Information fusion for automotive applications – An overview

Christoph Stiller, Fernando Puente León, and Marco Kruse

***Abstract*—This article focusses on the fusion of information from various automotive sensors like radar, video, and lidar for enhanced safety and traffic efficiency. Fusion is not restricted to data from sensors onboard the same vehicle but vehicular communication systems allow to propagate and fuse information with sensor data from other vehicles or from the road infrastructure as well. This enables vehicles to perceive information from regions that are hardly accessible otherwise and represents the basis for cooperative driving maneuvers. While the Bayesian framework builds the basis for information fusion, automobile environments are characterized by their a priori unknown topology, i.e., the number, type, and structure of the perceived objects is highly variable. Multi-object detection and tracking methods are a first step to cope with this challenge. Obviously, the existence or non-existence of an object is of paramount importance for safe driving. Such decisions are highly influenced by the association step that assigns sensor measurements to object tracks. Methods that involve multiple sequences of binary assignments are compared with soft-assignment strategies. Finally, fusion based on finite set statistics that (theoretically) avoid an explicit association are discussed.**

Keywords: Multimode information fusion, information exchange between automobiles, cognitive cooperative automobiles, Bayesian data fusion, data association and tracking, finite set statistics

I. Introduction

Since its invention by Carl Benz in 1886, the automobile has conquered virtually any populated region on our planet. While the benefit of individual mobility provided by automobiles becomes available to an ever growing population, the resulting vast increase of total milage travelled on our planet demands for technological improvements to prevent an associated boost in traffic fatalities, congestions, and use of environmental resources.

Cognitive capabilities in automobiles and in the infrastructure are widely considered a key technological component to enhance safety, flow, and efficacy of traffic systems. Cognitive automobiles acquire information from their environment by video, radar, and lidar sensors. Based on an interpretation of this information, they build a mental model of the real world and are able to plan and conduct automated driving maneuvers or to assist humans in their driving task. As the potential roadmap of automotive sensors and functions depicted in Fig. 1 shows, the trend towards an increasing number of sensors and sensor-based functions is not new to the automotive domain. Early driver assistance functions focussed on vehicle dynamics stabilization. For this purpose, vehicles were equipped with odometry and inertial sensors to acquire internal vehicle quantities as, for example, the velocity of the individual wheels and acceleration or angular velocity of the vehicle. This information allows to detect extreme driving situations and to support the driver in their stabilization. Prominent examples for vehicle dynamics stabilization systems are anti-lock braking system (ABS), electronic stability program (ESP), or anti-skid control (ASC). New sensing technologies like sonar, radar, lidar, or video extend the sensed information beyond the ego-vehicle state to environmental information. This enables a wide field of new functions, such as, e.g., lane departure warning (LDW), parking aids based on sonar, radar, or video or the adaptive cruise control (ACC) that automatically adjusts velocity to keep a comfortable distance to predecessing vehicles. Despite impressive recent merits in research in this field, the uncertainty of environment information is far too high as to allow automated driving in the near future.[1] Thus the functional spectrum is restricted to information, warning, and comfort enhancement while the final responsibility stays with the driver. In order to enhance vehicle safety, first functions to mitigate collision hazards have cautiously been implemented. Automated emergency braking (AEB) engages a strong braking when an immanent and inevitable collision is detected. Due to uncertainties in the processing chain, this action does not aim to avoid the collision but just to reduce the kinetic energy of the impact. In order to

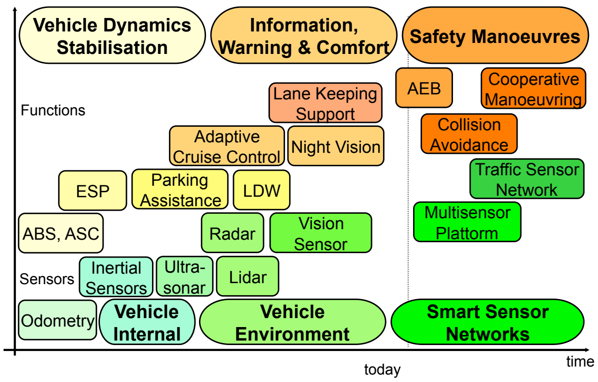


Fig. 1. Potential evolution of automotive sensors (green) and functions (orange).

C. Stiller is with Institut für Mess- und Regelungstechnik, Karlsruhe Institute of Technology (KIT), Engler-Bunte-Ring 21, D-76131 Karlsruhe, Germany; F. Puente León and M. Kruse are with Institut für Industrielle Informationstechnik, Karlsruhe Institute of Technology (KIT), Hertzstr. 16, D-76187 Karlsruhe, Germany, e-mail: {stiller, puente, kruse}@kit.edu.

[1] In the 2007 DARPA Urban Challenge, a competition of autonomous vehicles in a mock-up urban-like setting, the eleven finalists were involved in six accidents on a ca. 100 km parcours [1]. For comparison, human drivers produce an accident about every 10^5 km.

further extrapolate this trend towards safe traffic, the enhancement of reliability and certainty of the perception system is of prevailing importance. As a first step, diverse information of different sensors will be combined to a consistent and plausible scene representation. Such multi-sensor platforms will then allow to recognize selected critical situations with a level of plausibility that is even sufficient to engage evasive steering onto a last-second trajectory.

The combination of information from sensors onboard different vehicles and on the infrastructure through communication systems will finally yield traffic sensor networks opening up a totally new spectrum of functionalities with unprecedented benefits [2]. First of all, cooperative sensing and cooperative maneuver planning will considerably improve traffic safety. Furthermore, such technology enables coordinated traffic trajectories, which avoids sharp acceleration/deceleration and idling. Based on this information, speed can be harmonized with both the traffic light cycles and the traffic situation, thus yielding improved traffic flow as well as fuel and CO_2 savings of up to 14%. Up to 25% of fuel and the vast majority of traffic space can be saved through tight convoy driving of vehicles on highways. To foster international research collaboration in this field, the Grand Cooperative Driving Challenge (GCDC) has been launched. In a scientific competition, international teams that lead the field shall compete for the best cooperative driving strategies and demonstrate technical feasibility and benefits [3].

In the long run, the trend towards intelligent and communicating infrastructures will further improve on this situation. A wide spectrum of potential improvements are expected from roads, intersections, traffic lights and signs that transmit their occupancy and other status information. Nevertheless, due to the broad variety of stakeholders (policy makers, road operators, infrastructure constructors, vehicle manufacturers, vehicle insurances, various road users, etc.) that need to agree on a concerted action, the introduction of infrastructure systems on an international level may be somewhat more difficult as compared with systems the are solely mounted on vehicles. While this overview mainly focusses on onboard technologies, we refer interested readers to a recent comprehensive special issue on intelligent transportation systems that addresses the broader scope [4].

A rich body of literature exists on information fusion methods for automotive applications. Most contributions are tailored to a specific sensor setup, and function. E.g., in an early work, object detections of an ACC radar have been included as additional observations in a vision-based lane recognition procedure to enhance the range of lane detection and the reliability of object-to-lane assignments [5].

