



**HAL**  
open science

# Traffic Sign Recognition: Benchmark of Credal Object Association Algorithms

Jean-Philippe Lauffenburger, Jérémie Daniel, Mohammed Boumediene

► **To cite this version:**

Jean-Philippe Lauffenburger, Jérémie Daniel, Mohammed Boumediene. Traffic Sign Recognition: Benchmark of Credal Object Association Algorithms. 17th International Conference on Information Fusion (FUSION), 2014, ISIF/IEEE, Jul 2014, Salamanca, Spain. pp.1-7. hal-01123466

**HAL Id: hal-01123466**

**<https://hal.science/hal-01123466>**

Submitted on 4 Mar 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Traffic Sign Recognition: Benchmark of Credal Object Association Algorithms

Jean-Philippe Lauffenburger, Jérémie Daniel  
Laboratoire MIPS-EA2332  
Université de Haute Alsace  
12 rue des frères Lumière  
68093 Mulhouse Cedex, France  
Email: firstname.lastname@uha.fr

Mohammed Boumediene  
Laboratoire LSI  
USTOMB  
Oran, Algérie  
Email: boumediene.m@gmail.com

**Abstract**—Static and dynamic objects detection and tracking is a classic but still open problem in Intelligent Transportation Systems. Initially formalized in the Bayesian framework, new methods using belief functions have recently emerged. Most of them have been essentially validated in simulations. This paper proposes an association and tracking framework devoted to Traffic Sign Recognition in a mono-sensor context. Potential signs are detected in the camera images. A credal association between new observations and already known objects is performed. Associated objects are tracked over time and in the image space using Kalman Filtering. This global tracking system has been used to evaluate in real-time on large datasets several state-of-the-art credal association methods. The main evaluation criteria is their capability to reduce false detections by keeping a high traffic sign detection rate.

## I. INTRODUCTION

Multi-target tracking (MTT) is a classic data processing step in perception-oriented applications. It requires several connected tasks in which the object association and tracking are probably the most important. Basically, association consists in finding the correlation between two sets, the one gathering the newly perceived objects and the other the known ones. The observation of these objects is usually done by sensors providing different features (number of objects in the scene, kinematic data, etc.). In a multisensor context, it consists in a spatial fusion by defining relationships between objects observed by different sensors. In this paper, a temporal fusion is proposed in which the association is performed over objects captured by a single sensor at successive times.

Initially, most of the data association methods were based on the Bayesian theory such as the seminal work of Reid [1], Bar-Shalom [2] or Blackman and Popoli [3]. Drawbacks of these methods are essentially that they require *a priori* sensor error models and have limited capabilities in managing association ambiguities or hypothesis rejection (appearance, disappearance).

To cope with these limitations, research studies have been focused on belief function-based MTT. The evidence theory [4], [5] and more precisely the Transferable Belief Model (TBM) [6] provide a well-suited formalization context. The TBM is a flexible framework in which the object appearance and disappearance can be straightforwardly managed while

taking account of sensors reliability and ignorance without any assumption on their error models. One of the critical step lies in the decision, i.e. the selection of the best associations between the objects of both sets. For data association with belief functions, pioneer works have been made by Rombaut [7] with the introduction of a combination rule between mass functions regarding the pairs of objects to be associated. Mourllion et al. [8] defined two decision rules after belief mass combination in order to solve hazardous pairwise local associations. More recently, Mercier et al. [9] proposed to match pairs of detected and known objects with respect to the maximum of combined pignistic probability. All of these studies encode attributes provided by sensors as pieces of evidence defined by a mass function supporting the association between a pair of objects [10], [11]. In [12], a review of the solutions from Mercier et al. [9] and Mourllion et al. [8] has been provided based on classical literature simulation configurations. To cope with suspicious local associations when evidence conflicts, the authors introduced assignment methods retaining pairwise associations with the highest pignistic probabilities.

