Generalized Fusion of Heterogeneous Sensor Measurements for Multi Target Tracking

Michael Munz, Klaus Dietmayer
Institute of Measurement, Control, and Microtechnology
University of Ulm, Germany
michael.munz@uni-ulm.de

Mirko Mählisch
Department Environment Perception (GR/PAP)
Daimler AG
Ulm, Germany

Abstract – In this contribution, a multi sensor fusion system for multi target tracking is proposed which is independent of the properties of the attached sensor. This system therefore supports attaching different sensors without the need for any adaption of the system. A previously published approach for such a sensor independent fusion system based on probabilistic models is examined. An extension of the fusion algorithm is presented, which greatly enhances both the state and existence estimation.

Keywords: Multi sensor fusion, multi target tracking, Dempster Shafer Theory

1 Introduction

In many works, sensor data fusion has been addressed using a large set of different techniques which are optimized for the sensor properties used in the fusion system. However, in many areas, a fusion system which supports different types of sensors, like radar, laser or image based sensors without adapting the fusion algorithm is desired. Such a reusable fusion system would lead to a high reduction of costs and time to market. In [1] and [2], a sensor independent fusion system is proposed allowing an attached sensor to be replaced by a different sensor without the need for any modification in the fusion system. This can be achieved by representing sensor specific properties using probabilistic models. A common and powerful tracking algorithm based on the Joint Integrated Data Association (JIPDA) [3] is used for the processing of the sensor data.

This paper is organized as follows. First, the basics of the used JIPDA algorithm and modeling of sensor specific properties are briefly reviewed. Then state and existence estimation of objects is shortly introduced. Afterwards, some limitations for the fusion process of heterogeneous sensor data like range measurements and image detections are analyzed. Finally, a powerful and adaptive solution as an extension to the JIPDA algorithm is proposed and examined.

2 Generalized fusion system

2.1 Overview

In this section, the basics of the sensor independent fusion framework of [1], [2] are briefly reviewed. The proposed system has a feature-based central level fusion architecture. Several independent sensor modules deliver extracted feature vectors to the fusion system, which will be called measurements in the following. Each sensor has to provide precise timestamps along with these measurements. The interface of each sensor to the fusion module consists of measurement vector $z$, corresponding spatial measurement uncertainties $R$ and the sensory existence measurements $p_S$ defined below. All tracks in the system at time step $k$ are represented by a state estimation $N(x, \hat{x}, P)_{k|k}$ modeled as a spatial Gaussian distribution with mean $\hat{x}$ and covariance $P$. In addition to this, a probability of existence $p(\exists x)_{k|k}$ is assigned to each track. The existence of a track is defined as the event that the track corresponds to a real object which is currently relevant and observable, for example a vehicle.

The fusion system uses the Joint Integrated Probabilistic Data Association (JIPDA) [3] algorithm which simultaneously calculates the track state update and an estimate of the track existence probability $p(\exists x)$ over time. These estimates are based only on the information provided by the sensor modules, allowing to integrate any sensory information in a sensor independent and probabilistic way. Any optimization for the sensor setup can be done in the sensor specific part of the models and is therefore transparent to the fusion kernel.

2.2 State and existence estimation

JIPDA is basically an algorithm for solving the probabilistic data association (PDA) problem [4]. This problem can be stated as follows: given a set of tracks $\{x_i\}, i = 1, ..., N$ and a set of measurements $\{z_j\}, j = 1, ..., M$, calculate the probabilities $\beta_{ij}$ that measurement $z_j$ was caused by track $x_i$. Simultaneously, JIPDA calculates the probability of track existence $p(\exists x)_{k|k}$...
given all measurements and all tracks in a joint way similar to the Joint Probabilistic Data Association (JPDA, [4]), but it also takes into account the possibility that a target is caused by a false positive measurement and that current measurements can be due to clutter.

