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

Rider Skill Identification by Probabilistic Segmentation into Motorcycle Maneuver Primitives

Nils Magiera, Herbert Janßen, Martin Ernst Heckmann, Hermann Winner

2016

Preprint:

This is an accepted article published in 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC 2016). The final authenticated version is available online at: https://doi.org/[DOI not available]

Rider Skill Identification by Probabilistic Segmentation into Motorcycle Maneuver Primitives

N. Magiera, H. Janssen, M. Heckmann, H. Winner

***Abstract*—As a result of the static and dynamic instabilities of a Powered-Two-Wheeler, the rider performs a highly demanding control task. Rider safety strongly depends on the individual abilities and skills of the rider. To improve the riders' skill level and reduce riding errors, safety trainings are well established. Additionally, safety systems and recently also advanced rider assistance systems help to avoid or mitigate accidents. While conventional rider training is limited to a small number of training scenarios in a controlled environment, safety systems are typically limited to specific situations (collision warning) or physical limits of the vehicle (ABS). We propose a method to identify riding errors based on a statistical rider model for cornering scenarios and estimate a personal rider skill score. The intention is to extend rider skill training beyond organized events and towards personal self-training. Automatic scoring of the cornering skill level has to take into account the high variability in speed and local curvature as well as the variety of different traffic situations that may be encountered during a ride. We suggest to split the complex driving task e.g. riding along a winding road, into smaller control tasks e.g. roll-into-corner, stable lean, and roll-out-of-corner and analyze them separately first, then their sequence and transitions. We evaluate various approaches based on Hidden Markov models that can split a complex task into smaller segments and show indicators for rider skill based on the best segmentation model.**

I. Introduction

Participating in public road traffic with a Powered-Two-Wheeler (PTW) is highly risky compared to other vehicles. Several studies and statistics show that the risk of getting involved in fatal crashes is over five times higher with regard to the number of registered vehicles or up to twenty times higher with regard to the travelled distance than for normal passenger car occupants. Beside fatal crashes, nearly 30.000 motorcycle accidents with injured traffic participants were officially recorded in 2014 in Germany. In 65 percent of the accidents, riding errors of the motorcycle rider were the primary reason for their occurrence [1]. Rider trainings to enhance individual riding skills as well as technical vehicle safety systems such as ABS are used in order to reduce the number of crashes. Both of these means try to prevent critical situations and facilitate their handling. An advantage of safety trainings is that they usually improve the general skill level of the rider by training specific scenarios, e.g. the vehicle-rider-interaction or riding stability. However, training events are limited to short time periods, a small number of training scenarios usually within a controlled environment, and pose a significant barrier to entry. In contrast, vehicle safety systems including advanced rider assistant systems (ARAS) assist the rider during everyday riding before or after detected critical events by smart interventions e.g. at tire friction limits. Unfortunately, such systems are not designed to interact with the rider and help to improve riding skills. Combining the information from modern ARAS with rider safety trainings in an online system could lead to a further reduction of the number of critical situations and accidents. The idea is to give motorcycle riders useful feedback about their individual skills and about the occurrence of critical situations for everyday riding and thus support self-improvement. The key requirement for such an application is a method to estimate a personal rider skill via a reliable personal model which can then be used to detect certain behaviors as indicators for riding errors.

N. Magiera is Research Associate and PhD-student at Technische Universität Darmstadt, Darmstadt, Germany (phone: +49-6151-16-24237; fax: +49-6151-16-24205; e-mail: magiera@fzd.tu-darmstadt.de)

H. Winner is Full Professor and Head of the Institute of Automotive Engineering at Technische Universität Darmstadt, Darmstadt, Germany (email: winner@fzd.tu-darmstadt.de)

H. Janssen is Principal Scientist at Honda Research Institute Europe GmbH, Offenbach, Germany (email: Herbert.Janssen@honda-ri.de)

M. Heckmann is Principal Scientist at Honda Research Institute Europe GmbH, Offenbach, Germany (email: Martin.Heckmann@honda-ri.de)

II. State of the Art

Human factors have become a common subject in the area of automotive research. However, the amount of relevant research work on modeling or characterizing driver skills and performance by recorded maneuver data is limited, especially for motorcycles. In general, the driver skill models proposed in literature consist of two steps [2]:

1. Extract predefined physical attributes and features directly linked with the driver's abilities and skills.
2. a) Estimate the driver skill as an output of classification algorithms that were trained on groups of drivers with driving skills assumed to be known.
   b) Estimate the driver skill from predefined feature mapping.

Zhang et al. [3],[4] and Tang [5] reported several proposals to identify lateral driving skills from lane change maneuvers. Double lane change maneuvers as well as a lane change in a curve were performed multiple times in a driving simulator by expert and novice drivers. The authors analyzed the frequency spectrum of the steering wheel angle during both maneuvers and reported that the frequency spectrum comprises a second high frequency peak for expert drivers. They proposed a method which uses the coefficients either of a discrete Fourier transform or of a wavelet transform of the steering wheel data to distinguish between

driver skill levels.

Chandrasiri et al. [2] modeled the longitudinal and lateral skills of drivers in a simulator study with 6 different curves. The test circuit was driven multiple times per driver. They extracted features via Principal Component Analysis (PCA) from multi-dimensional time-series data. In the second stage, a Support Vector Machine was trained to classify test subjects into "highly skilled" or "typically/low-skilled".

