A Framework for Multisensor Multiple Extended Target Tracking with Vessel Size Categorisation

Aman Kumar
Remote Sensing Technology Institute (IMF)
German Aerospace Center (DLR)
Oberpfaffenhofen, Germany
aman.kumar@dlr.de

Jaya Shradha Fowdur
Institute of Communications and Navigation
German Aerospace Center (DLR)
Neustrelitz, Germany
jaya.fowdur@dlr.de

***Abstract*—Traffic monitoring and surveillance form part of the basic aspects of several maritime applications. The observation regions around harbours have multiple modern radar sensors for adequate coverage, and the sensors produce measurements from potential targets that appear as point clouds. With measurements from multiple sensors, extra information about the target's shape can be gained in order to improve monitoring within the particular observation region. In this paper, we present a framework for processing radar measurements from multiple sensors which outputs the extended state estimates of the vessels over an observation region jointly with their approximate size range categories. We demonstrate the performance of the framework based on real-world radar video streams with respect to its real-time capability, and finally present our implementation on both, single and multiple targets, showing promising results accompanied by some discussion.**

***Index Terms*—data association, elliptical target tracking, radar fusion, size categorisation**

I. Introduction

The main mode of transport for global trade is dominated by ocean shipping, thereby necessitating the safety and security of onboard crew and goods [1], [2]. One of the key factors supporting this objective is having a reliable maritime situational assessment, monitoring, and surveillance. The marine radar is one of the most widely used sensors in maritime navigation, usually cooperatively with the Automatic Identification System (AIS). It provides information about the targets present within a specific region of interest, also known as the observation region, as well as the surroundings restricted to a perspective (line-of-sight). This restriction could be allayed by the use of multiple radar sensors available at vantage points to gain a more detailed information about the traffic present within the observation region.

Target tracking forms an essential part of many maritime applications to ensure a constant monitoring of vessels at sea [3]–[7]. Conventionally, it can be defined as the estimation of a target's properties of interest, such as its position and kinematic parameters. The latest radars nowadays are endowed with fine resolution causing multiple measurements to arise from the surface of vessels, thus allowing information relating to the vessel's shape to be estimated. This problem is known as Extended Target Tracking (ETT) [8]. The approach can particularly be beneficial in the maritime context when large vessels are equipped with only one or two AIS antennas, giving limited and sometimes erroneous information about the vessels, besides being affected by vulnerabilities such as jamming, spoofing, or extreme weather conditions [9]. When combined with the radar, traffic monitoring and representation could be enhanced [10]. We only focus on tracking targets from radar streams in this work. A vessel's point cloud-representation from radar measurements is often dense and noisy, making it appropriate to being modelled as an ellipse. The ellipse parameters, orientation and lengths of semi-axes, correspond approximately to the dimensions and heading of a vessel. In addition, the states and cardinality of multiple vessels can be jointly estimated using specific Multiple Target Tracking (MTT) approaches. Until now, multi-target trackers have been mostly implemented based on the data association [11]–[14], multi-hypothesis tracking [15], and random finite sets methods [16]–[18].

In this paper, we present a framework for vessel detection, tracking and size categorisation from radar video streaming. A data association approach is employed to associate a set of point clouds to a particular vessel, under the assumption that the point clouds are spatially distributed as Poisson point processes [19]. Based on this assumption and approach, the elliptical extent parameters of vessels can be simultaneously estimated together with their kinematic parameters. When multiple sensors are available, a vessel's extent information could be significantly improved especially when the sensors cover different perspectives in ports. As the first contribution, we consider such a setting to apply a multisensor version of the multi-ETT tracker over radar measurements from the port of Hamburg in Germany. The measurements are processed and converted to an appropriate Cartesian coordinate system, before being input to a target detection algorithm. An elliptical model based on the Multiplicative Error Model-Extended Kalman Filter (MEM-EKF)* [20] integrated within a Joint Probabilistic Data Association (JPDA)-tracker [19] has then been applied for the state estimation of detected potential targets, whereby a sequential updating approach has been adopted for processing measurements from each sensor as soon as they become available. The estimated extent parameters are used in the next step, as a second contribution, to categorise the vessels into specific size categories including a corresponding

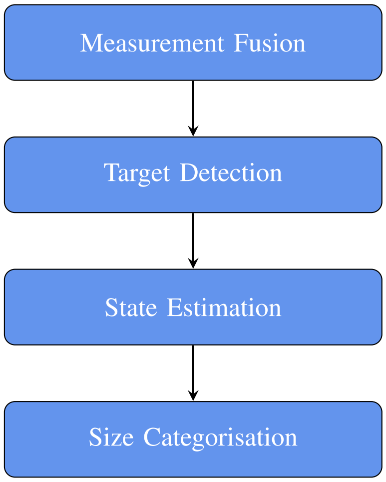


Fig. 1: Functional blocks of the framework.

confidence level based on the estimated semi-axes lengths.