A classical architecture that is followed by the broad majority of approaches is depicted in Figure 2. Input to the fusion loop structure is the raw data of all sensors. The extraction of features of interest, clustering of features and the detection of object hypotheses may be conducted separately for each sensor or in a concurrent treatment. Features or objects hypotheses are then associated to individual tracks, which represent the state estimates of all objects detected. The associated information is then fed into the individual state estimators, such as Bayesian filters, to update and predict the current states. The loop is closed through the predictions that are fed back to the detection and association modules. A dedicated track management module organizes validation, deletion or augmentation of tracks. Association is a critical module in this architecture, as it conducts an early decision that may seriously affect the fusion result.

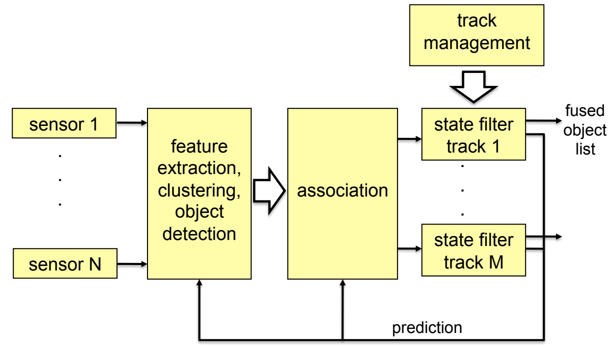


Fig. 2. Information fusion architecture.

Rather than making hard decisions, probabilistic data association (PDA) omits or reduces the deterioration from erroneous decisions. Extensions to the original joint probabilistic data association (JPDA) [6] include fast approximations (cheap joint probabilistic data association, CJPDA) or explicit modeling of object existence (joint integrated probabilistic data association, JIPDA) [7].

Motivated by highly cluttered sensor data, that prevents a reasonable detection of object features prior to a valid track hypothesis, so called ‘track-before-detect’ methods completely avoid an explicit association step [8]. Based on the theory of finite set statistics (FISST), such methods formulate probability distributions over object lists of variable dimensions.

A crucial step towards deployment will be in the design of generic fusion techniques that perform sensor- and function-independent. In this context, several approaches to ‘plug-and-play’ information fusion are reported, see e.g. [9].

The remainder of this article is organized as follows: Section II outlines various fields of information fusion for cognitive automobiles along with their potential benefits and challenges. While significant advantages are expected from information exchange between vehicles, both the additional time delay and the inherent accumulation of uncertainty bound the benefit of propagated information. Section III provides an overview on the major methods for information fusion applied in the automotive domain. Emerging from the Bayesian framework, multi-object detection and tracking methods are outlined. Various approaches to data association as well as approaches that avoid an explicit association step are discussed. Section IV summarizes our results and concludes the paper.

II. Diversity in automobile systems

Fusion techniques may, generally, be classified based on the level of information processing where fusion takes place [10], [11]. Automotive data fusion offers a particularly rich variety in potential exploitation of information diversity. Figure 3 illustrates some important goals that are discussed in the remainder of this section.

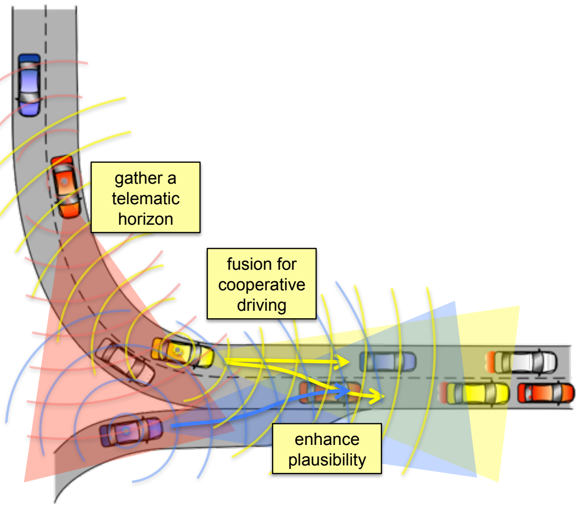


Fig. 3. Information fusion aims to yield harmonized driving trajectories, extend the vehicle view to a telematic horizon, and to enhance evidence of the information available to each vehicle.

A. *Information fusion for cooperative driving*

The ultimate goal of information exchange in traffic is to harmonize driving maneuvers. As illustrated in Figure 3, an augmentation of the knowledge base of vehicles by intended driving trajectories of other vehicles possibly followed by negations may yield driving decisions that improve the overall traffic flow and traffic safety. One of the rare functions in this field that has successfully been implemented on experimental vehicles and trucks is platoon driving. Despite the known benefits of such cooperative traffic operation (see e.g. [12]), such maneuvers require a whole set of innovative equipment. All participating vehicles need a reliable and fast communication device, a positioning or distance sensor, and high-dynamic actors that will only eventually be available to the market. Recently, an international series of friendly competitions called the *Grand Cooperative Driving Challenge* has been initiated to foster research activities in cooperative driving maneuvers [3].

B. *Information fusion for cooperative perception*

The importance of information exchange and data fusion increases with the growing equipment rates of vehicles and infrastructure with sensors and communication devices. As illustrated in Figure 3, vehicles may gather important information on traffic participants in their blind spot, at far distances and in occluded regions. Such a *telematic horizon* may significantly extend a vehicle's understanding of the current traffic situation. It is worth noting from the above figure that cooperative sensing does not require a 100% equipment rate, but provides benefits even at moderate rates. Preliminary experiments with cooperative perception between vehicles have been reported in [13], [14]. An important issue in this context is the spatio-temporal registration of data transmitted in the coordinate system of other vehicles. Since the uncertainty of the spatio-temporal alignment cumulates with the intrinsic uncertainty of the sensor information, this alignment must be conducted with high precision. An alignment strategy that combines the coarse localization information of a GPS system with the onboard sensor information is considered a promising solution (Fig. 4).

Let vehicle C' observe an object of interest in its coordinate frame at time t' and at position $\mathbf{X}' = (X', Y', Z')^{\mathrm{T}}$. Let further the position uncertainty of C' be expressed by the covariance $\Sigma_{\mathbf{X}'}$. This information is transmitted to and transformed by vehicle C to ego-coordinates $\mathbf{X} = (X, Y, Z)^{\mathrm{T}}$. This transformation requires knowledge on the relative pose of C' wrt C expressed through the rotations $\omega = (\omega_X, \omega_Y, \omega_Z)^{\mathrm{T}}$ about the X-, Y-, and Z-axis, and the translation $\mathbf{t} = (t_X, t_Y, t_Z)^{\mathrm{T}}$, respectively. Let their uncertainties be denoted by Σ_ω and $\Sigma_{\mathbf{t}}$ and let all uncertainty vectors be mutually uncorrelated. The coordinate transform yields the position estimate in the ego coordinate system

$$\mathbf{X} = \mathbf{R}\mathbf{X}' + \mathbf{t}\ , \tag{1}$$

where $\mathbf{R} = \mathbf{R}(\omega)$ denotes the rotation matrix associated with ω. The uncertainties in object position and relative pose accumulate to

$$\Sigma_{\mathbf{X}} = \mathbf{R}\Sigma_{\mathbf{X}'}\mathbf{R}^{\mathrm{T}} + [\mathbf{X}']_\times \Sigma_\omega [\mathbf{X}']_\times^{\mathrm{T}} + \Sigma_{\mathbf{t}} \tag{2}$$

$$\text{with}\quad [\mathbf{X}']_\times = \begin{pmatrix} 0 & -Z' & Y' \\ Z' & 0 & -X' \\ -Y' & X' & 0 \end{pmatrix}.$$

