Coupled Detection, Association and Tracking for Traffic Sign Recognition*

Mohammed Boumediene[1], Jean-Philippe Lauffenburger[2], Jérémie Daniel[2] and Christophe Cudel[2]

Abstract— **This paper tackles the problem of tracking-based Traffic Sign Recognition (TSR) systems. It presents an integrated object detection, association and tracking approach based on a spatio-temporal data fusion. This algorithm tracks detected sign candidates in order to reduce false positives. Regions Of Interest (ROIs) potentially containing traffic signs are determined from the vehicle-mounted camera images. An original corner detector associated to pixel coding ensures the detection efficiency. The ROIs are combined using the Transferable Belief Model semantics. The associations maximizing the pairwise belief between the detected ROIs and ROIs tracked by multiple Kalman filters are processed. The track evolution helps to detect false positives. Thanks to this solution and to a feedback loop between the tracking algorithm and the ROI detector, a false positive reduction of** 45% **is assessed.**

I. INTRODUCTION

Vision-based TSR systems are usually composed of detection and recognition (or classification) steps [1]. In each frame, the detection localizes ROIs which may contain signs. The second task is to classify the detected signs. This consists in identifying the pictograms included in the sign candidates.

When detection and recognition are independently processed, TSR systems have known limitations: (1) Multiple local detections for the same sign; (2) Misdetections due to temporary occlusions; (3) Wrong detections, usually called "*false positives*". These drawbacks can be lowered when tracking is used [2], [3]. Indeed, target tracking helps to take account of temporal redundant information of the scene, and thus allows TSR systems to track the signs over time [4], [5]. In these approaches, a traffic sign is usually confirmed when it has been tracked over successive frames. Another advantage of tracking is its ability to cope with temporary occlusions by predicting the sign position and size over time. While the interests of tracking for TSR systems are obvious, it has until now been under-utilized as illustrated in the complete review from Møgelmose et al. [1].

The present paper focuses on an integrated approach for traffic sign detection, association and tracking. First, an original detector from Boumediene et al. [6] based on both corner and edge orientation detection is implemented. Using only grey-scale images, the corners are detected to reduce the search area of signs in images. Secondly, a temporal data fusion algorithm is developed for object association and tracking. Generally, joint association and tracking is treated in the probabilistic framework. Several books such as [2] and [3] are dedicated to this topic. One can particularly mention the Global Nearest Neighbor (GNN), (Joint) Probabilistic Data Association Filters ((J)PDAF) [3] or the Multi-Hypothesis Tracking (MHT) [7] as main techniques devoted to the observation-to-track association. These solutions suffer from the limits of the Bayesian theory when dealing with uncertainty and imperfections. The seminal book from Blackman [8] argues that, as a generalization of the Bayesian inference, the evidential reasoning [9], [10] avoids limitations of the probabilistic framework: inappropriate ignorance, uncertainty and incompleteness modeling, necessity of measurement error types assumptions, etc. Several applications of this theory to joint association and tracking have been performed and comparisons with Bayesian tracking solutions can be found in [11], [12].

This paper proposes to perform the association with belief functions [9], [10] using the Transferable Belief Model (TBM) [13] as in previous studies [14], [15], [16], [17]. The detected traffic sign candidates are considered as information to be combined in order to reduce the influence of data imperfections (sensor inaccuracies, false positives, occlusions, etc.). The authors propose a Multi-ROI Tracking (MRT) algorithm which implements an object association method validated on a simple literature static case study in [16]. For the associated objects, a spatial data fusion for the prediction of their future positions in the images is integrated to reduce the detection search region and time. This prediction of the track state vector is performed using dynamically managed multiple Kalman Filters (KF) [18] followed by a feedback of the predicted sign locations to the detector. The performance of the overall system will be shown on its ability to deal with imperfections, i.e. the reduction of false positive detections thanks to the joint association and tracking. An extended version of this paper providing additional experimental validations as well as a comparison with state-of-the-art systems is available in [19].

The paper is organized as follows. Section II presents the system overview. Section III details the proposed MRT and its experimental validation is presented in Section IV. Finally, Section V concludes this paper.

II. SYSTEM OVERVIEW

The system is composed of a ROI detector coupled to a MRT algorithm (see Fig. 1). The *ROI detection* generates a set of traffic sign candidates (i.e. detected ROIs) which can

*This work was supported by a PROFAS research fellowship granted by the algerian and french governments.

