A LANE ASSESSMENT METHOD USING VISUAL INFORMATION BASED ON DYNAMIC BAYESIAN NETWORK

VOICHTA POPESCU¹, SERGIU NEDEVŞCHI¹, RADU DANESCU¹, TIBERIU MARITA¹

¹ Computer Science Department, Technical University of Cluj-Napoca, 28th Memorandumului Street, Cluj-Napoca, 400114, Romania, Voichita.Popescu@cs.utcluj.ro, Sergiu.Nedevschi@cs.utcluj.ro, Radu.Danescu@cs.utcluj.ro, Tiberiu.Marita@cs.utcluj.ro

The perception and the interpretation of the vehicle’s surrounding are the core modules in any Advanced Driving Assistance Systems and the first two levels in any autonomous driving vehicle. The task of the perception level is to provide objects assessment (objects’ position, form, classification, orientation, speed) using the primary sensorial information. The task of the interpretation level is to use the objects assessment in order to achieve situation assessment, i.e. identification of the instantaneous relations between the traffic environment elements as well as their evolution in time. The work presented in this paper lies within the interpretation layer, aiming to identify the ego-vehicle travelling lane by analyzing the relationships between the ego-vehicle and the surrounding traffic objects, which are detected by a stereovision based perception system. The proposed solution is to identify the ego-vehicle lane by matching the visually detected lane landmarks with the corresponding map landmarks, available through an original Extended Digital Map solution. Additionally, the visually detected vehicles are used in the lane identification process. Due to the proposed approach for lane identification, this task will be referred to as lane assessment in this paper. The used solution for the lane assessment is a probabilistic one, in the form of a Bayesian network, which uses as evidence the visual cues provided by the stereovision perception system. In order to incorporate the dynamic nature of the problem, the temporal dimension was introduced in the network, by proposing a dynamic Bayesian model. Experimental results illustrate both the efficiency of the proposed method even in complex situations, and also the improved efficiency of the dynamic approach over the static one.

Keywords: high-level situation assessment, Bayesian network, visual evidence, temporal filtering

INTRODUCTION

The research in the field of Advanced Driving Assistance Systems (ADAS) is oriented towards improving the safety, quality, mobility, and efficiency of driving. The first generation of ADAS was limited to mere passive assistance through warnings; for such systems the perception of the environment through various, complementary sensors was the main objective. As the ADAS evolve and incorporate active assistance, the importance of the proper and coherent traffic scene understanding has also increased.

In the driving context, the situation assessment refers to the comprehension of the current state of the traffic environment elements: the ego-vehicle, the road, the driver and the other road users. While the perception provides objects assessment (reconstruction of the sensorial detected objects in the scene: objects’ position, form, classification, speed, orientation etc.), the interpretation is responsible for assessing the instantaneous relation between the previously mentioned objects, as well as their evolution in time. Hence, there is an important connection between the perception and the interpretation of the environment: the outputs of the perception modules constitute the inputs for interpretation module. Due to the noisy and uncertain nature of the sensorial information, and due to the fact that not all quantities can be directly measured at all time instances, most approaches for situation assessment are based on probabilistic models.

In this paper, we propose a solution for identifying the ego-vehicle travelling lane (ego-lane number) using the objects assessment of the surrounding traffic objects, detected by the stereovision based perception system. The lane identification is done by analyzing the relationships between the objects in the traffic scene (ego-vehicle, lane markings, and other vehicles), which according to (Rendon-Velez, Horváth, & Opiyo, 2009) is the definition of situation assessment. Therefore, this task is referred to as lane assessment in the rest of the paper. The main idea of the solution is to use both static and dynamic environment objects and their relationship with the ego-vehicle in order to infer
the lane number that the ego-vehicle is travelling on. The static environment objects useful for lane identification are the lateral lane landmarks. In the current solution the visual lane landmarks are matched with corresponding map lane landmarks, available from a proposed Extended Digital Map. The dynamic environment objects useful for lane identification are the other detected vehicles. In the current approach the relative position and orientation of the detected vehicles are used in indentifying the ego-vehicle driving lane.

The used stereovision camera system (Nedevschi, 2009) provides the object information required in the proposed solution for lane identification:
• estimation of current and side lanes;
• the type of painted road objects (lateral lane markings, painted arrows);
• the curb information;
• the position, speed and orientation of the other vehicles.

This information is used as on the spot evidence in the proposed reasoning framework, in the form of a Bayesian network (BN). The training data for the constructed network is an a priori data from an Extended Digital Map. In order to take into consideration the time evolving nature of the driving environment a dynamic Bayesian network (DBN) is used to filter over time the frame-by-frame beliefs.

This problem has been initially addressed by the authors in (Popescu, Bace, & Nedevschi, 2011a), but with the purpose of improved global localization prior to intersections. The current solution however addresses only the issue of ego-lane assessment in various driving contexts, and therefore provides new and original contributions, such as:
• an improved, more complex and more general structure of the Bayesian network, which performs frame-by-frame ego-lane assessment using the visual information as evidence;
• a time filtering mechanism of the frame-by-frame beliefs, by adding the temporal dimension to the initial network and thus obtaining a dynamic Bayesian network;
• a mechanism for efficient DBN inference.