A motorcycle-specific method was proposed by Yoneta et al. [6]. Their target is to evaluate the turning skill of a motorcycle rider using time-series data from the vehicle's motion and the rider's head motion. A turning maneuver is detected if the yaw rate crosses a threshold for a certain minimum time Δt. In the subsequent analysis, three different scores to indicate the rider skill are calculated: a vehicle stability score, a turning performance score, and a head stability score. The scores are derived through filtering the data into different frequency bands and calculating their relations. Classes for low, medium, and high skill are defined for all three scores so that a final score can be measured on a scale with 27 steps. As this method has only been published as a patent, no results are publically available. We analyzed their method by re-implementing the proposed formulas and found that this method rewards corrective actions of the rider such as lean angle changes during steady cornering, although they are often the result of a riding error. The reason for this is that the corrective actions are in the same frequency range as entering and leaving the turn.

A statistical model for the rider's capabilities was introduced by Biral et al. in their work on an intelligent curve warning system [7],[8]. Their warning cost function is based on the distribution of x- and y-accelerations which were observed in the past for an individual motorcycle rider. The envelope of this distribution is used to represent the rider's limits. Similar other statistical models to characterize different drivers were reported in literature [9]-[12].

III. Definitions

A. Riding error

Various definitions of driving errors and erroneous behavior have been proposed. A very brief and general definition was given by Kleberg [13] who classified any deviation from a normal behavior as erroneous and unsafe. Reason [14] defined the term of erroneous behavior as the sum of incidents where a planned sequence of mental or physical activities will not lead to the intended result, considering the failure is not caused by external forces. Wang et al. [15] defined driving errors as the "deviation between the driving task required to be carried out and the dynamic response of the driver in a specified period of time in the operational driver-vehicle-road environment".

For our model, we assume that the desired task of the rider while riding on a winding road is to follow the road in a smooth way and act within the rider's personal limits. By the definitions mentioned above, a performed cornering maneuver can be classified as a riding error if we can detect sudden changes of the vehicle motion or if we can detect that the rider did not behave according to a learned model.

B. Riding Skill

In general, the formulation of 'Skill' is used to describe the proficiency to smoothly and adaptively handle complex tasks. Skill can be acquired through systematic and sustained training. Skill cannot be measured directly, thus typically indicators are used to describe it. We define riding skill as the ability or competence to perform complex riding tasks reproducibly, consistently, and without riding errors. Reproducibility and consistency describe riding skill in a positive way while riding errors indicate that the rider is overstrained and thus his skill level is too low for the riding situation.

IV. Driving Task Segmentation

A. Motivation for Maneuver Segmentation

In the referenced literature, the driving skill classification was commonly derived from repeated and very specialized maneuvers in a controlled and simulated environment. For our task, we have to account for naturalistic driving on public roads. The most obvious difference is that during a single naturalistic driving cycle usually no road segment is driven twice. Even though several data sets of the same driving cycle may be available, one has to account for variable environmental and traffic conditions. Apart from that, local maneuver models would only help to measure the skill level for a specific scenario and not for a general road segment. All these limitations prohibit the use of statistical models for a local maneuver as reference for error detection or skill classification.

Fig. 1 shows three exemplary cornering maneuvers (a) to (c) from a typical weekend trip of a motorcycle rider. The corresponding road parameters are summarized in Table 1.

Fig. 1 (a) and (b) both show a maneuver with a single curve, however the time duration differs significantly as a result of the different curvature and curve angle and length. Fig. 1 (c) represents an S-curve which is different from the previously mentioned single curve. Many more types of curves are found in our data, thus we propose to model maneuvers as a sequence of shorter recurring segments which can be concatenated to represent all types of curves.

B. Definition of Maneuver Primitives

The general idea to model complex context as a sequence of simpler segments can be found in various fields of research ranging from speech recognition to robotics.

TABLE 1
ROAD PARAMETER FOR CORNERING SCENARIOS IN FIG. 1

Maneuver	Curvature	Curve angle	Arc length	Time
(a) curve	-17/km	60 °	120 m	8 s
(b) curve	-6.5/km	100 °	350 m	15.5 s
(c) s-curve	20/km	40 °	60 m	4 s
	-40/km	220 °	150 m	11 s

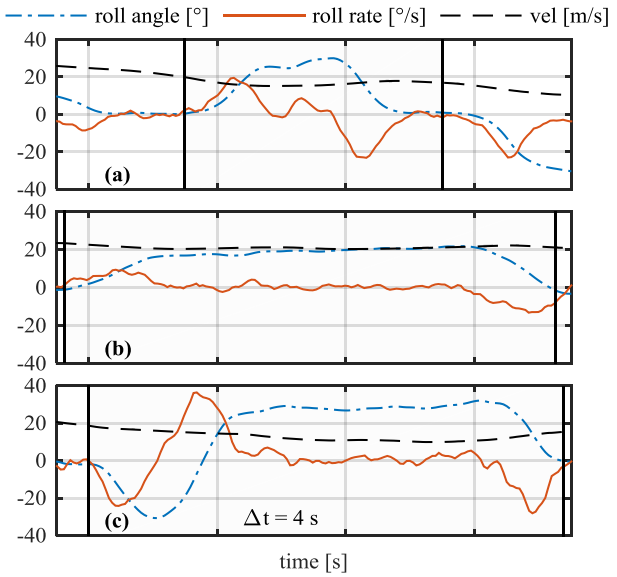


Fig. 1: Vehicle motion data for three exemplary cornering scenarios. Time axes are equally scaled. (a) represents a 60°-curve, (b) represents a 100°-curve and (c) an S-curve (40° left, 220° right)

We call our segments maneuver primitives as a reference to the term behavior primitive or sensory-motor primitive used in robotics or biology [23]. Studies viewing the human driver as the operator found that drivers tend to switch between multiple control strategies during a complex driving task [17]-[19]. For now, we restrict our analysis to continuous driving along a winding road without stops or traffic lights and with a focus on cornering scenarios. We propose to decompose cornering maneuvers based on their dynamics as this reflects the rider's control strategies in two classes: stationary and dynamic maneuver primitives. This includes a stationary primitive for straight road segments.

