Modelling Stop Intersection Approaches using Gaussian Processes

Alexandre Armand, David Filliat, Javier Ibanez-Guzman

Modelling Stop Intersection Approaches using Gaussian Processes

Alexandre Armand[1,2], David Filliat[1], Javier Ibañez-Guzmán[2]

***Abstract*— Each driver reacts differently to the same traffic conditions, however, most Advanced Driving Assistant Systems (ADAS) assume that all drivers are the same. This paper proposes a method to learn and to model the velocity profile that the driver follows as the vehicle decelerates towards a stop intersection. Gaussian Processes (GP), a machine learning method for non-linear regressions are used to model the velocity profiles. It is shown that GP are well adapted for such an application, using data recorded in real traffic conditions. It consists of the generation of a normally distributed speed, given a position on the road. By comparison with generic velocity profiles, benefits of using individual driver patterns for ADAS issues are presented.**

I. Introduction

In general, more than 40% of car accidents, and over 20% of road fatalities occur at road intersections [15]. Crossroads are points of convergence and represent the most complex area in road networks. At least 90% of accidents in intersection contexts are caused by driver errors [19]. These errors mainly occur because of a lack of situation understanding or because of the influence of other factors on the driver, such as tiredness, alcohol, distractions, etc. which increase the risk. The age and the experience of the driver are also important factors. Indeed, [7] presents statistics results which show that young and elderly drivers are more often involved in traffic accidents than middle-aged drivers.

The behaviour of a person on the road evolves over time, due to age, experience, etc. Further, the manner in which each person drives differs from one person to another, with the difference more accentuated by age divergence. Statistics show that driving speeds depend on gender and age as well as driver experience [15]. Road intersections particularly highlight drivers behaviour differences. [4] shows that drivers have their own manner to decelerate as they approach intersections.

In this paper, we propose to model driver deceleration patterns using Gaussian Processes as they approach to stop intersections. The pattern allows us to determine the manner individual drivers approach intersections. If their current profile is different of such patterns, it can be inferred that a risk exists. The manner a driver decelerates is represented by velocity profiles. The remainder of this paper is organised as follows. In Section II, a literature review related to velocity models at road intersections is proposed, which leads to the problem formulation. We also review the principles of heteroscedastic Gaussian Processes (GP) in Section III, in order to see the features that can be used. Section IV shows how GP are applied to infer the velocity profile patterns. Results and discussions are proposed in Section V and the conclusions complete the paper in Section VI.

[1] ENSTA ParisTech/ INRIA FLOWERS team, 828 Boulevard des Maréchaux, 91762 Palaiseau Cedex, France. alexandre.armand@ensta-paristech.fr, david.filliat@ensta-paristech.fr

[2] Renault S.A.S, 1 av. du Golf, 78288 Guyancourt, France. javier.ibanez-guzman@renault.com

II. Related Work and Problem Statement

Advanced Driving Assistance Systems at road intersections have been extensively studied in automotive and robotics domains. Nowadays, many approaches have been proposed in the literature for risk estimation at road intersections, such as Time To Collision (TTC) based methods [5], methods based on trajectory prediction, or on driver intention estimation [3], [2].

Some of the approaches available in the literature use velocity profiles as a way to decide whether or not the driver's behaviour is dangerous, or not appropriated to the context. In [21], the proposed system uses 3 velocity thresholds (based on constant decelerations) to estimate the risk when approaching to the intersection. The thresholds are arbitrary set and are not driver dependant. The authors in [9] use speed profiles within a Dynamic Bayesian Network, implemented in cooperative cars for risk assessment at road intersections. The cruise speed of the host vehicle is used for the estimation of the expected deceleration profile, however it is assumed that all driver react in the same manner. In [10], velocity profiles are also the core of the proposed approach. The expected velocity profile is generated online using the driver model, according to the observed velocity. Once again, the profile does not depend on the driver's patterns. It is shown in [4] that there can be significant behaviour differences at road intersections between 2 drivers, and that these differences have to be taken into account by ADAS. However to our knowledge, the literature does not provide any approach which uses velocity profiles fully adapted to the driver.

So far, driver pattern modelling and learning have not been extensively studied. In [13], a framework is proposed to model car following patterns and the manner the driver pushes the gas and brake pedals. Gaussian Mixture Models (GMM) are used as a probabilistic way to extract personal features, and the model is evaluated in driver identification. Such approaches, to our knowledge, have not yet been investigated in the context of road intersections.