Considering a structured environment, it is very useful for any ADAS to know the lane on which the ego-vehicle is travelling on. This information provides the followings opportunities:
• navigation assistance: lane selection before an intersection for navigation;
• ego-vehicle accurate global localization prior to an intersection (Popescu, et al., 2011a);
• enhanced environment representation (improving the visual environment representation with enhanced digital map information, provided that the precise ego-vehicle’s global positioning is known);
• enhanced environment understanding (for example, objects that are visually detected but not classified as vehicles can now be classified based on their relative position, speed and orientation to the ego-vehicle);
• collision assistance systems: knowing the ego-lane number, an improved traffic constellation of the parallel travelling vehicles can be obtained, thus mitigating lateral collisions.

The remaining of this paper is structured as follows: Section 2 presents related work in the field of situation assessment for intelligent vehicles, Section 3 is dedicated to a short introduction to Bayesian networks and Dynamic Bayesian networks; Section 4 presents the proposed system for ego-lane assessment, while Section 5 illustrates the efficiency of the proposed method through experimental results. Section 6 concludes the paper and illustrates future research and development directions.

**RELATED WORK AND CONTRIBUTIONS OF THE PROPOSED METHOD**

Even if situation assessment in the context of driving assistance has been under research for some time (Huang et al., 1994), (Forbes, Huang, Kanazawa, & Russell, 1995), the domain has gained significant importance in the second half of the last decade, mainly because the sensorial perception has reached a considerable maturity in this time interval. In the context of driving assistance, situation assessment tackles the
following more specific issues: traffic scene understanding (Choi et al., 2010), (Souza & Santos, 2011), (Weiss, Philipps, To, & Kirchner, 2005), lane change assistance (Schubert, Schulze, & Wanielik, 2010), (Althoff, Stursberg, & Buss, 2009), collision avoidance (Brannstrom, Coelingh, & Sjoberg, 2010), (Hillenbrand, Kroschel, & Schmid, 2005). For these problems, various mathematical and probabilistic fundaments have been considered: Markov chain (Althoff, et al., 2009), (Hillenbrand, et al., 2005), fuzzy logic (Weiss, et al., 2005), dynamic probabilistic networks (Forbes, et al., 1995), Bayesian networks (Choi, et al., 2010), decision graphs (Schubert, et al., 2010), and Markov logic networks (Souza & Santos, 2011). A more comprehensive survey about situation assessment in the field of ADAS and autonomous driving can be found in (Rendon-Velez, et al., 2009).

Due to its advantages (human-like reasoning under uncertain measurements, while providing a solid mathematical ground) Bayesian networks have been a preferred reasoning tool in many other applications related to driving assistance: localization (Smaili, Charpillet, & Najjar, 2008), action prediction and generation (Dagli, Brost, & Breuel, 2003), etc.

The specific problem of assessing the ego-vehicle’s lane is addressed in (Choi, et al., 2010), as well as in (Dao, Leung, Clark, & Huissoon, 2007). While the former uses similar technology (a vision system), the latter employs different technologies (GPS and inter-vehicle communication). Since the method proposed in this paper aims to improve the driving function using a stereovision system, the approach that we will relate to is the former (Choi, et al., 2010). In (Choi, et al., 2010) the authors compare a rule-based approach, based on if-then statements with a probabilistic approach, based on a Bayesian network. The elements used in assessing the current lane are: the type and color of the lateral lane delimiters, the sidewalks and the lane change event. This information is encoded in binary nodes, having two possible values: true or false. Even though the method proposed in this paper for ego-lane assessment is also based on a Bayesian network, the method is significantly different, and brings new, important contributions. Both the structure and the parameters of the network are entirely original, as is the parameter learning process. The proposed network encodes, besides the type of lateral lane delimiters, the type of painted arrows for both the ego-lane and the side-lanes. These clues are very important for road segments prior to intersections, with more than three lanes per driving direction.

The nodes in the network that encode the static environment information are discrete nodes, with values corresponding to the type of landmarks of the road segment in question. This information is assumed to be known from an Extended Digital Map (Popescu, et al., 2011a). In the proposed method, the initial parameters of the network are set using the Extended Digital Map information as training data set. Another original aspect is that the proposed Bayesian network encodes both information about the static environment (road landmarks), as well as information about the dynamic environment (other vehicles). The other detected vehicles provide significant cues about the position of the ego-vehicle on the road, and the proposed method exploits this information.

The final, but very important contribution of this paper is the temporal dimension added to the problem. Considering the dynamic nature of the domain, the problem of lane assessment must be evaluated as a whole, by integrating the frame-by-frame results. This is done in the proposed method by extending the Bayesian network with the time dimension, obtaining thus a dynamic Bayesian network. The importance and relevance of treating this problem as a time dependent one is illustrated in the experimental results section, where the Bayesian network results are compared with the dynamic Bayesian network ones, and the improved efficiency of the latter over the former is clearly illustrated. The dynamic Bayesian network approach arises inference and real-time execution issues which are discusses, and solutions are provided in the following section.
**BAYESIAN NETWORKS**

Bayesian networks are an intuitive means of modeling human-like decision-making and provide a method of implementing the power of probabilistic reasoning under uncertainty. The graphical representation is in the form of a directed acyclic graph: each random variable is represented by a node, and each node has a set of probable values called states. The nodes are connected with edges representing the relationships between the variables, and hence the conditional dependence between variables. In other words, the absence of an edge in the graph G implies conditional independence. The graph G encodes the conditional independence assumption: each variable is independent of its non-descendents given its parents in G. This is one of the most important properties of a BN. The graph G represents the qualitative part of the model. The quantitative part of the model is represented by the network’s parameters. These are the probabilities of the nodes: prior probabilities for nodes with no parents, and conditional probability tables (CPTs) for nodes with parents, for the case of discrete nodes. Together, the probabilities collectively quantify the joint probability distribution (JPD) associated with the variables in the graph. Due to the conditional independence property the JPD can be factored, i.e. decomposed into a product of conditional probability distributions over each variable, given its parents in the graph. The advantage of having the JPD in a factorial form is the possibility to evaluate all inference queries by marginalization.

