Towards a Unified Traffic Situation Estimation Model
– Street-dependent Behaviour and Motion Models –

Florian Kuhnt, Ralf Kohlhaas, Thomas Schamm and J. Marius Zöllner
Department of Technical Cognitive Assistance Systems
FZI Research Center for Information Technology
Karlsruhe, Germany
Email: {kuhnt, kohlhaas, schamm, zoellner}@fzi.de

Abstract—For Advanced Driver Assistance Systems and Autonomous Driving, estimating and predicting traffic situations becomes more and more essential. Many approaches focus on one specific application like vehicle state estimation from sensor data or road model estimation from environment perception. To integrate single approaches to one coherent system, one unified model is needed where the existing applied algorithms can be grounded to.

In this paper we propose a Unified Traffic Situation Estimation Model that describes the probabilistic dependencies between road elements. While its independence from time makes it usable for offline mapping tasks, we show that online prediction capabilities can be achieved by applying the model to a longitudinal vehicle state estimation problem: Using the Markov assumption and appropriate state spaces the general unified model can be specialized to an Interacting Multiple Model Filter. Finally, experiments show an improvement in state estimation and prediction over standard models, which only consider vehicle dynamics. Additionally the unified model allows the prediction of street related routes of vehicles.

I. INTRODUCTION

Over recent years, Advanced Driver Assistance Systems (ADAS) have evolved to an ensemble of assistants that altogether will be able to make autonomously driving cars possible. To really remove the driver from the loop the car has to take the full responsibility over every aspect of the situations. It is common sense that this step will be huge since currently the ADAS extends the driver’s abilities and together they perform better than one alone. Removing the driver from the loop will mean that the car has to replace all skills the driver had.

To achieve that, the autonomous car will need one coherent interpretation of the situation including static (markings, signs) and dynamic objects (other cars). Especially dynamic objects change over time: The movement of traffic participants mainly influences the evolvement of the situation. To get a coherent interpretation of the situation, objects have to be observed over time, their state has to be estimated and their future movements have to be predicted using the estimated states and preestimated model information. Especially interactions with infrastructure and other traffic participants influence the movements of objects a lot. While interactions with other traffic participants are a major influence of the behaviours of vehicles, the interactions with infrastructure are an important basis that has to be handled before. When focusing on infrastructure many aspects can already be shown.

The environment can only be observed by sensors that are governed by sensor noise. Thus, the current state can only be estimated probabilistically by considering sensor and model noise. The same is true for prediction. To achieve the best situation interpretation, one coherent probabilistic model is needed that can perform on the whole time-space continuum of the scene.

The paper is organized as follows. Section II will give an overview about related work. The general concept of the Unified Traffic Situation Estimation Model will be presented in Section III, including definitions, required characteristics and inference capabilities. Afterwards we specialize the generic model to a vehicle state estimation and route prediction application by reducing state spaces and using behaviour-dependent motion models in an Interacting Multiple Model Filter (Sec. IV). The given approach is evaluated considering vehicle state estimation and route prediction in Section V.

II. RELATED WORK

Estimating the current traffic scene and predicting its evolvement can be achieved in very different forms. Directly using the sensor data a very good short term prediction can already be achieved [1], [2].

For long-term prediction the influences of the infrastructure, especially at intersections is of major importance. Directly modeling influences can be omitted by observing the behaviours at one specific intersection and learning models for
exactly this intersection. This has already been shown using Hidden Markov Models [3, 4].

The more promising approach is to utilize a model of the underlying infrastructure. This model can include geometric information where vehicles usually drive (e.g. at the level of lanes [5] or corridors [6]) but also semantic information about traffic lights and right-of-way. It can either be loaded from a predefined map but also be created on-the-fly using environment perception [7].

Using an infrastructure model, prediction can be improved. Recent works vary from projecting the objects to the infrastructure model and separating longitudinal and lateral movement [8] over Case Based Reasoning [9] to creating a Dynamic Bayesian Network and learning generic behaviour models from unlabeled observations [10].

The infrastructure model is usually considered as a static given model without uncertainties. But neither maps can always be up-to-date nor environment perception can detect the infrastructure without noise. Probabilistic methods have to be applied to get a basic road model from observations [7]. This can be used to estimate and predict vehicle states in a subsequent step. But for an overall coherent estimation, infrastructure and vehicle state estimation has to be integrated in an overall probabilistic concept.