The robotics field has often used regression algorithms to learn mechanical models of robots. An extensive survey of online regression algorithms is proposed in [18].

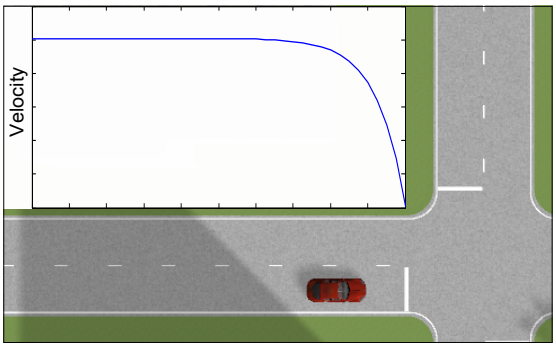

Fig. 1. A classical stop intersection and a corresponding speed profile.

Among all the algorithms presented, Locally Weighted Projection Regression (LWPR) appears as the most often used regression method in the robotics literature, providing accurate models. Gaussian Mixture Regression (GMR) based approaches are presented and gives similar results as LWPR. These approaches could be well suited for velocity profiles modelling, however they do no provide degrees of confidence on the outputs, such as a variance. Pattern repeatability for driver behaviour can differ very much, there is a wide spread. Within this context, LWPR and GMR do not cope very well with the driver pattern modelling needs. However, Gaussian Processes (GP) enable accurate non-linear regressions, and provide normally distributed outputs (mean and variance). This method seems to be more robust to pattern spread than the other cited methods. Further, in [14], a comparison between LWPR, GMR and GP is presented. The authors show that GP offers high learning precision and accuracy, and are rather simple to implement. This algorithm is used by [17] within a learning by demonstration framework. Heteroscedastic (input dependant variance) Gaussian Processes are used to learn mechanical models of robot arms from a couple of samples. The paper highlights GP capabilities to extract the "task constrains", that is the critical features of the learnt patterns. In ADAS research, GP have been used in [1] and [6] for trajectory predictions and were real-time capable.

It is proposed to apply Heteroscedastic Gaussian Processes for the learning of velocity profiles at the approach to road intersections. It is shown that this algorithm is well suited for our application. The rationale is that GP can highlight the areas were the driver always behaves in the same manner. In addition, they can be implemented in real time and infer continuous and consistent patterns.

III. Gaussian Processes Modelling

A. Basic Principles

In recent years, there has been an increased use of Gaussian Processes (GP) for regression and classification problems [16]. GP apply simple linear algebra while providing powerful tools to solve non linear problems. The aim is to recover a functional dependency $y_i = f(x_i) + \epsilon_i$ from n observed data points $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{n}$. Here, $\epsilon \in \mathbb{R}$ is a random noise that is independent, identically distributed for each observation. The training data set comprises y_i as the noisy output values at input locations x_i. The Gaussian Process regression consists in learning the predictive Normal distribution $p(y^*|x^*, \mathcal{D})$ of a new test output y^* given a test input x^*.

We simplify the notation by defining the $d \times n$ matrix X that collects all the training inputs $\{x_i\}_{i=1}^{n}$, d is the dimension of x_i. The same is done for the training outputs $\{y_i\}_{i=1}^{n}$ which are put into a Y vector of size n.

The covariance function (or Kernel) $k(.,.)$ is used to compute the terms of the covariance matrices K, K^* and K^{**}. This function defines the Gaussian Process which is written as $\mathcal{GP}(0, k(.,.))$. The Kernel depends on parameters θ known as hyperparameters which are determined in advance. The mostly used covariance function is the Squared Exponential given by:

$$k(x_i, x_j) = \sigma_n^2 \exp\left(-\frac{\|x_i - x_j\|^2}{2l^2}\right) \tag{1}$$

with $\theta = \{\sigma_n, l\}$ defines the noise level and the length scale. We also define $K \in \mathbb{R}^{n \times n}$ with $K_{i,j} = k(x_i, x_j)$, $K^* \in \mathbb{R}^{n^* \times n}$ with $K_{i,j} = k(x_i^*, x_j)$ and $K^{**} \in \mathbb{R}$ such as $K_{i,j}^{**} = k(x_i^*, x_j^*)$.

