Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Probabilistic, Discriminative Maneuver Estimation in Generic Traffic Scenes using Pairwise Probability Coupling

Stefan Klingelschmitt, Volker Willert, Julian Eggert

2016

Preprint:

This is an accepted article published in Intelligent Transportation Systems Conference (ITSC) 2016. The final authenticated version is available online at: https://doi.org/[DOI not available]

2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)
Windsor Oceanico Hotel, Rio de Janeiro, Brazil, November 1-4, 2016

Probabilistic, Discriminative Maneuver Estimation in Generic Traffic Scenes using Pairwise Probability Coupling

Stefan Klingelschmitt[1], Volker Willert[1] and Julian Eggert[2]

Abstract— **Future advanced driver assistance systems as well as autonomous vehicles are expected to further increase their areas of applicability. Reliable maneuver estimations are a prerequisite for many of the intended functionalities. Accordingly, maneuver estimation systems need to cover a wide range of scenarios. The majority of recently presented approaches are targeted at fixed scenarios. However, having specialized maneuver estimation systems covering each possible scenario is unrealistic. Therefore, we present an approach for tackling discriminative maneuver estimations in generic traffic scenes. It is based on reusable, partial classifiers that are combined online using a technique called *pairwise probability coupling*. As a result we are able to make discriminative maneuver estimations in generic traffic scenes. The benefits and applicability are presented on inner-city real-world data sets. Our evaluation indicates that the assembled probabilistic maneuver estimation is not only able to outperform generative models; it surpasses the performance of specially designed models due to the reduced complexities of the partial classifiers.**

I. INTRODUCTION

Automobile manufactures, academia and governments around the world have the goal of reducing the injuries and casualties caused by traffic accidents. The vast majority of accidents can be traced back to human error. E.g. 86% of the traffic accidents registered in Germany can be assigned to an incorrect behavior of the vehicle driver [1]. Therefore, technical systems assisting and supporting the driver will be crucial for reducing traffic accidents.

Recently, many Advanced Driver Assistance Systems (ADAS), including warning systems, partially automated and fully autonomous functions have been developed. Such systems typically take the current observed scene, predict it several seconds into the future and evaluate its criticality in order to provide the function they were designed for. Predicting the evolution of a traffic scene is a demanding task as traffic participants' future trajectories involve many uncertainties. A common approach is to decompose the prediction into two steps. First, the discrete prototypical spatio-temporal behavior is classified, e.g. turn right, turn left or go straight. This is also referred to as maneuver or intention estimation. Secondly, specialized prediction models for these discrete behavior alternatives are used for more accurate kinematic state predictions. Predicting future kinematic states of traffic participants is usually referred to as motion or trajectory prediction. Since the correct maneuver assessment is a prerequisite for the motion prediction in such twofold approaches powerful maneuver estimation systems are needed.

The maneuver estimation is typically affected by many uncertainties, like ambiguous behavior patterns or imperfect sensory measurements. Hence, making discrete choices will probably lead to many false maneuver estimations. This problem increases in conjunction with longer prediction horizons and increasing numbers of possible maneuver alternatives. Therefore a probabilistic treatment, quantifying the uncertainties of the executed maneuver estimation, seems reasonable. Additionally, it provides valuable information for subsequent systems.

A. Related Work

Recently, a variety of approaches addressing the problem of maneuver estimations have been presented. In this brief review, we focus on probabilistic approaches. [1] A more comprehensive overview on the subject of maneuver estimation can be found in [2]. The discussed probabilistic approaches can be divided into two main groups.

The approaches of the **first** group are based on building one model in a discriminative manner in order to distinguish between a fixed and predefined set of maneuver hypotheses. This model can be used to directly asses the current maneuver. [2] The authors of [3] and [4] use Bayesian networks for their maneuver estimation at intersections. Additionally, traffic light information and preceding vehicles are incorporated in [3]. In [5], a comparison of several discriminative classification methods is performed in order to reliably estimate yielding scenarios at an intersection. Maneuver estimation systems tested on highway scenarios are presented by [6] and [7]. The authors of [6] present a hierarchical framework based on single layer perceptrons, while [7] use a relevance vector machine. Due to the fact that the models of this group are adapted in a discriminative manner, they are well suited for distinguishing between different maneuver alternatives. However, they are in general highly specialized and cannot straightforwardly reused in different contexts, e.g. missing or additional possible maneuver alternatives.

The **second** group of approaches build a separate model for each maneuver alternative. The maneuver prediction is

[1]S. Klingelschmitt and V. Willert are with Control Methods and Robotics Lab, Technische Universität Darmstadt, D-64283 Darmstadt, Germany stefan.klingelschmitt, vwillert@rmr.tu-darmstadt.de

[2]J. Eggert is with Honda Research Institute Europe GmbH, D-63073 Offenbach am Main, Germany jeggert@honda-ri.de

[1]Also approaches are discussed that could easily be extended to probabilistic predictions, e.g. support vector machines.

[2]In this context *discriminative* means that these models are optimized for distinguishing/discriminating between a certain set of maneuver hypotheses. Not that they necessarily approximate posterior conditional distributions, directly.

978-1-5090-1889-5/16/$31.00 ©2016 IEEE

performed by a subsequent comparison, determining which maneuver model explains the observed data best. Learned, probabilistic models are introduced by [8] and [9]. In [8], maneuver-specific Gaussian Processes are used, while the authors of [9] present a scalable approach based on quadratic discriminant analysis. Handcrafted systems are presented by [10], [11] and [12]. They are based on comparing measured kinematic states with maneuver-specific indicator functions. These indicator functions can be dynamically adapted, as proposed by [11] and [12].