The aim of this article is to extend the review paper from Daniel and Lauffenburger [12] by an evaluation of the already tested credal assignment methods on a real test case from the Intelligent Transportation Systems (ITS) context: Traffic Sign Recognition (TSR). Vision-based TSR systems are usually providing good results when they contain tracking capabilities. Indeed, instead of processing each frame independently, tracking allows to take temporal information into account, thus allows TSR systems to be more robust to temporal sign occlusions and false positives. The contribution of this paper is to define a complete framework within the belief function context for traffic sign detection and joint association and tracking to improve their recognition. By managing track appearance, disappearance and tracking, it proposes a flexible and efficient solution to deal with wrong detections, occlusions and false positives, thanks to the specific data modeling mechanisms of the TBM (ignorance, conflict and imperfections modeling). The result is a real-time compliant complete tracking-based TSR system providing high performance, similar to Bayesian approaches [13]. The TSR considered couples a detection al-

gorithm, defining location of traffic signs, and a tracking filter following them temporally and spatially. Pieces of evidence, i.e. mass functions, supporting potential associations are modeled based on the data provided by the image sequences. The tracker is composed of an association step determining, at each sample time, the best detected-to-known and known-to-detected object pairings using the pairwise mass functions defined by the camera measurement. The tracking is then performed considering a spatio-temporal fusion by the use of dynamically managed multiple Kalman Filters (KF). Experimental results show that simple assignment methods based on the selection of the highest pairwise pignistic probability and requiring less computation time allow good association and tracking results with respect to the achieved false positive reduction and the detection rate.

The remainder of the paper is as follows: a brief recall of belief functions is proposed in Section II. Section III describes the credal association and tracking framework, consisting in a vision-based Traffic Sign Recognition system, used to evaluate the object assignment algorithms. Experimental results are described in Section IV while Section V concludes the paper.

## II. THEORETICAL BACKGROUND

In this paper, the model of uncertainty retained is the TBM introduced in [6]. The latter is based on two levels: the credal level in which the information or expert opinions are represented by belief masses and the pignistic level where probability measures are determined for decision making.

### A. Evidence Modelling

In the theory of belief functions, the knowledge (provided either by sensors or by expert opinions) is formalized on a finite set, the *frame of discernment*  $\Theta = \bigcup_{j=1}^k \{H_j\}$  where  $\{H_j\}$  are exclusive and exhaustive hypotheses. This knowledge is represented by a *basic belief assignment (bba)* mass function  $m(A)$  with  $A \subseteq \Theta$ . The *bba* is a mapping of  $2^\Theta$  to  $[0,1]$  verifying:

$$\sum_{A \in 2^\Theta} m^\Theta(A) = 1. \quad (1)$$

$m^\Theta(A)$  is the degree of belief supporting a proposition  $A$  and which could not be committed to any subset of  $A$  because of insufficient information. This is a key difference with Bayesian approaches which only express probabilities over singletons.

From the *bba*, the credibility and plausibility of each proposition can be computed (cf.  $Cr()$  and  $Pl()$  in (2)). The credibility is a pessimistic measurement since it only includes the belief masses of propositions  $B$  such that  $B \subseteq A$  (cf. (2a)). Contrary to this, the plausibility is an optimistic measurement including the masses of all propositions  $B$  such that  $B \cap A \neq \emptyset$  (cf. (2b)).

$$Cr(A) = \sum_{B \subseteq A} m^\Theta(B) \quad (2a)$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m^\Theta(B) = 1 - Cr(\bar{A}) \quad (2b)$$

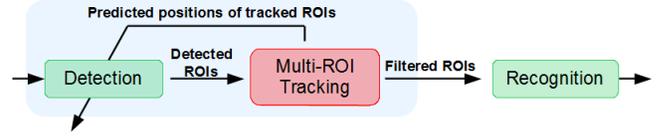


Fig. 1. Overview of the TSR system.

### B. Evidence Combination

The combination of the different *bba* can be done considering numerous operators [14]. Here, as sources are considered as reliable and independent, the conjunctive operator has been retained:

$$m_{\cap}^\Theta(A) = \sum_{A_1 \cap \dots \cap A_k = A} \prod_{i=1}^k m^\Theta(A_i) \quad (3)$$

Since it is based on intersection, when sources are discordant, a mass on the conflict  $m_{\cap}^\Theta(\emptyset)$  is generated.

### C. Pignistic Transformation

Decision making consists in the selection of the most relevant solution of  $\Theta$  with respect to the combination results. This is usually achieved through a transformation of belief masses into a probability measure so that masses on unions are transferred to singletons  $\theta \in \Theta$ . In the TBM framework, the most common mapping function is the pignistic probability [15] (cf.  $BetP$  in (4)). It is considered as a compromise between the credibility and plausibility as it distributes the mass of a set  $A$  over all its singletons  $H_j \in A$ , proportionally to the cardinality  $|A|$  of  $A$ .

$$BetP(H_j) = \sum_{\substack{A \in 2^\Theta \\ H_j \in A}} \frac{m(A)}{|A|(1 - m(\emptyset))} \quad (4)$$

The most appropriate solution among those of  $\Theta$  is then selected considering a decision criterion. A straightforward approach lies in the selection of the *BetP* maximum.