The predictive distribution at the query points x^* is a multivariate Gaussian distribution $\mathcal{N}(\mu^*, \Sigma^*)$ based on the training data with mean μ^*such as:

$$\mu^* = K^* \left(K + \sigma_n^2 I\right)^{-1} Y \tag{2}$$

and covariance Σ^*such as:

$$\Sigma^* = K^{**} - K^*(K + \sigma_n^2 I)^{-1} K^{*T} \tag{3}$$

The hyperparameters θ must be tuned to obtain a smooth and reliable regression. The tuning of θ is part of the learning process of GP. The maximum a posteriori estimate of θ occurs when $p(y|X, \theta)$ is a its greatest. The log marginal likelihood [16] is computed as:

$$\log p(y|X, \theta) = -\frac{1}{2} y^T K_y^{-1} y - \frac{1}{2} \log |K_y| - \frac{n}{2} \log(2\pi) \tag{4}$$

where $K_y = K + \sigma_n I$ and $|.|$ is the matrix determinant. We use a conjugate gradient multivariate optimization algorithm to locally minimize the log marginal likelihood and to get an optimized value of θ. Figure 2.a shows an example of a GP with an optimized value of θ.

B. Gaussian Processes using heteroscedastic variance

The Gaussian Process model described in the previous section assumes that the noise level σ_n is constant over the whole process. It is not a problem to estimate the mean value, however depending on the value of σ_n the variance might be either under or over-estimated. Therefore, to model driver patterns, it is judicious to use a more flexible noise model. Kersting introduced a heteroscedastic variance which

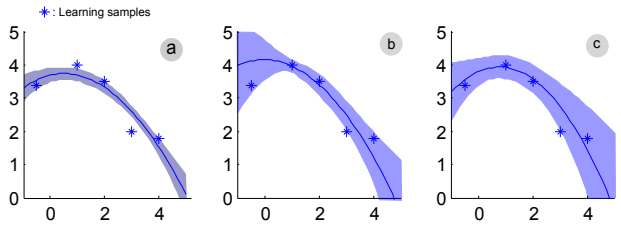

Fig. 2. Influence of the variance on GP regressions. a) Use of a standard homoscedastic variance. b) Use of a heteroscedastic variance. c) Taking into account input uncertainty.

is input dependent [8]. We replace the constant noise level σ_n by a input dependent function $r(x)$, thus the mean and variance functions become:

$$\mu^* = K^* \left(K + \mathbf{R}\right)^{-1} Y \tag{5}$$

$$\Sigma^* = K^{**} + \mathbf{R}^* - K^* \left(K + \mathbf{R}\right)^{-1} K^{*T} \tag{6}$$

where $R = diag(r)$, with $r = (r(x_1), ..., r(x_n))^T$, and $R^* \in \mathbb{R}$ such as $R^* = diag(r^*)$ with $r^* = (r(x_1^*), ..., r(x_n^*))^T$. R^* is usually learned from training data. Figure 2.b illustrates the influence of a heteroscedastic variance on a GP.

C. Gaussian Processes with input noise

In general, the GP model assumes that the inputs are noise free. However, in our application, the input which is the position of the vehicle is subject to noise (due to the localization sensor). Therefore, making the assumption of noise-free inputs would lead to bad regression performances. Mc Hutchon defined an input noise dependent version of GP in [12].

In standard GP, y is a noisy measurement of the output:

$$y = \tilde{y} + \epsilon_y \tag{7}$$

where $\epsilon_y \sim \mathcal{N}(0, \sigma_y^2)$ with a homoscedastic GP, and $\epsilon_y \sim \mathcal{N}(0, r(x)^2)$ with a heteroscedastic GP. We also assume that the inputs in the model are noisy:

$$x = \tilde{x} + \epsilon_x \tag{8}$$

where $\epsilon_x \sim \mathcal{N}(0, \sigma_x^2)$. We assume here that the input noise is constant. The output as a function of the input can be written as:

$$y = f(\tilde{x} + \epsilon_x) + \epsilon_y \tag{9}$$

Mc Hutchon and Rasmussen have defined a linear model for the input noise, thus the output becomes:

$$y = f(\tilde{x}) + \epsilon_x^T \partial_{\bar{f}} + \epsilon_y \tag{10}$$

where $\partial_{\bar{f}}$ is the derivative of the mean of a GP function. To obtain it, we run a first GP to predict the mean plus the derivative of the output for each input. The probability of an observation is given by $P(y^*|x^*, \mathcal{D}) \sim \mathcal{N}(f, \sigma_y^2 + \partial_{\bar{y}}^T \Sigma_x \partial_{\bar{y}})$ where $\Sigma_x = diag(\sigma_x^2)$. The covariance matrix K is not changed to compute the new mean and variance functions:

