# Powered-Two-Wheeler safety critical events recognition using a mixture model with quadratic logistic proportions

Ferhat Attal, Abderrahmane Boubezoul, Allou Samé and Latifa Oukhellou

#### Université Paris-Est - IFSTTAR

Abstract. This paper presents a statistical methodology that uses both acceleration and angular velocity signals to detect critical safety events for Powered Two Wheelers (PTW). The problem of recognition of critical events has been performed towards two steps: (1) the feature extraction step, where the multidimensional time trajectories of accelerometer/gyroscope data were modeled and segmented by using a specific mixture model with quadratic logistic proportions; (2) the classification step, which consists in using the k-nearest neighbor (k-NN) algorithm in order to assign each trajectory characterized by its extracted features to one of the three classes namely Fall, near Fall and Naturalistic riding. The results show the ability of the proposed methodology to detect critical safety events for Powered Two Wheelers.

## 1 Introduction

In recent years, road safety has become a priority for the governments of European countries. While statistics show a substantial decline in the number of fatalities on the road for four-wheeled vehicles, in the category of Powered Two-Wheelers (PTWs), the statistics show only a minor reduction with a constant decrease (approximately 2%). In France, PTW traffic represents approximately 2% of the overall traffic flow; however, it represents 40% of all serious injuries and approximately 20% of all deaths.

Driver behavior and driver errors are major causes of vehicular accidents. Therefore, observing and understanding driver behavior has attracted much attention from researchers. Many studies have been made to determine what factors are associated with critical events that occur before crashes. One method among these tools is naturalistic driving/riding studies (NDS/NRS). Thus, several vehicles are equipped with embedded sensors. Each participant drives his/her vehicle for an extended period of time. In this way, NRS data can provide knowledge about rider behavior, i.e., how the rider interacts with her/his vehicle. Additionally, by identifying critical events, useful contextual information can be provided to intelligent transportation systems (ITS) developed for PTWs, thereby improving their effectiveness. An event is defined as an undesirable riding event, such as hard braking, lane changing and sharp turning.

This paper focuses on automatic incident detection based on the same idea as [1]. The authors applied a robust outlier detection methodology based on the Mahalanobis distance to detect critical incidents. As this method is a threshold

based method, the main drawback of such approaches is the difficulty of determining the thresholds of regular and irregular riding behavior. In this work our aim is to distinguish between the regular and irregular riding behavior and to detect the switching time between these two profiles. In this paper, the problem of incident detection is performed via two steps: (1) the segmentation step, this step consists in segmenting in an unsupervised context the multidimensional time series of accelerometer/gyroscope data. Therefore, each segment is modeled by a regression model and logistic functions are used to model the transitions between segments. (2) the classification step, this step consists in classifying segments by using the k-nearest neighbor (k-NN) algorithm, each segment being characterized by its mean and its variance. This approach was applied on real experiments conducted by different subjects driving instrumented PTWs and its performance was assessed by performing comparison with another algorithm, the standard Multiple Hidden Markov Model (MHMM) [2]. To start with, the next section is dedicated to the proposed statistical model for the PTW events detection problem.

# 2 PTW safety critical events detection model

#### 2.1 PTW experimental data

In this work we have used the data coming from 3D Inertial Measurement Unit (accelerometers/gyroscopes) mounted on the PTW. In order to evaluate the reliability and the robustness of the proposed methodology, three kinds of scenarios were conducted: (i) Fall scenarios (F): Five falling scenarios were selected and replayed by a stuntman. A set of 8 trajectories were performed by the stuntman including five scenarios known: Standstill fall, Fall in a curve, Fall on a slippery straight road section, Fall with leaning of the motorcycle "intentional maneuver" and Fall in a roundabout. For more details on this experimentation the reader is invited to see [3]. (ii) Near fall scenarios (Nf): A set of 10 trajectories were performed during extreme cases of riding behavior including aggressive riding such harsh braking, accelerating and swerving. These extreme manoeuvres were carried out on track by a professional rider. The purpose of performing these manoeuvres is to study the robustness of the proposed algorithm in such limit handling behavior. (iii) Naturalistic riding scenarios (N): This experiment was performed in urban area near to the city of Paris, under different weather conditions (sunny, rainy and foggy). 11 trajectories were rode by five riders with different profiles and riding experiences. The participants were given an instruction to drive like they usually do. For more details on this experimentation the reader is invited to see [4].

The signals have been collected with a sampling frequency of 1 Khz. A filtering task was carried out by using the Wavelet Filter (WF) with a six level of wavelet decomposition and Daubechies mother wavelet (Db20) [4].

