2014 IEEE Intelligent Vehicles Symposium (IV)
June 8-11, 2014. Dearborn, Michigan, USA

A Lane Change Detection Approach using Feature Ranking with Maximized Predictive Power

Julian Schlechtriemen[1], Andreas Wedel[2], Joerg Hillenbrand[3], Gabi Breuel[4], Klaus-Dieter Kuhnert[5]

Abstract— **Risk estimation for the current traffic situation is crucial for safe autonomous driving systems. One part of the uncertainty in risk estimation is the behavior of the surrounding traffic participants. In this paper we focus on highway scenarios, where possible behaviors consist of a change in acceleration and lane change maneuvers. We present a novel approach for the recognition of lane change intentions of traffic participants.**

Our novel approach is an extension of the Naïve Bayesian approach and results in a generative model. It builds on the relations to the directly surrounding vehicles and to the static traffic environment. We obtain the conditional probabilities of all relevant features using Gaussian mixtures with a flexible number of components. We systematically reduce the number of features by selecting the most powerful ones. Furthermore we investigate the predictive power of each feature with respect to the time before a lane change event.

In a large scale experiment on real world data with over 160.781 samples collected on a test drive of 1100km we trained and validated our intention prediction model and achieved a significant improvement in the recognition performance of lane change intentions compared to current state of the art methods.

I. INTRODUCTION

Scene understanding and scene interpretation are crucial for the progress from assisted to semi-autonomous or even fully autonomous driving. Especially the task of risk assessment has to be thoroughly performed by an autonomous car as the driver is not in the loop. The intention recognition, whether an observed vehicle will change its lane in the near future is a special case of this challenge. In this paper we focus on estimating the probability for such a lane change event. Such probabilistic information is not only needed to determine the risk of the current situation, it might also be important for an adaptive cruise control system. It creates the possibility for an early reaction to cut-in maneuvers. This is not only desirable under comfort aspects but also helps to reduce fuel consumption by avoiding unnecessary acceleration.

Lane change recognition or prediction has already been investigated in various publications, see Sec. II. However, a large part of the described methods is only able to detect lane changes for the ego (system) vehicle. An often used feature in this context for example is the yaw rate, which is not directly measurable and can only be used as derived and noise affected feature for other traffic participants. On the other hand, more global information like the relation to the vehicle in front, the traffic flow on the neighboring lanes, the topology of the road, the distance to the next freeway exit ramp and other factors influence the behavior of human drivers and are not investigated in theses approaches. Another problem which has to be taken into account, is the uncertainty of the measured values which is significantly higher for the measurements of observed vehicles in comparison to the values relating to the ego vehicle. These aspects are implicitly handled with our proposed approach. We also investigate two important questions regarding lane change recognition:

- How early can a lane change be detected based on all the information available in a typical traffic scene?
- Which subset of all conceivable features shows the best trade-off between classification performance and a desirable small number of features?

In this paper, we propose a general bottom-up approach. In the first step, we define relational features for a defined set of surrounding vehicles. These include relative speeds, time gaps, time-to-collision and the required acceleration to avoid a collision with all related vehicles. On top, features like the absolute speed, the distances to the relative left and right lane marking, the curvature of the street, and the number of available lanes to the left and right are collected. After the collection step, we analyse the contribution of each feature to the classification results in different time-intervals before the lane change event.

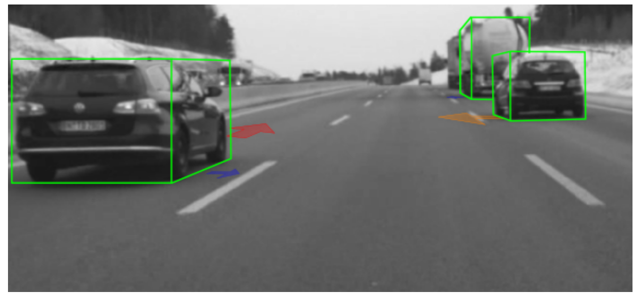


Fig. 1. Typical traffic scene on highways. Can a lane change be detected before lateral speed and offset to the lane center become significant?