Solving a BN means inferring the state of an immeasurable variable using the values of the measurable variables and the relationship between the variables (expressed here as conditional probabilities). While for small, simple networks the reasoning can be done in a simple way, by using Bayes’s rule and marginalization rule, efficient algorithms can be implied in order to allow real-time operations for larger models (Pearl, 1988).

**Dynamic Bayesian Networks**

In the traffic environment, we are dealing with dynamic events that cannot be detected based on a specific moment in time, but rather can be described through the observations of multiple states that yield the judgment of one complete event. Therefore, in this approach we propose the use of a dynamic Bayesian network (Murphy, 2002), (Mihajlovic & Petkovic, 2001), which is capable of describing a system that is evolving in time. The advantage of DBNs is the extension of the BNs with a temporal dimension while maintaining the BN’s capability of reasoning using uncertain evidence.

Figure 1 illustrates the general structure of the simplest form of DBN, which encodes the first-order Markov property. Using Murphy’s convention (Murphy, 2002), shaded nodes indicate the observed ones (i.e. their values are measurable), while the white nodes are hidden ones (i.e. immeasurable and to be estimated). In this approach we will consider DBNs that encode the first-order Markov property (Murphy, 2002), i.e. the state of the hidden variable at a discrete time instance ($X_t$) depends only on the state of the hidden variable at the immediately previous time instance ($X_{t-1}$), and not on the sequence of events that preceded it, equation (1).

$$P(X_t \mid X_{t-1}, X_{t-2}, \ldots, X_0) = P(X_t \mid X_{t-1}) \quad (1)$$

Similarly, the measurement at the current time instance ($Z_t$) depends only on the current state ($X_t$), and is conditionally independent of all previous states given the current state, equation (2):

$$P(Z_t \mid X_t, X_{t-1}, X_{t-2}, \ldots, X_0) = P(Z_t \mid X_t) \quad (2)$$
\[ P(Z_t | X_t, X_{t-1}, X_{t-2}, ..., X_0) = P(Z_t | X_t) \quad (2) \]

From the graph point of view, it means that the nodes in the DBN can only have parents in the current slice or in the previous slice. The arcs connecting the nodes in the same temporal slice are normal arcs, while the arcs connecting nodes from different time slices are called temporal arcs. Actually, a DBN that assumes the first-order Markov property is a BN structure that is replicated over several time slices.

Similar to a Bayesian filter, where the aim is to recursively estimate an unknown probability density function using an incoming measurement and a state process, in a DBN the aim is to infer the belief of the current hidden node using the incoming measurements as evidence for the observable nodes, and the state process. Hence, a DBN encodes the following two models:

1. the state transition model that quantifies how the system is supposed to evolve in time;
2. the observation model that quantifies the likelihood of making certain observations of the world given the actual state of the world.

As illustrated in Figure 2, the state transition model is encoded in the CPTs of the nodes linked by temporal arcs, while the observation model is encoded in CPTs of the observable nodes.

**Efficient Inference**

A DBN that assumes the first order Markov property represents the following two properties: \( P(X_t | X_{t-1}) \) and \( P(Z_t | X_t) \) (from equation (1) and equation (2)), in a compact way, using a parameterized graph. As a result, as long as the ego-vehicle’s representation of the world conforms to these properties, there is no need to maintain the history of percepts to estimate the current state, since the effect of these observations are accumulated in the belief state at time \( t-1 \) (Forbes, et al., 1995). This is very useful for an efficient network inference. There are several mechanisms for performing both exact and approximate inference in a DBN (Murphy, 2002).

Considering the problem domain, the real-time performance remains a constraint when choosing the inference mechanism. The simplest way to do exact inference in a DBN is to unroll the DBN for \( T \) slices and then apply an inference algorithm - such as Pearl’s algorithm (Pearl, 1988), to the obtained static Bayesian network.

In a DBN that assumes a first-order Markov property the time slices corresponding to the past can be removed while new slices are being added. Before a past slice is removed, its influence must be absorbed in the remaining part of the network. This process is called rolling up (Forbes, et al., 1995).

In this paper we propose a new method for rolling up the network based on a two-slice temporal Bayesian network (2TBN) (Murphy, 2002), Figure 2. The 2TBN will also assume the first order Markov property, as it can be seen in the transition model in Figure 2.

![Figure 2 2TBN](a) Initial model (b) Transition model)

The behavior we want to simulate is the following: at each frame a new slice is added to the front of the network, while a past slice is removed from the end of the network, but before removing it, its influence is stored in the remaining part of the network. We simulate this behavior by considering a 2TBN, with the property that the first slice stores the influence of the past. Hence, the CPTs of the first time slice are revised each frame, while the CPTs of the second slice remain unchanged. At each frame the new evidence is added to the second time slice, and the new belief is obtained. This belief is stored in the CPTs of the first time slice, assuring thus the storing of the past influence in the current network. Through this method, the unrolled DBN will be a 2TBN, with different parameters for the
first time slice of every new frame. This process of removing a past slice when a new slice is added prevents the DBN structure from blowing-up. Furthermore, the computational complexity of the DBN inference will be the computational complexity of the 2TBN, and will be discussed for the DBN proposed in this paper, in the following section.