The structure of the paper is as follows. In Sections II and III, the components in the framework employed are presented and explained. The results are included and discussed in the next section. The work is then summarised in Section V.

II. Measurement Processing

In this section, the first two components of the framework, shown in Figure 1, are thoroughly covered.

A. Measurement Fusion

Measurements from different radars are acquired separately based on the All Purpose Structured Eurocontrol Surveillance Information Exchange (ASTERIX) [21] protocol which is a standard protocol for data definition that supports transmission and exchange of surveillance related data. In the implementation, an ASTERIX video of category 240, corresponding to radar video transmission, is used. Since the radar-specific measurements have their individual local coordinate systems, an arbitrary point close to the centre of the observation region is chosen as the reference point, given that the geodetic positions of all the radars are known in advance. The measurements are then converted to an equivalent East North Up (ENU) Cartesian coordinate system from their original polar representation and are registered to a common frame.

B. Target Detection

To detect the potential targets (vessels) automatically, a blob detection algorithm is employed. Blob detection methods are typically used at detecting those regions in an image that have different properties, such as intensity or colour, taking into consideration their surrounding regions. A blob would ideally represent a specific region within an image so that all points within the blob share approximately similar properties.

The Determinant of Hessian (DoH) approach [22] is one of the common blob detection algorithms that detects blobs of different dimensions by computing the local maxima of the second derivatives' determinants of an image matrix. The approach has been chosen due to its fast computation and additionally, its detection speed being independent of the size of blobs. Points belonging to the detected blobs are extracted and these are the measurements that are fed to the MTT tracker.

III. MTT and Size Categorisation

This section describes the estimation and the categorisation processes, respectively.

A. State Estimation

At every observation step k, the kinematic state and the shape parameters of the target(s) are estimated simultaneously from the sensor measurements. The kinematic state $\mathbf{x}$ of a specific target i is represented as

$$\mathbf{x}_k^i = \left[t_e, t_n, \dot{t}_e, \dot{t}_n\right]^T \in \mathbb{R}^4 \tag{1}$$

where (t_e, t_n) are the ENU coordinates, $\dot{t}_e$, $\dot{t}_n$ are the corresponding velocity components and, T is the transpose operator.

The shape parameter vector $\mathbf{p}$ is represented as

$$\mathbf{p}_k^i = \left[\alpha, l_1, l_2\right]^T \in \mathbb{R}^3 \tag{2}$$

where α is the orientation and l_1 and l_2 are the major and minor axes respectively.

The temporal evolution of the kinematic state is modelled based on a nearly constant velocity model (omitting the target indices i for simplicity) as follows:

$$\mathbf{x}_{k+1|k} = \mathbf{F}_{\mathbf{x}}\mathbf{x}_{k|k} \tag{3}$$

$$\mathbf{C}^{\mathbf{x}}_{k+1|k} = \mathbf{F}_{\mathbf{x}}\mathbf{C}^{\mathbf{x}}_{k|k}\left(\mathbf{F}_{\mathbf{x}}\right)^{\mathrm{T}} + \mathbf{C}^{\mathbf{x}}_{\boldsymbol{\omega}} \tag{4}$$

Similarly, that of the shape parameter vector is given by

$$\mathbf{p}_{k+1|k} = \mathbf{F}_{\mathbf{p}}\mathbf{p}_{k|k} \tag{5}$$

$$\mathbf{C}^{\mathbf{p}}_{k+1|k} = \mathbf{F}_{\mathbf{p}}\mathbf{C}^{\mathbf{p}}_{k|k}\left(\mathbf{F}_{\mathbf{p}}\right)^{\mathrm{T}} + \mathbf{C}^{\mathbf{p}}_{\boldsymbol{\omega}} \tag{6}$$

where $\mathbf{F}_{\mathbf{x}} = \mathbf{I}_4$ and $\mathbf{F}_{\mathbf{p}} = \mathbf{I}_3$ are the respective system matrices such that $\mathbf{I}_d$ is an identity matrix of state dimension d. $\mathbf{C}^{\mathbf{x}}_{\boldsymbol{\omega}} = \mathrm{diag}\left(\sigma^2_{t_e}, \sigma^2_{t_n}, \sigma^2_{\dot{t}_e}, \sigma^2_{\dot{t}_n}\right)$ and $\mathbf{C}^{\mathbf{p}}_{\boldsymbol{\omega}} = \mathrm{diag}(\sigma^2_{\alpha}, \sigma^2_{l_1}, \sigma^2_{l_2})$ are the covariances of the kinematic state and shape parameter vector with zero-mean additive process noise.