[1]Julian Schlechtriemen is with Daimler AG, 71034 Böblingen, Germany julian.schlechtriemen@daimler.com
[2]Andreas Wedel is with Daimler AG, 71034 Böblingen, Germany andreas.wedel@daimler.com
[3]Jörg Hillenbrand is with Daimler AG, 70163 Sindelfingen, Germany joerg.hillenbrand@daimler.com
[4]Gabi Breuel is with Daimler AG, 71034 Böblingen Germany gabi.breuel@daimler.com
[5]Klaus-Dieter Kuhnert is with the Institute for Real-Time Learning Systems, University of Siegen, Germany kuhnert@fb12.uni-siegen.de

978-1-4799-3637-3/14/$31.00 ©2014 IEEE

To estimate the conditional probabilities of the different features, we analyze the respective distributions. In [1], the authors show that a single Gaussian distribution is often not precise enough. In order to meet the real-time and low memory-usage requirements of embedded systems, we chose to use the Gaussian mixture approach. We use the Schwarz Criterion [2] to ensure maximum entropy for our model selection.

The paper is structured as follows. In Section II related work is described. The perception model and the situation description parameters are derived in Section III. Finally, we give an overview on our experimental results in Section IV and conclude with Section V.

II. Related Work

The term "recognition of driving maneuvers" is strongly linked to terms like prediction, intention recognition and situation assessment. In this section, we review approaches addressing the "recognition of lane change maneuvers" before giving a short overview of our environment model and used features in the following section.

In [3], a Support Vector Machine(SVM) is used with a feature vector consisting of the lateral relative position of a car in the lane, the steering angle relative to the road curvature and the first derivatives of these two features. By using a Bradley Terry Model [4], a probabilistic output is generated from the result of the SVM. This probabilistic output is then processed by a Bayesian filter for the final lane change intention prediction. By using this filter the precision of the classification algorithm is significantly improved, mainly due to noise in the input data. In [5], Feed Forward Neural Networks (FFFN), Recurrent Neural Networks (RNN) and SVMs were compared using different combination of features consisting of lateral relations to the corresponding lane, steering angle, the Time To Collision (TTC) to the preceding car and the curvature of the road. It was shown, that the SVM achieved the best results followed by the RNN. Furthermore, the usage of the TTC to the Car in Front reduced the false positive and false negative rate and increased the prediction time. In [6], an object oriented Bayesian network is used for the recognition of lane changes. The feature set consists of the movement in relation to the assigned lane and a free-space representation. The approach predicts a lane change 0.6s earlier than a standard adaptive cruise control system. The authors point out that there is always a trade-off between an earlier detection and the false positive rate. In [7], a Case-based reasoning approach was used to detect cut-in maneuvers of vehicles into the lane of a system vehicle with a feature set based on the relative distance and velocity to the vehicle in front and the system vehicle. The final features were chosen using temporal abstraction and consisted of trend and level information of the distance and relative velocity. It was shown that the approach has the ability to detect lane change maneuvers until 2.3s before a car is in the target lane with a percentage of correct classifications of 79%.

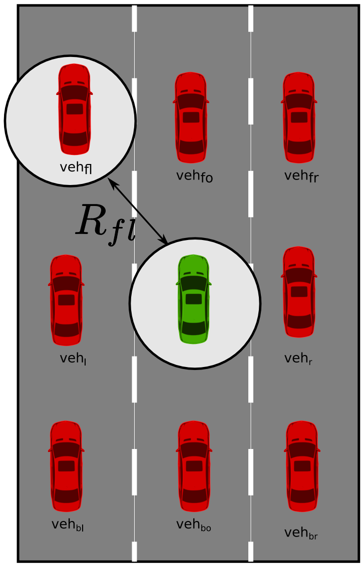


Fig. 2. Relation vectors R_r are generated to the cars which directly surround the observed vehicle o, for which we want to determine P_{LcL} P_{LcR} and P_{Flw}. In this example we show the vector R_{fl}, which is the relation between o and the vehicle positioned relative in (f)ront on the (l)eft lane.

The contributions of our paper are twofold. We consider our main contribution to be the investigation of the most relevant features for lane change recognition in highway scenarios. This is quite unlike previous strategies [3] [6] [7] where the feature set was chosen by experience. We investigate the relevance of the features against the time until a vehicle changes its lane assignment. The secondary contribution concerns the use of the Naïve Bayes Classifier using Gaussian mixtures for probability density estimation. We show, that this simple strategy is able to achieve the best results without the additional need of more sophisticated methods like Gaussian Filtering and the Hidden Markov Models. Because the term *prediction time* is not used consistently in the different approaches, we need to determine a time reference. All our time statements refer to the time-point when the center of a vehicle crosses the lane marking. This time-point is relatively earlier than the time reference used in [6] [7] but consistent with the research done in [3].

III. Proposed Approach

In the following section, we compare three classification algorithms in III-D, which need distribution estimates described in III-B. In III-C we explain the strategy used for the selection of the features, which were used as input for our classification algorithm, where the environment model consisting of all possible features is specified in III-A. The performance of the classifiers will be compared in Sec. IV.