The approaches of the second group can be more easily adapted to changing environments, e.g. missing or additional possible maneuver alternatives. However, due to their generative character they will often not be able to achieve the same amount of generalization capabilities and performance compared to models adapted in a discriminative manner ([13], [14]). This hypothesis is strengthened by the work of [15]. The authors compare three different maneuver estimation systems, based on a Support Vector Machine, a Dynamic Bayesian Network and Hidden Markov Models. Their evaluation showed that the system based on maneuver specific Hidden Markov Models is clearly outperformed by the systems based on one discriminative model. Similar results are presented in [16], where their LSTM Recurrent Neural Network based maneuver estimation outperforms several variants of Hidden Markov Models. Howsoever, the increased recognition performance could also be a result of the more elaborate temporal integration.

B. Contribution

Traffic scenarios that undergo a high variability form a high cognitive load on the drivers which provokes an increased number of human errors. An example for traffic scenarios with a high variability is inner-city traffic. E.g. in the year of 2014, $69,3\%$ of all registered accidents in Germany took place in urban environments [17]. Therefore, there is a strong demand for flexible, adaptable maneuver estimation systems in order to be able to apply them to scenarios with a large variability like generic inner-city traffic.

This paper addresses the question of how a flexible maneuver estimation system can be created while simultaneously maintaining the performance levels of systems specially targeted at restricted traffic scenarios (i.e. discriminative adapted models). We propose a framework that is based on learning discriminative models while maintaining the possibility to be easily adapted to generic traffic scenes, online. It is based on learning a pool of discriminative, probabilistic binary classifiers. Given the currently entered scene, the corresponding binary classifiers can then be selected capturing only the currently possible maneuver alternatives. The obtained probabilistic, binary predictions are combined online using a technique called *pairwise probability coupling*. This approach enables us to benefit from discriminative models while simultaneously retaining the ability to adapt to generic traffic scenes.

The reminder of this paper is organized as follows. In Section II we discuss the general problems arising in the context of probabilistic maneuver estimation for generic traffic scenes. Section III introduces the theoretic background of our approach. Its applicability to the problem of adaptable maneuver estimation is outlined in Section IV. Experimental results on two real-world data sets are shown in Section V.

II. PROBLEM STATEMENT

In this section we discuss the challenges of making maneuver estimations in highly variable environments. At first the necessary probability theoretical definitions are stated.

A. Probabilistic Maneuver Estimation

We define a discrete random variable B which can take on different values (events), each of them specifying one prototypical spatio-temporal behavior pattern/maneuver, as introduced in [18]. The temporal aspect addresses the circumstance that different maneuvers are not only characterized by different path alternatives (e.g. right turn, left turn) but also by their temporal course of execution (e.g. different execution speeds like right turn and yielding compared to an undisturbed right turn). All possible maneuvers (events) form the sample space $\Omega = \{b^1, b^2, \dots, b^n\}$. For probabilistic maneuver estimation, we are interested in the discrete, conditional probability distribution

$$p(\{b \in \Omega : B = b\}|\mathbf{x}), \tag{1}$$

given the relevant current scene evidence $\mathbf{x}$. The scene evidence typically includes state information of the subject vehicles and possibly present other dynamic scene entities as well as the relevant spatial layout and context information (e.g. presence of road signs). Concrete probability values are given by

$$p(B = b^j|\mathbf{x}). \tag{2}$$

For the further course of this paper, the formal notations given in (1) and (2), are abbreviated with the common lower case notations $p(b|\mathbf{x})$ and $p(b^j|\mathbf{x})$.

B. Varying Sample Spaces Ω

The sample space Ω contains different events for each encountered scene, as different traffic participants and road-layouts lead to different possible events. The examples in Fig. 1 clarify this circumstance. For each of the illustrated four scenes the blue subject vehicle has different maneuver alternatives, hence varying sample spaces Ω. For Fig. 1 a) Ω is mainly composed of the three path alternatives (left, right or straight) for an undisturbed, free drive. In b), the left turn can be considerably influenced by the green car, e.g. slowing down to yield, stopping completely to stop or making a fast left turn in front of the green vehicle. Likewise, a slower preceding vehicle, as in c), can influence every possible maneuver significantly. A maneuver can even be completely missing, e.g. the right turn option in d).

Let us assume that $\tilde{\Omega}$ is the union of all maneuvers of each possibly encountered scene, ever. Consequently, the sample space for a current scene is given by

$$\Omega = \tilde{\Omega} \setminus \Omega_{\subset}, \tag{3}$$

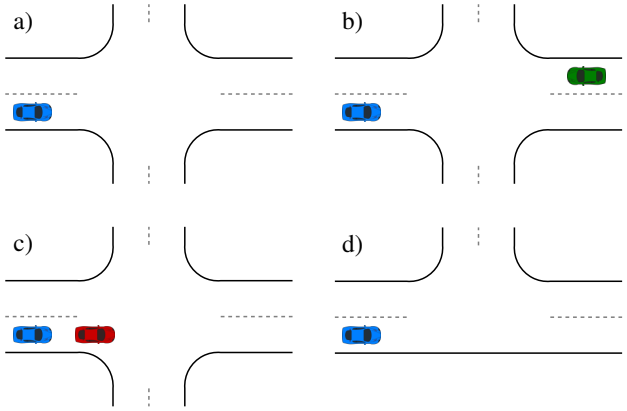

Fig. 1: 4 different scenarios, leading to 4 different sample spaces Ω.

with $\Omega_{\subset}$ containing all events which are incompatible with that scene (e.g. right turn for Fig. 1 d)).