The state updates are carried out based on the MEM-EKF* model while the JPDA filter tackles the measurements-to-vessel data association [19]. We further adopt a central-level tracking mechanism as the differently-sourced measurements are already on a common frame. Thus, the states were updated sensor-wise in a sequential manner to account for the different sensor uncertainties.

B. Track Management

The track management is implemented separately from the tracking based on a construct similar to the so-called M/N logic [23]. For a predefined N number of observation steps, the number of detections M determines the creation, confirmation and termination of a track. In the current implementation, track confirmation requires an $8/10$ constraint. Tracks are dropped when no measurements have been validated for $8/8$ constraint.

TABLE I: Size categories from estimated axis length

Category	Dimension [m]
Small (S)	< 50
Medium (M)	50-100
Large (L)	> 100

C. Size Categorisation

In the final step, the vessels are categorised on the basis of their estimated lengths while factoring in their uncertainties. Depending on the vessels present in the period of the streaming considered, three different categories were as defined in Table I.

The current major axis of target i, l_1 from $\mathbf{p}_k^i$, and its corresponding variance $\sigma_{l_1}^2$ from $\mathbf{C}_k^{\mathbf{p}}$ are taken into account in order to categorise the vessel size. The probabilities $\Theta(\mathbf{z}_i)_c$ of having target i assigned to a suitable category c are evaluated utilising a normalised exponential function [24]

$$\Theta(\mathbf{z}_i)_c = \frac{e^{z_{i,c}}}{\sum_{j=1}^{C} e^{z_{i,j}}} \tag{7}$$

where $C = 3$ represents the number of categories. The input vector $\mathbf{z}_i$ is a target-specific incremental counter that stores the number of occurrences of categorical assignments to the target throughout its existence. The categorical assignments, calculated at every k, are obtained by summing up the occurrences of the categories from a temporary vector defined as,

$$\boldsymbol{\lambda}_i := \begin{bmatrix} l_1, & l_1 + \sigma_{l_1}, & l_1 - \sigma_{l_1} \end{bmatrix}^T \tag{8}$$

For example, consider an estimated extent of 48m having a variance of 5^2m, yielding $\boldsymbol{\lambda} = [48, \quad 53, \quad 43]^T$, and being processed at the current observation step. If it is an existing target with its $\mathbf{z}_{i|k-1}$ counter at $[30, \quad 5, \quad 0]^T$ (representing S, M and L), the current counter $\mathbf{z}_{i|k}$ will then be incremented to $[32, \quad 6, \quad 0]^T$ before being input to (7). The target would be assigned to both small (S) and medium (M) size categories (see Table I), with their associated probability values $\Theta(\mathbf{z}_i)_S$ and $\Theta(\mathbf{z}_i)_M$.

This categorisation process is carried out at every observation step and for every target that has been estimated at that step. It is to be noted that the probabilities are henceforth used interchangeably as confidence levels in the remaining sections.

IV. Results

The whole framework was applied on two different ASTERIX streams covering the same observation region: the first one of a single target and the second one of two targets. The results obtained are presented and discussed in the following subsections.

A. Single Target

The stream comprises a target navigating in a Western direction. A sample image of the transformed and registered multisensor radar measurements is illustrated over a light background in Figure 2.

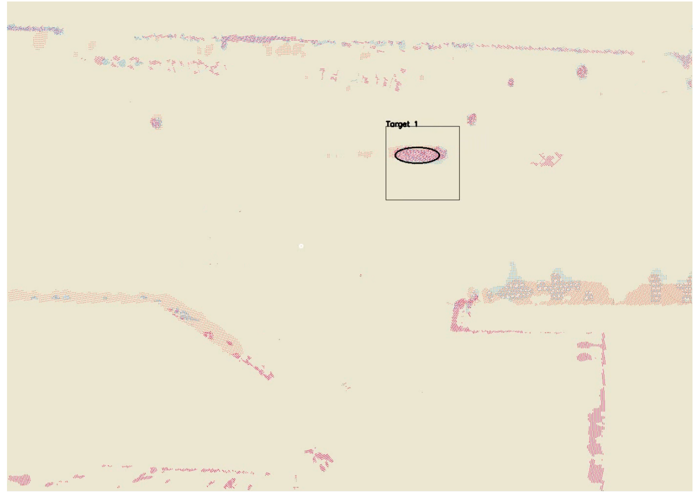


Fig. 2: Overlaid measurements from multiple sensors, where each colour corresponds to a specific sensor. The detected target (also labelled "Target 1") is enclosed within a rectangular box. The ellipse represents the estimated extended state.