A. Environment Model

All data used in the environment model is measured by an ego vehicle, which is equipped with a front facing stereo camera and several radar sensors to obtain a 360° field of view. It is obvious that the performance of the

TABLE I

DESCRIPTION OF THE EVALUATED FEATURES f FOR AN OBSERVED VEHICLE o

$\boldsymbol{R}$	$\boldsymbol{f}$	description	constraint
R_r	$d_{x,r}^{rel}$	longitudinal distance between o vs. related vehicle r	
	$v_{x,r}^{rel}$	longitudinal relative speed between o vs. related vehicle r	
	$a_{x,r}^{req}$	longitudinal decleration required for o to avoid a collision with the related vehicle r	constant acceleration
	$ttc_{x,r}$	time to a longitudinal collision between o vs related vehicle r	constant acceleration
	$\tau_{x,r}$	timegap between o and related vehicle r	
	$v_{y,r}$	lateral velocity of a related vehicle r	
R_{lane}	d_y^{ml}	lateral distance between the center of o and left marking	
	d_y^{mr}	distance between the center of o and right marking	
	d_{cl}	distance between center of o and centerline of assigned lane	
	$ttcr_y^{mr}$	time to cross the right marking of assigned lane for o	constant velocity
	$ttcr_y^{ml}$	time to cross the left marking of assigned lane for o	constant velocity
	a_y^{req}	required acceleration which is needed to stay in the current lane	constant acceleration
	ψ	angle of the observed vehicle relative to the direction of the lane	
	v_y	lateral speed of the observed vehicle relative to the lane	
	$nlane_r$	number of lanes on the right side of observed the vehicle	
	$nlane_l$	number of lanes on the left of observed the vehicle	
	t_l^m	type of marking left	(0 = dashed, 1 = solid)
	t_r^m	type of marking right	(0 = dashed, 1 = solid)
R_{infra}	c_0	curvature of the road	clothoid model
	d_x^a	distance to the next approach to the highway	
	d_x^e	distance to the next exit of the highway	
	v_x^a	speedlimit of the current highway section	

proposed algorithm strongly depends on the quality of the sensor measurements. Furthermore, the environment model is simplified by using a curvilinear coordinate system along the curvature of the road, in which all of the following measures are computed. To reduce the overall number of the possible features, which depend on relations to surrounding vehicles, we only take the direct neighbor vehicles into account, see Fig. 2.

Using the features from Tab. I, the feature vector F_{sit} is therefore defined as

$$F_{sit} = \begin{pmatrix} R_v & R_{lane} & R_{infra} \end{pmatrix}^\top \tag{1}$$

where R_v is the concatenation of the relations to the neighbouring vehicles,
$R_v = \begin{pmatrix} R_{fl} & R_{fo} & R_{fr} & R_l & R_r & R_{bl} & R_{bo} & R_{br} \end{pmatrix}^\top$. The feature vector F_{sit} is computed every *100ms* for every measured vehicle with sufficient measurement confidence of our system-vehicle and row-wise written into a database. Even if it is obvious that some of those features are highly dependent, we are interested in a subset of the information which performs best in the proposed classification algorithm. As above we denote the time where a lane marking is crossed by the center of a vehicle as the time point of a lane change. By an offline evaluation of the data, the time until the next lane change to the left lane t_{LcL} and a lane change to the right lane t_{LcR} were computed for every vehicle and later on used for labeling.

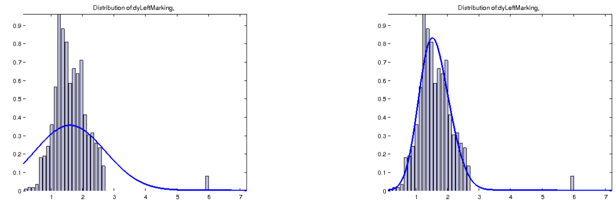
B. Distribution model

Although most implementations of the Naïve Bayes Classification algorithm (see III-D) assume a normal distribution of the variables, this is obviously not a valid assumption in our case. For example, $ttcr_y^{ml}$ for lane changing vehicles in the time interval 1s to 2s before the lane change event strongly depends on different driver profiles. Because the assumption of an univariate normal distribution is inappropriate for a single feature (see Fig. 5 for example), we propose to use a Gaussian mixture. This gives us the possibility to precisely estimate the probability density [8] without using a large amount of kernels, which would lead to a higher consumption of memory and processing time.