In this paper we propose an overall probabilistic bayesian concept that can then be specialized using assumptions to fit specific applications. For the case of longitudinal state and behaviour estimation including route prediction we show how the concept can be instantiated as an Interacting Multiple Model Filter.

III. CONCEPT

A. Definitions

Before we can look at the actual concept, the following traffic elements have to be defined (Fig. 1):

- **Street**: The street is the ground that can be driven on. It can consist of corridors and lanes [6].
- **Route**: The route is the sequence of street segments a vehicle will take or has taken to reach a target area.
- **Behaviour**: A Behaviour describes the interaction of a vehicle with the environment, e.g. braking because of a turn.
- **Trajectory**: A trajectory describes the actual dynamics of a vehicle while driving. These can be separated in longitudinal and lateral dynamics if correlations are neglected [8].
- **State**: A vehicle state is the current state of the vehicle’s position and dynamics at one time slice. It can be seen as a marginalization of the trajectory.
- **Measurement**: Parts of every vehicle’s state can be measured using sensors that underly measurement noise because of the physical measurement principle. This noise differs the measurement value from the actual state value.

B. Required Characteristics

Having defined the main elements of a traffic scene, we can now describe the main characteristics, a unified model has to address.

Firstly, there are dependencies between the different traffic scene elements. In reality these relationships can’t be preserved deterministically but have to be modelled in a probabilistic way in order to consider uncertainties: For example, every route can only define a probabilistic distribution over all possible behaviours on that route.

Secondly, there are two different ways of modelling dependencies: Causal and non-causal. Both can lead to valid solutions but causal dependencies are usually easier to describe, lead to a lower number of dependencies, and result in simpler models. We want to use as causal dependencies as possible to get a straight forward model.

Long-term decisions are quite static while short-term behaviours can change more frequently. In our overall concept we want to address this by modeling the time as one probabilistic facet. This is contrary to approaches using time slices, e.g. Bayesian Filters [10].

Finally, many common approaches use the Markov Assumption to simplify dependencies in time based models. This is an assumption that highly reduces complexity but is not needed in an overall concept.

This leads to following characteristics of our basic concept:

- Probabilistic models
- Causal dependencies
- Time as probabilistic facet
- No Markov Assumption

C. Probabilistic Model

Over years, Bayesian Networks have emerged to a standard in modeling probabilistic dependencies in a causal way [11], [12], [13]; The probabilistic directed acyclic graphical model describes the conditional probabilities between random variables in a directed acyclic graph. The conditional probabilities can be used to model causal and less-causal dependencies. Due to the Factorization Theorem and the Bayesian Theorem forward and backward inference can be achieved on both models but the causal models usually lead to simpler models. Furthermore, Bayesian Networks are also the base theory of many applied methods like Hidden Markov Models and Kalman Filters (Fig. 2).

We propose a causal Bayesian Model (Fig. 3) that brings the already defined traffic elements (Sec. III-A) into context and fulfills the required characteristics (Sec. III-B). The nodes represent random variables that represent a probability distribution over possible values. These state spaces will not be defined at this point. Instead, to construct the concept we focus on the dependencies and conditional probabilities between these random variables, that are described by the edges of the directed graph.

Firstly, given a probability distribution over all possible routes $P(R)$ and a distribution over all possible street layouts $P(S)$, the possible behaviours $P(B)$ can be derived using a traffic logic model $P(B|R, S)$. For example, if the route $R$ describes a selected corridor and the street layout $S$ describes the geometry of the corridor as a turn, the behaviour brake-to-turn will have a high probability.
Given a behaviour distribution $P(B)$ possible trajectories $T$ can be derived using $P(T|B)$. This is mainly a partition of possible trajectories that belong to one behaviour.

Given a trajectory over a specific time period, the actual state at time slice $t$ can be derived using the conditional probability function $P(X|T, t)$. If we see $T$ as an explicit description of a long term motion model the current state $X$ can easily be derived. This is contrary to the typical Kalman Filter motion model that implicitly describes the state in the current time slice given the previous state $P(X_t|X_{t-1})$. The explicit motion model has the advantage of being able to model more detailed motions like braking and accelerating again, related to a given environment.