THE PROPOSED METHOD FOR SITUATION ASSESSMENT

System Overview

Figure 3 illustrates the overall data flow from sub-objects and objects assessment stages to the proposed situation assessment stage, in the context of JDL data fusion model\(^1\) (Steinberg & Bowman, 2001). The first two stages process the raw sensorial information and provide the prerequisite visual information necessary for the situation assessment (e.g. the lane’s width, the type of lane markings, the type of painted arrows, the curb information, and the other detected vehicles information). Their detailed description can be found in (Nedevschi, 2009), (Danescu & Nedevschi, 2010), (Oniga, 2008); these are part of the work of the Image Processing and Pattern Recognition Group, from the Technical University of Cluj-Napoca.

The goal of this paper is to infer high-level situations based on the tracked individual units identified in the previous stages. Hence, this paper addresses the problem of situation assessment based mainly on the visual information provided by the stereovision camera system; the particular problem addressed is assessing the ego-lane number. For this, the relations between the ego-vehicle, the road and the other vehicles are being evaluated. Besides the stereovision camera system, we consider a second sources of information: an Extended Digital Map (Popescu, et al., 2011a), which contains detailed information about the road’s infrastructure (lane numbers per way, lanes widths, type of lateral lane markings and painted arrows).

---

\(^1\)Data fusion model introduced by the Joint Directors of Laboratories’ Data Fusion Sub-Panel (now Group)
the corresponding map landmarks for ego-lane assessment. Additional useful information for reasoning about the ego-lane number is the information about the other vehicles, i.e. their relative position, speed and orientation.

Because of the uncertain and noisy nature of the perception system, a probabilistic framework for reasoning was the best choice. A Bayesian network approach was chosen due to its applicability to the problem domain: infer the state of an immeasurable variable (ego-lane number) using the evidence of the measurable variables (landmarks, other vehicles), besides its advantages as the capability to perform inference on uncertain and incomplete data, in order to reason about the states of the hidden variables.

**The Proposed Bayesian Network**

Figure 4 illustrates the structure of the Bayesian network proposed for identifying the ego-lane number, by analyzing the relationship between the ego-vehicle, the detected road landmarks and the other detected vehicles. This approach is dedicated to structured road segments, with two up to six lanes per driving direction. Important visual cues about the ego-lane number are road landmarks such as: the type of lateral lane delimiters (e.g. double, single, interrupted, curb), the type of painted arrow (e.g. forward, left, right, left-forward, right-forward), but also the observable nodes, for which measurements are available from the visual perception system.

The considered observable system variables that are encoded as nodes in the network are:
- *LeftDelimiter (LD)* – lane left delimiter type;
- *RightDelimiter (RD)* – lane right delimiter type;
- *PaintedArrow (PA)* – lane painted arrow type;
- *LeftLanePaintedArrow (LL-PA)* – painted arrow type of the left side lane,
- *RightLanePaintedArrow (RL-PA)* – painted arrow type of the right side lane), and
- *OtherVehicles (OV)* – the ego-lane number according to the other visually detected vehicles – for example, a vehicle detected as travelling parallel and to the right of the ego-vehicle suggest that the ego-vehicle cannot be on the rightmost lane of the road, therefore the probability that the rightmost lane is the ego-lane is close to zero and the probability for all the other hypothesis (lanes) is computed as a function of the lateral distance between the vehicles. Hence, in the proposed network both static environment information, i.e. road infrastructure information (the first five nodes encode), as well as dynamic environment information (the sixth node) are used together for lane assessment.

Therefore, for the first five nodes, their state and initial parameters are defined using the *a priori* information from the Extended Digital Map (Popescu, et al., 2011a). The states of the *OtherVehicles* node are the possible lane numbers $L_1, L_2, ..., L_n$, where $n$ is the number of lanes per driving direction, of the road that the vehicle is travelling on; this information is available from the digital map.

![Figure 4 The proposed BN for ego-lane number assessment](image-url)
The lanes are numbered from left to right, i.e. $L_1$ is the leftmost lane, and $L_n$ is the rightmost lane.
The inference node is the *Ego-Lane* node. The states of this node are also $L_1, L_2, \ldots, L_n$. Through the relationship between the system variables encoded in the BN, and the on the spot measurements provided by the visual perception system we aim to infer the posterior probability distribution over the states of the inference node, i.e. the belief for each of the states $L_1, L_2, \ldots, L_n$.

In the traffic environment however, it is common that the ego-vehicle performs a lane change maneuver. In order to consider this behavior into the proposed network, a node was introduced in order to encode this event:

- **LaneChangeBehavior (LCB)** – encodes the possible ego-vehicle maneuvers relative to the lane, i.e. lane change left, lane keeping and lane change right. However, when considering lane maneuvers, two lanes are considered: the source lane, and the destination lane. In the proposed network, the destination lane is the same as the ego-lane; hence a new node is not required.

On the other hand, the source lane information is encoded in a new node:

- **SourceEgoLane (SEL)** – encodes the possible source lanes ($L_1, L_2, \ldots, L_n$). While for all the other nodes in the network, the evidence is provided by the visual measurement, for this node (SEL) the evidence is the belief of the *Ego-Lane* node, at the previous time instance.

Therefore, in the proposed network, the root node represents the hypothesis of interest, which is unobservable (the ego-lane number), while the leaf nodes represent items of evidence, which are observable (visual measurements). The arcs represent the direction of influence; therefore, the ego-vehicle being on a certain lane is the cause for the road landmarks to have a certain state. For example, if the ego-vehicle was travelling on the leftmost lane of a road segment with the configuration illustrated in Figure 5, the left delimiter would be double, the right delimiter would be interrupted, and the painted arrow would point left.