Table 2 shows an overview of the defined primitives and their connections. The primitive representing straight driving (S) is characterized by small roll angles close to zero as well as by small high-frequency changes of the roll rate around zero as the rider maintains dynamic stability. The second and third primitive represent parts of the cornering maneuver where the rider intends to hold the vehicle at a stable positive (SR) or negative (SL) roll angle. These segments are characterized by roll angle values that significantly deviate from zero as well as by roll rate values with small but high-frequency variations around zero. The remaining primitives describe the dynamic parts of the cornering

TABLE 2
MANEUVER PRIMITIVE DEFINITIONS

ID	Name	State No.	Previous state No.
S	straight driving	1	6,9
SR	stationary right	2	4,5,8
SL	stationary left	3	5,7,8
RI	rolling right in	4	1
LR	rolling left to right	5	3,7,8
RO	rolling right out	6	3,8
LI	rolling left in	7	1
RL	rolling right to left	8	2,4,5
LO	rolling left out	9	2,5

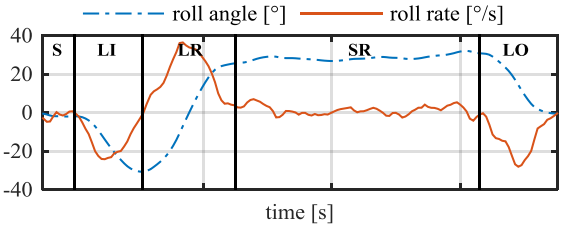


Fig. 2: Exemplary segmentation of the cornering scenario shown in Fig. 1 (c) into maneuver primitives

maneuver. They are defined as rolling to the right (RI, LR, RO) or rolling to the left (LI, LR, LO) and are all characterized by high values of the roll rate. By our proposed definition of maneuver primitives, we also imply an ordered sequence. Transitions between stationary maneuver primitives (S, SL, SR) can only be made by a previous change to one of the dynamic maneuver primitives (RI, RO, LR, LI, LO, RL). Allowed transitions are listed in Table 2. Fig. 2 shows an example for the segmentation of a single turn maneuver into 5 maneuver primitives.

V. PROBABILISTIC DATA SEGMENTATION MODELS

In this section, we introduce the general probabilistic model we use to segment the recorded data according to the definition of maneuver primitives (section IV.B). Subsequently, we present two different concepts how the proposed GMM-HMM (see definition below) structure can be applied to our problem.

A. Hidden Markov model with Gaussian mixture emissions

Inspired by speech- and gesture recognition methods, we consider Hidden Markov models with Gaussian mixture models (GMM-HMM) as a suitable solution for our problem. A Hidden Markov model (HMM) is a probabilistic model for a process with unobserved hidden states wherein it is assumed that each hidden state only depends on its predecessor [20]. As a consequence, an HMM is capable of capturing the sequential properties of a process through transition probabilities between consecutive states. While the HMM structure describes the sequence and ordering of states, Gaussian mixture models are used to describe the HMM output emission. Multivariate Gaussian probability density functions are utilized to estimate the probability that a continuous observation O_i (i: time step) is emitted from an unobservable hidden state S_j (j: state no.). In our case we consider maneuver primitives as hidden states and the recordings of the motorcycle dynamics as their observable emissions. The sequential and temporal properties of maneuver primitives are implicitly described in the state transition matrix A of the HMM. The recordings of the motorcycle dynamics are mapped to the non-observable maneuver primitives via a multivariate Gaussian distribution. Finding the best representation of our recorded data can be formulated as a search for the most likely sequence of maneuver primitives. In the case of the proposed HMM structure, this sequence can be computed using the Viterbi algorithm.

B. *Method I - Hierarchical Stepwise HMM (HS-HMM)*

Applying the first method, the most probable sequence of maneuver primitives is found by a stepwise classification of each data sample. We propose to structure the GMM-HMM as follows:

- Each stationary primitive is modeled as a single state in the HMM structure (i.e. 3x 1 state).
- Each dynamic primitive is modeled as sequence of two successive states - representing the rising and falling edge of the roll rate signal (i.e. 6x 2 states).
- Change points between stationary and dynamic maneuver primitives are modeled as independent states with zero self transition probability (i.e. 20x 1 state).