## III. CREDAL TSR

In a previous study [12], several decision techniques have been compared. The limitations of global methods and the advantage of local ones have been shown on literature examples. In this paper, the aim is to evaluate these approaches in the ITS context. A vision-based Traffic Sign Recognition system is introduced in this section and constitutes the evaluation framework of the association solutions.

### A. TSR Description

This monosensor system combines detection and tracking to perform a spatio-temporal tracking of traffic signs in order to be more robust to occlusions and false positives. It is based on the coupling of a Region Of Interests (ROIs) detection algorithm and a Mutli-ROIs tracker (cf. Figure 1). In a first step, the system uses a straightforward image processing approach [16] to detect potential areas containing

signs. These ROIs, considered as objects, are then tracked spatially in the image referential frame as well as temporally over the sequence. The spatial data fusion is performed using dynamically managed Kalman Filters [17]. The feedback of the position predictions defines search areas for the detector and enhances its performance. A confidence value, calculated for each tracked object, is finally used to select those which have to be used for traffic sign recognition.

The multiple target trajectories estimation is based on the spatial context, where a single target is identified by its state vector generally containing the target kinematics. The proposed Multi-ROIs Tracking framework is composed of two main functions: filtering and data association (see Figure 2). The aim of filtering is to estimate the present and future kinematic quantities of all tracked targets (ROIs) at each time  $k$  considering the new sensor observations. The data association assigns a correct observation to each track. It first requires, through the gating, to deal with improbable observation-to-track pairings to reduce the association problem complexity. For each track, using the predicted state and the predicted error covariance, a gate is defined around the predicted position. The association defines at time  $k$  the most probable relationships between the new observations and the tracks. This is usually done considering imperfect measurements (observations) in a highly dynamic context since the number of tracks and targets vary from time to time. Association contradictions and ambiguities can then occur and lead to tracking performance reduction. This particular point is treated based on state-of-the-art credibilistic association solutions described in [12] and briefly recalled in Section III-C. Finally, *Track Maintenance* analyzes observation-to-track associations to initiate new tracks, i.e. new KF, and/or delete existing ones. This is done through the track temporal analysis and the track trajectory length evolution over time. Indeed, road signs do not appear and disappear suddenly in images, thus are visible on several successive frames. Considering the track state vector, a confidence score in each track can be determined. The latter is used to define if the track is considered for recognition or not. Consequently, the consideration of temporal information helps to decrease the false positive rate of the TSR system while keeping a high detection rate.

### B. Credal Association and Tracking

Let us first consider the association step from Figure 2. We briefly recall hereafter the object association formalism in the belief functions context. Extensive descriptions can be found in [9], [12], [13]. Consider a set of objects  $X_i$  (detected ROIs) detected in the camera images from the front driving scene and a set of known objects  $Y_j$  (tracked ROIs) at a processing time  $k$  [12]. These sets are respectively denoted  $\Theta_{Y_j} = \{X_1, X_2, \dots, X_N, *\}$  and  $\Theta_{X_i} = \{Y_1, Y_2, \dots, Y_M, *\}$  with  $N$  the number of targets,  $1 \leq i \leq N$  and  $M$  the number of tracks,  $1 \leq j \leq M$ .  $*$  defines either the track appearance or disappearance. The belief level in a given association is described by a pairwise normalized mass function  $m_{i,j}$ , leading to  $N \times M$  belief functions when considering all assignment

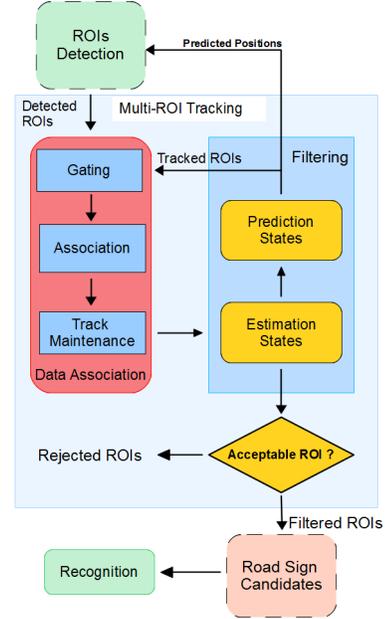


Fig. 2. Block-diagram of the TSR system.