The use of Gaussian mixtures results in the question which number of kernels is optimal. This again depends on the distribution of the features for which the probability density function should be estimated. One option is to select the number of kernels after visual data inspection. The second way is the scoring of different models for the distribution estimate. In case of Gaussian mixture models this is a trade-off between the maximization of the likelihood function (which leads to a larger amount of kernels and therefore to the risk of overfitting) and the number of kernels. For the scoring of the models, we use the Bayesian Information Criterion (or Schwarz Criterion) [2]

$$BIC = -2 * \hat{l}_i(D) + k_i * \ln(n) \tag{2}$$

where $\hat{l}$ is the log-likelihood of the data D according to the current model i with n samples in D and k_i free parameters in i. The value of k_i is, in case of a Gaussian mixture the sum of $c-1$ probabilities, c means and c variance estimates, where c is the number of Gauss kernels. The model with the smaller BIC should be preferred. To increase the robustness of our approach against outliers, we use the DB-Scan algorithm [9] to estimate one single Gaussian mixture distribution for every value range where a continuous distribution exists (Alg. 1). This is required because the expectation maximization (EM)



(a) EM-Algorithm executed for all values

(b) EM-Algorithm executed for DB-SCAN clusters

Fig. 3. Difference between the models with the lowest BIC for EM Algorithm used for all data or for the DB-SCAN clusters

[10] algorithm is designed to work with missing values and cannot distinguish between wrongly assigned numerical values and the surrounding values of missing data, see Fig. 3.

Algorithm 1 Estimation of the Gaussian Mixture Distribution

procedure ESTIMATEGMDISTRIBUTION(D, u)
  [K, P_k] = dbscan(D)
  for $i \leftarrow 1, k$ **do**
    $[C_i, \Sigma_i, P_i] = EM_{BIC}(D_K)$
    result.push($C_e, \Sigma_e, P_e * P_k(i)$)
  end for
  return(result)
end procedure

procedure $EM_{BIC}(D)$
  $BIC = \infty$
  for $j \leftarrow 1, k_{max}$ **do**
    $[C_j, \Sigma_j, P_j] \leftarrow$ EM(D,j)
    $[BIC_i] \leftarrow$ computeBIC(C_j, Σ_j, P_j, D)
    if $BIC_i < BIC$ **then**
      $BIC \leftarrow BIC_i$
      $[C, \Sigma, P] \leftarrow [C_j, \Sigma_j, P_j]$
    end if
  end for
end procedure

C. Feature Selection

It is useful to find a small subset of features which maximize the predictive power, in order to implement a classification algorithm which can be executed in real-time and optimizes the results of the Naïve Bayes Classifier. To select this subset of features, which are useful to build a good predictor, different techniques can be applied [11]. Another problem which has to be solved is, how early a lane change manoeuver of an opponent vehicle can be detected before it crosses the lane markings based on all the potential features in F_{sit}. We propose the use of Area under the Curve AUC of the Receiver Operating Characteristic (ROC), which is useful for skewed distributions, because it is insensitive to changes in class distributions [12]. We define the classification task to the recognition of three maneuver classes: lane-following Flw, lane change to the left lane LcL and lane change to the right lane LcR. To handle the problem, that ROC graphs can only handle two-class problems, a ROC graph is generated for every class against the remaining two classes. The area under the curve for multiclass problems can be calculated according to [12] by

$$AUC_{total} = \sum_{c \in M} AUC_c * p(c) \tag{3}$$

where M is the aggregate of the maneuvers LcL, LcR and Flw. To get a statement over the time, we take data at different time intervals before a lane marking is crossed. We can now denote a function for a feature vector F which describes the predictive power of the classifier C at a time instance t_m before a lane change maneuver

$$AUC_t^C(t_m) = AUC_{total}^C(F^{t=t_m}) \tag{4}$$

D. Classification Algorithm

The Naïve Bayes algorithm is a generative learning algorithm. Its naivety is reasoned in the assumed conditional independence between the different features. While it has a higher asymptotic error than discriminative models, when the number of training samples increases, it approaches its asymptotic error much faster than a discriminative classification model [13]. Besides the classification performance on real world applications is often surprisingly good [14]. In the following two subsections we give an overview over the compared algorithms for probability estimation, where decision making can be realized for example by *the winner takes all*.