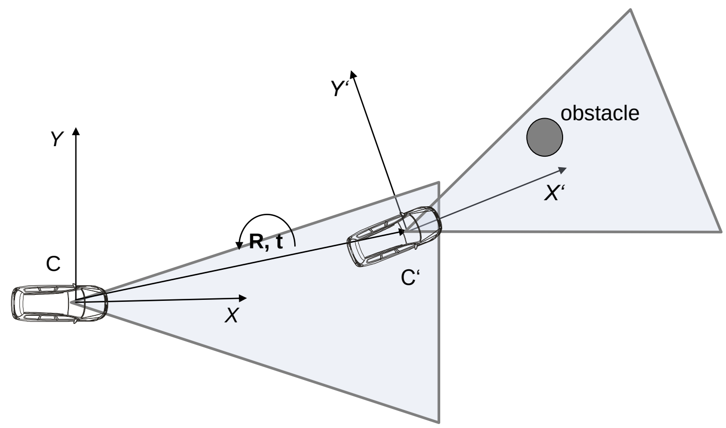


Fig. 4. Uncertainty in object position as observed by C' accumulates with uncertainty in coordinate transform $\mathbf{R}$,$\mathbf{t}$ for cooperative perception.

In practice, the second term becomes dominant for distant objects. As information sent by other vehicles may thus be deteriorated by an additional time delay and a pose uncertainty, the information that is selected for communication and the reference frame for this information must be carefully chosen. Clearly, the requirements on accuracy and latency time depend on the specific function considered. A demonstrated in the European project Prevent, e.g. warning of a local danger on the road like an obstacle or a construction zone may allow implementation with a precision in the meter domain and a latency of some seconds [15]. In contrast, collision avoidance and mitigation functions may require centimeter precision and millisecond latency.

When position information is exchanged between infrastructure and vehicles, uncertainty accumulation may be avoided through geo-referencing of information [13], [16]. Several European projects like Safespot, CVIS, COOPERS, and Intersafe have implemented geo-referenced roadside sensing and communication. Operation in geo-referenced coordinates allows the vehicles to avoid uncertainty accumulation of both received and transmitted information. It has been shown that some early information (e.g., on recommended driving route or speed to avoid a stop at a red light) and warning (e.g., of a potential collision with oncoming traffic to a left turning vehicle) could be provided to drivers. Furthermore, automated blockage of dangerous maneuvers (such as a colliding left turn) is investigated in [17], [18], [19].

III. Methods for automotive information fusion

If at least two distinct pieces of evidence providing information about the same entity of interest are available, the question arises, how to combine them to obtain a “better” knowledge about said entity. In automotive applications, this problem is encountered when measurements collected at different times or by different sensors are to be fused. Bayesian statistics provide an elegant answer to this question.

A. Bayesian tracking

Under this paradigm, the system state x_k encapsulates all relevant information about the state of the world at time t_k. Depending on the application, this may include the car’s dynamic state, road geometry [20], [21], the driver’s biomedical condition, the degree of the driver’s distraction [22], discrete driving modes (like “accelerating,” “standstill,” “going backwards,” etc.) [23], maneuvering intentions [24], and many more.

Unfortunately, this state is usually not perfectly observable, but only indirectly inferable from error-afflicted measurements z_k collected by sensors. Thus the system state is interpreted as a random variable X_k and all the knowledge is incorporated into its posterior probability density $f_{X_k}\big(x_k \big| \mathtt{z}^{(k)}\big)$ conditioned on the set $\mathtt{z}^{(k)} = \{z_1, \ldots, z_k\}$ of all currently available measurements. If required, state estimations can be obtained from this density, usually using its maximal, expected or median value. In many scenarios, the assumption that the states

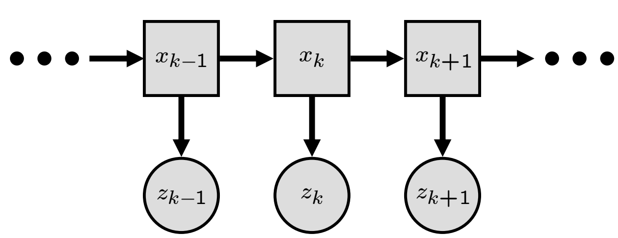


Fig. 5. Dependencies in a hidden Markov chain. An arrow expresses that its head is dependent on its tail.

and the measurements can be described by a hidden Markov model (HMM), as depicted in Fig. 5, is feasible. This implies that the state x_k conditioned on its direct predecessor x_{k-1} is assumed to be independent of previous states or measurements

$$f_{X_k}\left(x_k \middle| x_{k-1}, \mathtt{x}^{(k-2)}, \mathtt{z}^{(k-1)}\right) = f_{X_k}(x_k|x_{k-1}) \ , \tag{3}$$

and measurements only depend on the current state

$$f_{Z_k}\left(z_k \middle| x_k, \mathtt{x}^{(k-1)}, \mathtt{z}^{(k-1)}\right) = f_{Z_k}(z_k|x_k) \ . \tag{4}$$

The r.h.s. of (3) is the Markov transition density and describes the temporal development of the system state. In single-object tracking, e.g., it is fully determined by the chosen motion model and plant noise. The r.h.s. of (4) is called likelihood and is given by the sensor model, i.e., it describes the knowledge about the measuring principle, the sensor field-of-view, sensor noise, etc.

With these assumptions and Bayes’ theorem

$$f(x,z) = f(z|x)\, f(x) = f(x|z)\, f(z) \tag{5}$$

as well as marginalization

$$f(z) = \int f(x,z)\,\mathrm{d}x \,, \tag{6}$$

the posterior density $f_{X_k}\big(x_k \big| \mathtt{z}^{(k)}\big)$ can be expressed as

$$f_{X_k}\Big(x_k \Big| \mathtt{z}^{(k)}\Big) = \frac{f_{Z_k}(z_k|x_k)\, f_{X_k}\big(x_k \big| \mathtt{z}^{(k-1)}\big)}{\int f_{Z_k}(z_k|x_k)\, f_{X_k}\big(x_k \big| \mathtt{z}^{(k-1)}\big)\,\mathrm{d}x_k} \,, \tag{7}$$

$$f_{X_k}\Big(x_k \Big| \mathtt{z}^{(k-1)}\Big) = \int f_{X_k}(x_k|x_{k-1})\, f_{X_{k-1}}\Big(x_{k-1} \Big| \mathtt{z}^{(k-1)}\Big)\,\mathrm{d}x_k \ . \tag{8}$$

These two equations are known as the Bayes recursion. The prior density $f_{X_k}\big(x_k \big| \mathtt{z}^{(k-1)}\big)$ is predicted from the previous posterior density with the Chapman-Kolmogorov equation (8) utilizing the state transition density. It is then updated with (7) using the new measurement and the sensor’s likelihood. This two-step processing is also known as tracking, dynamic state estimation, or state filtering (cf. Fig. 2). It is worthwhile to note that upon the arrival of a new measurement z_k, $f_{X_k}\big(x_k \big| \mathtt{z}^{(k)}\big)$ can be computed solely based on the assumed models for system and sensor ($f_{X_k}(x_k|x_{k-1})$ and $f_{Z_k}(z_k|x_k)$) and the previous posterior $f_{X_{k-1}}\big(x_{k-1} \big| \mathtt{z}^{(k-1)}\big)$, without reconsidering any past measurements. This prevents the computational complexity, necessary to obtain $f_{X_k}\big(x_k \big| \mathtt{z}^{(k)}\big)$, to grow with time. However, upon the arrival of the first measurement, an initial prior $f_{X_0}(x_0)$ has to be provided. Usually this is chosen to be a very uncommitted density, like a broad uniform density or a normal density with a large variance, to express a lack of prior knowledge.