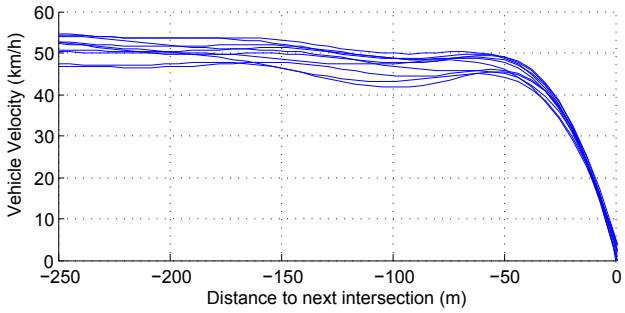

Fig. 3. An example of 10 learning velocity profiles for a 50 km/h cruise speed.

$$\mu^* = K^* \left(K + R + \mathbf{P}\right)^{-1} Y \tag{11}$$

$$\Sigma^* = K^{**} + R^* - K^* \left(K + R + \mathbf{P}\right)^{-1} K^{*T} \tag{12}$$

with $P = diag(\Delta_{\bar{y}} \Sigma_x \Delta_{\bar{y}}^T)$ and $\Delta_{\bar{y}} = \{\partial_{\bar{y},i}\}_{i=1}^n$.

Figure 2.c illustrates the consideration of a noisy input for a GP. It is noticeable that the variance is larger over the x axis.

IV. Pattern Extraction

The vehicle velocity profile as it is driven towards an intersection with a compulsory stop sign is to be modelled using Gaussian Processes. For this purpose, it is assumed that the different velocity profiles of the vehicle, as it approaches that type of intersection, follow a normal distribution of the distance/ speed ratio. The model developed in Section III-C is used as it provides the best estimation of the variance. The derivation of the model uses data acquired as described in the next section.

A. Data Acquisition

The experiment had to be done with real data recorded in a passenger vehicle. Indeed, this would not make sense to use simulated data to model individual driving patterns. A passenger vehicle, a Renault Espace with an automatic gearbox was used for the experimental part driven in open roads. For this purpose, data available in the CAN bus was recorded, namely the vehicle velocity as well as the actuation of the driver in the vehicle. The position of the vehicle was recorded from the position estimated from an automotive type GPS receiver running at 1 Hz. The localization of the vehicle with respect to a road intersection was made using the development software for navigation systems known as ADAS-RP (Nokia). The standard navigation map was extended to include the position of the vehicle stop point prior to entering the intersection.

The experiment to acquire the data which should enable the system to learn the velocity trajectory was as follows:

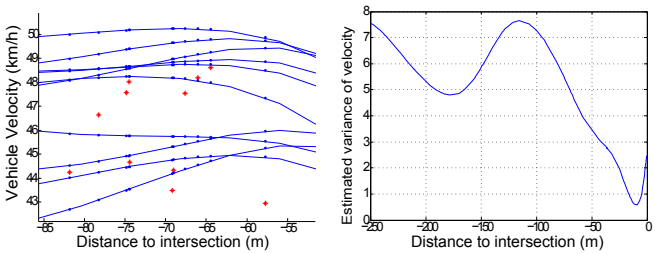

Fig. 4. On the left, the red dots are data from the 10 learning profiles, used for regression (blue curves). The blue dots are the projections of the red dots on the regressed learning profiles. The blue dots are used for the estimation of the output dispersion for every input of the data set $\mathcal{D}$. On the right, the estimated variance function of the distance to the road stop.