Therefore, based on this parent-child relationship, the *a priori* probabilities for the child nodes are computed. The *a priori* probability for the parent node is equally distributed among the states of the node. These are the initial parameters of the network. When the measurements come from the visual system, they set the evidence for the observable nodes. Using an inference mechanism this evidence is propagated throughout the network and the *a posteriori* probability distribution for all the nodes, including the hidden nodes, is obtained.

The two stringent problems regarding Bayesian network, learning and inference, will be discussed next.

**Learning the BN**

The construction is the most time consuming and difficult aspect when using a Bayesian network approach. Both the structure and the parameters of the network can be automatically learned. In this paper we consider the structure of the network to be fixed, and this structure is the one shown in Figure 4. The parameters (the initial probabilities and CPTs) of the nodes in the network, however, are learned using as training data set the information from the Extended Digital Map.

Therefore, a new BN is automatically constructed for each road segment, according to its infrastructure. In this approach we consider that as the vehicle is travelling it is aware of its position of the map (i.e. through a map-matching procedure, which is outside the scope of this paper, the ego-vehicle knows the road segment on which it is currently on). Therefore, the information about the road infrastructure (number
of lanes per driving direction, lanes’ widths, types of lateral delimiters and painted arrows) is available from the Extended Digital Map. This information is used as training data set for learning the network’s parameters.

Consider $D$ the training data set. Parameter estimation means computing $\theta = P(X=s)$, the probability that a node $X$ in the network has a certain state $s$, $s \in D$. Since the training data set is complete, the usual criterion for parameter estimation is Maximum-Likelihood Estimation (MLE), which is based on computing the probabilities that best match the data in $D$. This is done by maximizing the likelihood $P(D|\theta)$, i.e. finding $\theta^*$ such that $P(D|\theta^*)$ is maximum. According to (Zhang, 1996), equation (3) provides the answer, for the discrete case:

$$\theta^* = \frac{\text{# cases (} X = s \text{ and parent}(X) = k \text{)}}{\text{# cases (parent}(X) = k \text{)}}$$

That is, the value $\theta^*$ that maximizes the likelihood $P(D|\theta)$ is equal to the number of cases in which the node $X$ has the state $s$, given that the parents of $X$ have the state configuration $k$, divided by the number of cases in which the parents of $X$ have the state configuration $k$.

For example, consider the node LeftDelimiter (LD) in our network; the training data set for the road segment illustrated in Figure 5 is the one in Table 1.

<table>
<thead>
<tr>
<th>ParentNode</th>
<th>ChildNode</th>
</tr>
</thead>
<tbody>
<tr>
<td>EgoLane (EL)</td>
<td>LeftDelimiter (LD)</td>
</tr>
</tbody>
</table>

By equation (3) the resulting CPT is the one in Table 2.

<table>
<thead>
<tr>
<th>ParentNode</th>
<th>ChildNode</th>
</tr>
</thead>
<tbody>
<tr>
<td>EgoLane (EL)</td>
<td>LeftDelimiter (LD)</td>
</tr>
</tbody>
</table>

The OtherVehicles node has an initial probability equally distributed over the set of hypothesis (lanes).

**Inference in the BN**

The process of computing the *a posteriori* probability distribution of variables, given the evidence is called probabilistic inference. Due to the fact that in a BN the complete JPD over all variables is specified, through the graphical model, all inference queries can be answered by marginalization (summing out over irrelevant variables). However, for a graph with $n$ binary nodes, the JPD has size $O(2^n)$ and then, summing over the JPD takes exponential time (Murphy, 2002). Therefore, this approach for inference is suitable only in simple BNs. There are, however, also inference algorithms that perform well in complex BNs, such as variable elimination, polytree or Pearl’s, variation message passing, relevance tree etc. In this approach, the used inference algorithm is the classic Pearl’s algorithm, whose detailed explanation can be found in (Pearl, 1988).

Therefore, we will not emphasize on this aspect, but rather we will explain how the evidence is brought to the network. For the road landmarks the evidence is straight-forward since the previous stages consisting in computer vision algorithms (sub-object assessment and object assessment) provide the type of landmarks, at each time instance. For the other vehicles however, the computer vision algorithms provide their relative speed and position, with respect to the ego-vehicle. This information is pre-processed in order to provide a probability distribution of the ego-lane over the set of possible states (the lanes of the road, going in the same direction). This probability distribution is used as soft-evidence in the network. More detailed explanations regarding this topic can be found in the authors previous work (Popescu, Bace, & Nedevschi, 2011b).
Adding the Temporal Dimension

The previously described network assesses the ego-vehicle’s driving lane by analyzing the relationship between the ego-vehicle, the road landmarks and the other vehicles. This analysis is done at each time instance, i.e. frame-by-frame, and hence a result is issued every frame. The shortcoming of this approach is the lack of time filtering of the time instance results, and therefore the method is more susceptible to the errors in the sensorial measurements. There are various methods for the temporal filtering of the time instance results, but considering the already existing BN we propose a filtering aproach in the form of a dynamic Bayesian network. The DBN is created by adding a time dimension to the previously created BN. The time dimension consists of the arcs that connect the inference node *Ego-Lane* form one time slice to another. This way, the state of the system at time \( t \) does not depend only on measurements at time \( t \), but also on the system’s previous state, at time \( t-1 \). Therefore, in the proposed dynamic Bayesian network (Figure 6), the ego-lane number is deduced from the time instance measurement about the environment, and from the previous time instance state.