The collected discrete observations constitute a sample of multivariate trajectories:

$$\mathcal{D} = \{ \boldsymbol{x}_i \}_{i=1:N},\tag{1}$$

where N is the total number of trajectories. Each trajectory  $\boldsymbol{x}_i = \{\boldsymbol{x}_{it}\}_{t \in \mathcal{T}_i}$  is supposed to be observed at the time vector  $\mathcal{T}_i = \{t_{i1}, \ldots, t_{iT_i}\}$ , where  $T_i$  represents the trajectory length. In this study, only the data recorded from the 3D-accelerometer/gyroscope were used. Therefore,  $\boldsymbol{x}_{it}$  is defined as:

$$\boldsymbol{x}_{it} = \{a_x, a_y, a_z, r_x, r_y, r_z\} \in \mathbb{R}^6,$$

where,  $a_x$ ,  $a_y$  and  $a_z$  are the longitudinal, the lateral and the vertical accelerations respectively. And  $r_x$ ,  $r_y$  and  $r_z$  are the roll, the pitch and the yaw angular velocities respectively.

#### 2.2 A Gaussian mixture model with quadratic logistic proportions

This section aims at describing the model that will be used for accelerometer/gyroscope signals segmentation. For this purpose a specific mixture model with time varying proportions is formulated, which is a specific case of the regression model with hidden logistic process (RHLP) [5].

#### 2.2.1 Model definition

The signal segmentation model used in this work assumes that each observation  $\boldsymbol{x}_{it}$   $(t = 1, ..., T_i)$  of the trajectory  $\boldsymbol{x}_i$  (i = 1, ..., N) is distributed according to the following mixture of two Gaussian distributions:

$$p(\boldsymbol{x}_{it};\boldsymbol{\theta}) = \pi(t;\boldsymbol{w})\mathcal{N}(\boldsymbol{x}_{it};\boldsymbol{\beta}_1,\boldsymbol{\Sigma}_1) + (1 - \pi(t;\boldsymbol{w}))\mathcal{N}(\boldsymbol{x}_{it};\boldsymbol{\beta}_2,\boldsymbol{\Sigma}_2)$$
(2)

where  $\mathcal{N}(.; \boldsymbol{\beta}, \boldsymbol{\Sigma})$  is the Gaussian probability density function with mean  $\boldsymbol{\beta}$  and covariance matrix  $\boldsymbol{\Sigma}$ . The parameters  $(\boldsymbol{\beta}_{\ell}, \boldsymbol{\Sigma}_{\ell})_{\ell=1,2}$  are the means and covariance matrices of the Gaussian components, and  $\pi(t; \boldsymbol{w})$  is the first component proportion defined by:

$$\pi(t; \boldsymbol{w}) = \frac{\exp\left(w_0 + w_1 t + w_2 t^2\right)}{1 + \exp\left(w_0 + w_1 t + w_2 t^2\right)},\tag{3}$$

whose parameter is  $\boldsymbol{w} = (w_0, w_1, w_2) \in \mathbb{R}^3$ . These quadratic mixture proportions allow a specific segmentation of the trajectories. They are particularly suitable for segmentation problems with reswitching transitions, especially encountered in the case of riding behaviour.

The parameter  $\boldsymbol{\theta} = \{\boldsymbol{w}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2\}$  of this model is estimated by maximizing the log-likelihood through the Expectation-Maximization algorithm, as detailed in [5]. After this step, a set of features is calculated based on the parameter  $\boldsymbol{\theta}$  estimated for each trajectory. The set of features  $\mathcal{F}$  is calculated as follow:

$$\mathcal{F} = \{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \|\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2\|, diag(\boldsymbol{\Sigma}_1), diag(\boldsymbol{\Sigma}_2), \|diag(\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2)\|\}$$

where  $\|.\|$  is the norm associated the Euclidean distance, and  $diag(\Sigma)$  is the vector of the diagonal elements of  $\Sigma$ . It should be noticed that the segments

were rearranged according to temporal order which corresponds to the real riding activity. This set of features is used as learning database for classifying segments by using the k-nearest neighbor (k-NN) algorithm. By varying the value k between 1 and 10, we selected the value k = 3. The aim of this step is to judge if the transition between two contiguous segments is a consequence of an event occurring, which can be seen as an important change between the mean and the variance from one segment to another.

#### 3 Results and discussion

As mentioned above, the PTW events detection methodology is evaluated on a real database. This database is constituted by 29 trajectories in which we have 8 falling trajectories representing 5 scenarios, 10 near-falling trajectories representing 6 scenarios and 11 naturalistic riding trajectories.