To fuse information obtained at different times, all that is necessary is to calculate the posterior density for each time instance when sensor data is acquired. When information from different synchronized sensors shall be fused, usually the assumption that their generated measurements are independent from each other conditioned on the state is made, i.e.,

$$f_{Z_{k,1}Z_{k,2}}\left(z_{k,1}, z_{k,2} \,|x_k\right) = f_{Z_{k,1}}\left(z_{k,1} \,|x_k\right) \cdot f_{Z_{k,2}}\left(z_{k,2} \,|x_k\right) \ . \tag{9}$$

If this holds, fusion can be achieved by sequentially performing an update step according to (7) for each sensor, where the ordering of the sensors is insignificant. Otherwise the joint likelihood $f_{Z_{k,1}Z_{k,2}}(z_{k,1}, z_{k,2} \,|x_k)$ has to be modeled and one single update step can be performed. Of course, if all data to be fused originates from the same time instance, and tracking of the state over time is not required, the prediction step in (8) can be omitted.

The challenge in propagating the posterior density stems from the fact that the involved integrals in (7) and (8) cannot be solved analytically for many models used in real-world applications. To remedy this problem, either numerical approximation techniques are applied or further constraints that lead to an analytic solution are imposed on the models. Examples for the former are particle system- or grid-based approaches [25], [26], [27], [28], whereas the Kalman filter [29] is the most prominent example of the latter.

Automotive sensors are typically operated asynchronously and often possess different processing times. Hence, the order in which sensor measurements become available may differ from the order of raw data acquisition by the sensors. While naive forward and backward prediction results in sub-optimal results in Bayesian filtering, a simple buffering of information until the data is available in the order of its acquisition introduces unnecessary dead-time, which may severely degrade the performance of the overall control circuit. Several compromises are applied between those two extremes: retrodiction of measurements into pseudo-measurements that are aligned in time as well as asynchronous tracking systems that employ every measurement upon availability to validate and initiate its tracks have been proposed [30].

B. Detection and tracking

So far, nothing was said about the types of mathematical descriptors used for the system state x and the measurements z.

In most vehicle internal sub-systems (e.g., ESP), a fixed set of scalar parameters is the most natural way to describe the system at hand. Usually there is only uncertainty about the localization of the system in state space, not about its existence, and the cardinality of the data received per time step is fixed and known. For this type of application, random vectors $\mathbf{X}$ and $\mathbf{Z}$ are appropriate descriptors and various implementations/approximations of the Bayes recursion for a vast range of models are well-studied.

In some applications the state either includes inherently discrete elements, e.g., if one assumes that a car has a finite number of driving modes [23], or only a priori chosen discrete values of a continuous variable shall be considered, as for the height of a pedestrian in [31]. In practice, interacting multiple models are employed in these cases. This means the discrete elements of the state vector are modeled as a Markov Chain whose evolution is independent of the continuous state elements [32]. Also, with discrete state elements, binary and multi-class classification can be cast in a Bayesian way, see e.g., [22].

Many current and future automotive applications, however, require to perceive the environment of the automobile with an a priori unknown number and topology of objects. Hence, the dimension of the system state is a priori unknown, as it comprises a variable number of relevant entities in the vicinity. Especially for such applications fusion of heterogeneous sensors (possibly across sensor carriers) becomes vital, not so much as to decrease uncertainty in localization, but to validate object number and topology [24] as well as increasing the telematic horizon.

Often the raw sensor data cannot be utilized for fusion, due to bandwidth constraints. Instead, an a priori unknown number of object hypotheses, referred to as detections from hereon, whose union then forms the actual measurement, have to be extracted therefrom, utilizing detection and/or classification schemes. For extremely low SNR, discrimination between noise and true object detection in one measurement might be hardly possible, resulting in either high false alarm or high missed detection rates. In such scenarios, *track-before-detect* (TBD) methods are proposed that employ the raw sensor data as measurement [33], [34].

A natural way to handle the unknown and variable number of objects/detections is to describe the system state and the measurement as random finite sets (RFSs) that involve one per object/detection. Thus Eqs. (7) and (8) express probability distributions on finite sets $\mathtt{X}$ and $\mathtt{Z}$ of variable cardinality [35], [36], [8]. Because the number of objects varies, the multi-object Markov density $f_{\mathtt{X}_k}(\mathtt{x}_k \,|\mathtt{x}_{k-1})$ has to be capable of modeling object dis-/appearance. If the sensor delivers object detections, the multi-object likelihood $f_{\mathtt{Z}_k}(\mathtt{z}_k \,|\mathtt{x}_k)$ has to incorporate models for false alarms and missed detections due to imperfect detectors. The possible interrelations between object states at different time instances and measured detections are depicted in Fig. 6.

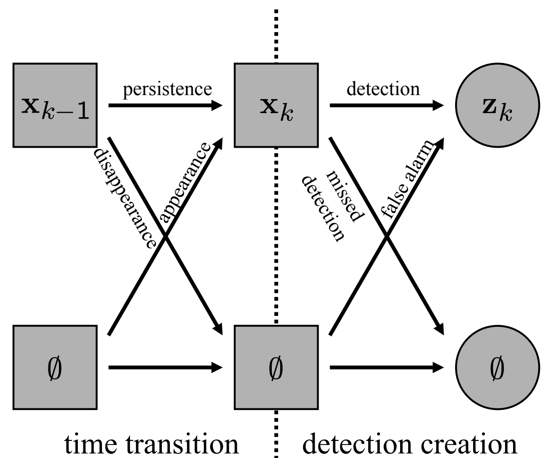


Fig. 6. Possible causes for time varying and unknown numbers of objects and detections.

C. Explicit data association tracking

Many existing multi-object tracking algorithms are extensions of well-known single-object trackers, like the Kalman filter, to multi-object problems. The general divide-and-conquer approach is to partition the timely ordered set of detections $\mathtt{z}^{(k)} = \{\mathtt{z}_1, \ldots, \mathtt{z}_k\}$ into subsets $\mathtt{t}^i \subseteq \mathtt{z}^{(k)}$, called tracks, and use the single-object Bayes recursion for each track (cf.

Fig. 2). This process of partitioning is called data association and is the main source of differences between multi-object trackers. Even if one only allows binary assignment decisions, the optimal assignment of detections to tracks is a combinatorial problem, and the number of feasible solutions grows exponentially with the number of time steps, detections and tracks. Hence, a major issue for multi-target trackers is to cope with this vast number of possible assignment sequences.

A basic technique to reduce the number of feasible assignments encountered in nearly every implementation of a multi-object tracker based on Gaussian noise models is gating. The rationale behind gating is that an association between detection and track is only admissible, when the squared Mahalanobis distance between the track's predicted measurement $\mathbf{z}^-$ and received detection $\mathbf{z}$ is below a certain threshold, which can be derived from the χ^2-distribution.

