Interactive Scene Prediction for Automotive Applications

Andreas Lawitzky\textsuperscript{1}, Daniel Althoff\textsuperscript{1}, Christoph F. Passenberg\textsuperscript{1}, Georg Tanzmeister\textsuperscript{1,3} 
Dirk Wollherr\textsuperscript{1,2} and Martin Buss\textsuperscript{1}

\textbf{Abstract}—In this work, a framework for motion prediction of vehicles and safety assessment of traffic scenes is presented. The developed framework can be used for driver assistant systems as well as for autonomous driving applications. In order to assess the safety of the future trajectories of the vehicle, these systems require a prediction of the future motion of all traffic participants. As the traffic participants have a mutual influence on each other, the interaction of them is explicitly considered in this framework, which is inspired by an optimization problem. Taking the mutual influence of traffic participants into account, this framework differs from the existing approaches which consider the interaction only insufficiently, suffering reliability in real traffic scenes. For motion prediction, the collision probability of a vehicle performing a certain maneuver, is computed. Based on the safety evaluation and the assumption that drivers avoid collisions, the prediction is realized. Simulation scenarios and real-world results show the functionality.

\section{I. INTRODUCTION}

Traffic reports [1] have shown that on average 93 people died each day on U.S. roads. However, this number has declined over the last two decades thanks to several safety systems in the vehicles. For a further decrease of the death toll, an important contribution is the development of additional driver assistance systems or even one step further completely autonomous driving systems. These systems shall help the driver of the vehicle to move the vehicle in a safe way, ideally also in a comfortable and fuel-conserving manner. Today, there are already lots of driver assistance systems integrated in vehicles which help the driver to avoid collisions or to save fuel such as the blind spot monitor or the intelligent shifters.

However, most assistant systems only consider the current state of the vehicle or the current traffic situation, but lack prediction of future situations. These systems are therefore hardly able to detect a dangerous situation at an early stage or to find the optimal gear with regard to the future motion. Thus, it is desirable to develop systems which are able to predict how surrounding traffic participants will behave.

This information could also be used by a driving assistant system, warning the driver about an upcoming dangerous situation, or for autonomous driving to generate safe and comfortable trajectories.

The goal of the approach presented in this work is to enhance the driving safety of vehicles in structured environments. A more reliable prediction of future movements of vehicles shall be achieved by explicitly taking into account the mutual influence of traffic participants. We call this mutual influence interaction of the traffic participants. This interaction is important as explained with Fig.\textsuperscript{1}, which shows three vehicles on a highway. The red vehicle is overtaking the yellow vehicle at a high velocity, which in turns is approaching a much slower truck at the same time. Thus, the yellow vehicle has to either brake or change lanes in order to avoid a collision with the truck. With this knowledge, the driver of the red vehicle can adjust his prediction and react on the dangerous situation ahead of time. Therefore, a prediction without consideration of the interaction would be less reliable.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Fig1.png}
\caption{Three vehicles are interacting in a highway scene.}
\end{figure}

In this work, a novel framework is presented which predicts the motion of vehicles on highways considering the interaction of the traffic participants, fulfilling the required time constraint for on-line computation. Although this work is explained with vehicles driving on highways, the approach can be used for all mobile robotic applications in dynamic and structured environments, where structured means that the motion space of the participants can be discretized without great losses.

\section{II. RELATED WORK}

In the following, different approaches for motion prediction and threat assessment are explained. In general these approaches can be grouped into learning based motion prediction, model based motion prediction, and motion prediction with a cognitive architecture. While learning based approaches for motion prediction learn from observation of the past movements of vehicles in order to predict the future motion, cognitive architectures try to reproduce human behavior and model based prediction use motion models. Former approaches ignore or insufficiently incorporate the influence of mutual interaction of the traffic participants.