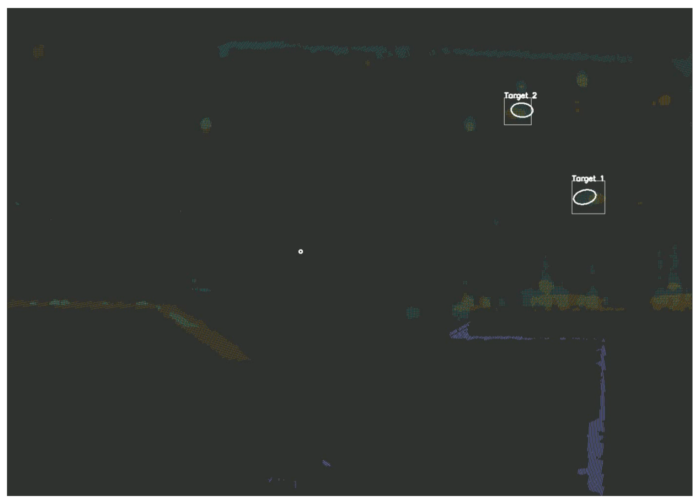
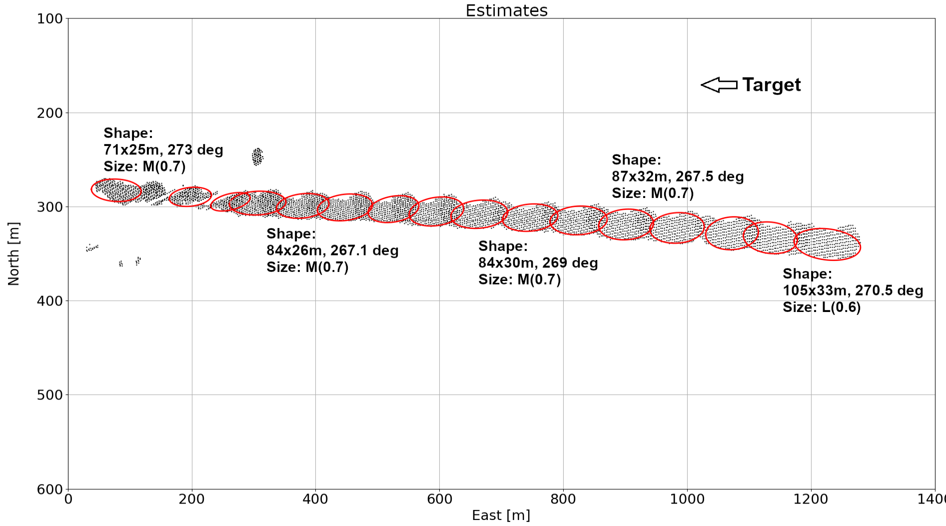


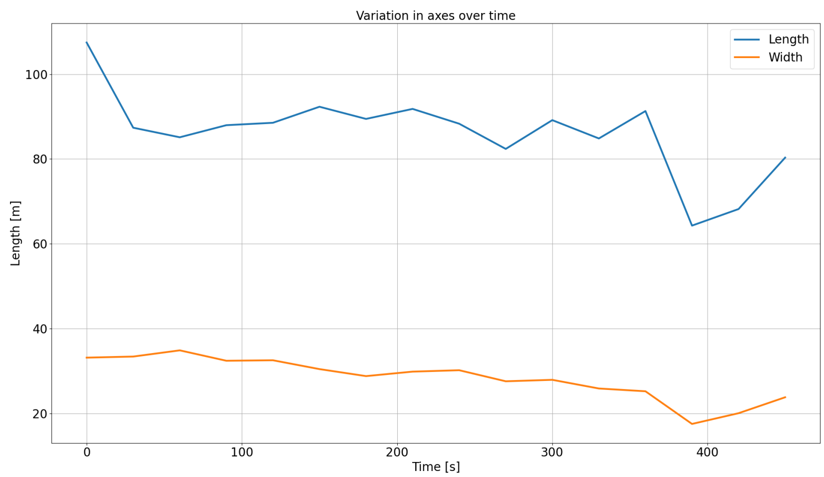
Fig. 3: Two labelled targets as detected and estimated from the multisensor measurements.