possibilities. The belief function  $m_{i,j}$  encodes a piece of evidence supporting the association between  $X_i$  and  $Y_j$  defined over the frame  $\Theta_{i,j} = \{y, n\}$ .  $m_{i,j}(\{y\})$  represents the belief in the association,  $m_{i,j}(\{n\})$  in the non-association whereas  $m_{i,j}(\Theta_{i,j})$  quantifies the ignorance in the given association. Each mass function  $m_{i,j}$  can be expressed on a common frame  $\Theta_{X_i}$  (or  $\Theta_{Y_j}$ ) by transferring  $m_{i,j}(\{y\})$  to  $m_j(Y_j)$ ,  $m_{i,j}(\{n\})$  to  $m_j(\bar{Y}_j)$ <sup>1</sup> and  $m_{i,j}(\Theta_{i,j})$  to  $m_j(\Theta_{X_i})$  [9]. To gather all the knowledge about the associations, the  $N \times M$  belief functions are combined using the conjunctive rule (3) giving  $m_{\cap}^{\Theta_{X_i}}$  and  $m_{\cap}^{\Theta_{Y_j}}$ . Finally, the pignistic probabilities  $BetP_{i..}$ ,  $BetP_{..j}$  can be computed using (4) [12].

Formally, the association problem consists in defining the relationships between  $\Theta_{X_i}$  and  $\Theta_{Y_j}$  by considering the following assignment hypotheses at time  $k$ :

- only one target can be associated with a track and *vice versa*,
- several tracks can appear or disappear.

The final goal of the credal TSR presented in this article is to find the correct assignments between detected potential traffic signs  $X_i \in \Theta_{Y_j}$  and already known ones  $Y_j \in \Theta_{X_i}$  and to track them in the image sequence. For the track state estimation with noisy measurements, the well-known Kalman filter [17] is used here. At time  $k$ , the detector extracts from frame  $k$  a set  $\mathcal{Z}$  of sign candidates which represents the  $N$  targets  $X_i$ . The set  $\mathcal{X}$  contains the  $M$  tracks to be associated with the elements of  $\mathcal{Z}$ . For simplicity reasons, observations, i.e. detected objects, defined by a  $p$ -dimensional vector  $z_i(k) \in \mathcal{Z}$  with  $i = 1, 2, \dots, N$ , and the tracked target state vector  $x_j(k) \in \mathcal{X}$  with  $j = 1, 2, \dots, M$ , are

<sup>1</sup> $\bar{Y}_j$  describes the complementary hypothesis of  $Y_j$  such that  $\bar{Y}_j = \{Y_1, Y_2, \dots, Y_{j-1}, Y_{j+1}, \dots, Y_M\}$

considered linearly dependent. Also, assume a linear state vector evolution. The track state dynamic model and the observation model are respectively given by:

$$\begin{aligned} X_j(k) &= FX_j(k-1) + w(k) \\ Z_i(k) &= HX_j(k) + v(k), \end{aligned} \quad (5)$$

with:

- $X_j(k) \in \mathbb{R}^l$  is the  $j^{\text{th}}$  track state vector with  $j \in [1, 2, \dots, M]$  at time  $k$ ,
- $F \in \mathbb{R}^{l \times l}$  state matrix function of the motion model retained,
- $Z_i(k) \in \mathbb{R}^p$  is the  $i^{\text{th}}$  observations given by the detector with  $i \in [1, 2, \dots, N]$  at time  $k$ ,
- $H \in \mathbb{R}^{p \times l}$  is a time invariant measurement matrix,
- $w(k)$  and  $v(k)$  are respectively the state noise and observation noise defined as i.i.d processes and mutually uncorrelated.  $w(k)$  is a  $l \times 1$  Gaussian noise with zero mean and covariance  $Q$  such that  $w \sim \mathcal{N}_l(0, Q)$ .  $v(k)$  is a  $p$ -dimensional Gaussian noise defined by  $v \sim \mathcal{N}_p(0, R)$ .

The number of tracks  $M$  as well as observations  $N$  are not constant over time and are related to the association results<sup>2</sup>. Especially, observations provided at time  $k$  can be due to already tracked targets, new targets or false detections.  $Z_i(k)$  contains the position  $(x, y)$  and size  $s$  of observation  $i$  in the image frame  $k$ :

$$Z_i(k) = [x, y, s]^T. \quad (6)$$

A track  $j$  is represented by its state vector  $X_j(k) = [x, y, s, v_x, v_y, v_s]^T$  where  $[v_x, v_y, v_s]$  define their relative velocities between two successive frames and  $v_s$  is the scale change. Here, a nearly constant velocity model is assumed [13] with  $F$ ,  $H$ ,  $Q$  and  $R$  are taken such as:

$$F = \begin{pmatrix} I_3 & T_e \cdot I_3 \\ 0_{3 \times 3} & I_3 \end{pmatrix}, \quad (7)$$

with  $T_e = 1$  when every successive frame is processed.