C. Probability Theoretical Considerations and Problems

The first group of approaches discussed in I-A directly estimate the desired conditional probability $p(b|\mathbf{x})$, as defined in Section II-A. The second, more generative, group of discussed approaches naturally estimate the likelihood function $p(\mathbf{x}|b)$, as it is investigated which model explains the observed evidence $\mathbf{x}$ best. Probabilistic maneuver estimates can then be obtained using the Bayes' Rule

$$p(b|\mathbf{x}) = \frac{p(\mathbf{x}|b)p(b)}{p(\mathbf{x})} \propto p(\mathbf{x}|b)p(b). \tag{4}$$

The distribution $p(b)$ represents the prior probability of a certain maneuver, but is usually assumed to be uniform to not bias the maneuver estimation. The prior distribution $p(\mathbf{x})$ is most often used to ensure that the conditional probability estimates obtained from likelihood functions guarantee $\sum_{b\in\Omega} p(b|\mathbf{x}) = 1$.
This probability theoretical difference of the two groups of approaches explains why the second group can be more easily adapted to different possible maneuver alternatives ($\Omega \neq \tilde{\Omega}$). As the discrete maneuver alternative b is in the conditional of the likelihood $p(\mathbf{x}|b)$, an adaptation to the currently possible sample space can simply be acquired by not making likelihood predictions for $\bar{b} \in \Omega_{\subset}$. To the contrary a model directly estimating $p(b|\mathbf{x})$ cannot straightforwardly be adapted to generic traffic scenes. In order to use such discriminative models in generic traffic scenes, two possibilities come into mind:

1) Normalization: One model could be adapted, discriminating between all theoretically ever possible maneuvers $\tilde{\Omega}$. This approach suffers from two major drawbacks. First, the complexity of the needed classification algorithm rapidly increases with increasing $|\tilde{\Omega}|$. Furthermore, models trained for estimating $p(b|\mathbf{x}), b \in \tilde{\Omega}$ would also make probability estimates for $\bar{b} \in \Omega_{\subset}$. These can be removed from the resulting posterior conditional distribution by normalizing

$$p_{\text{norm}}(b|\mathbf{x}) = \frac{p(b|\mathbf{x})}{\sum_{b'\in\Omega} p(b'|\mathbf{x})}. \tag{5}$$

However, it is not guaranteed that the internal discrimination function of a model estimating $p(b|\mathbf{x}), b \in \tilde{\Omega}$ also yields valid predictions for the reduced sample space Ω. Additionally, special care has to be taken with regard to the evidence vector $\mathbf{x}$ as not all values might be existing in the current scene (e.g. the curve radii of a not present left turn or a relative velocity to a not existing preceding vehicle). Missing values pose problems for the majority of classification algorithms.

2) Specialization: For each possibly encountered scenario a specialized model can be created. However, this possibility becomes quickly infeasible since the number of models grows exponentially. In order to have an appropriate model for each theoretically occurring scene

$$M_{\text{spec}} = 2^{|\tilde{\Omega}|}, \tag{6}$$

models would be needed for the maneuver estimation of a single traffic participant.

In Section IV, we present a framework for making discriminative maneuver estimation that does not suffer from the addressed normalization and missing values problem. In addition, the number of needed models scales much more favorably compared to having specialized models for each possibly encountered scenario. First, however, the theoretical background is explained.

III. Pairwise Probability Coupling

The idea of assembling complicated multi-class classification problems from partial, more simpler classifiers has been widely discussed in pattern recognition literature. The need for such techniques is due to the fact that many classification algorithms are naturally restricted to binary problems. Thus, the question of how these can be used for multi-class problems needs to be addressed. Especially, when probability estimates of class memberships are of interest, straightforward voting schemes can no longer be applied.

First, we state some additional probability theoretical formalisms. In order to make binary, one-vs.-one predictions for multi-class classification problems, $\binom{|\Omega|}{2}$ binary subsets $\Omega^{k,l} \subset \Omega$ are instantiated. For each ob the $\Omega^{k,l}$ subsets a binary classifier is used determining the binomial distribution

$$p(\{b \in \Omega^{k,l} : B^{k,l} = b\}|\mathbf{x}) = p^{k,l}(b|\mathbf{x}), \tag{7}$$

with $B^{k,l}$ being a discrete random variable that can take on the respective events $B^{k,l} \in \Omega^{k,l} = \{b^k, b^l\}$.
We denote the concrete probabilities of the events in $\Omega^{k,l}$ happening as

$$p^{k,l}(B^{k,l} = b^k|\mathbf{x}) = p^{k,l}(b^k|\mathbf{x}), \tag{8}$$
$$p^{k,l}(B^{k,l} = b^l|\mathbf{x}) = p^{k,l}(b^l|\mathbf{x}). \tag{9}$$

Additionally, in (7), (8) and (9) we applied the same common lower case abbreviation, as introduced in Section II-A. These estimates form the basis for the subsequent probability coupling.