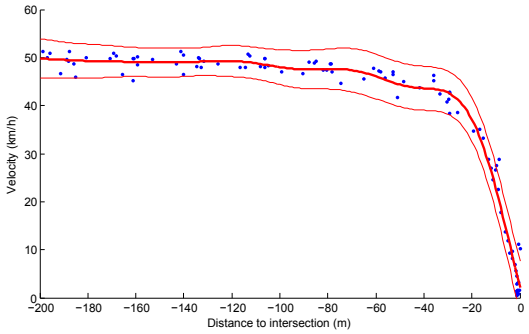

Fig. 5. An example of driver pattern modelled using GP (in red with 95% of confidence) with the training dataset $\mathcal{D}$ (blue dots).

The vehicle was driven by four users independently at five different velocities (30, 40, 50, 60 and 70 km/h). For each speed, 10 runs were executed to determine the manner in which the driver decelerated. An example is shown in Figure 3. We obtained $n = 10$ data sets $\mathcal{D}l_i = \{Xl_i, Yl_i\}_{i=1}^{n}$ with Xl_i the learning input vector which corresponds to the distance to the road stop, and Yl_i the associated learning velocity vector. We also define the full data set $\mathcal{D} = \{X, Y\}$ that gathers the n learning data sets $\mathcal{D}l_i$, with $X = \{x_j\}_{j=1}^{m}$ and $Y = \{y_j\}_{j=1}^{m}$. The size of the vectors X and Y is $m = \sum_{i=1}^{n} (sizeof(Xl_i))$.

B. Estimation of the variance corrections matrices

To generate the variance correction matrix R introduced in the part III-B, it is necessary to estimate the output variance $r_i(x_i)^2$ for each input x_i of the learning dataset $\mathcal{D}$. To do so:

- We modelled a simple Gaussian Processes (described in part III-A) $\mathcal{GP}l_i$ for each of the n learning data sets $\mathcal{D}l_i$. (See fig. 4.1).
- For each learning input x_i from the dataset $\mathcal{D}$, n output predictions y_i^* are computed with the n GP $\mathcal{GP}l_i$.
- The empirical variance $s_j = \frac{1}{n-1}\sum_{i=1}^{n}\left(\bar{y_j} - y_{ij}^*\right)^2$ is computed with $\bar{y_j}$ the mean of the n predicted values y_{ij}^*.
- We get the vector $S = \{s_j\}_{j=1}^{m}$ and the $m \times m$ matrix $R = diag(S)$ as the variance correction matrix.

Figure 4.2 illustrates the estimated variance of the velocity at any distance to the stop intersection, given a velocity and a driver.

It is also necessary to determine the correction matrix P related to the input noise σ_x, described in the part III-C. We used $\sigma_x = 5$ m to model the GPS noise.

- A first Gaussian Process $\mathcal{GP}d$ that uses the data set $\mathcal{D}$ is run to compute the predicted points $Y^* = \{y_i^*\}_{i=1}^{m}$ at every input point $\{x_i\}_{i=1}^{m}$ from $\mathcal{D}$. The correction matrix R is used to compute the parameters of $\mathcal{GP}d$, according to III-B.
- The output derivatives $\Delta^{Y^*} = \left\{\partial_i^{y^*}\right\}_{i=1}^{m}$ are then computed thanks to Y^*, according to III-C.
- We therefore create the second correction matrix $P = diag(\Delta^{Y^*}\Sigma_x\Delta^{Y^*T})$ with $\Sigma_x = diag(\sigma_x)$.

C. Learnt velocity profile

Finally, the main Gaussian Process $\mathcal{GP}$ was applied to the training data set $\mathcal{D} = \{X, Y\}$ for each driver independently, thus modelling each driver behaviour at intersections. The correction matrices R and P were taken into account. The full non linear regression for a driver driving at 50 km/h is shown in Figure 5.

V. Results And Discussions

In this section, results of velocity profile modelling using Gaussian Processes are presented. We used real data acquired on open roads, as explained in the previous section.

A. Gaussian Processes based modelling

Figure 5 illustrates a speed profile corresponding to the pattern of an average driver. Ten runs were used for the training step, which represent about 150 points for the GP-based regression. Two other learnt patterns (at different cruise speeds) are shown in the Figure 6. We can notice that:

- The mean is continuous and smooth, in addition it coherently matches with the training samples.
- A driver cannot behave exactly in the same manner at every approach. Indeed, there is a significant uncertainty on the moment he starts decelerating and on the deceleration rate. The variance provided by the GP takes these uncertainties into account in a straightforward manner.
- The use of a heteroscedastic version of GP enables a good estimation of the variance according to the training samples. If homoscedastic GP had been used, the covariance matrix would have been set from the σ_n parameter (see Section III-A) which does not depend on the training samples. The choice of this parameter would have been done with difficulty, and the variance generated by GP would have been either over or under-estimated. The driver pattern would not have been accurately modelled.