JIPDA based tracking of multiple targets in automotive scenarios was already published in [5]. The proposed approach was extended for the requirements of a generic fusion system in [1], [2]. One essential advantage of these approaches over the original JIPDA algorithm in [3] is the integration of a sensory individual existence probability \( p_{\exists x}(z_j) \) for each measurement \( z_j \). This probability is obtained by feature-dependent detector statistics [5], providing the possibility to account for detectors which are able to assign a posterior probability \( p(\exists x|z) \) for the detection as a confidence measurement:

\[
p_{\exists x}(z_j) = p(\exists x|z). \tag{1}
\]

Additionally, each sensor has to provide its current probability of detection \( p_D(x_i) \). This probability can depend on the state of the observed object \( x_i \).

### 2.3 Limitations of the JIPDA existence estimation approach

The JIPDA-based data fusion and existence estimation approach requires that all sensors are able to give evidence about the following binary frame of hypotheses \( \Omega \):

\[
\Omega = \{\exists x, \not\exists x\}, \tag{2}
\]

where \( \exists x \) is defined to be the event that the object is a real and relevant object, which is observable for the sensor fusion system. Each sensor attached to the fusion system has to provide the posterior probability \( p(\exists x|z) \) for this event.

The presented approach works fine fusing sensors which have similar target detection properties in terms of the possibility to distinguish between existing and non-existing objects. Hence, they make evidence about the same binary frame of hypotheses.

However, when fusing data of heterogeneous sensors this does not always apply. For example, let’s assume a sensor combination consisting of a laser scanner and a video camera. For an example view of the sensor raw data and the resulting environmental model, see figure 1.

The video camera is using an image classifier for vehicles and the laser scanner module extracts line and box features out of the scan points. A good video classifier will assign a high posterior probability \( p(\exists x|z) \) to an object of the relevant class “vehicle” and a low value for any other objects.

A laser scanner raw data often cannot provide features which are good enough to reliably distinguish between vehicles and other rectangular objects. Thus, a laser scanner classifier will fail in much more cases than the video classifier, resulting in a low posterior probability. Applying the JIPDA based approach to the sensor data will lead to an apparent contradiction in existence estimation. This contradiction can even cause a chaotic fusion result as for example in Figure 2. In such cases, one cannot predict if the oscillating probability of existence will tend towards one or zero.

Besides the existence estimation, the state estimate will also be influenced by this effect as the JIPDA assignment weights depend on the posterior probability of the measurements. Therefore, assignment weights between laser scanner measurements and tracks are calculated too low and the hypothesis that the object does not exists will receive a higher value. In the extreme case, this can lead to the ignorance of all measurements of the laser scanner sensor due to the low posterior probabilities.

One possible solution is to use a high posterior proba-
Figure 1: Automotive tracking scenario. Top left: video detections (magenta) and projected laser segments (blue, green, red, yellow). Top right: bird-eye view of laser scanner data. Bottom left: tracking result as 3D boxes (green) projected onto video image. Bottom right: bird-eye view of the resulting environment model (green) and FOVs of the sensors (laser: blue, video: red).

ability $p_T(z_j)$ for the laser scanner in order to get high measurement assignment weights and neglect the existential evidence in the existence update. Although this approach shows good performance, it totally neglects existence measurements of the second sensor. Therefore, a sophisticated solution for integrating any existence evidence for all sensors is developed.

The reason for this problem lies in the representation of the frame of hypotheses $\Omega$. From the point of view of the video sensor, the world consists of two types of image regions: those which are representing a relevant object (e.g. a vehicle) and those which don’t. The latter could be regions of one or many other objects or simply empty space.

The view of the laser scanner is a different one. One has two types of point clouds or segments: those which belong to a relevant object and those which belong to any other object. But there is hardly any segment in empty space nor a segment caused by two or more objects. Therefore we have the following three different elementary object hypotheses:

- **N**: No object exists
- **O**: Any object, but not relevant
- **R**: A relevant object (e.g. a vehicle)

We now refer to a set of elementary hypotheses as a proposition, which means the disjunction of the elementary hypothesis contained in the proposition.