The approaches in [2] and in [3], [4] aim at estimating the future position of a vehicle by its past movement. For that purpose, a database of motion primitives for different car actions is constructed preliminarily by recording trajectories. By observing the motion of a vehicle and the measurement of its current state, trajectories are assigned from the database.
to the car with a certain likelihood. Therefore, the former
is using an extension of the Longest Common Subsequence
(LCS), the Quaternion-based Rotationally Invariant LCS to
compare the trajectories. The latter uses clustered trajectories
in order to obtain typical motion patterns, where each cluster
consists of several trajectories. These approaches allow one
to estimate the future motion of a vehicle, assuming it
follows the representative trajectory. Unfortunately, for good
results the databases in these approaches have to contain
an infeasible number of different trajectories in order to
cover the amount of possible maneuvers. Additionally, the
interaction of traffic participants is not considered as in
following papers.

In contrast to the previously introduced approaches, the
one presented in [5] and [6] is able to learn new motion
patterns on-line while predicting the future movement of an
object. For that purpose, the authors use Growing Hidden
Markov Models. This approach worked well in their exper-
iment on a parking lot.

In [7], a combined optimization of the motions of all
vehicles in a particular traffic scene is performed using a cost
function which punishes unsafe maneuvers of the vehicles.
Additionally, this cost-function considers secondary goals
like comfort and fuel consumption. However, this approach
is based on the assumption that the traffic participants are
centrally controllable.

In [8], the cognitive architecture Adaptive Control of
Thought Rational (ACT-R) is introduced. The goal of this
framework is to model human driving behavior including
human perception, e.g. the steering motions are predicted in
consideration of the delay, caused by cognition and human
constraints regarding motor capabilities. The main focus of
this framework is the comprehension of the way a vehicle
is driven, but not to predict the actions of drivers, especially
under mutual influence.

An approach using model based motion prediction for
situation and threat assessment is presented in [9]. For that
purpose, the authors introduce two parameters: Predicted Ob-
ject Minimum Distance (PMD) and Predicted Time to Object
Minimum Distance (TPMD), which serve as a measure for
the threat of a situation, outperforming the commonly used
Time to Collision parameter in their experimental results.

A framework for designing an accurate vehicle motion
model is described in [10] and extended in [11]. This model
includes the expected driver input, enabling an accurate
long term prediction, using cost functions which depend on
the individual driving style and intention of the driver.
Information about the individual driving style is obtained
by previous observation. The authors introduce an approach
for the classification and recognition of maneuvers, which
a driver likes to perform. The expected driver input can be
computed by minimizing the cost function. It is possible to
predict the future trajectory with respect to the postulated
goals, the interaction of drivers is not considered.

In [12] an approach for vehicle maneuver identification
and driver’s intention prediction is presented. In order to
recognize early which maneuver a driver is performing, a
Dempster-Schafer reasoning system is utilized, based on the
theory of evidence. The authors use measurements like PMD,
TPMD, Time to lane crossing (Tlc), or Distance from vehicle
in path to identify the maneuvers. Although this approach
helps to identify maneuvers the ego vehicle driver is likely
to do, it does not estimate the future trajectories of the other
vehicles, which makes it unsuitable for situation assessment.

In [13], the motion of a vehicle is described by a differ-
cential equation and abstracted by a Markov chain which
consists of discrete states called reachable sets. They repre-
sent possible states which can be reached in a certain time
interval. The transition probabilities of the Markov chain are
obtained from the vehicle dynamics. The input is discretized
and also abstracted by a Markov chain. From observation
and heuristics the computation of the transition probabilities
is achieved. In contrast to the former approaches, interaction
of traffic participants is not neglected. For example, if a
slower vehicle arises in the same lane, the probability of a
lane change will increase. The safety assessment and motion
prediction approach of [14] uses a dynamic motion model as
well. With this model, the future trajectories can be computed
for known control inputs. If this is available for all traffic
participants, their trajectories can be checked for collisions.
But as the inputs are unknown, the prior probability distri-
bution has to be estimated with goal functions which model
the drivers’ behaviors. Monte Carlo sampling is used to find
an approximate solution for several trajectories. Further [15]
obtains a posterior distribution of the future inputs, assuming
that drivers try to avoid collisions. With this posterior dis-
tribution, threat assessment can be done. However, constant
velocities are assumed for the vehicles except the currently
considered one. Thus, it is not considered that an obstacle
may react on an upcoming vehicle.