In total, we use 29 HMM models with 1 or 2 hidden states (35 hidden states in total) to represent 9 maneuver primitives and the transitions between them. The observation vector $O_{I,i}$ is a 4x1 vector containing the current roll angle φ, the roll rate $\dot{\varphi}$, and the roll acceleration $\ddot{\varphi}$ of the motorcycle at a time step i. The fourth element of the observation vector is defined as the time passed since the last zero crossing of the roll rate $\Delta t_{i,\dot{\varphi}=0}$. Data preprocessing and filtering is described in Section VII.

$$O_{I,i} = [\varphi_i \quad \dot{\varphi}_i \quad \ddot{\varphi}_i \quad \Delta t_{i,\dot{\varphi}=0}]^T \tag{1}$$

C. *Method II - Batch HMM (B-HMM)*

Applying the second method, the sequence is discovered from batches of data with variable length which are derived from a heuristic pre-segmentation model beforehand. The general idea behind method II is that processing the data in predefined batches rather than stepwise (sample per sample) can help to improve the classification of maneuver primitives by taking statistical features into account, see (2). We decided to perform the pre-segmentation by a simple and efficient threshold for the roll rate $\dot{\varphi}$ of the motorcycle. It is assumed that crossing a certain threshold indicates the riders' intention to change the desired orientation and motion of the motorcycle, which is also in line with [6]. However, the pre-segmentation based on the roll rate $\dot{\varphi}$ leads to an over-segmentation of the data. We thus propose to use the GMM-HMM to classify whether the segment belongs to the previous maneuver primitive or if it is the initial part of a new primitive. As a result of the pre-segmentation of the data, the structure of the HMM for method II is rather simple compared to method I:

- Stationary primitives as well as dynamic primitives are modeled as single states.

This can be understood as a combination of 9 single state HMMs. Due to the predetermined segment boundaries, the second method is very similar to common classification approaches. The observation vector $O_{II,i}$ has the size 9x1 and consists of statistical features for the roll angle φ and the roll rate $\dot{\varphi}$.

$$O_{II,i} = [\varphi_1 \; \varphi_{end} \; \bar{\varphi} \; \sigma(\varphi) \; \max(\varphi) \; \min(\varphi) \; \max|(\dot{\varphi})| \; \bar{\dot{\varphi}} \; \sigma(\dot{\varphi})]^T \tag{2}$$

VI. Performance Evaluation Criteria

Measuring the performance of time series segmentation is a difficult task and only a few criteria can be found in literature. In general, the segmentation of times series can be treated as a classification problem. To evaluate the performance of our algorithms, we use two criteria.

The first criterion describes the match between two time series. Each time step is classified separately and we compare whether the results of our algorithm correspond with reference annotations from a human annotator. Consequently, the stepwise match m_i between algorithm and manual annotation is calculated using the following rule.

$$m_i = \begin{cases} 0 \; if \; l_{human,i} \neq l_{algorithm,i} \\ 1 \; if \; l_{human,i} = l_{algorithm,i} \end{cases} \tag{3}$$

As a final result, we derive the overall match c as the arithmetic mean of c_i for the complete test sequence:

$$m = \frac{1}{n}\sum_{i=1}^{n} m_i \tag{4}$$

The second quality criterion is based on a method that Gensler and Sicks proposed [16]. This method measures whether the data segmentation is done at the right time or sample. Segmentation zones derived from the annotations of human annotators allow us to evaluate the performance around characteristic points of the time series instead of on all points. These characteristic points are called segmentation centers and represent the point in time when label a transitions to label b. As we expect temporal adjacency, lower and upper boundaries are defined wherein the segmentation can be considered good enough for our application. For our evaluation, we use fixed distances between boundaries and segmentation centers ranging from $\Delta t_{Zone} = 0$ s to $\Delta t_{Zone} = 1.0$ s. Based on the list of segmentation zones, a confusion matrix can be determined to evaluate the classification performance.

The top graph of Fig. 3 visualizes positive classification results. A segmentation point is classified as True Positive

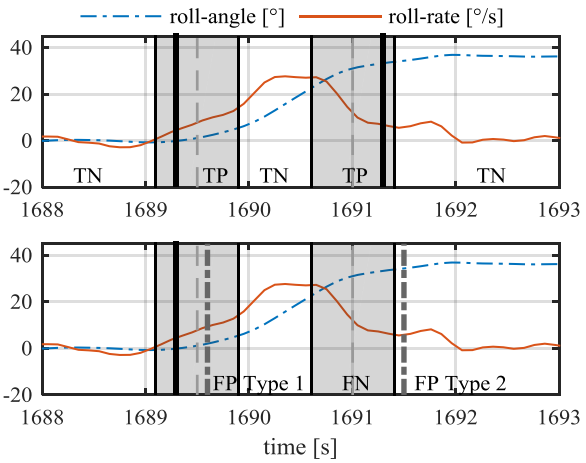


Fig. 3: Elements of the confusion matrix. Dashed grey lines: Ground truth segmentation center. Grey area: Segmentation zone. Black lines: True classification prediction (TP). Dash-dotted grey lines: False classification prediction (FP).

(TP) if it is found inside the corresponding segmentation zone. If no segmentation point is placed in the area outside of a segmentation zone, it is classified as True Negative (TN). In contrast to true classifications, the bottom graph of Fig. 3 shows negative results. If no segmentation is made by the algorithm inside a segmentation zone, a False Negative (FN) is counted. Vice versa, if a segmentation point is assigned outside a segmentation zone, this is considered as False Positive Type 2 (FP T2). Additionally, a False Positive of Type 1 is counted for each false prediction inside a segmentation zone.

VII. Experiments

In this section, we describe the data and the corresponding preprocessing to train and evaluate the proposed segmentation models. Next, the data annotation process is introduced and discussed. Finally, the results of the data segmentation that are achieved by both proposed HMM concepts are presented. All data processing is done using MATLAB and Kevin Murphy's HMM toolbox [21]. The theoretical foundation of the inference of the HMM with Gaussian mixture emissions is described in [22].