[1]Mohammed Boumediene is with Laboratoire LSI, Université des Sciences et de la Technologie Mohamed Boudiaf (USTOMB) and with Université de Mostaganem Abdelhamid ibn Badis (UMAB), Algérie boumediene.m@gmail.com

[2]Jean-Philippe Lauffenburger, Jérémie Daniel and Christophe Cudel are with Laboratoire MIPS-EA2332, Université de Haute Alsace, Mulhouse, France firstname.lastname@uha.fr

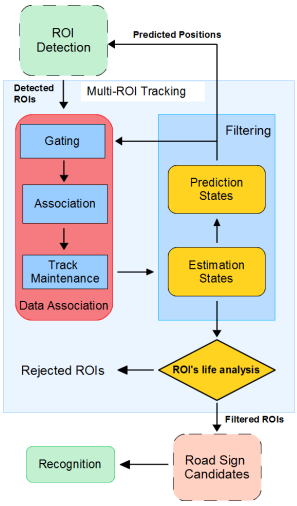


Fig. 1. Overview of the proposed system.

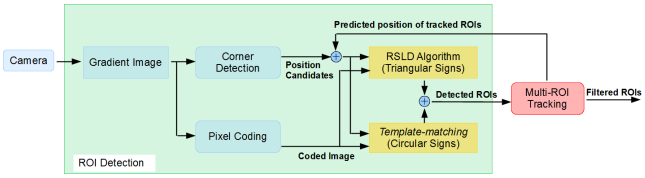


Fig. 2. Block-diagram of the ROI detector.

contain false positives. To cope with this problem, the MRT tracks the detected ROIs using a TBM-based association algorithm. The *Filtering* (or *Tracking*) provides the Detector - through a feedback loop - with the predicted position of ROIs for the subsequent frames. The filtered ROIs are then forwarded to the *Recognition*.

A. ROI Detection

Fig. 2 describes the ROI detector block-diagram introduced in [6] and extended in this paper. The aim is to select the candidate positions of traffic signs by detecting the corners related to the geometric shape and the pictograms of signs. It starts by computing the gradient images I_x and I_y according to the x and y directions respectively, followed by a *Corner Detection* and a *Pixel Coding*. The well-known Harris detector [20] is used to detect corners. The candidate positions help to reduce the search area in which the edge orientation is used to detect the presence of traffic signs.

The *Pixel Coding* extracts the gradient information that will be used to recognize the sign shapes. In the coded image, each pixel is defined according to its orientation gradient (1). Intermediate calculations of the Harris detector are used to define A, B and C such that $A = (I_x^2 > T)$, $B = (I_y^2 > T)$ and $C = (I_x I_y < 0)$. T is a threshold which allows to select a significant gradient. Five classes $\{0,1,2,3,4\}$ are defined depending on the pixel orientation: class 0 describes the non-edge pixels whereas class 1 and 2 correspond to the diagonal edges and 3 and 4 respectively to the horizontal and vertical pixels.

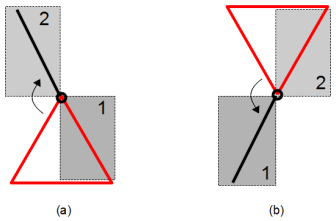


Fig. 3. (a) A vertical y-coordinate inversion applied around the top corner. (b) A vertical y-coordinate inversion applied around the bottom corner.

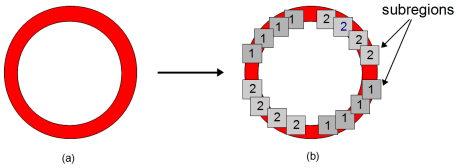


Fig. 4. Template for circular road signs detection. (a) Circular road sign. (b) Classes distribution of the template ().

$$Class = AB(2-C) + 3A(1-B) + 4B(1-A) \qquad (1)$$

The idea is to use the coded image to recognize the sign shapes at the candidate positions previously detected. The authors propose two approaches: the Ransac Symmetric Lines Detection (RSLD) algorithm for triangular signs [6] and a template-matching method for circular signs. The RSLD algorithm consists to straighten the two diagonal sides of the triangular shapes in order to detect one segment instead of two diagonal ones (see Fig. 3). A simple vertical inversion of the y-coordinate applied on coded pixels allows to transform the two diagonal sides of a triangular shape into a single segment. This inversion is applied around the candidate positions related to the top or the bottom of a triangular shape (cf. Fig. 3(a) and 3(b)).