An approach for obtaining probabilistic multi-class estimates has been presented in [19]. The idea of their proposed pairwise probability coupling is based on optimizing the

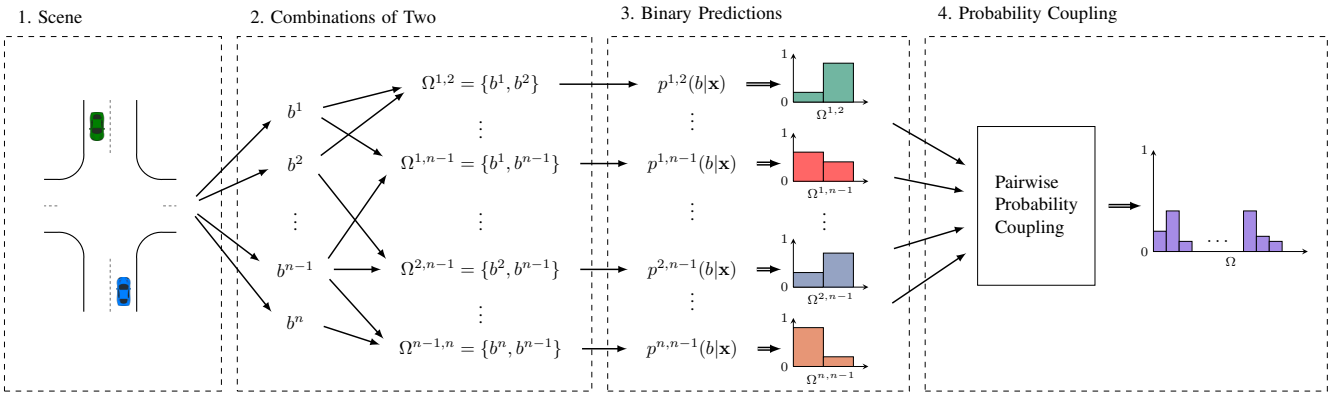

Fig. 2: Detailed system overview, see Section IV for more details.

approximated conditional probability distribution $p^*(b|\mathbf{x}) \approx p(b|\mathbf{x})$ of the desired target distribution with the objective

$$p^{k,l}(b^k|\mathbf{x}) = \frac{p^*(b^k|\mathbf{x})}{p^*(b^k|\mathbf{x}) + p^*(b^l|\mathbf{x})}. \tag{10}$$

This objective is based on the intuition that the probabilities $p^{k,l}(b^k|\mathbf{x})$ should take on the same value as $p^*(b^k|\mathbf{x})$ normalized to the subset $\Omega^{k,l}$.

The problem of finding appropriate probability values for $p^*(b|\mathbf{x})$ can be reformulated as

$$p^*(b^k|\mathbf{x}) = \sum_{l \in \Omega : l \neq k} \frac{p^*(b^k|\mathbf{x}) + p^*(b^l|\mathbf{x})}{n-1} p^{k,l}(b^k|\mathbf{x}), \tag{11}$$

$\forall k \in \Omega$, see [19]. This results in a linear equation system of the form

$$\mathbf{Q}\mathbf{p}^* = \mathbf{p}^*, \tag{12}$$

with

$$Q_{k,l} = \begin{cases} \sum_{s \in \Omega : s \neq k} p^{k,s}(b^k|\mathbf{x})/(n-1) & \text{if } k \neq l, \\ p^{k,l}(b^k|\mathbf{x})/(n-1) & \text{if } k = l \end{cases} \tag{13}$$

and $\mathbf{p}^*$ being a $(n \times 1)$ vector containing the n probability values of $p^*(b|\mathbf{x})$.

Furthermore, the authors of [19] showed that a unique, stationary probability distribution exists as long as $p^{k,l}(b|\mathbf{x}) > 0$. For practical implementations this constraint can easily be realized.

The discussed pairwise probability coupling is based on one-vs.-one predictions. Alternative techniques for obtaining multiclass predictions from binary classifiers are based on one-vs.-all predictions. One-vs.-all strategies have the advantage of needing fewer binary classifiers compared to one-vs.-one strategies. Nevertheless this strategy is is not further pursued due to the following reason. The objective of our approach is to make flexible maneuver prediction that can be adapted to the currently encountered scene. This means that the sample space Ω only contains maneuvers that are actually possible given the current scene. However, a one-vs.-all strategy does implicitly incorporate not possible maneuver alternatives $\Omega_\subset$, since these maneuvers are used as "all" examples during the training process. Hence, the binomial probability distribution of each binary one-vs.-all classifier k actually represents that b either corresponds to the associated behavior pattern b^k or that it corresponds to a behavior pattern of the set $\tilde{\Omega}_{\setminus k}$ without k. Thus, a one-vs.-all approach does consider all maneuver hypotheses, regardless of the current scene, and is therefore not further investigated.

IV. System Overview

In this section the approach for making discriminative maneuver estimations in generic traffic scenes is explained. The idea is to assemble the conditional probability distribution $p(b|\mathbf{x})$ for the sample space Ω of a currently encountered scene from partial, reusable classifiers, online. Therefore a binary, probabilistic maneuver estimation for each combination of two events in Ω is performed. The separately estimated class probabilities for each combination are then combined online using the pairwise probability coupling presented in Section III, resulting in the desired posterior conditional distribution $p(b|\mathbf{x})$.
The underlying binary classifiers of the proposed approach do only discriminate between two maneuvers k, l at a time. Therefore, they are reusable in different scenarios, where the overall number and types of maneuvers might vary, but the two maneuvers k and l are still possible events.

A prerequisite for such an approach is a pool of trained binary classifiers covering all combinations of two in $\tilde{\Omega}$, leading to

$$M = \binom{|\tilde{\Omega}|}{2}, \tag{14}$$

overall prototypical classifiers. This number scales much more favorably compared to that of specialized recognition systems, in (6).