A. Data Basis and Preprocessing

With regard to the objective of evaluating the cornering performance of different riders during naturalistic driving, the segmentation algorithms have to be tested on a dataset which represents a typical motorcycle route. Therefore, driving experiments were conducted by two different riders on an exemplary route which consists of 420 maneuver primitives and has an overall length of 28 km. To allow the evaluation of local scenarios, the route has been driven multiple times per rider. Rider A is a professional test rider while rider B is an experienced motorcycle rider with no professional background. The vehicle used for the driving experiments is a Honda NC700X equipped with an inertial measurement unit (IMU) and additional sensors, e.g. steering torque sensors, to record the input of the rider.

All data is recorded at a sample rate of 100 Hz. However, the IMU measurements are internally preprocessed by a Kalman filter with a frequency of 10 Hz. We do not expect high frequency components in the data and thus it has been low-pass filtered with an FIR-filter and downsampled to a frequency of 10 Hz.

B. Data annotation procedure

The recorded data of the driving experiments are annotated by two different groups which had different insight into the data. This is inspired by the fact that in the context of professional racing a race-engineer typically has insights into the data as well as into video recordings to help the rider improve his performance. On the other side, an instructor of safety trainings only has an outside view on the riding to evaluate the performance of the rider.

1) Setup I

In the first setup, one human annotator was inspecting the measurement data of roll angle and roll rate together with the recorded video. This can be compared to the previously mentioned case of a professional race engineer. The annotation marks are set onto the graphs of the measurement recordings. In total, five runs of the professional test rider and two runs of the experienced motorcycle rider are annotated in this manner.

2) Setup II

In the second setup, five experienced motorcycle riders were advised to annotate the video recordings only. This group had no insight into the recorded measurement data and thus can be compared to the instructor during safety trainings. The annotation marks were set in Adobe Prelude. All annotators could scroll back and forth to find the best segmentation point. The annotations were then exported to MATLAB and synchronized with the recorded measurement data. This procedure was done for two runs of the experienced motorcycle rider.

C. Training and evaluation data split

1) Setup I

Both segmentation models are trained with the five annotated datasets from the professional test rider. The annotated data recordings of the second run from the experienced motorcycle rider are used for evaluation only.

2) Setup II

Only two runs are annotated for the second setup. However, for each run five different annotations are made by the human annotators. This results in a similarly large training dataset as for Setup I. The annotations for the first run of all five annotators were chosen to train both the HS-HMM as well as the B-HMM. The final evaluation of the performance is done by comparing the output of the segmentation model and the individual five annotations for the second run of the experienced motorcycle rider.

D. Results

1) Setup I

Table 3 shows the results for the match between the human reference annotations and the annotations automatically derived from both proposed segmentation models. An overall match of about 89 % is achieved by method I and II. However, there are significant differences between how the match is split across the maneuver primitives. The HS-HMM generally performs better for stationary maneuver primitives while the B-HMM more accurately represents dynamic maneuver primitives. It can be noted that the HS-HMM has very consistent scores of around 86 to 91 % across all cornering primitives (2-9) whereas the B-HMM ranges between 81 and 95 %. Fig. 4 shows the result for the True Positive Rate (TPR) and False Positive Rate (FPR) calculated from the confusion matrix introduced in Section VI. It can be clearly observed that for

TABLE 3
MATCH BETWEEN EXPERT ANNOTATIONS AND ALGORITHMS

	Total	Maneuver primitive specific				
	m	m_1	m_2	m_3	m_{4-6}	m_{7-9}
HS-HMM	**0.886**	**0.945**	**0.864**	**0.909**	0.893	0.889
B-HMM	0,884	0.917	0.805	0.807	**0.960**	**0.941**

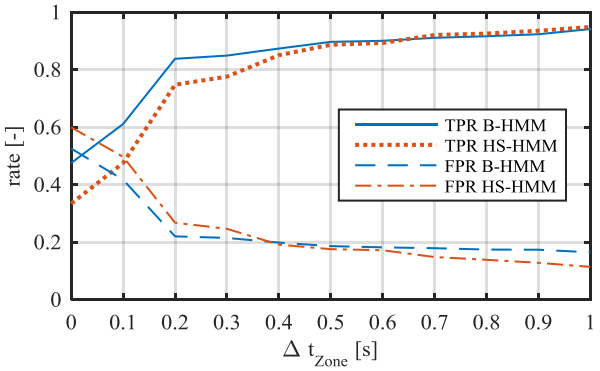


Fig. 4: TPR and FPR for segmentation models for Setup I

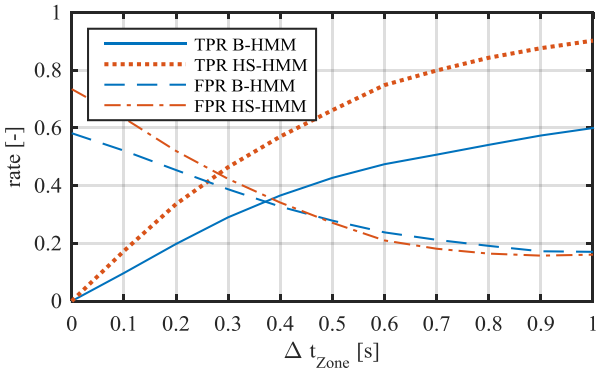


Fig. 5: TPR and FPR for segmentation models for setup II

segmentation zone confidence intervals Δt_{Zone} of 0.1 to 0.3 s the B-HMM with heuristic segmentation points outperforms the fully probabilistic HS-HMM by 5 to 10 % in terms of the TPR. However, when increasing the segmentation zone width, there is no noticeable difference for the TPR between HS-HMM and B-HMM. Both models converge against a TPR of 96 %. Remarkably, the purely probabilistic HS-HMM has a 5% lower FPR compared to the B-HMM.