Measurements from three sensors are shown in different colours, each corresponding to a particular sensor. Note that the true sensor locations lie outside the frame of the observation region in the Northeast, North and South directions approximately. The acquired target is enclosed within a rectangular box and an ellipse has been constructed from the estimated shape parameters. Figure 4 shows the estimated extent of the target throughout its trajectory together with the variation of its axes length, plotted at an interval of every 15s.

Due to the absence of reference (ground truth) trajectories, it is not possible to determine the exact performance of the filter. However, as it can be seen in Figure 4, the estimations were clear and consistent, especially in the first half of the trajectory as the target had adequate coverage from the sensors. The decrease in size accounts for the sensors' perspectives as



(a) The red ellipses are the extent estimates over time, and the black dots the measurements. The estimated shape parameters and the size category of the vessel with their confidence level in round brackets are also output at regular intervals.



(b) Variations in the lengths of the estimated elliptical axes. The blue line corresponds to the length of the vessel while the orange one to its width.

Fig. 4: Estimations of a single target.

the target moves away from them.

B. Multiple Targets

A sample image of two vessels moving in opposite directions from the second radar stream considered is shown in Figure 3, under a dark background setting. Similar to the former stream, the measurements are colour-coded according to their respective sensors.

The tracking and categorisation results throughout the vessels' trajectories within the observation region are shown in Figure 5, also plotted within regular intervals of 15s. Measurements for the green target were quite occluded, noisy and affected by buoys in the vicinity, especially in the second part of its trajectory. Nonetheless, the estimates held up throughout the vessel's visibility in the observation region. The red target was consistent owing to its clearer visibility throughout. The estimated categories were also consistent for the vessels in both steams.

The framework was also tested for real-time capability on a Linux-based 1.9GHz Quad Core system, using a Python implementation. The average performance for the second stream is as follows: for 500s of data stream, the framework took 510s for execution. In general, the performance was influenced by the number of vessels and also the vessels' dimensions, as larger vessels would yield higher number of measurement points.

V. Summary

In this paper, a framework for multiple extended target tracking from multiple sensors along with vessel size categorisation has been presented and described. Measurements from multiple radars are captured and put together before

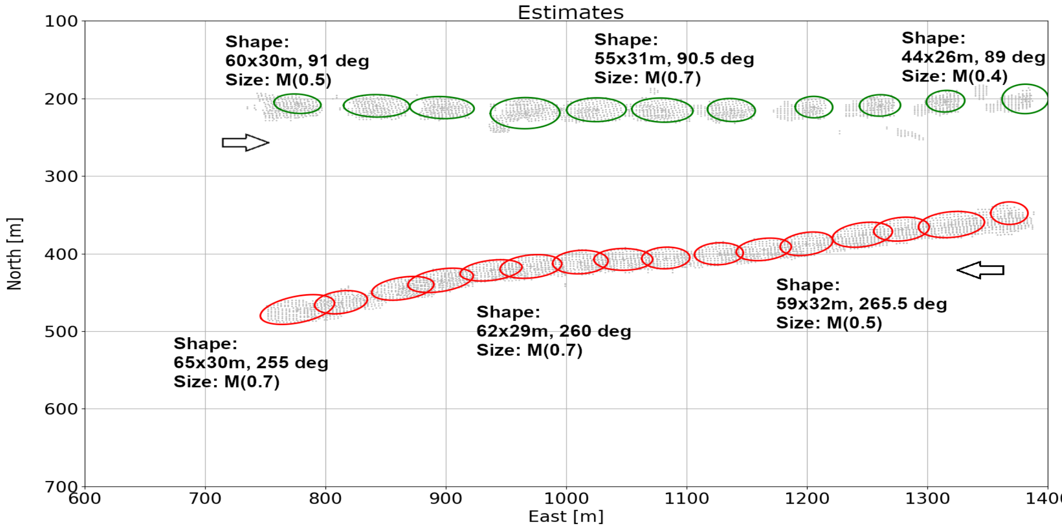


Fig. 5: Extent estimates of the two targets shown in green and red as they navigate in opposite directions, marked by the arrows. The extent parameters, size categorisations having the highest of the confidence levels are included.

the detection and tracking processes are carried out. A JPDA tracker based on the MEM-EKF* measurement model has been employed for the vessels' extended state estimation. In addition, the estimated axes lengths and variances are used to determine a vessel's size categorisation accompanied by its confidence level.