$$H = \begin{pmatrix} I_3 & 0_{3 \times 3} \end{pmatrix} \quad (8)$$

$$Q = \begin{pmatrix} 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & \sigma \end{pmatrix} \quad (9)$$

with  $\sigma = \text{diag}(\sigma_x^2, \sigma_y^2, \sigma_s^2)$ .

$$R = \begin{pmatrix} \sigma_x & 0 & 0 \\ 0 & \sigma_y & 0 \\ 0 & 0 & \sigma_s \end{pmatrix} \quad (10)$$

### C. Association Methods

Beside the numerous probability-based multi-target tracking solutions (see Section I), several researches tackle this problem in the theoretical context of the belief functions. For the assignment of detected to known objects dealing with hypothesis rejection, we particularly focus on the investigations from [8]

<sup>2</sup>Given the new observations and according to the data association results at time  $k$ , the existing KFs corresponding to tracks already followed are updated and/or new filters are created for appearing tracks.

and [9] recently evaluated and extended in [12]. This section briefly recalls the decision rules and main features of these association methods. Readers interested in a deeper description may refer to [12] which provides a detailed presentation and illustration examples.

A common point of the approaches from [8] and [9] is that they are bi-directional, taking account of two points-of-view at time  $k$ : perceived to known objects (observation-to-track) and known to perceived (track-to-observation). They attempt to find suitable relations between  $\Theta_{Y_j}$  and  $\Theta_{X_i}$ . The similarity of these solutions resides in the decision measure used: the maximum of Pignistic Probability [6]. Mercier et al. [9], with the so-called *Joint (Classified) Pignistic Probability* (J(C)PP) methods, proposed to select the assignments globally satisfying the association problem, i.e. maximizing the normalized pignistic probability product. [12] showed that, due to the  $(N+1)^M$  solutions to be considered, the computation time increases quickly with the number of tracked objects. Moreover, it may lead to different results depending on the association direction. Mourllion et al. [8] proposed to consider the conflict during decision taking by using a non-normalized pignistic probability decision criterion. Moreover, they allowed non-associations or multiple associations between objects depending on the conflict level:

$$H = \arg(\text{BetP}(H_j) \geq \text{Threshold}), \quad (11)$$

with:

$$\text{Threshold} = k_s(1 - \text{BetP}(\emptyset)). \quad (12)$$

In [12], a simple example highlighted the incorrect associations of Mercier's JPP and JCPP approaches, and more generally the weaknesses of global cost optimization methods. To cope with these drawbacks, the authors proposed simple approaches (LPP and GCPP) consisting in keeping, for each object, the highest unnormalized pignistic probability associations.

After a theoretical comparison of these solutions in [12], the next section proposes to bench 3 credal assignment methods in the framework of the TSR system.

## IV. RESULTS

### A. Test Conditions

The tracking-based TSR depicted in Section III, including the JCPP association method from Mercier as well as the GCPP and LPP from Daniel and Lauffenburger, has been tested on real data sequences. The detected traffic signs (either triangular or circular) representing the set of perceived objects  $\Theta_{Y_j}$  were obtained through the principle from [16] extended to the circular sign detection by a template-matching process [18]. A 12-bit black and white Sensata Technologies mono-camera is used. In this context, a maximum of  $M = 10$  targets can be tracked by the system. The characteristics of the tracking filter have been set according to  $(\sigma_x, \sigma_y, \sigma_s) = (2, 2, 3)$ .

TABLE I  
TEST SEQUENCE FEATURES

	Features
Number of frames	12 500
Length (km)	9.3
Motorway (%)	26.3
City (%)	20.2
Number of signs	42
Number of circular signs	31
Number of triangular signs	11
Min Vehicle Speed (km/h)	0
Max Vehicle Speed (km/h)	110
Speed < 90 km/h (%)	20.9
Occlusions (partial/total)	14(13/1)
Lane change	8(3)



Fig. 3. Multiple detection case. When the signs are close to the camera, the template-matching algorithm can detect two circles (the inner and the outer circle) for the same sign leading finally to a set  $\Theta_{Y_j}$  of 2 detected objects.

The performance was evaluated on a sequence of 12 500 frames corresponding to various driving conditions (see Table I). 31 circular and 11 triangular signs appear in 1 452 frames with a non-neglectable number of occlusions.