The complete process for making probabilistic maneuver estimates using binary classifiers is illustrated in Fig. 2. A step-by-step explanation is given in the following:
First, the information on the currently encountered scene needs to be collected. This includes identifying dynamic objects, abstract context information and gathering information on the road layout using digital maps.
Second, based on the acquired scene information the n possible prototypical maneuvers are identified and selected. These form the sample space Ω with $b^i, i = 1, \ldots, n$. This

rule-based selection process is based on structural information like digital map data, e.g. drivable path alternatives and different temporal execution patterns provoked by other scene entities. Subsequently, all $\binom{|\Omega|}{2}$ binary combinations $\Omega^{k,l}$ are instantiated.

Third, for each combination, the corresponding binary classifier is selected. The relevant scene information, extracted in the first step, forms the feature vector $\mathbf{x}$. The class probabilities $p^{k,l}(b|\mathbf{x})$ for each combination are estimated using their respective classifier, separately.

Fourth, the individual binomial distributions are combined using the presented pairwise probability coupling. This results in the entire probability distribution $p(b|\mathbf{x})$ over the currently possible maneuvers $b \in \Omega$.

V. Experimental Results

Due to differing numbers of present traffic participants and road layouts intersection scenarios are particularly variable. Thus, we choose such intersection scenarios to evaluate our proposed approach. Two different evaluations for two different data sets are performed. The first evaluated scenario focuses on investigating the reusability of the partial, binary classifiers in changing environments. We demonstrate this on varying intersection layouts. For the second evaluation we focus on investigating the performance on a more complicated scenario with several dynamic traffic participants resulting in a larger number of possible maneuver alternatives. First, however, the used evaluation metrics are explained.

A. Evaluation Metrics

Maneuver estimation is affected by many uncertainties. In Section I, we argued that it is reasonable to treat the estimation problem probabilistically. Hence, it is also reasonable to evaluate the probabilistic output of the evaluated models. This can be done using ROC curves. These provide an objective metric for comparing the performances over the complete spectrum of *false positive* and *true positive* rates. As we have no special application in mind we use the meta metric for the ROC curves, the Area Under the Curve (AUC).

The AUC values are determined for different estimated Time To Intersections (*TTI*). This means that the maneuver estimation for a $TTI = 2s$ is triggered when the estimated *TTI* first fell below 2 seconds. In [3], we showed that assuming a constant velocity for the *TTI* calculation leads to conservative estimations for intersection approaching scenarios and thus it is also used in this paper.

B. Scenario I

In order to validate the reusability of the underlying binary classifiers, we use the data extracted in [20]. The data set consists of 3-and 4-way intersection approaches in an urban environment. No other dynamic scene entities which could possibly interfere with the intended maneuver were present. The four different types of possible intersection approaches are illustrated in Fig. 3.

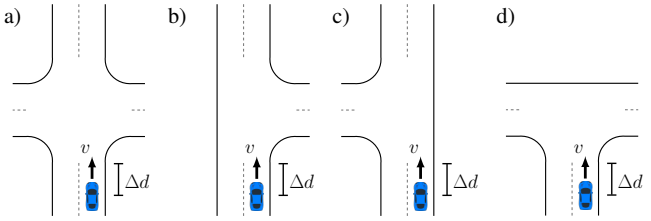

Fig. 3: 4- and 3-way intersection types present in the data set of [20] with the according possible maneuver alternatives (r = right turn, l = left turn, s = straight crossing). In total 112 cases (a) → 36 cases, b) → 20 cases, c) → 29 and d) → 27 cases) are examined.

1) Evaluated Models: We test the binary classifiers trained on 4-way intersection approaches for their reusability on 3-way intersections. The proposed pairwise probability coupling approach needs $\binom{|\Omega_a|}{2} = 3$ binary classifiers for handling 4-way intersections. For the three possible maneuver alternatives of Ω_a (right turn, left turn and straight crossing), the resulting binary subsets are identical to the sample spaces of the 3-way intersection crossings Ω_b, Ω_c and Ω_d. Hence, each 3-way intersection approach can be assessed using the respective binary classifier, originally created for 4-way intersection approaches.

We compare three models. The first model $\mathcal{M}_{\text{spec}}$ serves as reference. It is based on having a specialized classifier for each possible 3-way intersection b), c), d) in Fig. 3. These are trained using a 4-fold cross validation on the 3-way intersection approaches of the extracted data.

The second model $\mathcal{M}_{\text{4-way}}$, is based on the proposed pairwise probability coupling approach. It is trained only on the 4-way intersection approaches. Hence, the underlying binary classifiers of this model have not been confronted with any 3-way intersections.

The third model $\mathcal{M}_{\text{4/3-way}}$ is, again, based on the proposed pairwise probability coupling approach. However, additional training data from the 4-fold cross validation of the 3-way intersection data is used.

For such turning scenarios in free drive, logistic regression is successfully applied in [21] and [3]. Therefore, it is also the basis for all classifiers needed to determine the conditional distribution $p(b|\mathbf{x})$ in this paper.