An approach for motion prediction and risk assessment
for lane-crossing scenes can be found in [16] and [17]. The
authors separate the intention of drivers and the expectation
of their motion. In order to avoid high complexity, Dynamic
Bayesian Networks are used to model the motion of vehicles
near intersections with the intention and expectation as
hidden variables. Interference between these variables is
assumed to be a measure for risk. Although this approach
tailor-made for crossings and uses a different theoretic
background, it has parallels with the presented one as it con-
siders the mutual influence of traffic participants. Likewise
it distinguishes between the high-level intention of drivers
and their situation-aware driving. This work also differs from
works like [18], where the control input is implicitly chosen
in a cooperative manner based on global knowledge of the
desired goal.

In contrast to the presented algorithms, which do not
consider the interaction of drivers sufficiently, the following
novel algorithm focuses on the interaction of traffic partici-
pants in order to achieve more reliable motion prediction.

III. PROBLEM STATEMENT

This framework aims to predict the future motions of
vehicles on highways. Like [15], this framework is based
on the postulate that drivers will not perform maneuvers with high collision risks if safer options are possible. As shown before, the common approach to situation prediction implies independent prediction of each traffic participant. In contrast, the approach needs to be extended in order to respect the interaction of the drivers. Therefore, it is necessary to distinguish between the intention and the expected behavior of the drivers. The intention expresses which trajectory drivers want to drive having some high-level goal in mind, while the expectation incorporates knowledge about the current traffic scene. The first can be approximated with the presented common techniques, the latter is estimated with this novel approach. To achieve this, perfect perception of the traffic participants is assumed. With this, the framework will provide an improved estimation of the future motion, making use of the characteristics of structured environments like highways.

All traffic participants are able to perform different motion trajectories, superimposed in \( m \). The future motion of the \( v \)-th vehicle can be modelled with a distribution function \( f_i \), that assigns trajectories a certain probability. The desired output of the motion prediction is therefore the most accurate estimate of \( f_i \).

IV. APPROACH

The work assumes that the future motion of traffic participants can be estimated as a combination of the intention of each driver and the driver’s local risk assessment to perform a maneuver. A situation based approach is presented in the following that unifies the intention and the threat estimate of all drivers, leading to an interaction-aware motion prediction.

Given a traffic scene, the probability that a collision will occur anywhere in the whole scene is

\[
P(C) = \int \text{ind}(C|m)f(m) \, dm
\]  

with \( C \) the event of a collision. The function \( f(m) \) combines the distribution functions \( f_i \) to a continuous probability distribution for the infinite number of possible combinations of future motions of all vehicles in the road scene. These future trajectories determine how traffic situations will evolve in the future. The index-function \( \text{ind}(C|m) \) equals one if at least one collision between two vehicles occurs, else zero.

With this information, this approach should improve the prediction of the situation. Fig. 2 illustrates that the possible trajectories a driver can perform is only restricted by the dynamic constraints. The number of possible trajectories is infinite even for a single vehicle. It is therefore not feasible to solve the combination of possible trajectories for all vehicles as in (1).

A. Discretization of the Continuous Movement Space

Making use of the fact that highways can be seen as a structured environment, the infinite number of possible movements a driver is able to perform can be approximated by a limited number of different maneuvers. Some examples for the discrete maneuvers on the highway include lane changes, acceleration, maintaining the speed, deceleration, and combinations as illustrated in Fig. 3.

Fig. 2: The possible future positions of this vehicle on a highway are only restricted by the dynamic constraints. The black circles, called “Circles of Forces”, define the area which a vehicle is able to reach in a particular time. The green arrows show possible future trajectories in this area.