2) Setup II

Table 4 shows the results for the match between the annotations (only from video recordings) of each human annotator and the segmentation by our models. The HS-HMM clearly outperforms the B-HMM. The mean match amounts to 82.3 % and 53.4 %, respectively. To better understand these values, Table 4 also shows the results if human annotators are compared individually. The comparison shows that annotations among human annotators are only to 83 % consistent. The evaluation of the segmentation zones for Setup II is shown in Fig. 5. Again, the HS-HMM performs noticeably better than the B-HMM in terms of the TPR. Interestingly, the False Positive Rate for both algorithms is quite similar for segmentation widths of 0.3 s or higher.

E. Discussion of the results

Our experiments lead to following conclusions. First, we note that recordings of the motorcycle dynamics help annotators to find recurring patterns in the data and thus an evaluation of the rider skill will be more precise. With regard to the recordings, these patterns are peak values, zero crossings, or linear trends for roll angle, roll rate, and roll acceleration. Both last mentioned quantities could not be easily seen in video recordings. An analysis of the segmentation points supports this hypothesis. The results are visualized in the histogram in Fig. 6. It shows the probability distribution of segmentation points at a specific roll rate $\dot{\varphi}$. It is clearly observable that the variance of the distribution for video annotation only is much higher than for annotations based on data and video recordings. We also note that the high variance of the segmentation points together with the fixed threshold-based pre-segmentation in the case of B-HMM leads to a significant drop of performance compared to the HS-HMM.

TABLE 4
MATCH BETWEEN HUMAN VIDEO ANNOTATIONS AND HMM PREDICTIONS

	Sub.1	Sub.2	Sub.3	Sub.4	Sub.5	HS-HMM	B-HMM
Sub.1	1	0.842	0.857	0.815	0.836	**0.826**	0.536
Sub.2		1	0.847	0.832	0.812	**0.803**	0.535
Sub.3			1	0.829	0.832	**0.861**	0.543
Sub.4				1	0.802	**0.812**	0.527
Sub.5					1	**0.806**	0.528

A second outcome is that when combining measurement data and video for annotation, the pure probabilistic HS-HMM performs equally well as the B-HMM when TPR, FPR, match, and their distribution across the maneuver primitives are considered together. It also has to be taken into account that some cornering scenarios are difficult to interpret. In this case, the probabilistic segmentation point of the HS-HMM can be classified as false even though the point is plausible. To avoid this problem in future evaluations, each ground truth segmentation center should be provided with its own segmentation zone width instead of a generally fixed one. For our analysis of the rider skill, we will use the segmentation based on the HS-HMM. We think that the purely probabilistic model is more flexible for future integration of further maneuver primitives, e.g. following other traffic participants, braking, or stopping.

The third and main result is that our automatic probabilistic segmentation using the HS-HMM is comparable to human segmentation performance (Table 4).

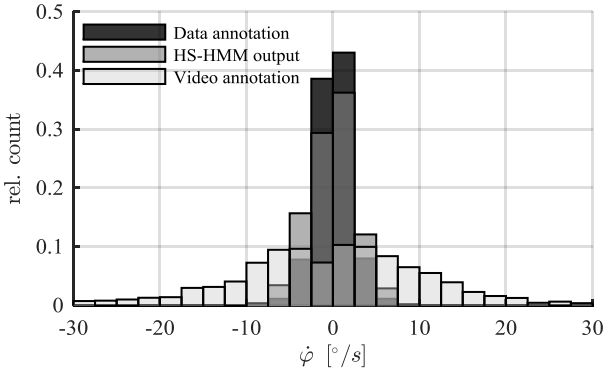


Fig. 6: Probability distribution of segmentation points at a specific roll rate $\dot{\varphi}$

VIII. Statistical Riding Skill Indicators

The segmentation of the cornering maneuvers into smaller recurring maneuver primitives enables to extract local scores for each type of primitives through models that are independent of time and locality. The statistical distribution of those primitive specific scores is defined as a riding skill indicator (RSI) which allows comparing riders of different skill levels. In this section, we shortly introduce two indicators which can be extracted through scores from dynamic or stationary maneuver primitives. A deeper analysis of scores for maneuver primitives as well as skill indicators is the topic of ongoing research.

A. *Rider skill indicator for stationary maneuver primitives*

In stationary cornering segments (SR & SL), the rider tries to keep the motorcycle dynamically stable around a certain roll angle depending on the road curvature and vehicle speed. Considering road construction regulations, we assume the curvature in a stationary segment to be constant. Hence, variations of the roll angle mainly result from a change of the vehicle speed, wrong trajectory planning, or other sudden instabilities. The latter two are observable as high frequency variations and peak values of the roll rate. In contrast, increasing or decreasing the velocity with a constant acceleration will result in a constant offset of the roll rate.

Based on these heuristics, we propose to use the high-pass filtered signal of the roll rate during stationary segments. For this evaluation, the roll rate is filtered using an FIR filter of order 10 with a rectangular window as we constrain our analysis to stationary segments with a minimum duration of 1 s. To avoid transient filter artifacts at segment boundaries, the signal is filtered before segmentation. The cut-off frequency is set to 0.2 Hz allowing slow periodical changes, e.g. due to decreasing velocity. The score $S_{SR,SL}$ for stationary primitives is defined as the standard deviation of the high-pass filtered roll rate $\dot{\varphi}$ (5).

$$S_{SR,SL} = \sigma(HP(\dot{\varphi}, f_c)) \tag{5}$$

Fig. 7 shows the cumulative distribution of scores for stationary segments (SR and SL) of single-curve scenarios.