1) Naïve Bayes Algorithm: The probability of a sample $Z = f_1, f_2 ... f_n$ with the features f belonging to a maneuver m with $M = \{m | m \in \{Flw, LcL, LcR\}\}$ at a timepoint t is

$$p(m|Z_t) \propto p(Z_t|m)p(m) \tag{5}$$

and

$$p(Z_t|m) = \prod_i^n p(f_i|m) \tag{6}$$

2) Hidden Markov Model and Gaussian Filtering: Besides this, the probability $p(m|Z_{0:t})$ is also of interest. It can be computed using a Hidden Markov Model (HMM) [15] by setting the outputs of the Naïve Bayes algorithm as emissions of the HMM, see Fig. 4.

$$E^m = p(m|Z_t) \tag{7}$$

The prediction step is then defined as:

$$\bar{p}(S_t^k) = \sum_M p(S_t = S^k | S_{t-1} = S^m) p(S_{t-1}^m) \tag{8}$$

and the update step:

$$p(S_t^k) = \eta p(E_t^\Sigma | S_t = S^k) \bar{p}(S_t^k) \tag{9}$$

with:

$$p(E_t^\Sigma | S_t = S^k) = \sum_M p(S_t = S^k | E_t^m) p(E_t^m) \tag{10}$$

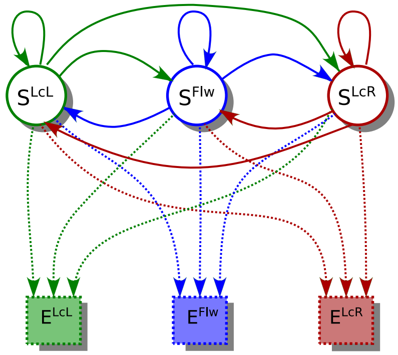


Fig. 4. Hidden Markov Model for the three maneuver classes

The conditional state transition probabilities of Eqn. 8 are stored in the matrix $\boldsymbol{T}$ and the conditional emission probabilities of Eqn. 10 in the matrix $\boldsymbol{E}$. For the proposed approach of Bayesian filtering in [3], $\boldsymbol{E}$ is set to the identity matrix $\boldsymbol{I}$.

IV. Results

We base our investigations on a database consisting of feature vectors F_{sit} collected from 13 hours on German highways. Our experimental vehicle was equipped with a sensor fusion using automotive sensors consisting of a front-facing stereo camera, front- and back-facing long range radar sensors, and two sensors at the left and right side of the vehicle. With a cycle-time of 100ms, a feature vector for every observed and tracked vehicle using a single track model [16] was written to the database. The database consists of a total of 160781 samples. Labeling vehicles $2s$ before a lane change event as positives samples results in a-priori class probabilities of $P_{LcL} = 0.0051$, $P_{LcR} = 0.0037$ and $P_{FLW} = 0.9911$.

Due to the problem that positive examples are very rare, evaluation metrics like $accuracy$, $precision$ and the F_1 measure are not well-conditioned to evaluate the performance of the classifier [17], because even a classification algorithm predicting every measured sample of a vehicle as *Flw* would lead to an accuracy of 99,1% (corresponding to the probability of P_{Flw}). Some authors have therefore evaluated the performance of their classification algorithm on nearly equally distributed classes [3], [7]. In order to make use of the non-equal distribution of our data set, we propose the use of a balanced precision measure defined according to the definition of balanced accuracy in [18] as

$$precision^{bal} = \frac{TPR}{TPR + FPR} \tag{11}$$

and the balanced F_1 measure defined as

$$F_1^{bal} = 2 * \frac{precision^{bal} * recall}{precision^{bal} + recall} \tag{12}$$

where TPR denotes the *true positive rate* and FPR the *false positive rate*. Both measures are independent from class skew and are directly comparable to the results of $precision$ and F_1 of test data with equal distributed classes.

TABLE II

Predictive Power of features f and their Spearman's rank correlation coefficient ρ to the chosen features

f	t_{max}	$\rho\left(v_{x,fo}^{rel}\right)$	$\rho\left(v_y\right)$	$\rho\left(d_y^{cl}\right)$
$v_{x,fo}^{rel}$	2.2	1	-0.05	-0.07
v_y	2.0	-0.05	1	0.15
$ttcr_y^{ml}$	1.8	0.03	-0.65	-0.19
a_y^{req}	1.8	0.05	-0.96	-0.17
d_y^{cl}	1.0	-0.07	0.15	1
d_y^{mr}	1.0	-0.09	0.14	0.88
d_y^{ml}	1.0	0.06	-0.12	-0.88
$ttcr_y^{mr}$	0.8	-0.06	0.69	0.2
$a_{x,fo}^{req}$	0.8	0.62	-0.06	-0.06

We divide our evaluation into the two parts: feature evaluation and classification performance.