B. Approximation of the Collision Probability

For the approximation of (1) and simplification of the notation, some sets are introduced. The set \( \mathcal{V} := \{v_1, v_2, \ldots \} \) with \(|\mathcal{V}| = v \) contains all \( v \) traffic participants of a road scene. Each vehicle \( v_i \) can perform maneuvers from the set \( \mathcal{M}_i := \{m_{i,1}, m_{i,2}, \ldots \} \) with \(|\mathcal{M}_i| = m_i \). In order to obtain the possible future evolutions of a road scene, the sets \( \mathcal{M}_i \) are permuted,

\[
\mathcal{S} := \mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \times \mathcal{M}_v = \prod_{i=1}^{v} \mathcal{M}_i, \quad (2)
\]

Hence, the set \( \mathcal{S} \) consists of all \(|\mathcal{S}| \) possible future evolutions of the current traffic scene i.e. all possible combinations of maneuvers. The set \( \mathcal{S} \) contains all future scenes \( s \in \mathcal{S} \), where \( s \) is a \( v \)-tuple with \( s = (s_1, s_2, \ldots, s_v) \). An example for a maneuver combination \( s \) is illustrated in Fig. 4.

Fig. 3: The arrows show example trajectories for the allowed maneuvers. Red arrows label braking maneuvers. Green arrows show trajectories which keep their speed and the blue ones indicate an acceleration of the vehicle.

Fig. 4: The red arrows demonstrate one out of all possible maneuver combinations for this road scene. As each vehicle within this scene has three different opportunities to act, there are \( m_1 \cdot m_2 \cdot m_3 = 27 \) different maneuver combinations.

Each element \( s_i \) of the tuple represents a certain maneuver \( m_{i,j} \) of vehicle \( v_i \) so \( \mathcal{S} = \{(s_1, s_2, \ldots, s_v) \mid \forall i \in [1,v]: s_i \in \mathcal{M}_i \} \). With the limitation of possible maneuvers, the integral in (1) is approximated with a sum

\[
P(C) \approx \sum_{s \in \mathcal{S}} P(C|s)f(s),
\]

where \( f(s) \) is the discrete probability distribution. \( P(C|s) \) is the risk that a collision occurs in a scene and can be determined with a stochastic collision checker [19] or with a stochastic reachability analysis [13] for all pairs of vehicles in the scene. As each element \( s_i \) of the tuple \( s =
maneuvers to the corresponding collision probabilities \( P_f \) when the maneuvers of the vehicles are permuted. After each
estimation is to set

\[
S_{q,p} := \{ (s_1, s_2, \ldots, s_q, \ldots, s_v) \in S \mid s_q = m_{q,p} \}
\]

with \( |S_{q,p}| = \frac{|S|}{m_q} \), where \( S_{q,p} \) consists of all maneuver combinations \( s \) which include the maneuver \( m_{q,p} \) of vehicle \( v_q \).

Hence, the conditional collision risk is

\[
P(C|m_{q,p}) = \frac{1}{f_q(m_{q,p})} \sum_{s \in S_{q,p}} P(C|s) f(s). \tag{4}
\]

With this information the prior intention estimation is adjusted to obtain an interaction-aware distribution.

C. Applicability of the Globally Optimal Solution

If a communication frame was available the overall collision risk of situation \( P(C) \) could be cooperatively optimized. But in contrast to \([7]\), this framework covers the case that no centralized cooperation is expected. Instead, it is assumed that the drivers locally optimize their trajectories based on the estimation of the intention of the surrounding drivers.

D. Implementation

To improve the prediction of the future movement of the traffic participants the algorithm adapts the prior intention estimations \( f_i(m_{i,j}) \) of the maneuvers by computing the collision probabilities \( P(C|m_{i,j}) \) and taking the interaction of traffic participants explicitly into account.

1) Calculation of the Collision Probability \( P(C|m_{i,j}) \):

Instead of setting up a table of all maneuver permutations, an iterative approach is used, avoiding the need to store the high amount \( |S| \) of all permutations. This approach computes the values \( f(s) \) and \( P(C|s) \) at the same time, when the maneuvers of the vehicles are permuted. After each permutation, the value of the product \( f(s)P(C|s) \) is added to the corresponding collision probabilities \( P(C|m_{i,j}) \) of the maneuvers \( m_{i,j} \), which are part of the combination \( s \).