1) Multi-hypothesis tracking: In multi-hypothesis tracking (MHT) [37], [38], [39], [40], [41], [42], [43] a set of tracks is referred to as a feasible hypothesis Ω_k^l, if they are mutually disjoint (or compatible)

$$\mathtt{t}^i \cap \mathtt{t}^j = \emptyset \qquad\qquad i \neq j \ , \tag{10}$$

cover all received detections

$$\bigcup_{i:\mathtt{t}^i \in \Omega_k^l} \mathtt{t}^i = \mathtt{z}^{(k)} \ , \tag{11}$$

and (usually) if each track contains at most one measurement per time step k

$$\left|\mathtt{t}^i \cap \mathtt{z}_k\right| \leq 1 \ . \tag{12}$$

Basic MHT enumerates all feasible hypotheses, assigns a score to them, and outputs the hypothesis (and derived state estimates for each track) with the highest score. This means that association uncertainties at the current time step can be resolved with later measurements, because no irreversible assignments are made immediately.

Whereas hypothesis-oriented MHT [38] propagates and scores all hypotheses and expands them to new ones upon arrival of a new measurement, track-oriented approaches [41] propagate the tracks and construct hypotheses from scratch at each time step, which usually results in simpler and faster implementations.

Typically [37], [39], [44] the log-likelihood ratio $LLR_i = \ln \frac{f\left(H_{\mathrm{T}} \middle| \mathtt{t}^i\right)}{f\left(\overline{H_{\mathrm{T}}} \middle| \mathtt{t}^i\right)}$ for the assumption H_{T} that all detections of a track originate from the same true object is used as score for individual tracks, whereas the sum of scores of all contained tracks constitutes a hypothesis' score. Note that the probability of track $\mathtt{t}^i$ representing a true object can be obtained from LLR_i by $f\left(H_{\mathrm{T}} \middle| \mathtt{t}^i\right) = \frac{\exp(LLR_i)}{1+\exp(LLR_i)}$.

Due to the aforementioned explosion in the number of hypotheses, a strategy is mandatory that concentrates tracking on the most prominent hypotheses. Most common approaches to multi-target tracking delete tracks/hypotheses with low scores/probabilities (gating), or merge "similar" tracks. For MHT, this can efficiently be implemented by N-scan pruning. In this technique, all hypotheses are maintained in a rooted forest of depth k such that each tree represents one object and each path from a root to a leave represents a possible association sequence for this object up to time step k. Two tracks at step k have a common parent node if and only if they share all but the latest detection. N-scan pruning subsequently reduces branches in each tree retaining only the branch at layer $k - N$ that optimizes a predefined criterion, e.g., maximal or average probability/likelihood of connected leave nodes, while all other branches are removed. Hence, at time instance k irreversible decisions are only conducted on assignments N steps in the past.

2) Probabilistic data association: Whereas MHT tries to resolve association uncertainties by deferring them until more data is available, the class of probabilistic data association (PDA) algorithms protects itself against false associations by soft-assignments between detections and tracks. Rather than assigning a unique element of $\mathtt{z}^{(k)}$ per time step to each track $\mathtt{t}^i$, a set of weighting coefficients $\beta_k^{i,l}$ expresses the probability of detection $\mathbf{z}_k^l$ contributing to track $\mathtt{t}^i$. Since these so-called association probabilities are not needed in later iterations of the filter, they are not stored in practical implementations, and the term "track" usually refers to the state estimates rather than the assigned measurements in the specific literature.

To determine the association probabilities, all possible mappings of current detections to tracks (called joint association events) are listed together with their associated probabilities. The probabilities of single association events $\beta_k^{i,l}$ are then obtained by marginalizing over all joint events containing this association. This explicit listing of all possible joint events is the main drawback for this class of algorithms, as it has a complexity exponentially growing with the number of detections and tracks. The final posterior density $f\left(\mathbf{x}_k^j \middle| \mathtt{t}^j\right)$ of track j results from a weighted average of the temporary updates with each individual detection, taking the association probabilities $\beta_k^{i,l}$ as weights

$$f\left(\mathbf{x}_k^j \middle| \mathtt{t}^j\right) = \sum_{l=1}^{L} \beta_k^{i,l} \cdot f\left(\mathbf{x}_k^{i,l} \middle| \left\{\mathtt{t}^i, \mathbf{z}_k^l\right\}\right) \ . \tag{13}$$

While the original PDA method [45] was designed for single-object tracking with clutter, it was later extended to multi-object tracking for a known (JPDA) [46] and unknown (JIPDA) [47], [9] number of objects, by considering joint events as well as the events of missed detections, false alarms and explicit modeling of object existence probability. To account for uncertain motion models of tracks in automotive applications, JPDA has also been combined with *interacting multiple model* (IMM) filters [48].

D. Implicit data association tracking

While the methods in the last subsection can be interpreted as bottom-up approaches to multi-object tracking that extend single-object trackers with data association capabilities, finite set statistics (FISST) based approaches follow a top-down derivation that avoids explicit data association. This is not to say that specific methods presented so far cannot be cast or motivated in a FISST framework, indeed this has been done numerous times [49], [50], [51].

The innovation of this paradigm is to model both the system state and the measurements as random finite sets (RFSs) $\mathtt{x}_k$ and $\mathtt{z}_k$ and directly apply the Bayes recursion to these set-valued random variables and solving the data association problem implicitly. The necessary mathematical tools are found in stochastic geometry [35], [36], for which Mahler coined the term finite set statistics [8].

As with the Bayes recursion for vector-valued variables, the set-valued equivalent usually is not solvable analytically. Hence, tractable approximations are implemented in practice.

In contrast to explicit data association methods, which model and keep track of object identities in an explicit way, most FISST-based methods inherently marginalize over all association possibilities and output a set of object hypotheses per time step. These may not possess object identities, although this can be augmented in numerous ways [52], [53], [54].

1) (Cardinalized) probability hypothesis density filter: For set-valued random variables, moments can be used for description and tracking, just as for their vector-valued counterparts. However, the first moment of an RFS $\mathtt{X}$ is not set-valued but a density function $v(\mathbf{x})$ on the state space of a single-object, specifying the density of objects in said state space. For this reason it is also referred to as probability hypothesis density (PHD) or intensity. It can be defined by the property that its integral over a region $\mathtt{s}$ of a single-object state space is equal to the expected number of objects in that region

$$\int_{\mathtt{s}} v(\mathbf{x})\,\mathrm{d}\mathbf{x} = \mathrm{E}\left\{|\mathtt{s} \cap \mathtt{X}|\right\} \quad . \tag{14}$$

So instead of calculating the complete posterior density $f_{\mathtt{X}_k}\left(\mathtt{x}_k \,\middle|\, \mathtt{z}^{(k)}\right)$ of the set-valued system state, the PHD filter [55], [56], [57] only propagates its first moment, the eponymous PHD. The locations of likely existing objects are then extracted from the PHD's local maxima. Despite the seemingly rather drastic loss of information caused by the collapse of the posterior to its first moment, the filter was shown to perform surprisingly well in several practical applications [58], [59], [60], [61]. Its low computational complexity, which is linear in the number of targets and detections, is one of its most beneficial properties, whereas its main deficiency is the inaccurate estimate of object number, which results from the implicit assumption of a Poisson distribution, thus embodying only a single parameter to capture mean and variance. To alleviate this drawback, Mahler proposed the cardinalized PHD filter, which allows for arbitrary cardinality distributions. However, this extension also increases the complexity of the filter significantly, growing cubically in the expected number of objects.