In this terminology, the video sensor can distinguish between the proposition \{N,O\} and \{R\} but not between \{N\} and \{O\}. Whereas the laser scanner module can separate \{N\} from \{O,R\} but only very poorly \{O\}
The proposition which the sensor can classify will be referred to as the frame of perception \( F \). It is clear that the frames of perception of both sensors are different but overlapping. In consequence, the fusion of the existence data of both sensors cannot be done using the common JIPDA algorithm because JIPDA assumes that all sensors have the same binary frame of perception. Furthermore, the case that a sensor is not able to give any existence measurement cannot be handled with the current method. The binary existence modeling is not appropriate in these cases and has to be extended to multiple elementary hypotheses.

In this contribution, a novel JIPDA based data fusion approach is therefore proposed in order to solve these problems. One solution of the above problems is a different modeling of the frame of hypotheses which allows to assign probability mass over sets of hypotheses instead of the elementary hypotheses. The processing of such kind of evidence can be done using the Dempster Shafer Theory of Evidence (DST).

## 3 Dempster Shafer based JIPDA

### 3.1 Basics of the Dempster Shafer Theory of Evidence

In this section the basic elements of the Dempster Shafer Theory of Evidence (DST) [6] are briefly introduced. More details about DST are given in many publications, for example [7] and [8].

The Dempster Shafer Theory of Evidence (DST) is a generalization of the probability theory. All calculations can be applied to probability functions as well. The frame of discernment \( \Omega \) is defined as the set of elementary hypotheses \( a_i \):

\[
\Omega = \{a_i\}, \quad i = 1, ..., n. \tag{3}
\]

A basic belief assignment (BBA) \( m \) is a mapping from the power set \( 2^\Omega \) of the frame of discernment to the interval \([0, 1]\) with the following properties:

\[
m(\emptyset) = 0, \tag{4}
\]

\[
\sum_{A \subseteq \Omega} m(A) = 1. \tag{5}
\]

The mass \( m(A) \) is therefore the certainty which one can assign to the proposition \( A \). This allows us to assign probabilities to propositions which are unions of elementary hypotheses \(^1\).

All subsets \( A \) of \( \Omega \) with \( m(A) > 0 \) are called focal elements. If all focal elements of a BBA are elementary events, this BBA will be denoted as a Bayesian BBA. This Bayesian BBA is a probability function. If a Bayesian BBA has only two focal elements it will be called a binary BBA.

Two BBAs \( m_1 \) and \( m_2 \) can be combined using the Dempster Rule of combination as follows:

\[
m_1 \oplus_2 m_2(A) = \frac{\sum_{X \cap Y = A} m_1(X)m_2(Y)}{1 - \sum_{X \cap Y = \emptyset} m_1(X)m_2(Y)} \quad \forall A \in \Omega \tag{7}
\]

In the case of binary BBAs one can show that the above Dempster Rule of combination is equivalent to the Bayes rule [9] which shows the probability generalization nature of the DST.

The degree of belief of a BBA \( m \) for a proposition \( A \) is defined as

\[
Bel_m(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B). \tag{8}
\]

The degree of plausibility is defined as

\[
Pl_m(A) = \sum_{B \cap A \neq \emptyset} m(B). \tag{9}
\]

The following equation holds:

\[
Pl_m(A) = 1 - Bel_m(2^\Omega \setminus A). \tag{10}
\]

The plausibility \( Pl_m(A) \) is therefore the sum of all probability mass assigned to propositions which are not contradicting \( A \). The interval \( U(A) = Pl_m(A) - Bel_m(A) \) will be referred to as uncertainty interval.

Belief, Plausibility and uncertainty interval can be visualized as shown in Figure 3.

For a binary BBA \( m \) with \( m(A) = p \) the following holds:

\[
Bel_m(A) = Pl_m(A) = p. \tag{11}
\]

This is reasonable because the probability theory claims that there is no uncertainty about probability assignments.