The *state transition model* in the DBN is defined by the CPTs of the temporal nodes. These CPTs define how much the previous state and how much the current measurements actually influence the belief of the current state. In this approach the transition matrix was set to a slightly more relaxed identity matrix, i.e. the CPTs of the *Ego-*\( \text{Lane} \) nodes have a value of 0.9 on the main diagonal and 0.1 / \((n-1)\) in the rest.

In a DBN the problems with inference refers generally to the time complexity, considering the ever-growing structure of the network. However, due to the first-order Markov property employed by the proposed network, two slices are enough to have the equivalent of the unrolled network with any number of slices, as it has been explained in section 3.

Therefore, according to (Murphy, 2002) the time complexity of exact inference in the proposed DBN is between \( \Omega(n^2) \) and \( O(n^2) \), where \( n \) is the number of lanes per driving direction for the particular road segment, and also the maximum number of values for a node.

LANE CHANGE DETECTION

The lane change detection is done according to (Risack, Möhler, & Enkelmann, 2000), by checking the ego-vehicle’s position relative to the visually detected lane. Figure 7 illustrates the elements required for computing the ego-vehicle’s position: \( cw \) = ego-vehicle’s width, \( lw \) = lane’s width, \( yo \) = lateral offset. The last two values are provided by the stereovision perception system. The ego-vehicle is considered as approximately parallel to the lane delimiters due to the fact that the steering angle is small.

The position of the ego-vehicle on the lane is computed according to the position of the front wheels relative to the lane boundaries. Therefore, the position of the *left* wheel relative to the *left* lane boundary is given by equation (4):

![Figure 6 The proposed DBN for time filtering the frame-by-frame belief of ego-lane number](image)
Elements required for lane change computation

\[
d y_{\text{left}} = \frac{lw}{2} + \left( y_0 - \frac{cw}{2} \right) \tag{4}
\]

Similarly, the position of the right wheel relative to the right boundary is given by equation (5):

\[
d y_{\text{right}} = \frac{lw}{2} - \left( y_0 + \frac{cw}{2} \right) \tag{5}
\]

The ego-vehicle is keeping the lane if both \(dy_{\text{left}} > 0\) and \(dy_{\text{right}} > 0\); otherwise, if \(dy_{\text{left}} < 0\) then the ego-vehicle is changing the lane to the left, and respectively, if \(dy_{\text{right}} < 0\), the ego-vehicle is changing the lane to the right.

**EXPERIMENTAL RESULTS**

The proposed method was implemented in C++ using the SMILE (Decision Systems Laboratory) library for the Bayesian network implementation. The on-the-spot visual information is provided by the SCABOR (Nedevschi, 2009) stereovision perception system, through the CAN bus. Figure 8 illustrates in the left side a snapshot of one of SCABOR’s output, with the visually detected cues (lane, lane delimiters, and other vehicles). On the right side of the figure, a snapshot from the proposed situation assessment system, which illustrates the detected traffic elements, as well as results of the ego-lane assessment method as the probability distribution over the lanes of the road (upper right corner) are shown. The execution time of the ego-lane assessment is between 15-30 ms, on an Intel Core2 Duo PC, with 2.66 GHz processor and 2 GB of RAM.

The detailed map information has been a priori manually introduced in the Extended Digital Map, for a large number of test case road segments. The experiments were performed in the following way: the vehicle equipped with the stereovision perception system and a standard GPS receiver travelled in the down-town of the city of Cluj-Napoca, Romania. Based on the GPS data the ego-vehicle position and orientation on a road segment of the map was obtained. Knowing the road segment and the travelling direction, the detailed road information was extracted from the Extended Digital Map and used in the ego-lane assessment procedure.

In order to evaluate the proposed method several study case situations have been considered. The simplest one contains three lanes per driving direction, while the most complex one is for a road segment with six lanes per driving direction.

**Figure 7** Elements required for lane change computation

\[
d y_{\text{left}} = \frac{lw}{2} + \left( y_0 - \frac{cw}{2} \right)
\]

\[
d y_{\text{right}} = \frac{lw}{2} - \left( y_0 + \frac{cw}{2} \right)
\]

**Figure 8** (a) Snapshot from the stereovision perception system that detects the following road elements: the lane, the type of lane markings, other vehicles. (b) Snapshot from the reasoning system that assesses the ego-lane number using the visual information as evidence.
For road segments with more than three lanes per driving direction this is a challenging task, especially because the type of lateral lane delimiters no longer represents a discriminatory criterion. For such cases the type of painted arrow of the ego-lane, as well as the type of painted arrows of the side lanes become relevant visual cues for lane assessment.

Besides the static environment information, the dynamic information (evidence of the other vehicles) is very important information when assessing the ego-lane. The information about the dynamic environment is highly useful in cases such as: road segments with more than three lanes per driving direction, or road segments with painted road signs (lane delimiters, painted arrows) of poor visibility or quality.

In the following, a qualitative evaluation of the proposed ego-lane assessment method is presented. The method has been evaluated for the following test cases:

(1) The first case is for a road segment with three lanes per driving direction, and one lane in the opposite direction. The ego-vehicle is travelling on the middle lane ($L_2$), as it is correctly identified by the proposed method. Figure 8 illustrates a snapshot from this particular case. Figure 9 (a) shows the results of the frame-by-frame reasoning, i.e. the instantaneous results of the BN, while Figure 9 (b) illustrates the results of the dynamic approach. It should be noticed how the time instance results for the correct lane are emphasized by the temporal filtering. Figure 9 (c) compares the BN and the DBN results: for the overall sequence of approximately 320 frames the belief for the correct lane increases from an average of $P(L_2^{BN}) = 0.72$ to $P(L_2^{DBN}) = 0.92$.