Expressive information for distinguishing turning maneuvers are mainly composed of the distance to the intersection Δd and the velocity v of the subject vehicle, both indicated in Fig. 3. Additionally, whether the subject vehicle has right of way strongly influences the temporal execution of the intended maneuver. The priority information is encoded in a binary variable w. The scene evidence contains the following features

$$\mathbf{x} = [v, \Delta d, w]^T. \tag{15}$$

2) Evaluation: The resulting AUC curves for a *TTI* range from $0s$ to $2.5s$ are shown in Fig. 4. It can be seen that the model $\mathcal{M}_{\text{4-way}}$, never confronted with a 3-way intersection approach during training, almost reaches the performance of the specialized model $\mathcal{M}_{\text{spec}}$. This result emphasizes the

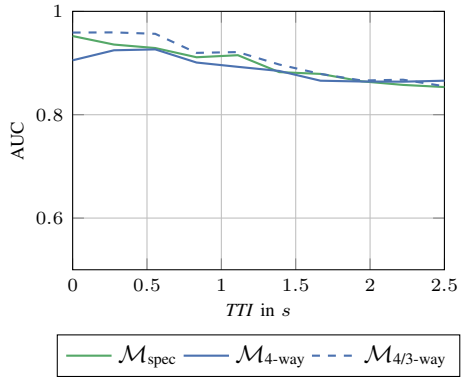

Fig. 4: AUC evolution for Scenario I.

possibility of reusing the same binary, partial classifiers in different scenarios. Furthermore, the results for $\mathcal{M}_{\text{4-way}}$ indicate that the generalization capabilities of reusable partial classifiers can exceed the performance of specialized systems. This is due to the reason that more training data, from different scenarios, can be used. E.g. in this evaluation the binary classifiers can be trained with data from 3-and 4-way intersection approaches.

C. Scenario II

The second evaluated scenario focuses on comparing the pairwise probability coupling approach on a more complicated scenario with the alternatives for making maneuver estimations in generic traffic scenes, discussed in Section II-C.

The used data set was recorded by rooftop mounted cameras, filming a three-way-crossing near Gothenburg, Sweden. Vehicle traces were generated using a Kalman filter approach [22]. Further information regarding the sensor setup and the detailed layout of the intersection can be found in [23].

Scenarios are extracted for a subject vehicle approaching the intersection, as illustrated by the blue vehicle in Fig. 5 b). The blue vehicle has two possible spatial path alternatives, a left turn and a straight crossing path. The temporal execution might be influenced by a possibly preceding vehicle, represented by the vehicle in red, or an oncoming vehicle with priority, represented by the vehicle in green.
For the left turn path, we define 5 prototypical spatio-temporal behavior patterns/maneuvers:

- "free drive",
- "influenced by preceding vehicle",
- "slow approach and yield to an oncoming vehicle",
- "full stop and yield to an oncoming vehicle",
- "turn shortly in front of an oncoming vehicle".

Additionally, 2 prototypical spatio-temporal behavior patterns/maneuvers are defined for the straight intersection path:

- "free drive",
- "influenced by preceding vehicle"

Hence, the union of maneuvers for the investigated scenarios form the sample space $\tilde{\Omega}$ with $|\tilde{\Omega}| = 7$. A detailed breakdown of the possible maneuver alternatives is shown in 5 a).

However, as the majority of defined possible maneuvers depend on the presence of either a preceding vehicle or an oncoming vehicle they will obviously not be possible if one of them or both are not present in a scene. Therefore, the presented scenarios are well suited for evaluating our proposed approach as not all encountered scenes lead to $\Omega = \tilde{\Omega}$.

1) Evaluated Models: In order to obtain the desired posterior probability distribution $p(b|\mathbf{x})$ using the proposed approach we need $\binom{7}{2} = 21$ prototypical binary classifiers. The model based on pairwise probability coupling is from now on referred to as $\mathcal{M}_{\text{coup}}$.

As we intend to benchmark our system we create three other models for estimating $p(b|\mathbf{x})$.
The first model is based on having specialized classifiers for all possible combinations of $\tilde{\Omega}$, as discussed in II-C. While this becomes quickly infeasible, for the restricted investigated test scenarios it is still manageable. The individual classifiers are summarized as $\mathcal{M}_{\text{spec}}$.
The second benchmark model is based on creating one classifier for the complete sample space $\tilde{\Omega}$ and normalizing over the possible events $b' \in \Omega$, as shown in (5). This model is referred to as $\mathcal{M}_{\text{norm}}$.
Additionally, a third model is created pursuing a generative approach based on probabilistic state predictions, as explained in Section II-C. For this purpose, we create nonlinear regression models for each maneuver based on commonly used Gaussian Processes ([8], [24]), with radial basis function kernels. The likelihood $p(\mathbf{x}|b)$ of the observed subject vehicle's velocity v_1 belonging to maneuver b, given the remaining scene evidence $\mathbf{x}_{\setminus v_1}$ is determined by

$$p(\mathbf{x}|b) \propto p(v_1|\mathbf{x}_{\setminus v_1}, b). \tag{16}$$

Subsequently, the desired conditional probability distribution can be acquired using (4), assuming a uniform prior. This system is referred to as $\mathcal{M}_{\text{gen}}$.

The classifiers used to determine $p(b|\mathbf{x})$ of models $\mathcal{M}_{\text{norm}}$ and $\mathcal{M}_{\text{spec}}$, as well as the binary classifiers for determining the distributions $p^{k,l}(b|\mathbf{x})$ are based on probabilistic support

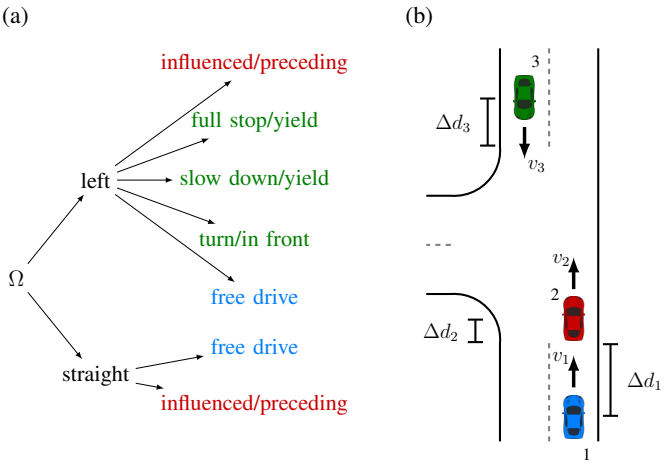

Fig. 5: a) The unfolding of the maneuver alternatives, starting from the spatial paths to the concrete prototypical temporal execution. b) The spatial-layout and the abstract position of the three possibly present traffic participants as well as the used features.