2) (CB)MeMBer: FISST-based approximations of the Bayes recursion that propagate the full posterior density of the system state usually assume that the system state can be approximated well with a multi-Bernoulli random variable. This is simply the union of a finite number of sets, each of which is either the empty set with probability $1 - r^i$ or with probability r^i a singleton Bernoulli random variable (usually assumed mutually independent of all other singletons), whose location in single-object state space is distributed according to a probability density function $p^i(\mathbf{x})$. This representation, which is also used, e.g., in the JIPDA filter, corresponds to the classical track representation in multi-target tracking algorithms with an associated probability of existence r^i for each track.

Unfortunately, this kind of representation is not closed under the Bayes update step for most models, so the resulting density after one iteration is strictly not multi-Bernoullian. Nevertheless, multi-Bernoulli approximations can be found.

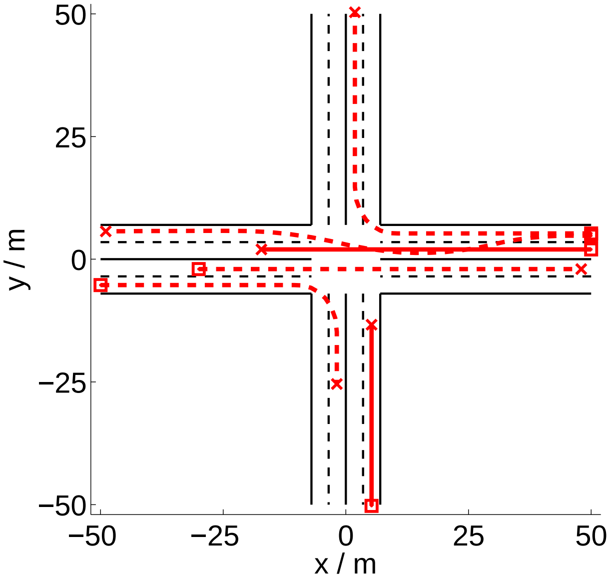
Whereas the JIPDA filter creates one Bernoulli component per track before the update and uses explicit marginalization over all association possibilities between detections and tracks, the multi-target multi-Bernoulli (MeMBer) filter creates one component per old track, hypothesizing a missed detection, and one track per detection comprised of contributions from all existing old tracks. With this approach, a recursion that has a complexity linearly dependent on the number of tracks and detections can be derived. It is worth noting that the original proposition by Mahler [8] has a significant bias in the estimation of the target number, which can be circumvented with the correction termed cardinality balanced MeMBer of Vo [62]. Although this filter showed advantageous properties in simulations, no real-world application has been reported yet.

A simulated traffic scene is depicted in Fig. 7 (a). There are up to six automobiles in the scene. One of these is equipped with a lidar sensor and a self-localization system, while another one, entering the scene later on, is equipped with a radar and a self-localization system. As Fig. 7 (b) shows, a moderate rate of false alarms and missed detections are taken into account. The fusion results employing all available sensor data and a CBMeMBer filter are shown in Fig. 8.

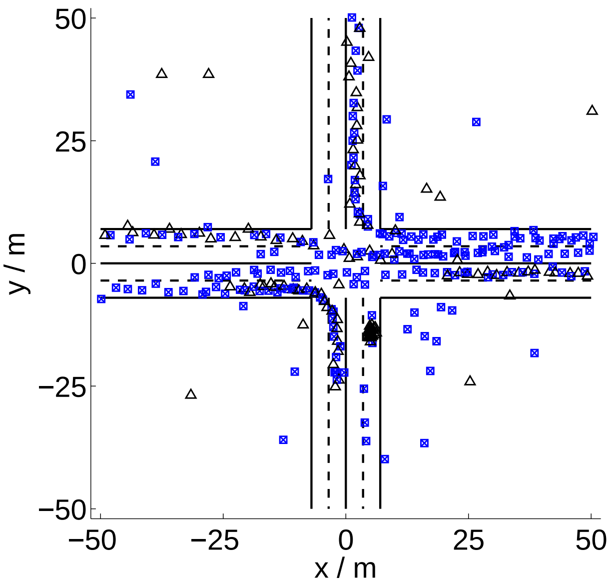
IV. Summary and Conclusion

Vehicle environment sensing offers unprecedented chances to enhance safety and efficiency in the automotive domain. While environmental sensors like radar, video, and lidar are successively mounted in automobiles and the traffic infrastructure, the inherent uncertainty of the information provided by each of these sensors prohibits to engage safety-relevant automated driving functions that rely on such information. Fusion of the information from all sensors onboard a vehicle and – when augmented with a communication system – with the sensors mounted on other vehicles and the infrastructure aims to exploit diversity to yield a more plausible and reliable representation of the driving environment. At the not so far end of this development, cooperative perception and traffic operation will significantly improve use of resources (including fuel and traffic space) and automotive safety.

While information fusion from data onboard the same vehicle mainly aims at an environment representation that combines the ranges and field-of-view of all sensors and plausibilizes information in overlapping areas, information exchange between vehicles or between vehicles and the infrastructure opens the potential of a telematic horizon and



(a)



(b)

Fig. 7. Simulation of a traffic scene. (a) True trajectories. Squares mark initial, crosses final automobile position. Automobiles with sensors are indicated by continuous lines. (b) Position components of simulated measurements.

cooperative planning of driving maneuvers. Nevertheless, the benefit of communicated data is not unlimited, as it involves additional time delay and accumulated uncertainty. These effects increase with the number of sensing and communication devices involved.

Numerous methods have been applied to information fusion in the automotive domain. This article has briefly reviewed the Bayesian framework, which can be extended to multi-object detection and tracking. Data association is a critical procedural step in information fusion, as a too early hard decision may yield poor estimates, whereas late decisions vastly increase the computational burden. While unique assignments between measurements and tracks are imposed

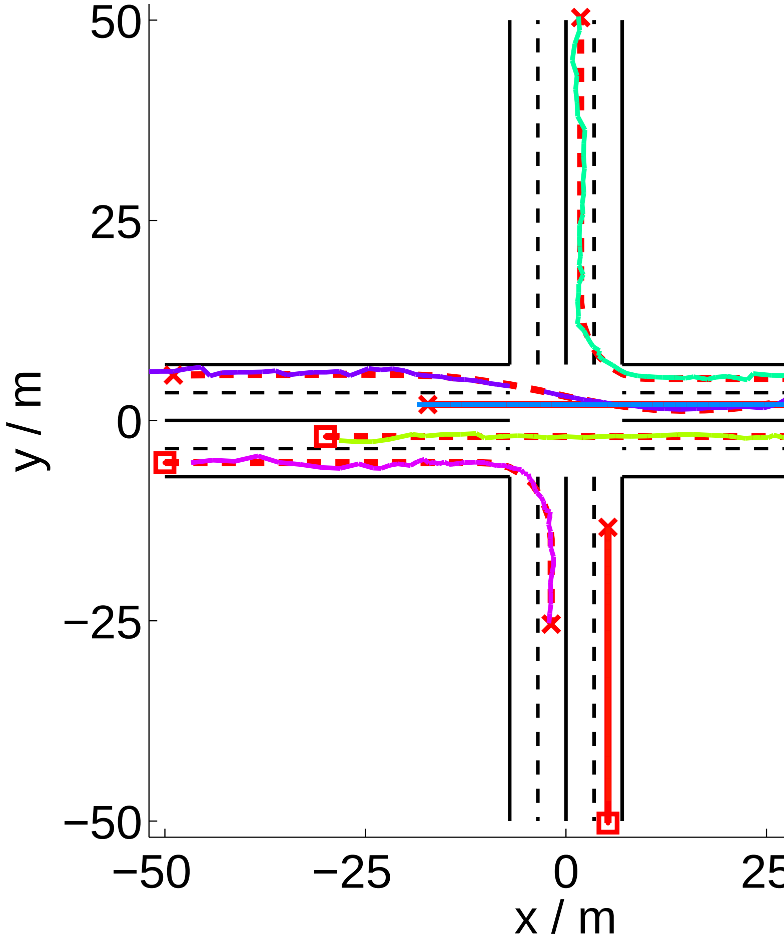


Fig. 8. Estimated trajectories after fusing all measurements of the simulation in Fig. 7 with a CBMeMBer filter.