A. Feature Evaluation

In the first experiment, we approximated the probability density functions for every feature and analyzed its predictive power. For this, we used single variable classifiers and computed $AUC_t^{C_f}(t_m)$ values for every single variable classifier C_f for the time interval $[0s, 15s]$, see also III-C. By selecting only values bigger than $AUC_t^{C}(t_{max}) > AUC_{min}$, we were able to compute a timepoint t_{max} for which a specific feature loses its predictive power for our lane change recognition algorithm. For the result in Tab. II, we set $AUC_{min} = 0.7$ and sorted the features according to their classification contribution.

One can clearly see from the results, that the relative velocity to the front vehicle, $v_{x,fo}^{rel}$, contributes significantly already 2.2 seconds before the lane change event and the lateral velocity of the vehicle with respect to its lane, v_y, contributes 2.0 seconds prior to the lane change. The following features $ttcr_y^{ml}$ and a_y^{req} are clearly correlated to the lateral velocity and therefore it is not surprising, that they contribute similarly. The lateral displacement d_y^{cl} contributes significantly up to 1.0 second before the lane change event. All remaining features show less predictive power.

Therefore, we focus in the following on the three most valuable features $v_{x,o}^{rel}$, v_y, and d_y^{cl} as input for our classification algorithm. Fig. 5 depicts the probability density functions for the three classes LcL, Flw, and LcR for these chosen features.

B. Classification Performance

In the second experiment, we evaluated the classification performance when combining the three features $v_{x,o}^{rel}$, v_y, and $d_y^c l$. Tab. III shows the result using the Naïve Bayes algorithm using cross-validation with a *winner takes all* strategy. One can clearly see, that a balanced performance score should be used in order to compare the results with results from other papers. As explained above, the precision does increase using the balanced score which can be explained with the a-priori class probabilities from the samples. The accuracy shows

TABLE III

EVALUATION OF THE NAÏVE BAYES ALGORITHM USING 2-FOLD CROSS-VALIDATION

$fold$	$precision_{bal}^{LcL}$	$recall^{LcL}$	$F1_{bal}^{LcL}$	$precision_{bal}^{LcR}$	$recall^{LcR}$	$F1_{bal}^{LcR}$	$accuracy_{bal}$
1	0.99	0.77	0.87	0.99	0.95	0.97	0.89
2	0.99	0.73	0.84	0.99	0.85	0.92	0.85

	$precision^{LcL}$	$recall^{LcL}$	$F1^{LcL}$	$precision^{LcR}$	$recall^{LcR}$	$F1^{LcR}$	$accuracy$
1	0.43	0.77	0.55	0.17	0.95	0.29	0.99
2	0.37	0.73	0.49	0.39	0.85	0.53	0.99

a small decrease as expected, because of the class skew with $P_{Flw} = 0.991$ and therefore a dominating behaviour of Flw in the evaluation. We then compared the Naïve Bayes algorithm with Gaussian Filtering and the HMM approach, see also III-D. For this, we reduced the output of the Naïve Bayes algorithm via the winner takes all strategy to a binary output. This output for each class has then been used to train the Emission Matrix $\boldsymbol{E}$ of our HMM:

$$\boldsymbol{E} = \begin{array}{c} \\ S^{Flw} \\ S^{Lcl} \\ S^{Lcr} \end{array} \begin{array}{c} \begin{array}{ccc} E^{Flw} & E^{Lcl} & E^{Lcr} \end{array} \\ \begin{pmatrix} 0.97 & 0.01 & 0.02 \\ 0.35 & 0.60 & 0.05 \\ 0.20 & 0.03 & 0.77 \end{pmatrix} \end{array} \tag{13}$$

and accordingly, we determined the state transition matrix:

$$\boldsymbol{T} = \begin{array}{c} \\ S_{t-1}^{Flw} \\ S_{t-1}^{Lcl} \\ S_{t-1}^{Lcr} \end{array} \begin{array}{c} \begin{array}{ccc} S_t^{Flw} & S_t^{Lcl} & S_t^{Lcr} \end{array} \\ \begin{pmatrix} 0.99 & 0.0003 & 0.0003 \\ 0.14 & 0.86 & 0 \\ 0.15 & 0 & 0.85 \end{pmatrix} \end{array}. \tag{14}$$

Fig. 6 shows the result from the Naïve Bayes Classificator with the results of the Gaussian Filtering using $recall$ and $precision^{bal}$ at different timepoints before a lane change using the winner takes all strategy. As can be seen in the figure, the Naïve Bayes Algorithm shows better results without the use of Gaussian Filtering and even the use of a HMM in Fig. 6 doesn't improve the classification performance. Reasons for the missing benefit of the HMM and the Gaussian Filtering can be expected by the already Kalman-Filtered lane and vehicle data, such that a filter in the domain does not extract additional information. The predictive power of the proposed Naïve Bayes Algorithm is depicted in Fig. 7 using the Receiver Operating Characteristic.