(2) The second case illustrates a road segment with five lanes per driving direction. Figure 10 illustrates the road landmarks configuration, in the Extended Digital Map; note that for five lanes there are three distinct types of painted arrows.

Figure 9 Results for the ego-lane assessment for a road segment with three lanes per driving direction ($L_1$, $L_2$, $L_3$)
(a) BN results (b) DBN results (c) BN versus DBN results for the ego-lane $L_2$

Figure 11 presents two snapshots of the visual perception system of this specific road: Figure 11 (a) illustrate a snapshots from the beginning of the sequence, where only the lateral lane markings are visually detected and classified as interrupted, and Figure 11 (b) illustrates a snapshot from a later frame, in which the painted arrows for the ego-lane and the side lanes are detected. These particular road landmarks make the difference in assessing the ego-vehicle lane. The road segment is 65 m long, and the proposed method correctly, clearly and uniquely identifies the second lane $L_2$ as the ego-lane starting with frame 19, which is
approximately 30 m before the intersection. This is clearly seen in Figure 12 (b), which illustrates the results of the proposed dynamic network.

Figure 10 Road landmarks configuration of the road in question (case 2). This information is available from the Extended Digital Map.

Figure 11 Snapshot from the stereovision perception system of the road segment with five lanes per driving direction (a) Frame 1 - interrupted lateral lane markings (b) Frame 21 - interrupted lateral lane markings, left painted arrow for the own lane, left painted arrow for the left side lane, forward painted arrow for the right side lane (the arrow pointing left is deteriorated affecting the visual recognition).

Figure 12 (a) illustrates the frame-by-frame results provided by static network. The improvement of the dynamic approach (Figure 12 (b)) relative the static one is very clear. The static approach gives an equal probability distribution among the central lanes \( (L_2, L_3, L_4) \) due to the same type of lane delimiters, and only provides spikes for the correct lane, at several time instances (frames 21, 23, 25 etc.) when the painted arrows are detected, continuing afterwards with insecure results. The dynamic approach, however, uses these spikes in order to assure an increase of the probability of \( L_2 \) for the following frames. Furthermore, the dynamic Bayesian network approach is able to handle the time instance’s incorrect visual detections/classifications cases (as the painted arrow in the right side lane in Figure 11 (b)), which occasionally occur, and filter over time the frame-by-frame outputs, providing robust results.

Very similar results have been obtained for all the lanes of this road segment, and in other several road conditions: good or poor road conditions (crowded traffic, rain).

Figure 12 Results for the ego-lane assessment for a road segment with five lanes per driving direction \( (L_1, L_2, L_3, L_4, L_5) \); \( L_2 \) being the actual ego-lane (a) BN results (b) DBN results.
The third case is for a road segment with six lanes per driving direction. The results of the BN and DBN ego-lane assessment are illustrated in Figure 13 (a) and Figure 13 (b), respectively. Again, it can be noticed how the DBN acts like a filter on the frame-by-frame results. The side-lane painted arrows are important visual cues in uniquely identifying the ego-lane; they help discriminate between the center lanes \(L_2, L_3, L_4\), and \(L_5\). Lanes \(L_2\) and \(L_4\) have the same lateral landmarks, therefore the distinction between lanes \(L_3\) and \(L_4\) is done with the help of the information about the other detected vehicles. The mis-assessment in frames 80-84 in Figure 13 is caused by the incorrect classification of a group of objects on the right sidewalk as “vehicle”. To eliminate this problem in the future, we intend to analyze the behavior of the detected vehicles (speed, persistence over frames) and only those that have a speed higher than a certain value and are persistent for several frames will be considered and used as evidence in the Bayesian network.

A more detailed discussion about the influence of the other detected vehicles in the lane identification process is presented in relation to the following test case.

In the forth study case the influence of the information regarding the other vehicles in the ego-lane assessment process is discussed. The road segment considered has five lanes per driving direction, and the ego-vehicle is travelling on the third lane, \(L_3\) (Figure 14). The sequence has approximately 130 frames. Based on the type of lateral lane delimiters (interrupted), the network can identify that the ego-vehicle is on one of the central lanes of the road \((L_2, L_3 \text{ or } L_4\)) but still cannot distinguish the proper one. In this situation the information provided by the other vehicles becomes very relevant. Even more, the results of the dynamic approach improve considerably the results of the static approach.

Figure 15 illustrates the probability distribution of the ego-lane over the five lanes, based on the information provided by the other detected vehicles, which are: \(LeftVehicle1\) and \(LeftVehicle2\) (the vehicles travelling on lanes \(L_1\) and \(L_2\), respectively); and \(RightVehicle1\) (the bus on the rightmost lane \(L_5\)). The dotted lines in the graphic represent the probability distribution based on each of the three detected vehicles, respectively, while the continuous line represents the average of these probabilities. This is the information that is used as evidence for the \(OtherVehicles\) \(OV\) node in the network. As it can be noticed in Figure 15, the lanes \(L_3\) and \(L_4\) have an equal, highest probability \(P(L_3) = P(L_4) = 0.275\). This evidence is propagated throughout the network, influencing the posterior probability distribution of the hypothesis \((L_1, ..., L_5)\). As a result, the posterior probability distribution of the lane \(L_2\) decreases while the posterior probabilities of the lanes \(L_3\) and \(L_4\) increase, between frames 1-57, as it can be seen in Figure 16. Figure 16 (a) and (b) illustrates the BN and DBN results, respectively, for this particular case.