in multi-hypothesis tracking, soft-assignments are conducted in probabilistic data association techniques. Recent methods based on finite set statistics totally avoid an explicit association step. Nevertheless, practical implementations of such methods, like the probability hypothesis density filter or joint integrated probabilistic data association filters, can be seen as extensions to classic tracking filters. Although this article has outlined many important trends in automotive information fusion, many details are left to the cited literature. Furthermore, automotive information fusion is a field of highly active research and as such possesses a dynamic state-of-the-art.

REFERENCES

[1] Special Issue on the 2007 DARPA Urban Challenge, Part I–III, Journal of Field Robotics 25 (8–10).

[2] S. E. Shladover, Cooperative (rather than autonomous) vehicle-highway automation systems, IEEE Intelligent Transportation Systems Magazine 1 (1) (2009) 10–19.

[3] Grand Cooperative Driving Challenge, http://www.gcdc.net/ [online] (2011).

[4] Special Issue on Intelligent Transportation Systems, Information Fusion 12 (1) (2011) 1–70.

[5] Z. Zomotor, U. Franke, Sensor fusion for improved vision based lane recognition and object tracking with range-finders, in: IEEE Intelligent Transportation Systems Conf., 1997, pp. 595–600.

[6] J. Dezert, Y. Bar-Shalom, Joint probabilistic data association for autonomous navigation, IEEE Trans. Aerospace and Electronic Systems 29 (4) (1993) 1275–1286.
[7] Y. Bar-Shalom, Multitarget-mulitisensor tracking: Applications and advances, 1st Edition, Vol. 1, Artech House Publishers, 1996.
[8] R. Mahler, Statistical Multisource-Multitarget Information Fusion, Artech House Publishers, 2007.
[9] M. Munz, M. Mählisch, K. Dietmayer, Generic centralized multi sensor data fusion based on probabilistic sensor and environment models for driver assistance systems, IEEE Intelligent Transportation Systems Magazine 2 (1) (2010) 6–17. doi:10.1109/MITS.2010.937293.
[10] B. Dasarathy, Sensor fusion potential exploitation—innovative architectures and illustrative applications, Proceedings of IEEE 85 (1997) 24–38.
[11] H. Ruser, F. Puente León, Informationsfusion – eine übersicht, Technisches Messen 74 (3) (2007) 93–102.
[12] B. van Arem, C. van Driel, R. Visser, The impact of cooperative adaptive cruise control on traffic-flow characteristics, IEEE Trans. Intelligent Transportation Systems 7 (4) (2006) 429–436. doi:10.1109/TITS.2006.884615.
[13] K. Tischler, B. Hummel, Enhanced environmental perception by intervehicle data exchange, in: Proc. IEEE Intelligent Vehicles Symposium, 2005, pp. 313–318.
[14] C. Stiller, G. Färber, S. Kammel, Cooperative cognitive automobiles, in: Proc. IEEE Intelligent Vehicles Symposium, Istanbul, Turkey, 2007, pp. 215–220.
[15] G. Mitropoulos, I. Karanasiou, A. Hinsberger, F. Aguado-Agelet, H. Wieker, H.-J. Hilt, S. Mammar, G. Noecker, Wireless local danger warning: Cooperative foresighted driving using intervehicle communication, IEEE Trans. Intelligent Transportation Systems 11 (3) (2010) 539–553.
[16] N. Kämpchen, M. Clauss, Y. Guenter, R. M. Schreier, M. Stiegeler, K. Tischler, K. Dietmayer, H. P. Grossmann, H. Kabza, H. Neumann, A. L. Rothermel, C. Stiller, Vernetzte Fahrzeug-Umfelderfassung für zukünftige Fahrerassistenzsysteme, in: M. Maurer, C. Stiller (Eds.), Proc. Workshop Fahrerassistenzsysteme, Freundeskreis Mess- und Regelungstechnik Karlsruhe e.V., Walting, Altmühltal, 2005, pp. 139–150.
[17] G. Toulminet, J. Boussuge, C. Laurgeau, Comparative synthesis of the 3 main European projects dealing with cooperative systems (CVIS, SAFESPOT and COOPERS) and description of COOPERS demonstration site 4, in: 11th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2008, pp. 809–814. doi:10.1109/ITSC.2008.4732652.
[18] P. Lytrivis, G. Thomaidis, I. Karaseitanidis, A. Amditis, Situation refinement for in-vehicle platforms in vehicular networks, in: 13th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2010, pp. 204–209. doi:10.1109/ITSC.2010.5625007.
[19] B. Roessler, K. Fuerstenberg, First European STREP on cooperative intersection safety, INTERSAFE-2, in: 13th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2010, pp. 422–427. doi:10.1109/ITSC.2010.5625101.
[20] M. Thuy, F. Puente León, Lane detection and tracking based on lidar data, Metrology & Measurement Systems XVII 3 (2010) 311–322.
[21] C. Lundquist, T. B. Schön, Joint ego-motion and road geometry estimation, Information Fusion (2010) doi:10.1016/j.inffus.2010.06.007.
[22] A. Sathyanarayana, P. Boyraz, J. H. Hansen, Information fusion for robust 'context and driver aware' active vehicle safety systems, Information Fusion (2010) doi:10.1016/j.inffus.2010.06.004.
[23] A. Ndjeng Ndjeng, D. Gruyer, S. Glaser, A. Lambert, Low cost IMU-odometer-GPS ego localization for unusual maneuvers, Information Fusion (2010) doi:10.1016/j.inffus.2010.06.006.
[24] A. Pérez Grassi, V. Frolov, F. Puente León, Information fusion to detect and classify pedestrians using invariant features, Information Fusion (2010) doi:10.1016/j.inffus.2010.06.002.
[25] A. Doucet, N. De Freitas, N. Gordon, Sequential Monte Carlo methods in practice, Springer-Verlag, 2001.
[26] B. Ristic, S. Arulampalam, N. Gordon, Beyond the Kalman Filter: Particle Filters for Tracking Applications, Artech House Publishers, 2004.
[27] H. Eberhardt, V. Klumpp, U. D. Hanebeck, Density trees for efficient nonlinear state estimation, in: Proceedings of the 13th International Conference on Information Fusion (Fusion 2010), Edinburgh, United Kingdom, 2010.
[28] V. Klumpp, U. D. Hanebeck, Dirac mixture trees for fast suboptimal multi-dimensional density approximation, in: Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008), Seoul, Republic of Korea, 2008. doi:10.1109/MFI.2008.4648009.
[29] R. E. Kalman, A new approach to linear filtering and prediction problems, Transactions of the ASME Journal of Basic Engineering 82 (1960) 35–45.