Finally, the state distribution $P(X)$ can be transferred to a measurement distribution $P(m)$ using a measurement model $P(m|X)$.

Summing up, the required a-priori conditional probabilities are:

- Behaviour depending on Route and Street $P(B|R, S)$
- Trajectory depending on the Behaviour $P(T|B)$
- State depending on the Trajectory and the time slice $P(X|T, t)$
- Measurement depending on the State $P(m|X)$

These conditional probabilities can be modeled in any form. To simplify calculations, discrete or Gaussian distributions can be used. Depending on the complexity it won’t be possible to model all parameters by hand. Hence, parameters are typically learned based on observations using expectation maximization [14].

The joint probability function for the whole Bayesian Network is

$$P(m, X, T, B, R, S, t) = P(m|X) \cdot P(X|T, t) \cdot P(T|B) \cdot P(B|R, S) \cdot P(R) \cdot P(S) \cdot P(t)$$

Using the Bayes Rule any joint probability distribution can be calculated from any number of evidences on every node. To be more precise we give some examples in the following.

**D. Inference**

Having this concept defined we want to show that it can be used to localize and predict vehicle states and estimate past and future vehicle routes, but it is also usable for mapping street layouts.

**Localization:** Evidences on street layout $P(S)$, the vehicle’s route $P(R)$ and current and past measurements $P(m|t \leq t_0)$ can improve the current vehicle state estimation $P(X_{t_0})$. This means for example for estimating the ego vehicle state a planned route from the navigation system can be integrated.

**Route estimation:** From evidences on measurements $P(m)$ and the street layout $P(S)$, the routes of vehicles $P(R)$ can be inferred. This is very useful when observing other vehicles to predict if their route crosses the ego route.

**State prediction:** Given a known route $P(R)$ of vehicles and a street layout $P(S)$, future states $P(X|t > t_0)$ can be inferred.

**Mapping street layout:** The task of mapping is to estimate the street layout $P(S)$ from given observations $P(m)$. Therefore trajectories over a longer time horizon have to be derived from single states so that vehicle motions can lead to street probability, compare [7].

Since there are very different goals when focusing on prediction against focusing on mapping the probabilistic model has to be specialized to fit the purposes of the application.

**IV. Specialization**

Since the Mapping step is mainly an offline step that shall be as precise as possible, it is senseful to see the time as one of the random variables. If we focus on Localization, Route Estimation and State Prediction, we work on one specific time step $t_0$, know the past and want to estimate the current state and predict future states. It is crucial to have a very efficient and fast inference algorithm to be able to predict states before they happen. Exemplarily we show how the basic model can be specialized to perform for a longitudinal state estimation and trajectory and behaviour prediction using an Interacting Multiple Model Filter.
A. Dynamic Bayesian Model

Bayesian Filters are Dynamic Bayesian Networks (DBN) that include the Markov Assumption, meaning the current state only depends on the previous state. This highly reduces complexity.

Fig. 4. The unified model can be transformed to a Dynamic Bayesian Model. Dependencies over time are described by dashed arrows.

The presented Bayesian Model (Fig. 3) can easily be transformed into a DBN: The straight-forward way is to remove the time node and duplicate the basic structure for every time slice. Additionally every node is now also dependent on the same node in the previous step (Fig. 4). Of course not every dependency from time slice to time slice is really necessary. In the following, we neglect the noise on the street model and therefore assume the street model as constant. Hence the nodes can be replaced by one constant instance of the static street model. Furthermore if we only keep one time-dependency in the higher levels (Behaviour-to-Behaviour) and one dependency in the lower levels (State-to-State), the whole structure becomes very manageable (Fig. 5).

B. State Spaces

To further reduce complexity it is necessary to use state spaces that fit the purpose of the random variables. Discrete state spaces are very effective. Gaussian models simplify continuous state spaces but complexer models are necessary if we want to describe multi-variate continuous random variables.