For the *Template Matching*, the authors propose a simple solution based on edge orientation in order to recognize the circular shape. The coded image is used to define this template which is represented by pixel classes distribution. Fig. 4 shows the defined template where subregions are located on the sign border circle. Each subregion, (4×4 pixels) should contain specific coded pixels. The matching process consists to measure the correlation between the proposed template and image windows at the candidate positions. A subregion is valid if the number of pixels belonging to the corresponding class is higher than a threshold $T_c = 4$. The circular shape is recognized if 85% of the subregions are valid. Since the road sign scale in the image frame depends on its distance from the camera, a set of templates having different size are used. The detection can generate several ROIs around the same sign. Therefore, a grouping process is used to select only a single ROI related to each sign by exploiting the sign position and scale in the images [21].

B. Region of Interest Tracking

The MRT algorithm (see Fig. 1) is composed of three functions: the *Filtering*, the *Data Association* and the *ROI life*

analysis. The *Filtering* performs target tracking by predicting the future position of the tracked ROIs in the frames. The *Data Association* assigns one target provided by the detector to a track and manages the track appearance, disappearance and update. The *ROI life analysis* uses the temporal context to validate/reject the tracks transmitted to the *Recognition*. Traffic sign candidates provided by the *Detection* define the targets of the MRT and the tracked ROIs define the tracks to be filtered to reduce false positives.

For the track *State Estimation*, the KF [18] is the straightforward solution. The filter predicts the track state vector (considered as an additional search region in the next frame) from their previous states. These predictions are transmitted to the detector as *a priori* information to influence future searches. Indeed, predicted positions are added to the position candidates defined by the corner detector in order to take account of past detection results.

The *Data Association* consists of three subtasks: *Gating*, *Association* and *Track Maintenance*. The *Gating* eliminates the improbable target-to-track pairings to reduce the computational complexity of the *Association*. The latter defines, at a time k, the relations between targets and tracks. Beside probabilistic methods, the *Association* is performed through a data fusion process where the targets and tracks are data to be combined. The *Association* selects the target-to-track relations required to update the existing tracks. However, traffic signs do not appear and disappear suddenly in frames, as they are usually visible through several successive frames. Therefore, the analysis of the track temporal evolution helps to detect false positives.

III. MULTI-ROI ASSOCIATION AND FILTERING

A. TBM Background

The TBM describes a model of uncertain reasoning and decision making based on a credal and a pignistic level.

1) Credal level: The credal level provides the tools for imperfect data modelling and combination. Consider a problem for which all the discrete solutions $H_j, j = 1,2,...,k$, with k the number of possible hypotheses, define the *frame of discernment* Θ:

$$\Theta = \{\{H_1\},\{H_2\},...,\{H_k\}\} = \bigcup_{j=1}^{k} \{H_j\} \tag{2}$$

Θ is the ignorance, i.e. the union of all hypotheses. Its corresponding *referential subset*, a power set denoted 2^Θ of 2^k disjunctions of H_j is such that:

$$2^\Theta = \{\emptyset,\{H_1\},...,\{H_k\},...,\{H_1,H_2,H_3\},...,\Theta\}, \tag{3}$$

where $\emptyset$ represents the impossible hypothesis interpreted as the conflict. A proposition $A = \{H_1,H_2\}$ refers to the disjunction meaning that either $\{H_1\}$ or $\{H_2\}$ can be the solution to the problem. Each proposition $A \in 2^\Theta$ from a source i is characterized by its *basic belief mass* (*bbm*) $m_i^\Theta(A)$ which represents its veracity:

$$m_i^\Theta : 2^\Theta \to [0,1],\ \sum_{A\in 2^\Theta} m_i^\Theta(A) = 1 \tag{4}$$

When the k answers are exclusive and exhaustive, the solution to the problem is one of the hypotheses of Θ. In this case, a mass on $\emptyset$ is not allowed. Practically, in Multi-Target Tracking (MTT) applications, this assumption is too restrictive since targets can appear and/or disappear over time. When it is impossible to explicitly define all solutions, the non-exhaustivity can be treated by adding to Θ an hypothesis $\{*\}$ representing all unknown propositions [22]. This singleton allows the new discernment frame Θ_{eow} (known as "extended open world" (eow)) to become exhaustive (cf. (5)). A non zero value of $m^{\Theta_{eow}}(\emptyset)$ is then only linked to the sources unreliability or to their discordance.