The particularity of this application is that the detection concerns non-moving objects of the scene allowing the assumption of a steady-state vehicle motion w. r. t. to the objects [13]. Moreover, sign positions are by nature chosen to avoid any misdetection or identification. Theoretically, it should therefore be easy for TSR systems to detect them. Practically, troubles arise for instance when signs are close to the camera. This may lead to the situation presented in Figure 3, where two objects are detected for the same sign.

### B. Mass Functions Computation

Trying to match pairs of objects in a frame  $k$  first requires the definition of a similarity measure  $d_{ij}$  between both object sets  $\Theta_{X_i}$  and  $\Theta_{Y_j}$  and the appropriate mass functions  $m_{ij}$ . Basically, it is considered that when  $X_i$  corresponds to  $Y_j$ ,  $d_{ij}$  tends to zero (the objects are close in the referential space considered, i.e. the image frame). When  $d_{ij}$  is high, the objects are not corresponding. The interpretation on the mass functions is straightforward: a low value of the similarity measure supports the association between  $X_i$  and  $Y_j$  ( $m_{ij}(\{y\})$ ), whereas a large value supports its contrary ( $m_{ij}(\{n\})$ ). The

remaining ignorance in the association due for instance to sensor imperfections is transferred to  $\Theta_{i,j}$ . The pairwise mass functions  $m_{ij}$  are therefore defined by a monotone decreasing function of  $d_{ij}$ :

$$\begin{cases} m_{ij}(\{y\}) = \alpha \exp^{-\gamma d_{ij}^\beta} \\ m_{ij}(\{n\}) = \alpha(1 - \exp^{-\gamma d_{ij}^\beta}) \\ m_{ij}(\Theta_{i,j}) = 1 - \alpha, \end{cases} \quad (13)$$

where  $0 < \alpha < 1$  represents the degree of confidence in the corresponding association,  $\gamma \in \mathbb{R}^*$  and  $\beta \in \mathbb{N}^*$  can be fixed to a low value [19]. The mass functions are determined with  $\alpha = 0.9$ ,  $\gamma = 0.01$  and  $\beta = 2$ . To take account of the detection and also the tracking performance,  $d_{ij}$  has been selected to be the Mahalanobis distance between a pair of objects  $X_i$  and  $Y_j$  such that:

$$\begin{aligned} \tilde{z}_{ij} &= (Z_i(k) - H\hat{X}_j(k)) \\ S &= (H\hat{P}_j(k)H^T + R) \\ d_{ij}^2 &= \tilde{z}_{ij}^T S^{-1} \tilde{z}_{ij}, \end{aligned} \quad (14)$$

where  $S$  is the covariance matrix and  $\tilde{z}_{ij}$  the innovation of the filter. The support in an association is thus related to the distance between the object pair and at the same time to the prediction and tracking.

### C. Algorithm Performance Comparison

The focus is placed on the cross evaluation of 3 algorithms: the JCPP<sup>3</sup> from Mercier [9] and the local association approaches from Daniel et Lauffenburger [12] (LPP and GCPP). The assignment principle from Mourllion et al. [8] (see Section III-C) is not considered due to its particular ability to provide multi-associations. For comparison purposes with classical probabilistic methods, the Global Nearest Neighbor (GNN) selection approach has also been tested.

The performance is assessed in the ability of the TSR system to reduce the False Positive Per Frame (FPPF) while providing a high Detection Rate Per Frame (DRPF). The required Computation Time Per Frame (CTPF) is also evaluated. Since false positives may appear in any frame independently of the presence of signs, a long test sequence (12 500 frames) is considered in order to provide representative results. The performance criteria are defined such that:

$$\begin{aligned} FPPF &= \frac{\text{Number of false positives}}{\text{Total frame number}} \\ DRPF &= \frac{\text{Number of true positives}}{\text{Ground truth}} \cdot 100 \end{aligned} \quad (15)$$

The FPPF informs about the wrong detections and should tend to 0 whereas the DRPF is the proportion of matched objects with respect to the hand-labeled ground truth.

The results are summarized in Table II. No algorithm is outperforming and all credal solutions give similar results compared to the GNN. In addition, it is difficult to define a heuristic for the mass function model parameters (13). Indeed,

<sup>3</sup>In [9], the authors mentioned that JPP and JCPP provide the same results and only differ in the computation time. That is the fastest method has been retained here.

it appeared during the tests that the TSR is not very sensible to these parameters. For instance, fixing  $\gamma = 0.1$  and  $\beta = 1$  gives close results ( $DRPF \approx 86\%$ ,  $FPPF \approx 0.15$ ).