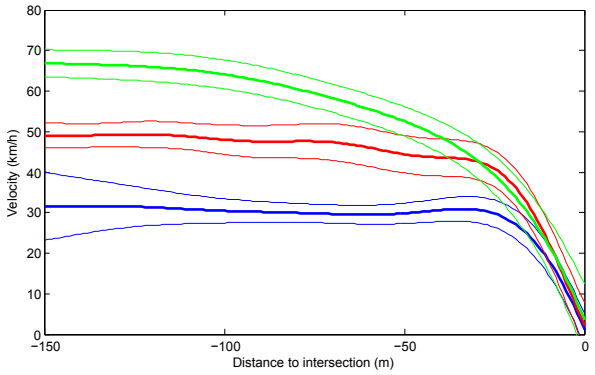

Fig. 6. Example of a driver's approaches modelled with GP for 3 speeds (30, 50 and 70 km/h). The three profiles are from the same driver.

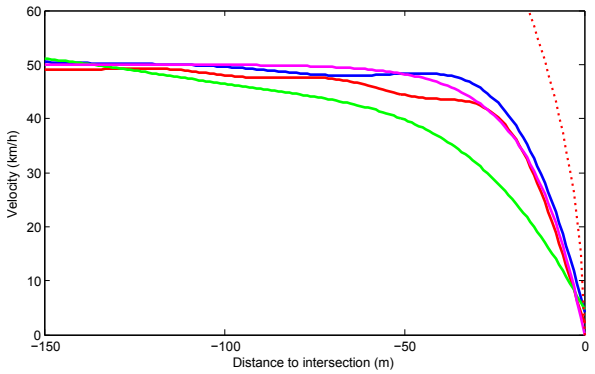

Fig. 7. Comparison of 3 different drivers models with a generic profile (relaxed in green, average in red, sporty in blue and generic in magenta)

- The normally distributed output provided by the GP-based approach allows to easily check if a given set of velocity and distance to intersection matches with the driver's manner to approach to an intersection. A probability that the driver behaves as usual in the same conditions can be easily estimated (through the likelihood computation for example).

In the Figure 7, the velocity profiles of 3 different drivers, generated with GP are shown (for clearness, the pattern of the fourth driver has not been shown). It is noticeable that the proposed modelling method allows to highlight differences between driving styles.

B. Benefit of personal patterns compared to generic profiles

1) Comparison: The literature provides several studies which analyse the deceleration behaviours, as presented in [11]. The authors explain that it is difficult to develop a generic model of deceleration profiles because several parameters influence the driver behaviour, such as his driving style, or also the type of vehicle.

In the Figure 7, we compare 3 driver patterns modelled with GP (relaxed, average and sporty) with a generic profile. This generic profile is set with a 2.4 m/s² deceleration rate, which is the average rate at 50 km/h according to [20]. This is a typical generic profile that may be used for ADAS applications. We can notice that:

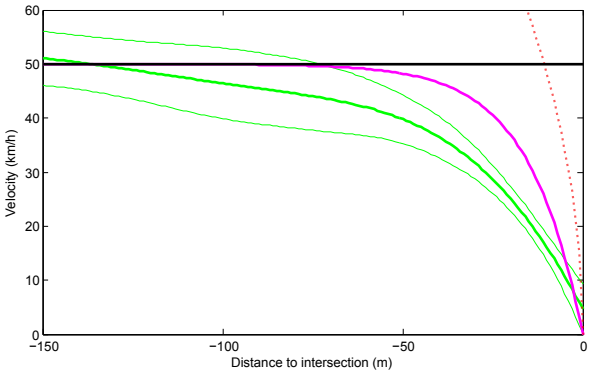

Fig. 8. Comparison of the use of a relaxed driver profile with generic average profile (generic profile in magenta, GP profile in green, the run in black and the 9m/s² threshold in red dots).