$$\Theta_{eow} = \Theta \cup \{*\} \tag{5}$$

The combination gathers the different sources *bbm* m_i^Θ. Beside the numerous operators [23], the conjunctive combination rule is the most straightforward [22]:

$$\begin{cases} m_\cap^\Theta(A) &= \sum\limits_{A_1\cap...\cap A_p=A} \prod\limits_{j=1}^{p} m_j^\Theta(A_j) \\ m_\cap^\Theta(\emptyset) &= \sum\limits_{A_1\cap...\cap A_p=\emptyset} \prod\limits_{j=1}^{p} m_j^\Theta(A_j) \end{cases} \tag{6}$$

2) Pignistic level: Solving the association problem consists in making a decision in Θ. Nevertheless, after the combination (see (6)), masses are placed either on singletons ($|A| = 1$) or on unions ($|A| > 1$). A transformation from 2^Θ to Θ is then necessary. Among the available transformations, the pignistic one [14] is commonly used. It equally shares the mass of A on its singletons. The pignistic probability (*BetP*) is given by (7), with $m_\cap^\Theta(\emptyset) < 1$:

$$BetP(H_j) = \sum_{\substack{A\in 2^\Theta \\ H_j\in A}} \frac{m_\cap^\Theta(A)}{|A|\,(1-m_\cap^\Theta(\emptyset))} \tag{7}$$

B. Object Association in the TBM

Consider two extended open worlds $\Theta_{i,.}$ and $\Theta_{.,j}$, where $i = 1,...,n$, with n the number of targets at time k, and $j = 1,...,m$, with m the number of tracks at time k. This leads to an $n \times m$ dimension association problem to be solved at each time step. The frame $\Theta_{i,.}$ contains the m possible target(i)-to-track(j) associations denoted $\{Y_{(i,j)}\}$, and $\{Y_{(i,*)}\}$ represents the appearance of target(i). The frame $\Theta_{.,j}$ is composed of the n possible track(j)-to-target(i) associations denoted $\{X_{(j,i)}\}$, and $\{X_{(j,*)}\}$ represents a track disappearance:

$$\begin{array}{ll} \Theta_{i,.} & = \big\{\{Y_{(i,1)}\},\{Y_{(i,2)}\},...,\{Y_{(i,m)}\},\{Y_{(i,*)}\}\big\} \\ \Theta_{.,j} & = \big\{\{X_{(j,1)}\},\{X_{(j,2)}\},...,\{X_{(j,n)}\},\{X_{(j,*)}\}\big\} \end{array} \tag{8}$$

In order to define if a target X_i is associated with a given track Y_j, three belief masses are used:

- $m_j^{\Theta_{i,.}}(\{Y_{(i,j)}\})$: belief in "$X_i$ is associated with Y_j",

- $m_j^{\Theta_{i,.}}(\overline{\{Y_{(i,j)}\}})$: belief in "$X_i$ is not associated with Y_j",
- $m_j^{\Theta_{i,.}}(\Theta_{i,.})$: ignorance about the association.

The belief masses $m_i^{\Theta_{.,j}}(\{X_{(j,i)}\})$, $m_i^{\Theta_{.,j}}(\overline{\{X_{(j,i)}\}})$ and $m_i^{\Theta_{.,j}}(\Theta_{.,j})$ are generated for the track-to-target associations in the same way. In this paper, the masses are described according to [24]:

$$\begin{cases} m_j^{\Theta_{i,.}}(\{Y_{(i,j)}\}) = \alpha \exp^{-d_{ij}^2} \\ m_j^{\Theta_{i,.}}(\overline{\{Y_{(i,j)}\}}) = \alpha(1-\exp^{-d_{ij}^2}) \\ m_j^{\Theta_{i,.}}(\Theta_{i,.}) = 1-\alpha \end{cases} \quad (9)$$

where:

- $0 < \alpha < 1$ represents the source *a priori* reliability,
- d_{ij} is a dissimilarity measure between target(i) detected at time k and track(j) already known at that time.