[30] X. Lin, Y. B. Shalom, T. Kirubarajan, Multisensor-multitarget bias estimation for general asynchronous sensors, IEEE Trans. Aerospace and Electronic Systems 41 (3) (2005) 899–921.
[31] S. Wu, M. Bansal, J. Eledath, Pedestrian localization by appearance matching and multi-mode filtering, in: IEEE Intelligent Vehicles Symposium, 2009, pp. 172–178. doi:10.1109/IVS.2009.5164273.
[32] X. Rong Li, V. P. Jilkov, Survey of maneuvering target tracking. Part V. Multiple-model methods, IEEE Trans. Aerospace and Electronic Systems 41 (4) (2005) 1255–1321. doi:10.1109/TAES.2005.1561886.
[33] S. J. Davey, M. G. Rutten, B. Cheung, A comparison of detection performance for several track-before-detect algorithms, in: 11th International Conference on Information Fusion, 2008, pp. 1–8. doi:10.1109/ICIF.2008.4632251.
[34] B.-N. Vo, B.-T. Vo, N.-T. Pham, D. Suter, Joint detection and estimation of multiple objects from image observations, IEEE Trans. Signal Processing 58 (10) (2010) 5129–5141. doi:10.1109/TSP.2010.2050482.
[35] D. J. Daley, D. Vere-Jones, An introduction to the theory of point processes I: Elementary theory and methods, 2nd Edition, Springer-Verlag, 2004.
[36] D. J. Daley, D. Vere-Jones, An introduction to the theory of point processes II: General theory and structure, 2nd Edition, Springer-Verlag, 2008.
[37] R. W. Sittler, An optimal data association problem in surveillance theory, IEEE Trans. Military Electronics 8 (2) (1964) 125–139. doi:10.1109/TME.1964.4323129.
[38] D. B. Reid, An algorithm for tracking multiple targets, IEEE Trans. Automatic Control 24 (6) (1979) 843–854.
[39] S. S. Blackman, Multiple-Target Tracking with Radar Applications, Artech House Publishers, 1986.
[40] I. Cox, S. Hingorani, An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking, IEEE Trans. Pattern Analysis and Machine Intelligence 18 (2) (1996) 138–150. doi:10.1109/34.481539.
[41] S. Blackman, R. Popoli, Design and Analysis of Modern Tracking Systems, Artech House Publishers, 1999.
[42] W. Koch, Fixed-interval retrodiction approach to Bayesian IMM-MHT for maneuvering multiple targets, IEEE Trans. Aerospace and Electronic Systems 36 (1) (2000) 2–14. doi:10.1109/7.826308.
[43] S. S. Blackman, Multiple hypothesis tracking for multiple target tracking, IEEE Aerospace and Electronic Systems Magazine 19 (1) (2004) 5–18. doi:10.1109/MAES.2004.1263228.
[44] Y. Bar-Shalom, Dimensionless score function for multiple hypothesis tracking, IEEE Trans. Aerospace and Electronic Systems 43 (1) (2007) 392–400. doi:10.1109/TAES.2007.357141.
[45] Y. Bar-Shalom, E. Tse, Tracking in a cluttered environment with probabilistic data association, Automatica 11 (5) (1975) 451–460. doi:10.1016/0005-1098(75)90021-7.
[46] T. Fortmann, Y. Bar-Shalom, M. Scheffe, Sonar tracking of multiple targets using joint probabilistic data association, IEEE Journal of Oceanic Engineering 8 (3) (1983) 173–184. doi:10.1109/JOE.1983.1145560.
[47] D. Musicki, R. Evans, Joint integrated probabilistic data association: JIPDA, IEEE Trans. Aerospace and Electronic Systems 40 (3) (2004) 1093–1099. doi:10.1109/TAES.2004.1337482.
[48] C. Hoffmann, T. Dang, Cheap joint probabilistic data association filters in an interacting multiple model design, Robotics and Autonomous Systems 57 (3) (2009) 268–278.
[49] S. Challa, B.-N. Vo, X. Wang, Bayesian approaches to track existence – IPDA and random sets, in: Proceedings of the Fifth International Conference on Information Fusion, Vol. 2, Annapolis, MD, 2002, pp. 1228–1235. doi:10.1109/ICIF.2002.1020953.
[50] M. R. Morelande, S. Challa, A multitarget tracking algorithm based on random sets, in: Proceedings of the Sixth International Conference on Information Fusion, Vol. 2, Cairns, Queensland, 2003, pp. 807–814. doi:10.1109/ICIF.2003.177322.
[51] R. Chakravorty, S. Challa, Multitarget tracking algorithm – Joint IPDA and Gaussian mixture PHD filter, in: 12th International Conference on Information Fusion, Seattle, WA, 2009, pp. 316–323.
[52] K. Panta, D. E. Clark, B.-N. Vo, Data association and track management for the Gaussian mixture probability hypothesis density filter, IEEE Trans. Aerospace and Electronic Systems 45 (3) (2009) 1003–1016.
[53] L. Lin, Y. Bar-Shalom, T. Kirubarajan, Track labeling and PHD filter for multitarget tracking, IEEE Trans. Aerospace and Electronic Systems 42 (3) (2006) 778–795. doi:10.1109/TAES.2006.248213.

[54] K. Panta, Multi-target tracking using 1st moment of random finite sets, Ph.D. thesis, University of Melbourne (2007).
[55] R. P. Mahler, Multitarget Bayes filtering via first-order multitarget moments, IEEE Trans. Aerospace and Electronic Systems 39 (4) (2003) 1152–1178. doi:10.1109/TAES.2003.1261119.
[56] T. Zajic, R. P. Mahler, Particle-systems implementation of the PHD multitarget-tracking filter, in: Proceedings of SPIE, Vol. 5096, 2003, p. 291. doi:10.1117/12.488533.
[57] B.-N. Vo, W.-K. Ma, The Gaussian mixture probability hypothesis density filter, IEEE Trans. Signal Processing 54 (11) (2006) 4091–4104.
[58] E. Maggio, M. Taj, A. Cavallaro, Efficient multitarget visual tracking using random finite sets, IEEE Trans. Circuits and Systems for Video Technology 18 (8) (2008) 1016–1027. doi:10.1109/TCSVT.2008.928221.
[59] H. Sidenbladh, Multi-target particle filtering for the probability hypothesis density, in: Proceedings of the Sixth International Conference of Information Fusion, Vol. 2, 2003, pp. 800–806. doi:10.1109/ICIF.2003.177321.
[60] B. Balakwmar, A. Sinha, T. Kirubarajan, J. Reilly, PHD filtering for tracking an unknown number of sources using an array of sensors, in: IEEE/SP 13th Workshop on Statistical Signal Processing, 2005, pp. 43–48. doi:10.1109/SSP.2005.1628561.
[61] M. Kruse, F. Puente León, Mehrobjekt-Verfolgung mit dem PHD-Filter, Technisches Messen 78 (4) (2011) 190–195.
[62] B.-T. Vo, B.-N. Vo, A. Cantoni, The cardinality balanced multi-target multi-Bernoulli filter and its implementations, IEEE Trans. Signal Processing 57 (2) (2009) 409–423.