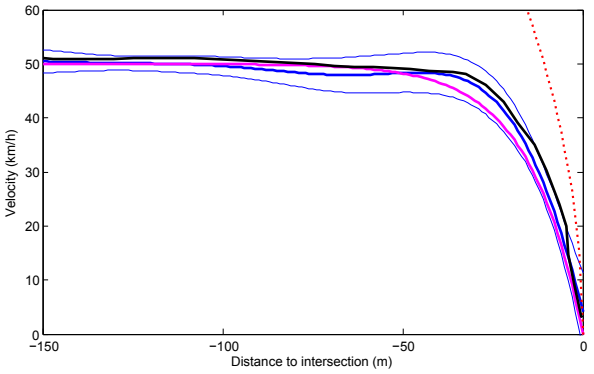

Fig. 9. Comparison of sporty driver profile with generic average profiles (generic profile in magenta, GP profile in blue, the run in black and the 9m/s² threshold in red dots).

- The personal patterns modelled with GP show particular features, specific to the driver, which cannot be presented in generic profiles. On the Figure 6, the 50 km/h profile shows that the driver usually decelerates in 2 steps, a first with a very low deceleration rate (he stops pushing the gas pedal) followed by a higher deceleration rate (starts pushing the brake pedal).
- Such generic profiles may over-estimate the deceleration rate of the driver. It is clearly visible on the Figure 7 that an average generic profile is not adapted to relaxed driving styles.

2) Perspectives: Knowing the driver patterns (in other words the habits) may help ADAS systems in terms of accuracy and performances.

- Frameworks which estimate the driver intention using velocity profiles can generate better data if individual profiles are used. For example, in [9] the system needs to estimate the driver intention to stop. With a relaxed driver, the intention not to stop can be detected earlier with individual profiles than with generic profiles. Indeed, the Figure 8 shows that the velocity leaves the profile envelope (generated by GP) whereas it still matches with an average generic deceleration profile. On the contrary, the deceleration of a sporty driver can

be interpreted as an intention not to stop with a generic profile (Figure 9). By knowing how the driver usually behaves when he intends to stop may help to estimate more precisely (and maybe in a more reactive way) his intention.

- Frameworks based on speed monitoring (like [21]) check at any time if the host vehicle velocity stays below critical thresholds. These methods can provide information and alerts only when the velocity goes over the threshold. By using profiles generated with GP, it is possible to detect unusual driver behaviours. For example, if at a given position the driver is usually decelerating, and if at the same position he has not started to decelerate, the probability that there is something wrong is high. Information can be generated early, even before the situation becomes dangerous.

C. Limitations

1) Computational cost: The main limitation of the proposed GP-based modelling method is the high computational cost mainly due to the inversion of the covariance matrix K. Thus the computational cost of using GP is directly linked to the number of training sample points. Therefore, for real time applications, a compromise has to be found between a large amount of training data (synonym of accurate model) and a reasonable computational cost. An alternative is to use Local Gaussian Processes (LGP) as presented in [17]. This approach consists in separating the pattern into several regions and in applying GP to each region.

2) Context: The proposed approach for stop intersections assumes that there is no other context element than the stop sign which interferes with the subject vehicle. For example, our model is not valid if there is a lead vehicle between the stop line and the subject vehicle. Velocity profiles can be modelled, taking into account permanent context elements for which the driver behaviour is always the same (such as element described in digital maps). However, to model the behaviour of the driver in front of other mobile entities (vehicles, pedestrians, etc), other variables than the velocity have to be taken into consideration.

VI. Conclusion

A pattern shaping method for the modelling of velocity profiles of individual drivers as they approach stop intersections has been presented. It showed that Gaussian Process is very well adapted for such a task. It allows to model accurately the driver patterns taking into account uncertainties that might exist due to the driver and the quality of the on-board sensors. The advantages of using personified driver patterns instead of generic profiles for ADAS applications have been addressed. If real time implementation constraints arise, execution optimisation can be made applying Local Gaussian Processes as they segment the trajectory to reduce the data needed during the training phase.

The proposed framework will be next applied to situation awareness monitoring in driving assistance functions.

References

[1] G. S. Aoude, "Threat assessment for safe navigation in environments with uncertainty in predictability," Ph.D. dissertation, Massachusetts Institute of Technology, 2011.