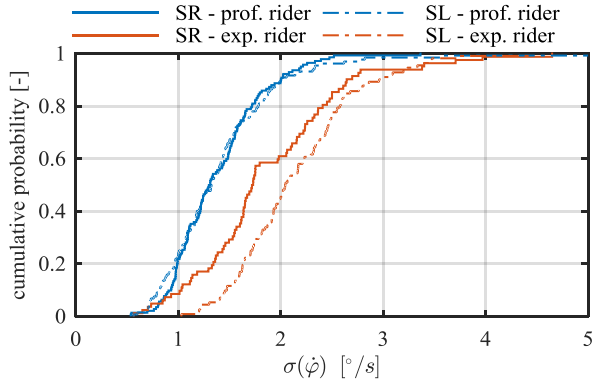


Fig. 7: Variance of the roll rate in stationary maneuver primitives (SL & SR) for single right and left curves. Constraints: length SR or SL > 1 s

B. *Rider Skill indicator for dynamic maneuver primitives*

For dynamic maneuver primitives, the assumption of a constant roll rate is not applicable, compare Fig. 2. From a macroscopic view, the signal is similar to prototypical shapes, e.g. it can be similar to a bell curve, a half period of a triangular oscillation, or a mixture of them. This characteristic can be explained as the result of the motorcycle dynamics (e.g. moment of inertia, tire properties) and the steering control of the rider trying to follow a path with increasing or decreasing curvature. Again, considering road construction regulations we assume that high frequency changes of the roll rate during maneuver primitives RI or LI occur due to wrong steering control of the rider, e.g. a wrongly timed or wrongly performed steering impulse. Hence, (5) can be applied in the same manner as for the stationary RSI. However, the cut-off frequency f_c needs to be adapted to the occurring frequency range in dynamic maneuver primitives. Based on a first analysis of roll rate in dynamic maneuver primitives, the high-pass filter cut-off frequency is set to 1.8 Hz. The main reason for this high cut-off frequency is to take into account different riding styles of both riders, e.g. smooth vs. sharp. The results for this evaluation of maneuver primitives RI and LI are shown in Fig. 8.

C. *Discussion*

Both distributions of scores (RSI) reveal differences between the professional test rider and the experienced motorcycle rider. It can be observed in Fig. 7 that the median of scores is lower for the professional test rider than for the experienced rider for which the gradient of the cumulative distribution is also more level. The first observation indicates a generally lower stabilization skill whereas the second observation can be interpreted as higher inconsistency. The inconsistency also becomes apparent by comparing SL and SR of the experienced rider. They are notably different, indicating that there is a difference in the left and right cornering behavior of the experienced rider. In contrast, the distribution of scores for the professional rider shows no significant difference between right and left cornering maneuvers. The results of this first analysis of maneuver primitives fit our expectation that a professional test rider has a higher riding skill. The interpretation of the

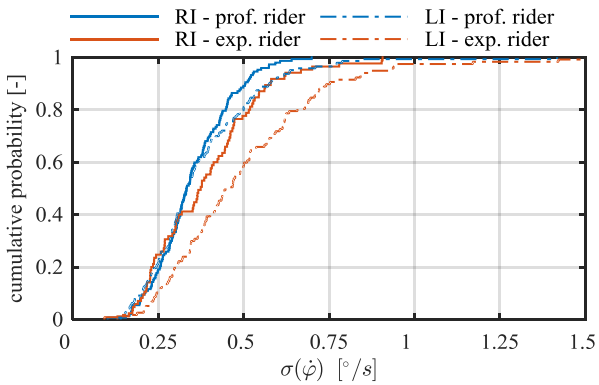


Fig. 8: Variance of the roll rate in dynamic maneuver primitives RI and LI for entering a right or left curve.

results for the evaluation of dynamic primitives, see Fig. 8, is slightly more difficult because the variance of the filtered roll rate can also depend on the individual riding style of the motorcycle rider, as mentioned above. However, it can be observed that the distribution of scores for entering a left cornering scenario (LI) by the non-professional rider clearly differs from the three other curves in its median and the gradient. This behavior again implies a lower consistency for the experienced motorcycle rider as the distributions of RI and RL are not concurring. On the other hand, they do for the professional rider. Due to the significant overlap between the RI for the experienced and RI and RL of the professional rider, no further statements can be made. We currently think that a more detailed evaluation of the frequency spectrum of dynamic maneuver primitives can lead to a better distinction between riders. By now, we have shown that the segmentation into maneuver primitives allows us to extract meaningful scores and RSI for local sub-segments of complex maneuvers which is an improvement compared to the state of the art method proposed by Yoneta [6].

IX. Summary and Outlook

In this paper, we proposed a definition of maneuver primitives to represent complex cornering scenarios through smaller recurring parts. Based on a general probabilistic HMM, we introduced two specific methods to find the most probable segmentation based only on the data for the roll dynamics of a motorcycle. Both models achieve a good match rate of about 89 %. Furthermore, above 80 % of segment boundaries are found in an interval of t = 0.3 s around the ground truth segmentation points. To achieve a good segmentation performance, it is necessary that measurement data and video recordings are known. Only in this case, a reliable segmentation can be performed automatically.

Based on the proposed segmentation in maneuver primitives, we introduced two riding skill indicators for stationary and dynamic cornering primitives. Both are based on the high frequency components of the roll rate in the specific segments. Future work will focus on a more detailed model to evaluate each specific maneuver primitive using unsupervised and supervised machine learning techniques and a more detailed representation of the primitives, e.g. the whole frequency spectrum. We will also focus on a detection of explicitly defined riding errors.