d_{ij} characterizes the level of confidence that target(i) corresponds to track(j). The choice of a monotone decreasing function of d_{ij} in (9) leads to the belief reduction in the association of target(i) to track(j) when the distance between the objects increases and *vice versa*. For the purpose of ROI tracking, d_{ij} is chosen as the Mahalanobis distance between track(j) and target(i) [3]. Thus, d_{ij} depends on the tracking filter covariance, so that the belief is determined with respect to the prediction and tracking performance. The *bbm*s are finally combined with the conjunctive rule (6) over each discernment frame and the corresponding pignistic probability matrices $BetP_{i,.}(.)$ and $BetP_{.,j}(.)$ are computed. More details about the combination and the association matrices computation can be found in [15] and [16].

In the pignistic level, decision making, i.e. the best target-to-track and track-to-target association selection must be found. Considering the pignistic matrices as cost matrices, the problem can be classically solved using the Hungarian or Munkres algorithm [27]. This optimization algorithm provides the solution maximizing the sum of beliefs but with the risk of non-optimal local associations [15]. Here, the decision is based on the Local Pignistic Probability (LPP) algorithm introduced in [16] (cf. Algorithm 1). Instead of looking for a globally sub-optimal solution leading potentially to suspicious associations [15], the LPP finds the best set of local assignments. It performs a successive selection of the association which maximizes the probability in each row of a pignistic matrix. An association is retained with respect to the results of the LPP applied successively on $BetP_{i,.}(.)$ and $BetP_{.,j}(.)$. Compared to state-of-the-art algorithms [16], the LPP has the advantage of generating decisions directly from the pignistic matrices with no extra-calculations. A major consequence of this feature is its ability to be used in real-time applications. Furthermore, the LPP is shown to be less conservative than other association solutions reducing ambiguities and/or association contradictions.

C. Track Filtering

KF is the straightforward solution usually retained for tracking applications. This section illustrates how it is applied to ROIs tracking.

Algorithm 1 Local Pignistic Probability

Require: Pignistic matrix $BetP$
Ensure: Best $\{Associations\}$ decision vector
  $\{Associations\} \leftarrow \{\}$
  for all lines of $BetP$ **do**
    $(i_{max}, j_{max}) = \arg\max_{(i,j)}(BetP(i,j))$
    $\{Associations\} \leftarrow (i_{max}, j_{max})$
    Remove i_{max} row and j_{max} column from $BetP$
  end for

From each frame k, a set of detected ROI which represents the n targets defined by the measurements $\mathrm{z}_i(k)$, $i = 1,...,n$ is provided by the detector. Let us consider that the state evolution and the measurement models are linear with respect to the track state vector $\mathrm{x}_j(k)$. For a validated track, a corresponding KF is created and updated. The state vector $\mathrm{x}_j(k)$ of a track(j) is:

$$\mathrm{x}_j(k) = [x, y, s, v_x, v_y, v_s]^T, \quad (10)$$

with $[x, y, s]$ respectively defining the position and size of track(j), $[v_x, v_y]$ its relative velocities in the image coordinate frame between two succesive frames [25] and v_s the scale change. In road scene sequences captured by an in-vehicle camera, road signs appear in frames as small objects moving away from their initial position with a growing size s as they come closer to the vehicle. Considering the high detection rate of the system (25 Hz), an effective dynamic model is the nearly constant velocity model [26] in which the displacement vector $[v_x, v_y]$ and the scale variation v_s of the traffic sign at time k are such that:

$$\begin{cases} x(k) = x(k-1) + v_x(k), \\ y(k) = y(k-1) + v_y(k), \\ s(k) = s(k-1) + v_s(k), \end{cases}$$
$$with: \begin{cases} v_x(k) \sim \mathcal{N}(v_x(k-1), \sigma_x^2), \\ v_y(k) \sim \mathcal{N}(v_y(k-1), \sigma_y^2), \\ v_s(k) \sim \mathcal{N}(v_s(k-1), \sigma_s^2). \end{cases} \quad (11)$$

where $\mathcal{N}(\mu, \sigma^2)$ defines a normal process centered on μ with a standard deviation of σ.