[2] G. Aoude, V. Desaraju, L. Stephens, and J. How, "Driver behavior classification at intersections and validation on large naturalistic data set," *in Proceedings of the IEEE Intelligent Transportation Systems Conference*, vol. 13, no. 2, pp. 724 –736, june 2012.

[3] H. Berndt, J. Emmert, and K. Dietmayer, "Continuous driver intention recognition with hidden markov models," *in Proceedings of the IEEE Intelligent Transportation Systems Conf.*, pp. 1189 –1194, oct. 2008.

[4] H. Berndt, S. Wender, and K. Dietmayer, "Driver braking behavior during intersection approaches and implications for warning strategies for driver assistant systems," *in Intelligent Vehicles Symposium, 2007 IEEE*, pp. 245–251, 2007.

[5] A. Berthelot, A. Tamke, T. Dang, and G. Breuel, "A novel approach for the probabilistic computation of time-to-collision," *in Proceedings of the IEEE Intelligent Vehicles Symp.*, pp. 1173 –1178, june 2012.

[6] T. M. K. Christopher, "Analysis of dynamic scenes: Application to driving assistance," Ph.D. dissertation, Institut National Polytechnique de Grenoble (INPG), 2009.

[7] P. J. Cooper, M. Pinili, and W. Chen, "An examination of the crash involvement rates of novice drivers aged 16 to 55." *Accident Analysis and Prevention*, vol. 27, no. 1, pp. 89–104, 1995. [Online]. Available: http://www.biomedsearch.com/nih/examination-crash-involvement-rates-novice/7718081.html

[8] K. Kersting, C. Plagemann, P. Pfaff, and W. Burgard, "Most likely heteroscedastic gaussian process regression," *In Zoubin Gharahmani, editor, ICML 07*, 2007.

[9] S. Lefevre, C. Laugier, and J. Ibanez-Guzman, "Risk assessment at road intersections: Comparing intention and expectation," *in Proceedings of the IEEE Intelligent Vehicles Symp.*, pp. 165 –171, june 2012.

[10] M. Liebner, M. Baumann, F. Klanner, and C. Stiller, "Driver intent inference at urban intersections using the intelligent driver model," *in Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 1162 –1167, june 2012.

[11] A. K. Maurya and P. S. Bokare, "Study of deceleration behaviour of different vehicle types," *International Journal for Traffic and Transport Engineering*, vol. 2, no. 3, 2012.

[12] A. McHutchon and C. E. Rasmussen, "Gaussian process training with input noise," *In NIPS*, 2011.

[13] C. Miyajima, Y. Nishiwaki, K. Ozawa, T. Wakita, K. Itou, K. Takeda, and F. Itakura, "Driver modeling based on driving behavior and its evaluation in driver identification," *Proceedings of the IEEE*, vol. 95, no. 2, pp. 427 –437, feb. 2007.

[14] D. Nguyen-Tuong, J. Peters, M. Seeger, and B. Schölkopf, "Learning inverse dynamics: a comparison," *In European Symposium on Artificial Neural Networks*, 2008.

[15] E. R. S. Observatory, "Traffic safety basic facts 2010 junctions," *SafetyNet, Project co-financed by the European Commission*, 2010.

[16] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[17] M. Schneider and W. Ertel, "Robot learning by demonstration with local gaussian process regression," *in Proceedings of IEEE Intelligent Robots and Systems (IROS)*, pp. 255 –260, oct. 2010.

[18] O. Sigaud, C. Salaun, and V. Padois, "On-line regression algorithms for learning mechanical models of robots: a survey," *Robotics and Autonomous Systems*, vol. 59, no. 12, pp. 1115–1129, December 2011.

[19] TRACE, "Analyzing human factors in road accidents," Tech. Rep., 2008. [Online]. Available: http://www.trace-project.org/publication/archives/trace-wp5-d5-5-v2.pdf

[20] J. Wang, K. Dixon, H. Li, and J. Ogle, "Normal deceleration behaviour of passenger vehicles starting from rest at all way stop controlled intersections," *Transportation Research Record*, pp. 158–166, 2005.

[21] J. Wang, D. Zhang, J. Liu, M. Lu, and K. Li, "Multi-objective driving assistance system for intersection support," *in Intelligent Transportation Systems (ITSC) 2010*, pp. 348–353, 2010.