V. CONCLUSIONS

This paper presented a novel approach for the recognition of lane change events utilizing a feature set which maximizes the predictive power. In a large scale experiment, we showed that our proposed approach is able to precisely detect lane changes of other traffic participants up to 2.2s before the lane assignment changes. We showed significant improvement in the precision of the classification problem compared to approaches, which focus only on the lateral behavior of vehicles [11]. By investigating the most significant features, we also showed, that the lateral velocity relative to the lane, the relative velocity of the preceding car and the lateral offset to the lane center are the most discrimant features for lane change recognition. Future work involves the calibration of the probabilities of the Naïve Bayes classifier [19] to obtain more interpretable and better weighted values between the manoveurs LcL and LcR.

REFERENCES

[1] G. John and P. Langley. Estimating continuous distributions in bayesian classifiers. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, UAI'95, pages 338–345, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc.

[2] G. Schwarz. Estimating the dimension of a model. *The annals of statistics*, (2):461 – 464, 1978.

[3] P. Kumar, M. Perrollaz, S. Lefèvre, and C. Laugier. Learning-based approach for online lane change intention prediction. In *IEEE Intelligent Vehicles Symposium*, Gold Coast, Australie, 2013.

[4] T. Huang, R. Weng, and C. Lin. Generalized bradley-terry models and multi-class probability estimates. *The Journal of Machine Learning Research*, 7:85–115, 2006.

[5] U. Dogan, J. Edelbrunner, and I. Iossifidis. Autonomous driving: A comparison of machine learning techniques by means of the prediction of lane change behavior. In *Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on*, pages 1837–1843, 2011.

[6] D. Kasper, G. Weidl, Thao Dang, G. Breuel, A. Tamke, and W. Rosenstiel. Object-oriented bayesian networks for detection of lane change maneuvers. In *Intelligent Vehicles Symposium (IV), 2011 IEEE*, pages 673–678, 2011.

[7] R. Graf, H. Deusch, M. Fritzsche, and K. Dietmayer. A learning concept for behavior prediction in traffic situations. In *Intelligent Vehicles Symposium (IV), 2013 IEEE*, pages 672–677, 2013.

[8] G. Mclachlan and D. Peel. *Finite Mixture Models*. Wiley Series in Probability and Statistics. TODO, October 2000.

[9] M. Ester, H. Kriegel, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. pages 226 –231. AAAI Press, 1996.

[10] J. Bilmes. A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. Technical report, 1997.

[11] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3:1157–1182, March 2003.

[12] T. Fawcett. An introduction to roc analysis. *Pattern Recogn. Lett.*, 27(8):861–874, June 2006.

[13] Andrew Y. Ng and Michael I. Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in neural information processing systems*, 14:841, 2002.

[14] H. Zhang. The Optimality of Naive Bayes. In Valerie Barr and Zdravko Markov, editors, *FLAIRS Conference*. AAAI Press, 2004.

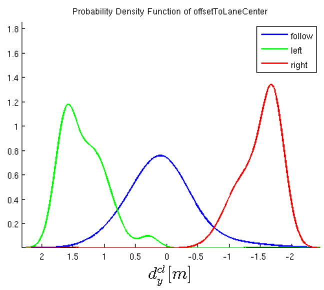
[15] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005.

[16] A. Barth and U. Franke. Where Will the Oncoming Vehicle be the Next Second? In *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 1068–1073, Eindhoven, June 2008. Springer.

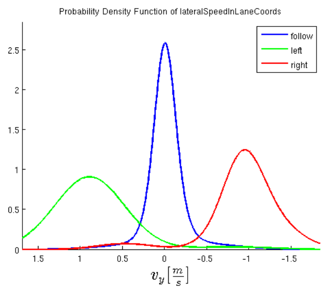
[17] M. Kubat and S. Matwin. Addressing the curse of imbalanced training sets: One-sided selection. In *In Proceedings of the Fourteenth International Conference on Machine Learning*, pages 179–186. Morgan Kaufmann, 1997.

[18] K.H. Brodersen, S. Cheng, K.E. Stephan, and J.M. Buhmann. The balanced accuracy and its posterior distribution. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 3121–3124, 2010.

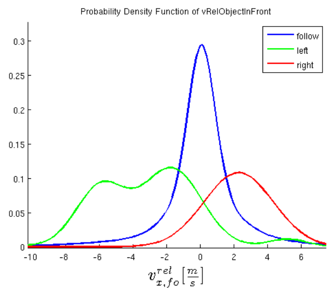
[19] B. Zadrozny and C. Elkan. Transforming classifier scores into accurate multiclass probability estimates. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 694–699. ACM, 2002.