The measurement vector $\mathrm{z}_i(k)$ is such that:

$$\mathrm{z}_i(k) = [x, y, s]^T. \quad (12)$$

The observation model linearly relates the state $\mathrm{x}_j(k)$ to a target measurement $\mathrm{z}_i(k)$ with the measurement matrix H (13). Finally, the state noise and the measurement noise v are i.i.d.[1] random noises assumed mutually uncorrelated.

$$\mathrm{z}_i(k) = H\mathrm{x}_j(k) + v(k) \quad (13)$$

A track goes through three steps: creation, update, and deletion [3]. The appearance decision (depending on

[1] independent and identically distributed

$BetP_{i,\cdot}(\{Y_{(i,*)}\})$) initializes a new KF. Its state vector $\mathrm{x}_j(k)$ is set to the measurement $\mathrm{z}_i(k)$ of the appearing track. The initial displacement $[v_x(0), v_y(0)]$ and size change $v_s(0)$ are set to 0. An existing track is updated if it is associated to a target. In addition, the update step helps to increment a count $c_{update}(j)$ related to track(j) which will be used in the ROI life analysis step to remove the false positives. The disappearance decision (related to $BetP_{\cdot,j}(\{X_{(j,*)}\})$) means that no target has been associated to the track considered. Therefore, the predicted state $\hat{\mathrm{x}}_j(k)$ of the tracked object is propagated without correction. Three consecutive disappearances lead to track deletion. The predicted state $\hat{\mathrm{x}}_j(k)$ is computed according to the previous state $\mathrm{x}_j(k-1)$ by (14). $\hat{\mathrm{x}}_j(k)$ is sent as *a priori* data to the *Detection* to define the position candidates of the signs in the incoming frame.

$$\hat{\mathrm{x}}_j(k) \quad = \quad F\mathrm{x}_j(k-1) \tag{14}$$

To deal with false positives, the evolution of the tracks is analyzed through a confidence score $C(j)$ for each track. $C(j)$ is defined by the ratio between the update count $c_{update}(j)$ and the length of the track (see (15)). $c_{update}(j)$ is the number of frames where track(j) is associated. The track length $length(j)$ represents the total number of frames where track(j) is present. If $C(j)$ is below a predefined threshold $T_{confidence}$, track(j) is considered as a false positive. When track(j) is valid ($C(j) > T_{confidence}$), it becomes a filtered ROI forwarded to the *Recognition*.

$$C(j) = \frac{c_{update}(j)}{length(j)} \tag{15}$$

IV. EXPERIMENTAL RESULTS

In this section, the results obtained during real-time tests performed with a fully equipped vehicle are presented. First, the system implementation and configuration are detailed; then, an analysis presenting the performance rates of the TSR with and without the use of the MRT are described. These results highlight the false positive reduction related to the joint association and tracking algorithm.

The vehicle-mounted camera is a 12-bit grey scale ECK-100 from *Sensata Technologies* providing VGA images (640×480 pixels, 25fps) of the vehicle front scene. The MRT system is written in C++ as a standalone application and uses the OpenCV 2.4 library. It runs on an Intel core *i*7 2.20*GHz* with 8*GB* RAM under Windows Seven 64 bits. The ground truth has been determined manually by defining, in each sequence, the positions of signs having between 20×20 and 60×60 pixels. The critical parameters of the application are: $T_{confidence} = 0.85$ (confidence threshold of the track evolution rate $C(j)$), $\alpha = 0.9$ (discounting factor for the *bbm*), $m_{max} = 10$ (maximal number of KF), $(\sigma_x, \sigma_y, \sigma_s) = (2, 2, 3)$ (standard deviations of the Kalman motion model).

The performance of the MRT is assessed by the *False Positives Per Frame* (*FPPF*), the *Detection Rate Per Frame* (*DRPF*) and finally the *Detection Rate Per Sign* (DRPS) (16). A detection is defined as correct if a sign is forwarded to recognition step more than $N = 3$ successive times.

TABLE I
TEST SEQUENCE FEATURES

Number of frames	12 500
Length (km)	10
Motorway (%)	33.6
City (%)	11.6
Number of circular signs	31
Number of triangular signs	11

$$FPPF = \frac{\text{Number of false positives}}{\text{Total Frame Number}}$$

$$DRPF = \frac{\text{Number of true positives}}{\text{Ground Truth}} \cdot 100 \tag{16}$$

$$DRPS = \frac{\text{Number of correct detections}}{\text{Total Sign Number}} \cdot 100$$

The focus is placed on the video sequence depicted in Table I and composed of various driving scenes (urban driving, highways, etc.). Fig. 5 (Top) shows a short sequence of circular sign detection with no feedback between the tracking and the detection. In frames (a) and (b), the signs are correctly detected. However, in frames (c) and (d), the corner detection fails and so the signs are missed twice. This drawback is solved by the prediction task of the MRT and the feedback of the tracked objects (see bottom of Fig. 5) providing the detector with potential search regions in which signs should be present.