In order to make a decision based on BBAs, the pignistic transformation [8] can be used:

\[
BetP_m(A) = \sum_{B \subseteq \Omega} \frac{|A \cap B|}{|B|} m(B). \tag{12}
\]

where \( |A| \) and \( |A \cap B| \) denotes the number of elementary hypotheses of \( \Omega \) in \( A \) and \( A \cap B \), respectively.

---

\(^1\)Although it is not essential for the DST that a BBA is based on probability assignments, probabilities can be used here.
This transformation converts the BBA into a probability function, where the mass assigned to a proposition $A$ will be distributed equally on the elements of $A$. This is reasonable because we do not have any additional information about the distribution of the probability mass to the subsets of $A$. Thereby, the principle of maximal entropy is used.

In general, $BetP_m(A)$ is bounded by $Pl_m(A)$ and $Bel_m(A)$:

$$Bel_m(A) \leq BetP_m(A) \leq Pl_m(A)$$  \hspace{1cm} (13)

For a binary BBA $m$ with $m(A) = p$ the pignistic transformation leads to:

$$BetP_m(A) = Bel_m(A) = Pl_m(A) = p$$  \hspace{1cm} (14)

This again shows that the DST is a generalization of the probability theory.

If we can only assure a certain probability of correctness $\alpha$ of a BBA $m$ we can discount this BBA prior to combine it with another BBA. The discounted BBA $m^\alpha$ is defined as follows:

$$m^\alpha(A) = \begin{cases} \alpha m(A), & A \neq \Omega \\ 1 - \alpha + \alpha m(\Omega), & A = \Omega \end{cases}$$  \hspace{1cm} (15)

This transfers the uncertainty about the correctness to the whole frame of discernment.

### 3.2 Modeling of existence probability with unified frame of perception

In order to be able to model the different frames of perception of several sensors, an extended frame of discernment has to be defined, where each sensor can make a decision about its own frame of perception $F$.

This leads to the following set of propositions:

$$2^\Omega = \{\emptyset, N, O, NO, R, NR, OR, NOR\}.$$  \hspace{1cm} (16)

Each of the propositions of $2^\Omega$ can now be assigned a probability. In the above two sensor example the video camera module would assign the following probabilities to the BBA based on the measurement $z$:

$$m(NO) = p_{FP}(z)$$  \hspace{1cm} (17)

$$m(R) = p_{TP}(z) = p_{\text{vehicle}}(z)$$  \hspace{1cm} (18)

with the frame of perception $F_C = \{R\}$.

The BBA of the laser scanner module will be the following:

$$m(N) = p_{FP}(z)$$  \hspace{1cm} (19)

$$m(OR) = 1 - p_{FP}(z) - p_{\text{vehicle}}(z)$$  \hspace{1cm} (20)

$$m(N) = p_{TP}(z) - p_{\text{vehicle}}(z)$$  \hspace{1cm} (21)

$$m(R) = p_{\text{vehicle}}(z)$$  \hspace{1cm} (22)

and the corresponding frame of perception $F_L = \{OR\}$. The belief functions of these BBAs are visualized in Figure 4.

![Figure 4: Example Beliefs of laser scanner (top) and camera (bottom) sensors.](image)

### 3.3 Modifications for the JIPDA algorithm

For combining JIPDA with DST, a novel method has been developed. In this JIPDA-DS approach, the existence estimation of tracks is now based on BBAs, which will replace the probability of existence $p(\exists x)$. During object tracking, the BBA of a track will therefore be continuously predicted and innovated.

In the JIPDA-based calculations of data association weights $\beta_{ij}$ and $p(\exists x)_{ijb}$, only the frame of perception $F$ of the current sensor has to be considered. Therefore, all BBAs can be reduced onto $F$, which will lead to a binary probability measurement of $F$. This can be done using the pignistic transformation (12).