Since our purpose is to estimate and predict the longitudinal position \( s \) of vehicles, the space of the State variable \( X \) will consist of the estimated one-dimensional longitudinal position \( s \) and velocity \( v \):

\[
\text{Range}_X = \mathbb{R}^2 = \{(s,v)\} \tag{1}
\]

Routes mainly differ at intersections. Approaching an intersection, we can derive the set of possible routes from the map. This is a finite number \( n_R \) of possible ways \( r_i \), the vehicle can take:

\[
\text{Range}_R = \{r_1, \ldots, r_{n_R}\} \tag{2}
\]

We assume that a finite number of behaviours \( B \) matches to every route \( r_i \); For example for a route turning right there are the behaviours brake-for-turn and brake-for-stop. Therefore the total number of behaviours is finite as well:

\[
\text{Range}_B = \{b_1, \ldots, b_{n_B}\} \tag{3}
\]

with usually

\[
n_B \geq n_R \tag{4}
\]

Depending on the driver and a lot of constraints like street layout, traffic density and occlusions the actual number of possible trajectories \( T \) matching one behaviour \( b_i \) could be huge. We show that using one simple trajectory model \( \tau_i \) per behaviour \( b_i \) already illustrates the potential of the proposed approach.

\[
\text{Range}_T = \{\tau_1, \ldots, \tau_{n_T}\} \tag{5}
\]

with

\[
n_T = n_B \tag{6}
\]

These trajectory models now describe how the continuous vehicle states \( X \) evolve over time. They can be interpreted as different motion models for the underlying process.

Fig. 5. By neglecting the noise on the street model and reducing the dependencies over time this simplified Dynamic Bayesian Model is derived.

Fig. 6. The chosen state spaces induce a Markov Chain at the higher level and a Kalman Filter at the lower level. An Integrated Multiple Model Filter can address this mixture between continuous state estimation and discrete model decision.

Thus, the lower part of the simplified DBN (Fig. 6) is equivalent to a bank of Kalman Filters that estimate the
current state $X_t$ from the previous state $X_{t-1}$, the current measurement $m_t$, and the motion model $T$. They share the same measurement model and have different motion models, defined by $T$. In the upper part we have a Markov Chain estimating the current distribution over all behaviours $B_t$ from the previous behaviours $B_{t-1}$. Since the behaviour estimate is also dependent on the currently estimated trajectory $T$ this part can be compared to a Hidden Markov Model.

Finally the whole concept reduces to a set of Kalman Filters with different Motion Models that will be estimated over time using a Markov Chain. This mixture between continuous state estimation and discrete model decision can be addressed by an Interacting Multiple Model Filter.

C. Interacting Multiple Model Filter

An Interacting Multiple Model Filter (IMM) [15] extends a Kalman-Filter by allowing multiple interacting motion models that are estimated over time using a Markov Chain for the interaction and the measurements for evaluation.

The Markov Chain includes a transition matrix $\pi$:

$$\pi^{ij} = P(r_t = j | r_{t-1} = i)$$

with initial probabilities

$$\pi^i_0 = P(r_0 = i).$$

(7)

The predicted probability of model $j$ can then be calculated from the probability distribution over all models from the previous time step:

$$\pi^{i|t-1} = \sum_{i=1}^{n} \pi^{ij} \pi^{i|t-1}_{t-1}$$

(9)

with

$$\pi^{i|t-1}_{t-1} = P(r_{t-1} = i | Y_{t-1}).$$

(10)

Every Kalman Filter is based on the two basic equations for the process model

$$x_t = A_t \cdot x_{t-1} + w_t$$

(11)

and the measurement model

$$m_t = H_t \cdot x_t + v_t.\)$$

(12)

While the measurement model is the same for every Kalman Filter in the IMM-Filter, we want to focus on how they differ in the process model: In Eqn. 11 it is assumed that the object moves with a motion described by $A_t$ and added white noise $w_t$ with covariance matrix $Q_t$.

The estimated state $\hat{x}_t$ with covariance $P_t$ is calculated using the prediction equations

$$\hat{x}_t = A_t \cdot \hat{x}_{t-1} + B_t \cdot u_t$$

(13)

and

$$P_t = A_t \cdot P_{t-1} \cdot A_t^T + Q_t.$$

(14)

The transition matrix $A_t$, control vector $u_t$ with transfer matrix $B_t$, and the process noise covariance $Q_t$ are the parameters that are specific for each Kalman Filter in the IMM-Filter. These have to be chosen carefully for each behaviour.

