2D Laser Based Road Obstacle Classification for Road Safety Improvement

Pierre Merdrignac, Evangeline Pollard, Fawzi Nashashibi

2D Laser Based Road Obstacle Classification for Road Safety Improvement

Pierre Merdrignac[1,2], Evangeline Pollard[1] and Fawzi Nashashibi[1]

***Abstract*— Vehicle and pedestrian collisions often result in fatality to the vulnerable road users (VRU), indicating a strong need of technologies to protect such persons. Laser sensors have been extensively used for moving obstacles detection and tracking. Laser impacts are produced by reflection on these obstacles which suggests that more information is available for their classification. This paper proposes a new system to address this issue. We introduce the design of our system that is divided in three parts : definition of geometric features describing road obstacles, multiclass object classification from an adaboost trained classifier and *track* class assignment by integrating consecutive classification decision values. During this study, we show how specific features adapted to urban obstacles enhance the state of the art method for person detection in 2D laser data. Hence, in this paper, we evaluate usefulness of each feature and list the best ones. Moreover, we investigate the influence of laser height for each class showing that classification performance depends on the sensor position. Finally, we tested our system on some laser sequences and showed that it can estimate the class of some road obstacles around the vehicle with an accuracy of 87.4%.**

I. Introduction

According to the statistics [1], more than 3000 people die daily due to dangerous driving. The main reason of accidents is the inability of road users to detect and perceive oncoming dangers before a sufficient amount of time so that reactions for accident avoidance can be taken. Much attention has been put in detecting pedestrians and predicting the possibility of collisions using sensors and computer vision techniques [2]. Pedestrian detection is known as a particularly hard problem because each human being is unique. It results in large appearance, shape and pose variation.

As mentioned in [2], most of existing pedestrian protection systems rely on exteroceptive sensors such as laser, camera or radar. Contrary to camera, laser range finder data are not subject to illumination changes. Moreover, static targets (i.e. a pedestrian waiting to cross the street) can be detected by means of a laser which is not always possible with radar sensors.

Laser sensors use infra red light to illuminate the environment and obtain a 2/3D point cloud representing its surrounding environment. In the past years, laser sensors were mainly used for obstacle detection and tracking for autonomous vehicles [3], [4], [5]. In this work, we are addressing road obstacle classification based on laser measurements. Once points are clustered in objects, one may like to characterize these clusters and recognize objects that are close to the vehicle. Thus, if a vehicle is able to distinguish between static (buildings, trees, posts,...) and moving objects (pedestrians, cyclists, vehicles,...) safer decisions can be taken. In the past years, machine learning such as Adaboost [6], [7], [8] or Support Vector Machines (SVM) [9] has been applied to tackle the classification problem.

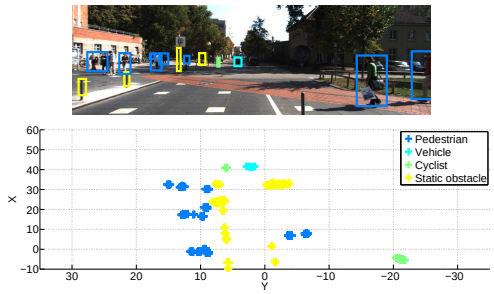


Fig. 1: Example of laser scan with labels associated to every laser impact. The corresponding image of the scene is displayed with bounding boxes around the objects.

Automotive sensors produce data at a certain time frequency. This information is exploited by adding tracking of the road obstacles [5], [10], [11]. Indeed, usual objects recognition methods in computer vision aim at classifying objects for each available data [12], [13] and do not integrate the information brought by consecutive observations. When similar approaches have been applied with laser data [6], [7], [8], the authors try only to get the best recognition performances for each data. In our paper, road obstacles are classified by considering time integration of consecutive classification scores for every track.

In this paper, we are introducing a new system estimating road obstacles classes from laser measurements. In this approach, 2D features are calculated for each detected object and a multiclass classifier is used to obtain decision values about each possible class. Finally, these decision values are integrated over time to refine classification. In this paper, 85 features are presented to describe the different road obstacles classes and a *strong* classifier is build using adaboost training. In the evaluation part, we provide insights about the best features by analyzing which one are the most important during the classification procedure.

The paper is organized as follows. Section II highlights the related work. In Section III, our method for multiclass object recognition from laser point cloud is described. The performances of the method are evaluated in Section IV

[1] Pierre Merdrignac, Evangeline Pollard, Fawzi Nashashibi are with RITS Project-Team, INRIA Rocquencourt, Domaine de Voluceau, B.P. 105, 78153, Le Chesnay, FRANCE. name.surname@inria.fr

[2] Pierre Merdrignac is with Institut VEDECOM, 77, rue des Chantiers, 78000, Versailles, FRANCE.