The existence estimation in the proposed fusion system is based on two different levels which are often used in DST: the credal level (lat. credo: i belief) using BBAs and the pignistic level (lat. pignus: bet) where only binary probability measurements exist. The transition...
into the pignistic level will only be performed if a temporary decision is required, as in the data association step.

Applying the pignistic transformation corresponds to reducing the problem on a sub-problem defined by $F$ for the JIPDA data association step. Because all JIPDA algorithm formulas are based on probability functions, the JIPDA algorithm can be applied with the prior probability of existence estimations in a weighted way. The mass calculated by JIPDA on the elements of $m_k$ for existence estimations has been developed in which version of the pignistic transformation which cannot be probabilities of existence have to be transformed back onto

$$p^F(\exists x_i)_{k|k-1} = BetP_{m_i}(F),$$

(23)

and the corresponding sensor existence measurement $j$

$$p^F(\exists x_j) = BetP_{m_{S,i}}(F).$$

(24)

After the JIPDA processing step, the posterior probabilities of existence have to be transformed back onto the credal level. This process corresponds to the inversion of the pignistic transformation which cannot be solved in a unique way [10]. Therefore, a new method for existence estimations has been developed in which a posterior BBA is calculated using the prior BBA and the sensorial BBAs by redistributing the probability mass calculated by JIPDA on the elements of $F$.

All $m$ measurements must be incorporated into the posterior existence estimation in a weighted way. The JIPDA association probability $\beta_{ij}$ denotes the probability that measurement $j$ was originated by track $i$. Therefore, the unified sensor BBA $m_{i,S}$ can be calculated by weighting the sensor BBAs using the association weights $\beta_{ij}$. This is done using (15). In addition to the BBAs of the measurements, each sensor delivers a BBA $m_{i,FN}$ for the case of missed detection of track $i$ which is discounted using $\beta_{i0}$.

The unified sensorial BBA $m_{i,S}$ is therefore calculated using:

$$m_{i,S} = m_{i,FN}^{\beta_{i0}} \oplus m_{i,S,1}^{\beta_{i1}} \oplus \cdots \oplus m_{i,S,m}^{\beta_{im}}$$

(25)

Then the prior BBA $m_{i,k|k-1}$ is combined with $m_{i,S}$ using the Dempster rule of combination (7) yielding the fused posterior BBA:

$$\tilde{m}_{i,k|k} = m_{i,k|k-1} \oplus m_{i,S}$$

(26)

This incorporates the sensorial knowledge about the distribution of the probability mass over the elementary hypotheses $N$, $O$ and $R$. The posterior probability of existence $p^F(\exists x_i)_{k|k}$ calculated by JIPDA must be equal to the bet probability $BetP_{\tilde{m}_{i,k|k}}(F)$ of the posterior BBA. Therefore, the probability mass calculated by (26) has to be redistributed.

In the following, the set of all elementary hypotheses which support a certain hypothesis $H$ will be referred to as $Sup_H$, the remaining symbols $2^O \setminus Sup_H$ will be named as $Con_H$. For $B \in Sup_H$ the following holds: $B \cap H \neq \emptyset$, for $B \in Con_H$ analogous: $B \cap H = \emptyset$.

We now define a belief transfer function $T_{\gamma}(H,m)$:

$$T_{\gamma}(H,m) = \left\{ \begin{array}{ll}
\gamma m(X) & \text{if } X \in Sup_H \\
m(X) + \sum_{Y \in Sup_H} m(Y)(1-\gamma) & \text{otherwise.}
\end{array} \right.$$  

(27)

$T_{\gamma}(H,m)$ therefore transfers part of the probability mass of an elementary hypothesis $A \in Con_H$ to $Sup_H$ (or in the opposite direction) according to $\gamma$.

In the following, we will abbreviate $p_{i}(F) := BetP_{\tilde{m}_{i,k|k}}(F)$. When applying $T_{\gamma}(F,m)$ on the fused posterior BBA $\tilde{m}_{i,k|k}$ (26), we gain the fused posterior track BBA $m^T_{i,k|k} = T_{\gamma}(Con_F, \tilde{m}_{i,k|k})$.