(a) $d_y^c l$

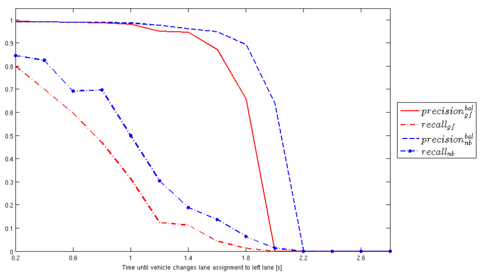


(b) v_y

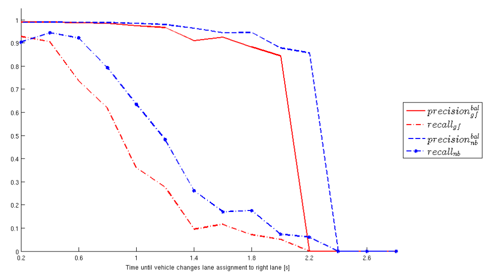


(c) $v_{x,o}^{rel}$

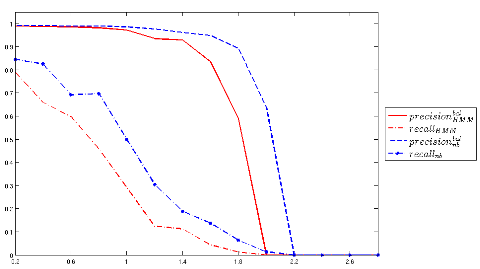
Fig. 5. Probability density functions of the different features, LcL (green), Flw (blue) and LcR (red). One can see, that a good seperation between the the three classes is possible for each of the chosen features



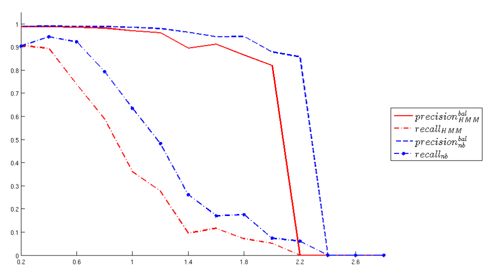
(a) $LcL(Gaussian filtering)$



(b) $LcR(Gaussian filtering)$

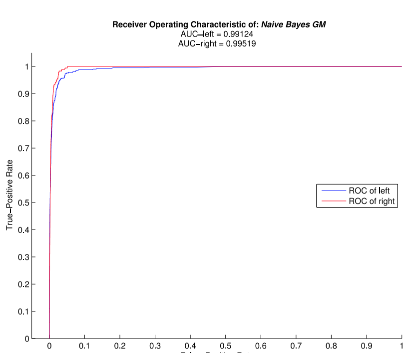


(c) $LcL(HMM)$

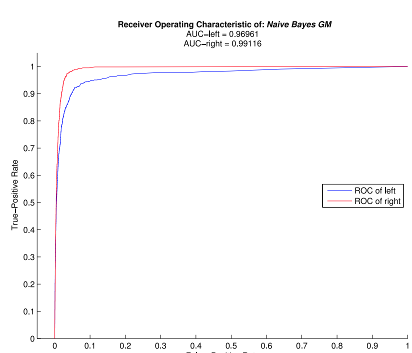


(d) $LcR(HMM)$

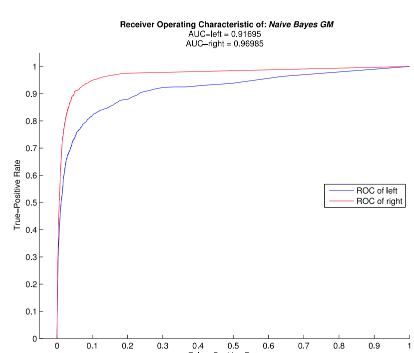
Fig. 6. Comparison between the Gaussian filtering (red) and the Naïve Bayes Algorithm (blue) in (a) and (b) for the recognition of the maneuver classes LcL and LcR plotted against the time before a lane change occurs. Same in (c) and (d) for the comparison between the HMM (red) and the Naïve Bayes Algorithm (blue). As can be seen in the graphs, the Naïve Bayes Algorithm is superior to the Gaussian Filtering and the results of the HMM in $recall$ and $precision_{bal}$



(a) ROC-0-1s



(b) ROC- 0-2s



(c) ROC- 0-3s

Fig. 7. Receiver Operating Characteristic showing the predictive power of our proposed Naïve Bayes Algorithm for samples in different time intervals before a lane-change