TABLE II  
TSR PERFORMANCE COMPARISON

	JCPP	LPP	GCPP	GNN
DRPF (%)	86.2	85.8	85.9	85.6
FPPF	0.1497	0.1483	0.1477	0.1508

Differences can be noticed in Table III when comparing the computation times. The most effective solution remains the GNN whose computation load is not affected by the number of objects to be associated. These experimental tests also confirm that Mercier’s approach is the most time consuming solution as underlined in [12] with simple simulations. Finally, it is important to note that the GNN performs only a track-to-observation association contrary to the other solutions.

TABLE III  
COMPUTATION TIME COMPARISON

CTPF (ms)	JCPP	LPP	GCPP	GNN
$n = m = 3$	0.5	0.142	0.156	0.0013
$n = m = 4$	3.55	0.39	0.39	0.0015
$n = m = 5$	28.16	1.227	1.228	0.002

Local optimization assignment methods such as LPP and GCPP are therefore interesting for TSR applications since they provide good real-time compliant association and tracking performance and consequently high detection rates and low false positives per frame rates. In [13], the authors evaluated particularly the LPP during extensive tests in real-time conditions on different driving sequences. ROC curves highlight the interest of the joint credal association and tracking for TSR. The paper also shows that the performance (detection rate, global computation time) are similar to recent state-of-the-art non-evidential tracking-based TSR systems.

#### D. Conflicting Evidences Case Study

This section explores the specific case of conflicting situations. With respect to the data model used, evidence conflicts when detected and tracked objects have close and low statistical distances depending on the tracking results, i.e. when the belief in several assignments between the same objects are high. Let us consider a typical conflicting case from the video sequence shown in Figure 4. The mass functions computed using (13) are depicted in Table IV. Multiple association possibilities between  $X_i$  and  $Y_j$  are highlighted and conduct to high conflict values. After combination and pignistic transformation, the pignistic matrices from Table V and Table VI are obtained. They show easily the difficulty in the assignments since the associations probabilities are lower than the conflict. In this case, all methods provide the same association ( $X_1 \Leftrightarrow Y_1$  and  $X_2 \Leftrightarrow Y_2$ ) and avoid any contradiction. Nevertheless in this case, the association  $X_2 \Leftrightarrow Y_1$  and  $X_1 \Leftrightarrow Y_2$  could also be plausible and it is

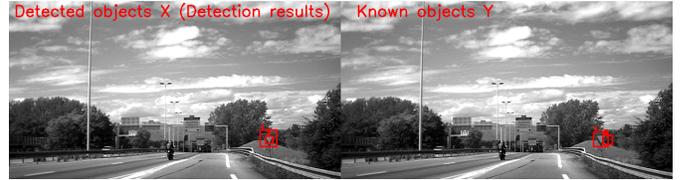


Fig. 4. Association when evidence conflicts. Due to the multiple detections of the same sign, 2 new detections are performed and 2 objects are already known in a close neighborhood, generating highly conflicting mass functions.

difficult to provide *a priori* reasons in favor of one particular association.

The selection of a normalized decision criteria like (4), by the redistribution of the conflict, can lead to counter-intuitive solutions as already shown by Zadeh’s famous example. That is why, it seems important to keep and use the piece of information stored by the conflict during decision taking for instance as in [8]. A different decision strategy would be to propagate multiple association hypotheses and defer the decision when evidence conflicts.

TABLE IV  
PAIRWISE MASS FUNCTIONS WHEN EVIDENCE CONFLICTS

$m_{ij}$	$Y_1$	$Y_2$
$X_1$	0.89	0.66
	0.01	0.24
	0.10	0.10
$X_2$	0.88	0.74
	0.02	0.16
	0.10	0.10

TABLE V  
 $BetP_{X_i}$  PIGNISTIC MATRIX

$BetP_{X_i}$	$Y_1$	$Y_2$	*	$\emptyset$
$X_1$	0.330	0.077	0.006	0.587
$X_2$	0.245	0.091	0.003	0.651

TABLE VI  
 $BetP_{Y_j}$  PIGNISTIC MATRIX

$BetP_{Y_j}$	$X_1$	$X_2$	*	$\emptyset$
$Y_1$	0.112	0.101	0.004	0.783
$Y_2$	0.191	0.279	0.042	0.488

## V. CONCLUSION

This paper investigated multi-object association algorithms based on the transferable belief model. Their aim is to select the most relevant associations among the possible ones between two sets of objects. These sets gather respectively the current sensor observations and the already known objects. Initially, this optimization was performed considering a global belief function optimization. Simple simulated conflicting case studies nevertheless highlighted non-optimal object pairs selection. Solutions supporting individual object associations

have therefore been defined and shown to be effective in the same simulated situations.