Two possible cases have to be separated:

1. $p^F(\exists x_i)_{k|k} > p_{i}(F)$:

   Probability mass has to be transferred from all elementary hypothesis of $Con_F$ to $Sup_F$.

   $$\gamma = \frac{1 - p^F(\exists x_i)_{k|k}}{1 - p_{i}(F)}.$$  

(28)

2. $p^F(\exists x_i)_{k|k} < p_{i}(F)$:

   Probability mass has to be transferred from all elementary hypothesis of $Sup_F$ to $Con_F$.

   $$\gamma = \frac{p^F(\exists x_i)_{k|k}}{p_{i}(F)}.$$  

(29)

This leads to a new BBA $m^T_{i,k|k}$ with $BetP_{m^T_{i,k|k}}(F) = p^F(\exists x_i)_{k|k}$. This BBA is used as the posterior existence BBA of the track $i$ in the next iteration.

If all BBAs are binary ones, the resulting posterior existence probabilities using this new method are identical with those achieved when using the approach in [2] because of (14). Therefore, the presented novel approach is a real generalization of JIPDA-based existence estimation. This allows us to use any already presented extension to JIPDA with this new Dempster Shafer based approach.

4 Results with real-world data

The proposed fusion system has been tested on sensor data from a research vehicle using a video camera and a laser scanner in urban scenarios. The JIPDA-DS fusion approach performs very good on real-world data. The system is able to track multiple objects in real-time and can separate very well between relevant and any other type of object $O$ if the necessary evidence is provided by the sensors. Examples for the BBAs of one real and one false track over time are shown in Figure 5. Both cases show that the uncertainty interval $U$ given by Plausibility and Belief is decreasing over time. In Figure 6 we can see other situations. The upper plot shows a track which first is only detected by the laser scanner. It is outside the field of view of the camera. The uncertainty interval is nearly
one which denotes the completely missing knowledge whether the track is a relevant object or not. After approximately 50 iterations, the target is in the field of view of the video camera and therefore the necessary additional existence information to distinguish between \( O \) and \( R \) is provided.

The lower plot shows a track which can only be detected by the laser scanner. Therefore, the probability of existence \( \text{BetP}(R) \) won’t exceed 0.5. Between time steps 35 and 50, the target was missed by the laser scanner a few times and therefore the probability of existence decreases in this time period and increases again afterwards.

The novel JIPDA-DS allows us to make use of what is often called negative evidence. This can be exploited if one sensor opposes a part \( A \) of a hypothesis \( H = \{AB\} \). The allocated mass is therefore transferred to its sub-hypotheses \( B \) due to the lack of positive evidence for \( A \). For example, prior mass is allocated to the hypothesis \( OR \). A sensor claims that the object cannot be of type \( R \) maybe due to the lack of any detection at a high probability of detection \( p_D \). The fusion process therefore leads to a transfer of probability mass from \( OR \) to \( O \).

5 Conclusion and Further Works

The presented approach of combining JIPDA for existence estimation with the DST has several novel and significant advantages over state of the art fusion systems. First of all, any kind of sensor can be incorporated in the fusion system regardless of what kind of object it can observe or classify. Using the presented JIPDA-DS algorithm, we can incorporate existence measurements of sensors which cannot give evidence about the same frame of perception. Furthermore, even sensor measurements without any information about existence can be fused by using an uncertainty interval of one. In addition to this, JIPDA-
DS provides an environment model which incorporates knowledge about several different kind of objects at the same time and can therefore solve a multi class tracking problem. Classification, existence estimation, data association and tracking is coupled via the JIPDA algorithm and can be implemented in a generic and sensor independent way. This leads us to a fusion system which can be modified by simply attaching a different sensor setup to the system according to the application requirements.

**Acknowledgment**

This work was supported by Daimler AG, Group Research & Advanced Engineering, Ulm, Germany.

**References**