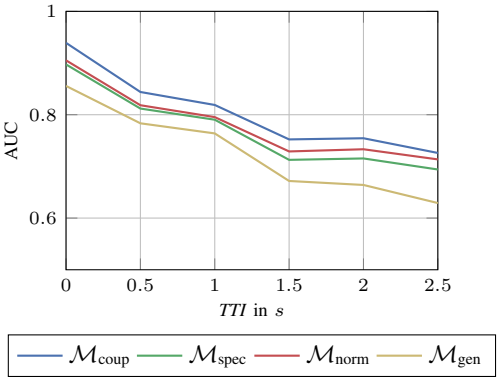

Fig. 6: AUC evolution for Scenario II.

vector machines with radial basis function kernels. The support vector machine is a discriminative kernel method making it a suitable candidate for comparison with the Gaussian processes, used for $\mathcal{M}_{\text{gen}}$. The kernels' length scales γ for the support vector machines as well as the Gaussian processes are set to the same value of 0.1 in order to improve comparability. Moreover, a neutral penalty parameter $C = 1$ is chosen for the support vector machines.

For the investigated scenarios the distances to the intersection and velocities, depicted in Fig. 5 are expressive information. Hence, the information for each traffic participant is given by

$$\mathbf{x}_1 = [v_1, \Delta d_1]^T, \mathbf{x}_2 = [v_2, \Delta d_2]^T, \mathbf{x}_3 = [v_3, \Delta d_3]^T. \quad (17)$$

The feature vector for each model is a result of the maneuver(s) and associated present traffic participants it is targeted at. E.g. the feature vector of $\mathcal{M}_{coup}$'s binary classifier discriminating between "left - free drive" and "left - influenced/preceding" is composed of $\mathbf{x}_1$ and $\mathbf{x}_2$. Another example would be the regression model of $\mathcal{M}_{gen}$ estimating the velocity expected with maneuver "left - slow down/yield"; its feature vector is composed of $\mathbf{x}_1$ and $\mathbf{x}_3$. An exception is constituted by $\mathcal{M}_{norm}$. Its feature vector is the union of all possible features $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$. Missing values as a consequence of not present traffic participants are substituted by default values.

2) Evaluation: On average, 440 cases could be extracted for each investigated *TTI*. [3] The extracted data is split into training and test data using a 4-fold stratified cross validation.

The resulting AUC curves for a *TTI* range from $0s$ to $2.5s$ are shown in Fig. 6. The model based on our proposed pairwise probability coupling $\mathcal{M}_{\text{coup}}$ shows the overall best performance. Model $\mathcal{M}_{\text{norm}}$ shows comparable performance in some areas but nevertheless does not reach the overall performance of $\mathcal{M}_{\text{coup}}$. It is expected that the performance deteriorates further for increasing $\tilde{\Omega}$. Surprisingly, the performance of $\mathcal{M}_{\text{spec}}$ does not reach the performance levels of our proposed approach. This can be caused by the high degree of specialization and the associated decreased number of training examples, as already observed in Section V-B. The generative approach shows the poorest performance, compared to the competing models adapted in a discriminative manner.

[3]Due to faulty traces or miss detections not every approach is usable for each investigated *TTI*; resulting in a slightly varying number of usable cases.

VI. CONCLUSION & FUTURE WORK

In this paper we presented a general approach for making single-entity-based maneuver predictions in generic traffic scenes. For this purpose we used a popular technique called pairwise probability coupling for obtaining multi-class class membership probabilities from probabilistic binary classifiers. We argue that we are able to assemble a probabilistic maneuver estimation from these discriminative binary and reusable classifiers. The benefits of our proposed approach have been demonstrated on real-world data sets. Our approach is not only able to outperform flexible, generative maneuver estimations, it surpasses specialized systems due to the reduced model complexities of the partial classifiers.

As a next step, we would like to evaluate the benefits of a temporal integration of successive states, e.g. using a deep learning architecture based on LSTM networks. Additionally, we will combine the herein proposed single-entity-maneuver estimation with our proposed interaction-aware situation assessment framework presented in [25]. Individual maneuver predictions form the basis of the therein composed decomposition. These can be provided by the approach presented in this paper.

ACKNOWLEDGMENTS

We would like to thank Professor Michael Felsberg at Linköpings Universitet for providing the data for the evaluated intersection approaches in Section V-C.

REFERENCES

[1] Statistisches Bundesamt, "Unfallentwicklung auf deutschen Straßen 2012," Tech. Rep., 2013.

[2] S. Lefèvre, D. Vasquez, and C. Laugier, "A survey on motion prediction and risk assessment for intelligent vehicles," *ROBOMECH Journal*, vol. 1, no. 1, 2014. [Online]. Available: http://dx.doi.org/10.1186/s40648-014-0001-z

[3] S. Klingelschmitt, M. Platho, H.-M. Gross, V. Willert, and J. Eggert, "Combining Behavior and Situation Information for Reliably Estimating Multiple Intentions," in *IEEE Intelligent Vehicles Symposium Proceedings*, 2014, pp. 388–393.

