar-Shalom and W. D. Blair (Eds.), *Multitarget-Multisensor Tracking: Applications and Advances*, vol. III, Artech House, 2001.

[30] K. Pattipati, S. Deb, Y. Bar-Shalom, and R. Washburn
A new relaxation algorithm and passive sensor network data association.
IEEE Transactions on Automatic Control, **37**, 2 (Feb. 1992), 197–213.

[31] A. Poore and N. Rijavec
A Lagrangian relaxation algorithm for multidimensional assignment problems arising from multitarget tracking.
SIAM J. Optimization, **3**, 3 (Aug. 1993), 544–563.

[32] R. L. Popp, K. R. Pattipati, and Y. Bar-Shalom
m-best S-D assignment algorithm with application to multitarget tracking.
IEEE Transactions on Aerospace and Electronic Systems, **37**, 1 (Jan. 2001).

[33] Y. Ruan and L. Hong
Feature-aided tracking with GMTI and HRR measurements via mixture density estimation.
IEE Proceedings—Control Theory Appl., **153** 3, 342–356.

[34] N. Srour, D. Lake, and M. Miller
Utilizing acoustic propagation models for robust battlefield target identification.
In *Proceedings of 1998 Meeting of the IRIS Speciality Group on Acoustic and Seismic Sensing*, Sept. 1998.

[35] G. Succi and T. K. Pedersen
Acoustic target tracking and target identification—Recent results.
In *Proceedings of SPIE Conf. on Unattended Ground Sensor Technologies and Applications*, vol. 3713, Orlando, FL, Apr. 1999, 3713-10–21.

[36] H. L. Van Trees
Detection, Estimation and Modulation Theory, Part IV, Optimum Array Processing.
New York: Wiley, 2002.

[37] B. V. Veen
Minimum Variance Beamforming, Adaptive Radar Detection and Estimation.
Edited by S. Haykin and A. Steinhardt, J. Wiley & Sons, Inc., 1992.

[38] X. Wang and H. Qi
Acoustic target classification using distributed sensor arrays.
In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, Orlando, FL, May 2002, IV-4186.

[39] S. Yan and Y. Ma
High-resolution broadband beamforming and detection methods with real data.
Acoustic Science and Technology, **25**, 1 (2004), 73–76.

[40] S-W. Yeom, T. Kirubarajan, and Y. Bar-Shalom
Track segment association, finite-step IMM and initialization with Doppler for improved track performance.
IEEE Transactions on Aerospace and Electronic Systems, **40**, 1 (Jan. 2004), 293–309.

[41] R. E. Zarnich
A fresh look at broadband passive sonar processing.
In *1999 Adaptive Sensor Array Processing Workshop* (ASAP '99), MIT Lincoln Laboratory, Lexington, MA, Mar. 1999, 99–104 .

Vishal Cholapadi Ravindra (S'03) was born in Manipal, Karnataka, India on May 14, 1980. He received the B.E. degree in electronics and comminications from the University of Mysore (P.E.S.C.E. Mandya), India in 2001. He received the M.S. degree in electrical engineering from the University of Colorado at Denver, in 2004. In August 2009, he received the Ph.D. degree in electrical engineering from the University of Connecticut, Storrs.

Since December 2009, he has been a Research Fellow at the National Aerospace Laboratories (NAL) in Bangalore, India. His research interests lie in estimation theory, target tracking, data fusion, automatic target recognition, and guidance and navigation for unmanned/micro aerial vehicles.

Yaakov Bar-Shalom (S'63—M'66—SM'80—F'84) was born on May 11, 1941. He received the B.S. and M.S. degrees from the Technion, Israel Institute of Technology, in 1963 and 1967 and the Ph.D. degree from Princeton University in 1970, all in electrical engineering.

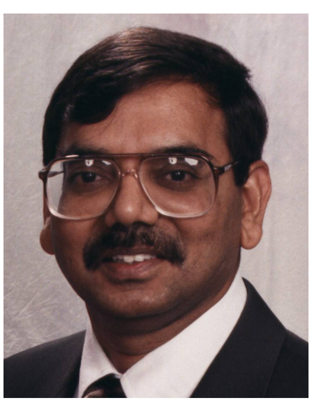
From 1970 to 1976 he was with Systems Control, Inc., Palo Alto, CA. Currently he is Board of Trustees Distinguished Professor in the Dept. of Electrical and Computer Engineering and Marianne E. Klewin Professor in Engineering at the University of Connecticut. He is also Director of the ESP (Estimation and Signal Processing) Lab.

His current research interests are in estimation theory and target tracking and has published over 370 papers and book chapters in these areas and in stochastic adaptive control. He coauthored the monograph *Tracking and Data Association* (Academic Press, 1988), the graduate texts *Estimation and Tracking: Principles, Techniques and Software* (Artech House, 1993), *Estimation with Applications to Tracking and Navigation: Algorithms and Software for Information Extraction* (Wiley, 2001), the advanced graduate text *Multitarget-Multisensor Tracking: Principles and Techniques* (YBS Publishing, 1995), and edited the books *Multitarget-Multisensor Tracking: Applications and Advances* (Artech House, Vol. I, 1990; Vol. II, 1992; Vol. III, 2000).

He has been elected Fellow of IEEE for "contributions to the theory of stochastic systems and of multitarget tracking." He has been consulting to numerous companies and government agencies, and originated the series of Multitarget-Multisensor Tracking short courses offered via UCLA Extension, at Government Laboratories, private companies and overseas.

During 1976 and 1977 he served as Associate Editor of the IEEE Transactions on Automatic Control and from 1978 to 1981 as Associate Editor of Automatica. He was Program Chairman of the 1982 American Control Conference, General Chairman of the 1985 ACC, and Co-Chairman of the 1989 IEEE International Conference on Control and Applications. During 1983–87 he served as Chairman of the Conference Activities Board of the IEEE Control Systems Society and during 1987–89 was a member of the Board of Governors of the IEEE CSS. He was a member of the Board of Directors of the International Society of Information Fusion (1999–2004) and served as General Chairman of FUSION 2000, President of ISIF in 2000 and 2002 and Vice President for Publications in 2004–08.

In 1987 he received the IEEE CSS Distinguished Member Award. Since 1995 he is a Distinguished Lecturer of the IEEE AESS and has given numerous keynote addresses at major national and international conferences. He is corecipient of the M. Barry Carlton Award for the best paper in the IEEE Transactions on Aerospace and Electronic Systems in 1995 and 2000 and the 1998 University of Connecticut AAUP Excellence Award for Research. In 2002 he received the J. Mignona Data Fusion Award from the DoD JDL Data Fusion Group. He is a member of the Connecticut Academy of Science and Engineering. He is the recipient of the 2008 IEEE Dennis J. Picard Medal for Radar Technologies and Applications.



Thyagaraju Damarla is working as an electronic engineer at the US Army Research Laboratory for the past 13 years. He received his B.S. and M.S. from Indian Institute of Technonology, Kharagpur, India and Ph.D. from Boston University.

He published more than 100 technical papers in various journals and conferences and has three US patents. His current interests include signal processing, and sensor fusion for situational awarenes.