D. Behaviour-dependent Motion Models

The different motion models are the central part of the IMM-Filter. One motion model is needed for every behaviour $b$. We consider the following behaviours at an intersection similar to [8].

1) drive-through
2) brake-for-turn
3) brake-for-stop

To achieve this we introduce two generic motion models, one for the drive-through behaviour and one that can be parametrized to model specific turn and stop behaviours. Please recall that the state space for $x$ is already reduced to the longitudinal position $s$ and velocity $v$ (equ. 1).

1) Constant Velocity Model: The Constant Velocity (CV) Model is used to model the drive-through behaviour. It is known from many basic filter applications and is based on the physical formula

$$s = s_0 + v \cdot \Delta t.$$  

(15)

The transition matrix $A_t$ is set to

$$A_t = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix}.$$  

(16)

Since no control input will be used, $u_t$ and $B_t$ are 0. We define the process noise covariance $Q_t$ as

$$Q_t = \begin{pmatrix} 0.5 \cdot \Delta t^2 \\ \Delta t \end{pmatrix} \cdot q \cdot \begin{pmatrix} 0.5 \cdot \Delta t^2 \\ \Delta t \end{pmatrix}^T.$$  

(17)

Then $q$ can be seen as the allowed acceleration of the vehicle. If the CV model is used as single model to fit the whole process, this value must be quite high to match every acceleration and braking maneuver. Since in this work this model is applied to the drive-through behaviour, the noise must be much smaller to match the small velocity changes that happen while driving through an intersection.

2) Target Velocity Model: For the brake-to-turn and brake-to-stop behaviour we introduce the Target Velocity (TV) model. The idea is to calculate the necessary acceleration $a_t$ for reaching a target velocity $v_T$ at a specific target position $s_T$. The target velocity is 0 in the brake-to-stop case and equals a specific curve velocity in the brake-to-turn case.

Assuming constant acceleration from current time slice $t$ to the future target state $x_T$ leads to

$$a_t = \frac{v_T^2 - v_{t-1}^2}{2 \cdot (s_T - s_{t-1})}.$$  

(18)

This calculated currently best acceleration $a_t$ can be used as control input. With the transition matrix $A_t$ from the CV model (equ. 16) and
\[
B_t = \left( \frac{0.5 \cdot \Delta t^2}{\Delta t} \right) ,
\]
\[
u_t = a_t ,
\]
the state prediction equation (Eqn. 13) depends on the velocity of the previous time slice quadratically. Hence, the system becomes nonlinear. Therefore, instead of simply using \( A_t \) in the covariance update function (Eqn. 14) we have to take the Jacobian matrix \( J_t \):
\[
P_t = J_t \cdot P_{t-1} \cdot J_t^T + Q_t .
\]
with
\[
J_t = \left( \begin{array}{c}
\frac{\partial \hat{s}_t}{\partial s_{t-1}} \\
\frac{\partial \hat{v}_t}{\partial v_{t-1}} \\
\frac{\partial \hat{v}_t}{\partial v_{t-1}} \\
\frac{\partial \hat{s}_t}{\partial s_{t-1}} \\
\end{array} \right) 
\]
(22)
\[
\frac{\partial \hat{s}_t}{\partial s_{t-1}} = 1 + \frac{\left( v_{t-1}^2 - v_{t-1}^2 \right)}{4 \cdot (s_t - s_{t-1})^2} \cdot \Delta t
\]
(23)
\[
\frac{\partial \hat{v}_t}{\partial v_{t-1}} = \Delta t + \frac{v_{t-1}^2 - v_{t-1}^2}{2 \cdot (s_t - s_{t-1})^2} \cdot \Delta t^2
\]
(24)
\[
\frac{\partial \hat{v}_t}{\partial v_{t-1}} = 2 \cdot (s_t - s_{t-1})^2 \cdot \Delta t
\]
(25)
\[
\frac{\partial \hat{s}_t}{\partial s_{t-1}} = 1 + \frac{v_{t-1}^2 - v_{t-1}^2}{2 \cdot (s_t - s_{t-1})} \cdot \Delta t .
\]
(26)

\( Q_t \) is defined in the same way as in the CV model (Eqn. 17) and \( q \) can be used to allow derivation from the estimated acceleration \( a_t \).