It should be noted that in frame 58 there is a spike in the probability of lane \(L_5\); this is due to a
vehicle coming from the opposite direction, which increases the probability for $L_3$. Even if in the static BN this event is treated like a singular event, in the DBN this evidence is further propagated into the following time instances; this is one of the advantages of the dynamic approach. As a result, the proposed dynamic Bayesian network correctly and uniquely identifies $L_3$ as the ego-lane, starting from frame 59 (which is actually approximately 60 m before the intersection). Starting with frame 76, this hypothesis is further sustained by the evidence of the painted arrows on the ego-lane and side lanes.

Therefore, even if in this case the detection was favorable, in a similar manner, a false singular event could influence the network to make the wrong inference. The conclusion is that in order to avoid this behavior, only the measurements from the vehicles that are persistent for more than a threshold number of frames should be used as evidence in the DBN. This is just part of the future improvements that the authors intend to work on.

Some remarks have to be made with respect to the spike in frame 58: (1) the vehicle approaching from a different direction is perceived as a spike in the BN because it is detected only for a single frame, and (2) the spike influences considerably the decision towards the correct lane, in the DBN.

(5) The fifth and final discussion is related to a lane change event (the ego-vehicle performs a lane change maneuver) on a road segment with three lanes. The vehicles is travelling on the middle lane ($L_2$) for the first 110 frames; then, for the following 40 frames, the ego-vehicle is performing a right lane change maneuver, moving from the middle lane ($L_2$) to the rightmost lane ($L_3$), on which it will remain until the end of the sequence. Figure 17 (a) and (b) show exactly this behavior: for the first 110 frames the probability...
for \( L_2 \) is the greatest; in the following 40 frames the ego-vehicle is changing to the right lane, \( L_3 \), and as a result the probability for \( L_2 \) decreases while the probability for \( L_3 \) increases; in the remaining frames the ego-vehicle is travelling on \( L_3 \) lane, hence the probability of this lane is the highest. The distinct right side delimiter, which is a curb, offers a significant cue in the ego-lane assessment.

Finally, a quantitative evaluation of the proposed method is performed. The results of the BN and, respectively, of the DBN have been categorized into the following three groups: correct and unique ego-lane assessment (class C1), correct but multiple ego-lane assessment (class C2) and incorrect ego-lane assessment (class I). While the first and last one are clear, the correct but multiple ego-lane assessment means that several road lanes (hypothesis) have an equal posterior probability, among which also the actual ego-lane. The evaluation has been performed separately depending on the complexity of the road infrastructure (number of lanes per driving direction). Real urban traffic scenarios were used for this test summing approximately 10000 image frames, in different road conditions (light traffic, heavy traffic, good road landmarks, and poor road landmarks). The comparative results of the BN approach versus the DBN approach are illustrated in Table 3.

**Table 3** BN versus DBN overall performance

<table>
<thead>
<tr>
<th>Method</th>
<th>BN</th>
<th>DBN[%]</th>
</tr>
</thead>
<tbody>
<tr>
<td># lanes</td>
<td>( C1[%] )</td>
<td>( C2[%] )</td>
</tr>
<tr>
<td>3</td>
<td>89.0</td>
<td>2.9</td>
</tr>
<tr>
<td>4</td>
<td>76.1</td>
<td>13.4</td>
</tr>
<tr>
<td>5</td>
<td>74.3</td>
<td>18.3</td>
</tr>
<tr>
<td>6</td>
<td>72.2</td>
<td>21.1</td>
</tr>
</tbody>
</table>

**CONCLUSION AND FUTURE WORK**

We have addressed the problem of ego-lane assessment by analyzing the relationship between the three elements of the traffic environment: the ego-vehicle, the road and other road traffic user. An original Bayesian network approach is proposed for ego-lane assessment for structured road environments. The network encodes rich information about the static environment (type of lane delimiters and painted arrows), as well as dynamic environment information (other vehicles). The road infrastructure is considered to be known a priori from an Extended Digital Map. The measurements (road landmarks and other vehicles) are performed by the stereovision perception system. Based on the measurements, the evidence for the network is computed, and the reasoning about the ego-lane number is performed. The results are more stable and considerably improved by the dynamic Bayesian network approach, which filters over time the frame-by-frame beliefs provided by the static Bayesian network.
Knowing the ego-lane number provides a better understanding about the traffic environment, enhances the world model, and facilitates driving assistance functions such as navigation, intersection assistance and collision mitigation. Future work includes the usage of continuous nodes instead of discrete nodes for encoding the other vehicles’ information in order to model more accurately the problem domain. Also we intend to perform further high-level situation and threat assessment, in the direction of increasing safety in the lane change maneuver.

**Acknowledgements**

This paper was supported by the project “Multiscalar and multimodal perception of dynamic 3D environment through dense stereovision, dense optical flow and visual odometry data fusion” contract no. PN-II-ID-PCE-2011-3-1086. Partial support comes from the “EXCEL” project, contract POSDRU/89/1.5/S/62557.

The Bayesian Network described in this paper was created using the GeNiE modeling environment developed by the Decision Systems Laboratory of the University of Pittsburgh (http://dsl.sis.pitt.edu/).

**REFERENCES**


Decision Systems Laboratory, U. o. P. GeNiE and SMILE from [http://genie.sis.pitt.edu](http://genie.sis.pitt.edu)