[4] S. Lefèvre, C. Laugier, and J. Ibañez-Guzmán, "Exploiting map information for driver intention estimation at road intersections," in *IEEE Intelligent Vehicles Symposium Proceedings*, 2011, pp. 583–588.

[5] E. Ward and J. Folkesson, "Multi-classification of Driver Intentions in Yielding Scenarios," in *18th IEEE International Conference on Intelligent Transportation Systems*, 2015.

[6] S. Bonnin, F. Kummert, and J. Schmudderich, "A Generic Concept of a System for Predicting Driving Behaviors," in *15th IEEE International Conference on Intelligent Transportation Systems*, 2012, pp. 1803–1808.

[7] B. Morris, A. Doshi, and M. M. Trivedi, "Lane Change Intent Prediction for Driver Assistance: Design and On-Road Evaluation," in *IEEE Intelligent Vehicles Symposium Proceedings*, 2011, pp. 895–901.

[8] Q. Tran and J. Firl, "Online maneuver recognition and multimodal trajectory prediction for intersection assistance using non-parametric regression," in *IEEE Intelligent Vehicles Symposium Proceedings*, 2014, pp. 918–923.

[9] A. Bender, J. R. Ward, S. Worrall, and E. M. Nebot, "Predicting Driver Intent from Models of Naturalistic Driving," in *18th IEEE International Conference on Intelligent Transportation Systems*, 2015.

[10] C. Rodemerk, H. Winner, and R. Kastner, "Predicting the driver's turn intentions at urban intersections using context-based indicators," in *IEEE Intelligent Vehicles Symposium Proceedings*, 2015, pp. 964–969.

[11] M. Liebner, F. Klanner, M. Baumann, C. Ruhhammer, and C. Stiller, "Velocity-Based Driver Intent Inference at Urban Intersections in the Presence of Preceding Vehicles," *Intelligent Transportation Systems Magazine, IEEE*, vol. 5, no. 2, pp. 10–21, 2013.

[12] D. Petrich, T. Dang, G. Breuel, and C. Stiller, "Assessing map-based maneuver hypotheses using probabilistic methods and evidence theory," in *IEEE 17th International Conference on Intelligent Transportation Systems*, 2014, pp. 995–1002.

[13] C. Bishop and J. Lasserre, "Generative or discriminative? getting the best of both worlds," *Bayesian Statistics*, vol. 8, pp. 3–24, 2007. [Online]. Available: http://research.microsoft.com/en-us/um/people/cmbishop/abstracts/GD2007.htm

[14] A. Y. Ng and M. I. Jordan, "On Discriminative vs. Generative classifiers: A comparison of logistic regression and naive Bayes," in *Neural Information Processing Systems Proceedings*, 2001.

[15] B. Tang, S. Khokhar, and R. Gupta, "Turn Prediction at Generalized Intersections," in *IEEE Intelligent Vehicles Symposium Proceedings*, 2015, pp. 1399–1404.

[16] A. Jain, H. S. Koppula, S. Soh, B. Raghavan, A. Singh, and A. Saxena, "Brain4Cars: Car That Knows Before You Do via Sensory-Fusion Deep Learning Architecture," 2016. [Online]. Available: http://arxiv.org/abs/1601.00740

[17] Statistisches Bundesamt, "Fachserie 8 Reihe 7 - Verkhersunfälle 2014," Tech. Rep., 2015.

[18] J. Eggert, F. Damerow, and S. Klingelschmitt, "The Foresighted Driver Model," *IEEE Intelligent Vehicles Symposium Proceedings*, 2015.

[19] T.-f. Wu and R. C. Weng, "Probability Estimates for Multi-class Classification by Pairwise Coupling," *Journal of Machine Learning Research*, vol. 5, pp. 975–1005, 2004.

[20] S. Klingelschmitt and J. Eggert, "Using Context Information and Probabilistic Classification for Making Extended Long-Term Trajectory Predictions," in *18th IEEE International Conference on Intelligent Transportation Systems*, 2015.

[21] S. Klingelschmitt, F. Damerow, and J. Eggert, "Managing the Complexity of Inner-City Scenes : An Efficient Situation Hypotheses Selection Scheme," in *IEEE Intelligent Vehicles Symposium Proceedings*, 2015, pp. 1232–1239.

[22] B. Johansson, J. Wiklund, P.-E. Forssén, and G. Granlund, "Combining shadow detection and simulation for estimation of vehicle size and position," *Pattern Recognition Letters*, vol. 30, no. 8, pp. 751–759, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167865509000452

[23] J. Cano, J. Kovaceva, M. Lindman, M. Brannstrom, M. L. J. Cano, J. Kovaceva, and M. Brannstrom, "Automatic Incident Detection and Classification at Intersections," in *Volvo Car Corporation*, 2009. [Online]. Available: http://www-nrd.nhtsa.dot.gov/Pdf/ESV/esv21/09-0234.pdf

[24] A. Armand, D. Filliat, and J. Ibañez-Guzmán, "Modelling stop intersection approaches using Gaussian processes," in *16th IEEE International Conference on Intelligent Transportation Systems*, 2013, pp. 1650–1655.

[25] S. Klingelschmitt, F. Damerow, V. Willert, and J. Eggert, "Probabilistic Situation Assessment Framework for Multiple, Interacting Traffic Participants in Generic Traffic Scenes," in *IEEE Intelligent Vehicles Symposium Proceedings*, 2016.