The framework has been implemented and tested on scenarios involving both single and multiple targets, and the results shown were realistic and promising. Despite this, due to the sequential updates of the filter combined with the sequential accounting of multisensors, the computational load increases with increase in the number of targets. Another factor that influences the performance would be the size of the vessels. Nevertheless, with three radars and up to five medium sized vessels, the framework performs close to real-time. In the next steps, we aim to optimise the current performance by improving on the measurement model used for the extent estimation.

References

[1] United Nations Conference on Trade and Development (UNCTAD), "Review of Maritime Transport 2020," Tech. Rep., 2021.

[2] European Maritime Safety Agency (EMSA), "Analysis of Marine Casualties and Incidents Involving Container Vessels," Tech. Rep., 2020.

[3] P. Braca, R. Grasso, M. Vespe, S. Maresca, and J. Horstmann, "Application of the JPDA-UKF to HFSW Radars for Maritime Situational Awareness," 01 2012, pp. 2585–2592.

[4] S. Maresca, P. Braca, J. Horstmann, and R. Grasso, "Maritime Surveillance Using Multiple High-Frequency Surface-Wave Radars," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, 12 2013.

[5] J. S. Fowdur, M. Baum, and F. Heymann, "A Marine Radar Dataset for Multiple Extended Target Tracking," 09 2019.

[6] ——, "Real-World Marine Radar Datasets for Evaluating Target Tracking Methods," *Sensors*, vol. 21, no. 14, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/14/4641

[7] G. Vivone, P. Braca, K. Granström, and P. Willett, "Multistatic Bayesian Extended Target Tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, 12 2016.

[8] K. Granström, M. Baum, and S. Reuter, "Extended Object Tracking: Introduction, Overview and Applications," *ISIF Journal of Advances in Information Fusion*, vol. 12, no. 2, Dec. 2017.

[9] International Association of Independent Tanker Owners (INTERTANKO), "Jamming and Spoofing of Global Navigation Satellite Systems (GNSS)," Tech. Rep., 2019.

[10] F. Katsilieris, P. Braca, and S. Coraluppi, "Detection of Malicious AIS Position Spoofing by Exploiting Radar Information," in *Proceedings of the 16th International Conference on Information Fusion*.

[11] Y. Bar-Shalom, F. Daum, and J. Huang, "The Probabilistic Data Association Filter," *IEEE Control Systems*, vol. 29, 2009.

[12] T. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Multi-Target Tracking using Joint Probabilistic Data Association," vol. 2, 01 1981, pp. 807 – 812.

[13] L. Svensson, D. Svensson, M. Guerriero, and P. Willett, "Set JPDA Filter for Multitarget Tracking," *Signal Processing, IEEE Transactions on*, vol. 59, pp. 4677 – 4691, 11 2011.

[14] D. Musicki and R. Evans, "Joint Integrated Probabilistic Data Association: JIPDA," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 40, pp. 1093 – 1099, 08 2004.

[15] S. Blackman, "Multiple Hypothesis Tracking for Multiple Target Tracking," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 19, pp. 5 – 18, 02 2004.

[16] B.-T. Vo and B.-N. Vo, "Labeled Random Finite Sets and Multi-Object Conjugate Priors," *IEEE Transactions on Signal Processing*, vol. 61, pp. 3460–3475, 07 2013.

[17] J. Williams, "Marginal Multi-Bernoulli Filters: RFS Derivation of MHT, JIPDA, and Association-based Member," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, pp. 1664–1687, 07 2015.

[18] R. Mahler, "An Introduction to Multisource-Multitarget Statistics and its Applications," 01 2000.

[19] S. Yang, K. Thormann, and M. Baum, "Linear-Time Joint Probabilistic Data Association for Multiple Extended Object Tracking," 07 2018, pp. 6–10.

[20] S. Yang and M. Baum, "Tracking the Orientation and Axes Lengths of an Elliptical Extended Object," *arXiv preprint arXiv:1805.03276*, 2018.

[21] EUROCONTROL, "ASTERIX Protocol," [Online]. Available: http://www.eurocontrol.int/asterix, [Accessed: 2021-09-23].

[22] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, pp. 346–359, 06 2008.

[23] Y. Bar-Shalom and X. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, 01 1995, vol. 11.

[24] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," *arXiv preprint arXiv:1811.03378*, 2018.