X. Acknowledgment

The presented work has been funded by the Honda Research Institute Europe GmbH and has also been supported by Honda R&D Germany by providing the motorcycle and measurement equipment. We express our gratitude to all contributors.

References

[1] Statistisches Bundesamt (Destatis), Ed, *Verkehr Verkehrsunfälle - Fachserie 8 Reihe 7 - 2014*. Wiesbaden: Statistisches Bundesamt, 2016.

[2] N. P. Chandrasiri *et al*, "Driving skill analysis using machine learning The full curve and curve segmented cases," in *2012 12th International Conference on ITS Telecommunications (ITST)*, pp. 542–547.

[3] Y. Zhang, W. C. Lin, Y.-K. S. Chin, "Driving skill characterization: A feasibility study," in *2008 IEEE Int. Conf. Robotics and Automation (ICRA)*, Pasadena, CA, USA, pp. 2571–2576.

[4] Y. Zhang, W. C. Lin, Y.-K. S. Chin, "A Pattern-Recognition Approach for Driving Skill Characterization," *IEEE Trans. Intell. Trans. Sys.*, vol. 11, no. 4, pp. 905–916, 2010.

[5] X. Tang, "Driving skill recognition: New approaches and their comparison," in *2009 Amer. Contr. Conf.*, St. Louis, MO, USA, pp. 420–425.

[6] K. Yoneta, H. Daimoto, A. Yamamoto, "Rider Characteristics Assessment Device and Straddle-Ridden Vehicle Provided Therewith" EP2517952A1, Oct. 31, 2012

[7] F. Biral, R. Lot, "An interpretative model of g-g diagrams of racing motorcycle," in: *Proc. 3rd Int. Conf. Mech. Eng. and Mechanics*. Beijing, P.R. China, Oct. 21–23, vol 2

[8] F. Biral, M. Da Lio, R. Lot, R. Sartori, "An intelligent curve warning system for powered two wheel vehicles," *Eur. Transp. Res. Rev*, vol. 2, no. 3, pp. 147–156, 2010.

[9] R. H. Osgouei, H. Lee, S. Choi, "Comparative evaluation of performance measures for human driving skills," *Intel Serv Robotics*, vol. 6, no. 4, pp. 169–180, 2013.

[10] R. H. Osgouei, S. Choi, "Evaluation of driving skills using an HMM-based distance measure," in *2012 IEEE Int. Workshop Haptic Audio Visual Environments and Games*, Munich, Germany, pp. 50–55.

[11] A. Aljaafreh, N. Alshabatat, M. S. Najim Al-Din, "Driving style recognition using fuzzy logic," in *2012 IEEE Int. Conf. Vehicular Electronics and Safety (ICVES 2012)*, Istanbul, Turkey, pp. 460–463.

[12] Y. L. Murphey, R. Milton, L. Kiliaris, "Driver's style classification using jerk analysis," in *2009 IEEE Workshop Comput. Intell. Vehicles and Vehicular Systems (CIVVS)*, Nashville, TN, USA, pp. 23–28.

[13] D. Klebelsberg, Ed., *Verkehrspsychologie*. Springer-Verlag 2013

[14] J. Reason, Ed, *Menschliches Versagen. Psychologische Risikofaktoren und moderne Technologien*. Heidelberg, Germany: Spektrum Akad. Verlag, 1994

[15] W. H. Wang, Q. Cao, K. Ikeuchi, H. Bubb, "Reliability and safety analysis methodology for identification of drivers' erroneous actions," *Int.J Automot. Technol*, vol. 11, no. 6, pp. 873–881, 2010.

[16] A. Gensler, B. Sick (2014): "Novel Criteria to Measure Performance of Time Series Segmentation Techniques", in: *Proc. LWA 2014 Workshops*, Aachen, Germany, Sept., 8-10, 2014, pp. 193-204

[17] G. Agamennoni, J. R. Ward, S. Worrall and E. M. Nebot, "Bayesian model-based sequence segmentation for inferring primitives in driving-behavioral data," in: *17th Int. Conf. Information Fusion (FUSION)*, Salamanca, Spain, Jul., 7-10, 2014, pp. 1-8.

[18] J.-H. Kim *et al*, "Modeling of Driver's Collision Avoidance Maneuver Based on Controller Switching Model," *IEEE Trans. Syst, Man, Cybern. B*, vol. 35, no. 6, pp. 1131–1143, 2005.

[19] R. Terada, H. Okuda, T. Suzuki, K. Isaji, N. Tsuru, "Multi-scale driving behavior modeling using hierarchical PWARX model," in *13th Int. IEEE Conf. Intell. Trans. Sys (ITSC 2010)*, Funchal, Madeira Island, Portugal, pp. 1638–1644.

[20] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[21] K. P. Murphy, *Hidden Markov Model (HMM) Toolbox for Matlab* (2005) [Online]. Available: https://www.cs.ubc.ca/~murphyk/ Software/ HMM/hmm.html

[22] K. P. Murphy (1998): *Fitting a Conditional Linear Gaussian Distribution* [Online], https://www.cs.ubc.ca/~murphyk/Papers/ learncg.pdf

[23] M.J. Mataric, "Sensory-motor primitives as a basis for imitation: linkin perception to action and biology to robotics," *Imitation in animals and artifacts*, pp. 391-422, MIT Press, Cambridge, MA, USA, 2002