V. EVALUATION

For evaluation the research vehicle CoCar [16] is used. It is equipped with a high precision GPS sensor (OXTS RT3003) that is used for ground truth localization data. This way using recorded trajectories we know at every time slice the current correct localization, every future localization as well as the route the vehicle is taking.

The system was tested on one chosen intersection. 4 different maneuvers have been recorded with 5 runs each, a total of 20 drives:

- 5 x Turn right
- 5 x Turn left
- 5 x Straight
- 5 x Stop

The system was instantiated with three different behaviours: *Straight, Turn* (with fixed target velocity) and *Stop* (target velocity is 0).

The estimated localization and prediction are compared to two basic Kalman Filters, one with a constant velocity model and one with a constant acceleration model. In all three filters we assume the same measurement noise of 0.1m. For every run, the root mean square error (RMSE) is calculated between the estimated value and the ground truth value from the high precision GPS sensor. The average RMSE over all 5 runs of one maneuver can be compared to those of the two reference filters (Fig. 7).

In localization the new approach reduces the error from between 0.05m and 0.25m to a value that is lower than 0.03m for all manoeuvres. This already matches the precision of 0.02m of the high precision GPS sensor that is used as reference. For evaluating the prediction results, we use the estimated velocity and acceleration (if available) to predict the state 2 seconds ahead. This value is then compared to the ground truth value at that time. The improvement is not that significant but the maximum reduces from about 3.5m (Constant Velocity) and 2.2m (Constant Acceleration) to lower than 1.5m. We can see that knowledge about street-dependent behaviours in the filtering process improves state estimation and prediction, especially in the case of the advanced behaviours *stop* and *turn-right*.

As proposed, the same model can not only be used for state estimation but also other elements up to the route of traffic participants can be inferred. Without considering interactions, only a probability distribution over driven routes can be estimated, especially in the case of the *brake-to-stop* behaviour, where knowledge about interactions with other vehicles will help reducing the distribution to blocked routes. Thus we do not evaluate the estimated route distribution but the estimated behaviour against the actually driven behaviour (Fig. 8). One second before the intersection the three behaviours right, straight and stop are correctly detected. Turning left is often
mixed with driving straight because of the fixed target velocity of the brake-to-turn behaviour model and the almost constant velocity when driving the left turn at this intersection. Of course, the basic concept allows much more explicit motion models to improve this.

Estimating current and future vehicle states as well as possible routes using a specialized IMM-Filter approach shows the potential of the inference possibilities provided by the unified model.

VI. CONCLUSIONS AND FUTURE WORK

In this paper a unified traffic situation model is presented consisting of a Bayesian model that includes measurements, driver’s plans and the street layout as probabilistic components. The model can describe the estimation of current and future vehicle state, prediction of vehicle routes and also the mapping of street layouts even in an offline processing step. Although the time is described as one probabilistic aspect, Bayesian filters can be derived introducing a time dependency and utilizing the Markov assumption.

Especially the idea of inducing motion models from behaviours, that are derived from possible routes and the given street layout, was applicable to an Interacting Motion Model Filter with behaviour-dependent process models. This shows how existing applied methods can be used to fulfill advanced estimation tasks. If existing methods can be grounded to the unified traffic situation model, this is the basis for combining them in a coherent way.

Evaluating the approach against standard models, which only consider vehicle dynamics, showed that an improvement is achieved using the context information from the street model. Additionally discrete route prediction is possible.

In this work we used a static street model. Future work will focus on estimating a probabilistic street model from preestimated map data and environment perception using the same unified traffic situation model. The behaviour-dependent trajectory models in the unified traffic situation model allow much higher precision than the models used in the IMM-Filter. These models could even be learned utilizing algorithms from the field of Deep Learning or Gaussian Processes. Another important step will be to extend the proposed model to also handle interactions between vehicles. The required adaptability to varying numbers of traffic participants leads to Object Oriented Bayesian Networks. The integration into the proposed model will be part of future work.

REFERENCES