In this section we present the results of the segmentation of different trajectories using the proposed approach GMMQLP (Gaussian Mixture Model with Quadratic Logistic Proportions) and MHMM. As we can see on the example presented on Figure 1, the proposed approach performs well the data segmentation compared to the MHMM. In the particular case of fall scenario, it can be clearly observed that the data is correctly segmented by the GMMQLP algorithm. We can also notice that the data is still correctly segmented by the MHMM algorithm with some "pseudo" transitions, Figure 1(g). In the case of near fall scenario, the same observation can be made for the two algorithms, with more "pseudo" transitions in the case of the MHMM algorithm. In the case of naturalistic riding scenario, the GMMQLP algorithm does not segment the data, which is not the case of the MHMM algorithm where many "pseudo" transitions can be observed, this result can be explained by the flexibility of the logistic process that govern the switching from one segment to another.

The Table 1 represents the obtained results in terms of correct segmentation rate for each scenario. These results are obtained by matching the segmentation results to the true labels (given by an expert). We can notice that the GMMQLP performs better segmentation than the MHMM. We have to recall that the aim

		Correct segmentation rate								
		F	Nf	Ν	Global correct					
		trajectories	trajectories	trajectories	segmentation rate					
	GMMQLP(%)	98.77	95.37	95.72	96.62					
	MHMM(%)	92.50	85.81	76.93	85.08					

Correct segmentation rat

Table 1: Correct segmentation rate obtained with GMMQLP and MHMM algorithms.

of this study is to develop an automatic incident detection approach. This approach aims is to distinguish between the regular and irregular riding behavior and to classify the segment into three classes (fall, near fall or naturalistic) riding ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8. Available from http://www.i6doc.com/en/.



Figure 1: Results obtained by applying the proposed GMMQLP model (middle) and the MHMM approach (bottom) on the acceleration time series measured during fall, near fall and naturalistic riding cases (from left to right). In a, b and c the red, blue and black signal represent  $a_z$ ,  $a_x$  and  $a_y$ , respectively. In d, e, f, g, h the estimated probabilities obtained with the two approaches are presented.

ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8. Available from http://www.i6doc.com/en/.

Predicted classes					Predicted classes				
		F	Nf	Ν			F	Nf	Ν
	F (%)	100	0	0		F (%)	100	0	0
Real	Nf (%)	0	90	10	Real	Nf (%)	0	70	30
classes	N (%)	0	0	100	classes	N (%)	0	0	100

Table 2: Global confusion matrix for the Table 3: Global confusion matrix for the<br/>GMMQLP algorithm.MHMM algorithm.

events. This classification task is performed by the k-nearest neighbor (k-NN) algorithm on the database  $\mathcal{F}$ , each segment being characterized by its mean and its variance, as stated before. The global confusion matrices for the GMMQLP and the MHMM algorithms are given in Table 2 and Table 3. AS expected we can observe that the confusions occur especially for near fall scenarios, less for GMMQLP algorithm than for MHMM algorithm. These confusions can be explained by the fact that in theses situations the features extracted from GMMQLP algorithm  $\mathcal{F}$  are more discriminative than the features extracted from the MHMM algorithm.

## 4 Conclusion and Further Work

In this paper we presented a new method for PTW events detection by using a specific mixture model with quadratic logistic functions. Among data collected from an instrumented motorcycle we have only used the 3D Inertial Measurement Unit (accelerometers/ gyroscopes) as input data. The obtained results show the effectiveness of the proposed algorithm to solve such problem. The increasing computational capabilities of on-board computers makes the proposed methodology suitable for realtime events detection problem, this idea will be investigated as a future work.

#### References

- [1] Eleni I Vlahogianni, George Yannis, and John C Golias. Critical power two wheeler driving patterns at the emergence of an incident. Accident analysis and prevention, 2013.
- [2] L.R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. volume 77 (2), pages 267–296, 1989.
- [3] Abderrahmane Boubezoul, Bruno Larnaudie, Stéphane Espié, and Samir Bouaziz. A simple fall detection algorithm for powered two wheelers. *Control Engineering Practice*, 21(3):286 – 297, 2013.
- [4] Attal Ferhat, Boubezoul Abderrahmane, Oukhellou Latifa, and Espie Stephane. Riding patterns recognition for powered two-wheelers users' behaviors analysis. In *Intelligent Transportation Systems (ITSC)*, 2013 16th International IEEE Conference. IEEE, 2013.
- [5] Faicel Chamroukhi, Allou Samé, Gérard Govaert, and Patrice Aknin. Time series modeling by a regression approach based on a latent process. *Neural Networks*, 22(5):593–602, 2009.