In Table II, the *FPPF*, *DRPF* and *DRPS* with/without the MRT are analyzed. The circular and triangular sign detector alone (denoted $\overline{MRT}$) provides respectively 87% and 91% detection rates per sign with 0.28 and 0.26 of *FPPF*. Adding the MRT reduces the false positives per frame by more than 45%. This lowers the amount of useless data to be processed in the *Recognition*. One can note that, even if the *DRPF* decreases with the MRT for triangular signs, no sign is finally lost.

The ROI detection requires an average of 40 ms to process a single frame. The MRT computation time depends on the number of detected (n) and tracked (m) ROIs but remains insignificant (less than 2 ms in the worst case). This shows that the complete system is real-time compliant, ensuring a high sampling rate.

Finally, even if a direct comparison of TSR performance rates is unrealistic [1], it can be shown that the MRT-based TSR provides interesting results with respect to state-of-the-art solutions [19].

V. CONCLUSIONS

An integrated detection, association and tracking algorithm based on spatio-temporal data fusion for traffic sign recognition has been presented. The idea is to reduce false positives by tracking sign candidates (ROIs) and processing the association with the Transferable Belief Model. An original detector extracts a set of ROI candidates which represents data to be combined. Then, the associations are

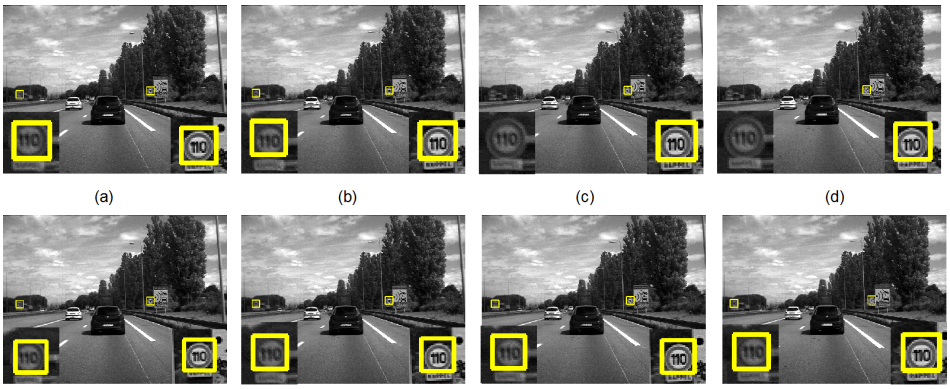


Fig. 5. Impact of the feedback on a sequence of images. (Top) Without feedback of the state prediction to the detector. (Bottom) With feedback to the detector.

TABLE II
PERFORMANCE RATES

	$\overline{MRT}$	MRT
Circular signs		
DRPS (%)	87	87
DRPF (%)	88	86.6(−2%)
FPPF	0.28	0.15(−46%)
Triangular signs		
DRPS (%)	91	91
DRPF (%)	64.2	55.4(−14%)
FPPF	0.26	0.14(−46%)

provided with respect to their pairwise belief maximum. The tracking of the ROIs is performed using dynamically managed multiple Kalman filters. Finally, the tracks are analyzed to detect false positives. The proposed solution feeds the predicted sign locations back into the detector as *a priori* knowledge to influence the detection in subsequent frames. The experimental results show the real-time capability as well as the effectiveness of this strategy in the reduction of false positives by 45%.

In the future, pedestrian and dynamic object detection and tracking applications will be considered. The association will also be enhanced by adding in the pairwise belief functions additional assignment primitives like for instance the correlation coefficients between targets and tracks.

REFERENCES

[1] A. Møgelmose, M. M. Trivedi, and T. B. Moeslund, Vision-based traffic sign detection and analysis for intelligent driver assistance systems: perspectives and survey, IEEE Trans. Intell. Transp. Syst., vol. 13, pp. 1484–1497, Dec. 2012.