The algorithm can be efficiently implemented:

- Initialize \( P(C|m_{i,j}) = 0 \) for all \( m_{i,j} \in M_i \) for all \( v_i \in V \)
- Permute all maneuver combinations \( s \in S \) and add the product \( f(s)P(C|s) \) of the occurrence probability

\[
f(s) = \prod_{i=1}^{v} f_i(s_i) = \prod_{i=1}^{v} f_i(m_{i,j}).
\]

and the collision risk of the current combination \( s \)

\[
P(C|s) = 1 - \prod_{i=1}^{v-1} \prod_{k=i+1}^{v} (1 - P(C|s_i,s_k)).
\]

2) Modification of the Prior Distributions: How drivers react on their risk estimation should be learned from data. For this interaction model, an exemplary approach distributing the probabilities linearly to the collision risk and the intention estimation is to set

\[
f^\text{new}_i(m_{i,j}) = f_i(m_{i,j}) g(P(C|m_{i,j}))
\]

for each maneuver \( m_{i,j} \) of each vehicle \( v_i \) with

\[
g(P(C|m_{i,j})) = 1 - \frac{P(C|m_{i,j}) - \min_{m_{i,k} \in M_i} P(C|m_{i,k})}{1 - \min_{m_{i,k} \in M_i} P(C|m_{i,k})}.
\]

Of course, \( f^\text{new}_i(m_{i,j}) \) needs to be normalized to one to get \( f^\text{norm}_i(m_{i,j}) \). In the result \( f^\text{norm}_i(m_{i,j}) \) the interaction of the drivers has been incorporated explicitly into the intention estimation \( f_i \) to obtain a more reliable motion prediction. The following shows results of the described algorithm.

V. RESULTS

In order to evaluate the presented approach the algorithm was exemplarily implemented in C++. Example simulation road scenes were created and real-world data has been used to test the approach. Therefore, trajectories for the maneuvers were generated and checked for collisions. The presented approach was integrated in the framework of \([20]\).

Minimum jerk trajectories are used to simulate human driving. For computing the trajectories, a road aligned coordinate system is used. As sketched in Fig. 5, this Frenet system consists of two coordinates \( s(t) \) and \( d(t) \) where the \( s \)-coordinate runs parallel to the road lane and the \( d \)-coordinate is perpendicular to the lane. Note that trajectories are represented more realistically in the Frenet frame. A more detailed look at the trajectory generation is given in \([21]\).

As collision checker for the minimum jerk trajectories the approach of \([19]\) is used to obtain the collision risk

<table>
<thead>
<tr>
<th>Permutations of the maneuvers</th>
<th>Table I: Example for the permutation of three vehicles with three possible maneuvers</th>
</tr>
</thead>
<tbody>
<tr>
<td>( v_1 ) ( m_{1,1} ) ( m_{1,1} ) ( m_{1,1} ) ( m_{1,1} ) ( m_{1,1} ) ( \ldots )</td>
<td>( v_1 ) ( m_{1,1} ) ( m_{1,1} ) ( f(s_1) ) ( f(s_2) ) ( f(s_3) ) ( \ldots )</td>
</tr>
<tr>
<td>( v_2 ) ( m_{2,1} ) ( m_{2,1} ) ( m_{2,1} ) ( m_{2,1} ) ( m_{2,1} ) ( m_{2,1} ) ( \ldots )</td>
<td>( v_2 ) ( m_{2,1} ) ( m_{2,1} ) ( m_{2,1} ) ( f(s_1) ) ( f(s_2) ) ( f(s_3) ) ( \ldots )</td>
</tr>
<tr>
<td>( v_3 ) ( m_{3,1} ) ( m_{3,2} ) ( m_{3,3} ) ( m_{3,1} ) ( m_{3,2} ) ( m_{3,3} ) ( \ldots )</td>
<td>( v_3 ) ( m_{3,1} ) ( m_{3,2} ) ( m_{3,3} ) ( m_{3,1} ) ( f(s_1) ) ( f(s_2) ) ( \ldots )</td>
</tr>
</tbody>
</table>

\[
\text{TABLE I: Example for the permutation of three vehicles with three possible maneuvers.}
\]
It tests if a collision will occur between two vehicles when they drive certain maneuvers. For the computation of $P(C|s)$, possible pairs of maneuvers have to be tested for collision.