In this paper, a vision-based Traffic Sign Recognition system is proposed as real evaluation context for the aforementioned state-of-the-art credal algorithms. It provides a general framework including object detection in the image frame, the association and finally the spatial (trajectory in the image frame) and temporal object tracking using dynamically managed Kalman Filters. Object appearances and disappearances are handled with a score function related to the object trajectory length. Results obtained through real experiments over long sequences showed equivalent performance between the credal algorithms. The latter performed similarly to the well-known Global Nearest Neighbor assignment solution included in this benchmark for comparison purposes.

The promising results obtained in association and tracking of static objects in the intelligent vehicles context allow to consider more complex situations. Future work will for instance concern dynamic objects tracking and classification in a multi-sensor context.

#### ACKNOWLEDGMENT

This work was funded by the Algerian and French Governments, through a PROFAS research fellowship.

#### REFERENCES

- [1] D. Reid, "An algorithm for tracking multiple targets," *IEEE Trans. Autom. Control*, vol. 24, no. 6, pp. 843–854, Dec. 1979.
- [2] Y. Bar-Shalom and X. Li, *Multitarget-multisensor tracking: principles and techniques*. Bradford, U.K.: YBS Publishing, 1995.
- [3] S. Blackman and R. Popoli, *Design and analysis of modern tracking systems*. Norwood, MA, USA: Artech House, 1999.
- [4] A. Dempster, "Upper and lower probabilities induced by a multivalued mapping," *Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 325–339, 1967.
- [5] G. Shafer, *A mathematical theory of evidence*. Princeton, NJ, USA: Princeton University Press, 1976.
- [6] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66, no. 2, pp. 191–234, Apr. 1994.
- [7] M. Rombaut, "Decision in multi-obstacle matching process using Dempster-Shafer's theory," in *International Conference on Advances in Vehicle Control and Safety*, 1998, pp. 63–68.
- [8] B. Mourllion, D. Gruyer, C. Royère, and S. Thérout, "Multi-hypotheses tracking algorithm based on the belief theory," in *International Conference on Information Fusion*, Philadelphia, PA, USA, 2005, pp. 922–929.
- [9] D. Mercier, E. Lefèvre, and D. Jolly, "Object association with belief functions, an application with vehicles," *Information Sciences*, vol. 181, no. 24, pp. 5485–5500, Dec. 2011.
- [10] T. Denoeux, N. El Zoghby, V. Cherfaoui, and A. Jouglet, "Optimal object association in the Dempster-Shafer framework," *IEEE Trans. Syst., Man, Cybern.*, 2014, accepted.
- [11] A. Dallil, M. Oussalah, and A. Ouldali, "Sensor fusion and target tracking using evidential data association," *IEEE Sensors J.*, vol. 13, no. 1, pp. 285–293, Janv. 2013.
- [12] J. Daniel and J.-P. Lauffenburger, "Multi-object association decision algorithms with belief functions," in *International Conference on Information Fusion*, Singapore, 2012.
- [13] M. Boumediene, J.-P. Lauffenburger, J. Daniel, C. Cudel, and A. Ouamri, "Multi-ROI association and tracking with belief functions: application to traffic sign recognition," *IEEE Trans. Intell. Transp. Syst.*, 2014, accepted.
- [14] J. Daniel and J.-P. Lauffenburger, "Fusing navigation and vision information with the transferable belief model: Application to an intelligent speed limit assistant," *Information Fusion*, vol. 18, pp. 62–77, Jul. 2014.
- [15] P. Smets, "Decision making in the TBM: the necessity of the pignistic transformation," *International Journal of Approximate Reasoning*, vol. 38, no. 2, pp. 133–147, Feb. 2005.
- [16] M. Boumediene, C. Cudel, M. Basset, and A. Ouamri, "Triangular traffic signs detection based on RSLD algorithm," *Machine Vision and Applications*, vol. 24, no. 8, pp. 1721–1732, Nov. 2013.
- [17] R. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [18] M. Boumediene, J.-P. Lauffenburger, J. Daniel, and C. Cudel, "Coupled detection, association and tracking for traffic sign recognition," in *IEEE Intelligent Vehicles Symposium*, Dearborn, Michigan, USA, June 2014.
- [19] T. Denoeux, "A k-nearest neighbor classification rule based on Dempster-Shafer theory," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, no. 5, pp. 804–813, May 1995.