[2] Y. Bar-Shalom and X.-R. Li, Multitarget-multisensor tracking: principles and techniques, YBS Publishing, 1995.

[3] S. Blackman and R. Popoli, Design and analysis of modern tracking system, Artech House, 1999.

[4] M. A. García-Garrido, M. A. Sotelo, and E. Martín-Gorostiza, Fast traffic sign detection and recognition under changing lighting conditions, in Proc. IEEE ITSC, 2006, pp. 811–816.

[5] A. Ruta, Y. Li, and X. Liu, Real-time traffic sign recognition from video by class-specific discriminative features, Pattern Recog., vol. 43, no. 1, pp. 416–430, Jan. 2010.

[6] M. Boumediene, C. Cudel, M. Basset, and A. Ouamri, Triangular traffic signs detection based on RSLD algorithm, Mach. Vis. Appl., vol. 24, no. 8, pp. 1721–1732, 2013.

[7] D. B. Reid, An algorithm for tracking multiple targets, IEEE Trans. Autom. Control, vol. 24, no. 6, pp. 843–854, Dec. 1979.

[8] S. Blackman, Multiple target tracking with radar applications, Artech House, 1986.

[9] A. Dempster, Upper and lower probabilities induced by a multivalued mapping, The Ann. of Math. Stat., vol. 38, pp.325–339, 1967.

[10] G. Shafer, A mathematical theory of evidence, Princeton University Press, 1976.

[11] B. Ristic, and P. Smets, The TBM global distance measure for the association of uncertain combat ID declarations, Inf. Fusion, pp. 276–284, 2006.

[12] P. Smets, and B. Ristic, Kalman filter and joint tracking and classification based on belief functions in the TBM framework, Inf. Fusion, vol. 8, pp. 16–27, 2007.

[13] P. Smets and R. Kennes, The transferable belief model, Artif. Intell., vol. 66, no. 2, pp. 191–234, 1994.

[14] A. Ayoun and P. Smets, Data association in multi-target detection using the transferable belief model, Int. J. Intell. Syst., vol. 16, pp. 1167–1182, 2001.

[15] B. Mourllion, D. Gruyer, C. Royère, and S. Théroude, Multi-hypotheses tracking algorithm based on the belief theory, in Proc. IEEE Conf. Inf. Fusion, vol. 2, 2005.

[16] J. Daniel and J.-P. Lauffenburger, Multi-object association decision algorithms with belief functions, in Proc. IEEE Conf. Inf. Fusion, 2012, pp. 669–676.

[17] J.-P. Lauffenburger, J. Daniel, and O. Saif, Multi-object association with belief functions: the dual decision algorithm, in Proc. IFAC ACATTA Workshop, Sep. 2013, pp. 193–198.

[18] R. E. Kalman, A new approach to linear filtering and prediction problems, J. Basic Eng., vol. 82, pp. 35–45, 1960.

[19] 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., Accepted, 2014.

[20] C. Harris and M. Stephens, A combined corner and edge detector, in Proc. AVC, 1988, pp. 147–152.

[21] D. Comaniciu and P. Meer, Mean shift: a robust approach toward feature space analysis, IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 1–18, 2002.

[22] P. Smets, Analyzing the combination of conflicting belief functions, Inf. Fusion, vol. 8, pp. 387–412, 2007.

[23] J. Daniel and J.-P. Lauffenburger, Fusing navigation and vision information with the transferable belief model: application to an intelligent speed limit assistant, Inf. Fusion, vol. 18, pp. 62–77, 2014.

[24] T. Denœux, A k-nearest neighbor classification rule based on Dempster-Shafer theory, IEEE Trans. on Systems, Man, and Cybernetics, vol. 25, no. 5, May 1995.

[25] S.-K. Weng, C.-M. Kuo, and S.-K. Tu, Video object tracking using adaptive Kalman filter, J. Vis. Commun. Image Represent., vol. 17, no. 6, pp. 1190–1208, Dec. 2006.

[26] C. G. Cifuentes, M. Sturzel, F. Jurie, and G. Brostow, Motion models that only work sometimes, in Proc. Conf. BMV, 2012, pp. 55.1–55.12.

[27] H. W. Kuhn, The hungarian method for assignment problem, Nav. Res. Log. Quart., vol. 2, pp. 83–97, 1955.