Note, that the presented approach is designed to work with arbitrary trajectories and collision checkers.

### A. Simulated Highway Scene

This example serves as a demonstration of the functionality of the presented algorithm. Additionally, the changes from the prior intention estimate to an interaction-aware distribution is shown. With the stochastic collision checker it is possible to compute the values of the collision risks $P(C|m_{i,j}, m_{k,l})$, $i \neq k$. The possible maneuvers of the vehicles and their prior estimated intention distribution are given in Fig. 6a.

(a) Seven vehicles $v_1, \ldots, v_7$ on a highway with three lanes. As an example the arrows indicate nine possible maneuvers of vehicle $v_4$. To distinguish between the maneuvers brake, keep velocity, and accelerate, the length of the arrows is varied. For clarity, the maneuvers of the others are not drawn.

(b) This table lists the probabilities, that drivers perform a certain maneuver. These values have yet not been modified.

(c) The vehicles drive at different initial speeds which are listed in this table.

### Fig. 7: Results of the collision checker and the presented algorithm.

In order to modify the prior distribution, the adjustment step of IV-D.2 is applied to the collision probabilities. The results are presented in Fig. 7b.

This scenario has been computed on an off-the-shelf Intel Core i7-2600 with 3.4 GHz in a multi-threaded C++ exemplary implementation. The $|S| = 629,856$ scenarios of this example have been calculated in less than 4.5 ms, so the computational time is very promising even for more complex scenarios. In Fig. 7c the most probable future actions $f^\text{nom}(m_{i,j}) > 10\%$ of the vehicles are illustrated by the arrows of the vehicles.

### B. Real-world Highway Scene

The presented algorithm has also been tested with recorded data from a German highway scene. Fig. 8 shows a frame of the trip that is evaluated with the novel approach. Again, the vehicles have been assigned an intention estimation analog the values in Fig. 6b. In future implementations these values...
In this paper, a novel approach for motion prediction and threat assessment of objects in structured environments, like highways or crossings, is presented. The approach is developed for driver assistance systems as well as for autonomous driving. The systems assess the danger of possible future trajectories. For that purpose, they require a reliable prediction of the movements of other traffic participants. Due to the fact that the drivers have a mutual influence on each other, it is necessary to consider the interaction of the traffic participants in order to obtain a reliable prediction. Taking the interaction explicitly into account, this framework offers reliability and differs from already existing approaches which do not consider the interaction sufficiently. In simulated and real-world scenarios, the computation time of an exemplary implementation of this approach is shown to meet the on-line requirements. Future work includes evaluation of the chosen interaction model, tests with inter-vehicle communication and different environment settings.

VI. CONCLUSIONS

In this paper, a novel approach for motion prediction and threat assessment of objects in structured environments, like highways or cross-roads, is presented. The approach is developed for driver assistance systems as well as for autonomous driving. The systems assess the danger of possible future trajectories. For that purpose, they require a reliable prediction of the movements of other traffic participants. Due to the fact that the drivers have a mutual influence on each other, it is necessary to consider the interaction of the traffic participants in order to obtain a reliable prediction. Taking the interaction explicitly into account, this framework offers reliability and differs from already existing approaches which do not consider the interaction sufficiently. In simulated and real-world scenarios, the computation time of an exemplary implementation of this approach is shown to meet the on-line requirements. Future work includes evaluation of the chosen interaction model, tests with inter-vehicle communication and different environment settings.

ACKNOWLEDGMENT

The authors gratefully acknowledge support of this work by the BMW Group Research and Technology, Munich, Germany. In part this work is supported by the DFG excellence initiative research cluster Cognition for Technical Systems (CoTeSys), by the Institute for Advanced Study (IAS), Technische Universität München, Munich, Germany and by the ERC Advanced Grant SHRINE.

REFERENCES