Finally, Section V concludes this paper.

II. Related Work

Moving obstacles recognition has been a topic of interest for the past years in robotics. In indoor environments, it has first been tackled by assuming that all moving targets were pedestrians, so that, the remaining issue was to deal with target tracking from laser measurements [14]. Arras *et al.* [6] were the first to apply learning methods to recognize persons in 2D laser data. In this work, the authors define a set of 14 features that can be extracted from laser segments and learn a model describing people in indoor environment by applying an adaboost strategy. Performances of this method are evaluated from a confusion matrix comparing classification of person and non person segments in a corridor environment and an office environment. These results are very promising with more than 90% of correct detection and less than 5% of false alarms. However, performances are decreased in a more cluttered environment (here, the office) and when the classifier learned in the office is used in the corridor environment. Thus, directly applying this method to outdoor environment would be too difficult because the clutter is far more important than in an indoor environment.

To achieve perfect recognition in outdoor environments, using a single 2D laser scanner seems too much limited. Hence, some researchers equipped their vehicles with cameras and used a combination of vision recognition methods together with detection on 2D range data [15], [4]. Other researchers decided to use 3D laser scanner to obtain more dense representation of the environment [9], [16], [7] and [8]. Indeed, some laser systems are made of multiple laser sensors scanning the environment over multiple layers and producing 3D point clouds of the surrounding of the vehicle. In [9], the authors calculate features at a point level in order to assign a label to each point of the point cloud. Then, SVM classifier is used to recognize between "vehicle" and "non-vehicle" data. In [16], the authors define two kinds of descriptors in laser data : *segment descriptors* to describe the appearance of an object and *holistic descriptors* to describe more general properties such as speed and acceleration. In this work, *segment descriptors* are calculated from various appearance descriptors inspired by computer vision methods such as spin images [12] and histogram of oriented gradients (HOG) [13]. This last approach may face a computational burden due to the high number of data required to calculate *segment descriptors*. Layer-based approaches have been proposed in [7] and [8]. Both rely on adaboost to train the classifier. These approaches propose simpler solutions that should reduce the computation cost with still acceptable performances, but, they did not consider how consecutive observations can be integrated to obtain a better estimation of each track class.

Within this paper, we decided to focus on classification with 2D lasers because we believe that more information can be extracted from these measurements than what was done in the previous works. The main contribution of the work is the introduction of a new system to dynamically estimate the class of road obstacles. This is made possible with 1) the definition of a new feature set fitted to road obstacles, 2) the design of a classifier using machine learning approach, and, 3) the integration of classification decision values given by successive observations.

III. Multiclass object recognition with laser data

A. Problem statement

In this paper, we are considering a 2D or a 3D laser points cloud has been correctly segmented and consecutive observations of the same objects are associated. Let us define $\mathscr{C} = \{C_i; i = 1, ..., c\}$, a set of c classes of road obstacles and a *track* as a temporal series of estimated states of a dynamic target. A track T_k, at a time step k, is defined in eq. (1) where $\hat{x}_k$ is the estimated position and velocity and $\hat{C}$ is the estimated class of T_k. At time k, a new observation is given by the laser sensor, called here a *cluster* S^k. Moreover, in this paper, we are focusing on estimating the class $\hat{C}$ of T_k. We are assuming $\hat{x}_k$ is estimated using a standard tracking method such as Kalman filtering.

$$T_k = \{\hat{x}_k, \hat{C}\} \tag{1}$$

In [17], the authors provide a wide review on statistical pattern recognition. Following [17], a *cluster* $S \in \mathscr{S}$, where $\mathscr{S}$ is the set of possible *segments*, is described by a vector of l feature values $\mathbf{x} = [x_1, x_2, ...x_l]$ where x_i is a real value given by applying feature function $f_i : \mathscr{S} \rightarrow \mathbb{R}$ to S. A Bayesian decision rule is applied to $\mathbf{x}$ in order to classify S. Here, a supervised learning approach has been chosen to design the classifier, i.e., the decision rule is learned from a training set with labeled *clusters*.

In our approach, we defined a large set of features describing road objects and applied adaboost method to learn a *strong* classifier offline from a predefined feature set.

Finally, in this system, every *cluster* S^k is classified from statistical pattern recognition and classification results are fused to estimate $\hat{C}$.

B. Geometric features definition

In this work, we propose an extending set of 2D geometric features from considerations about moving obstacle classes (pedestrians, cyclists and vehicles) encountered in urban environments such as :

- Pedestrians legs can be observed as two *sub-clusters* in the pedestrian point cloud.
- Bicycles are roughly linear *clusters*.
- Depending on the point of view, vehicles can be either observed with one line (usually its back or its front) or two perpendicular lines.

In Table I, a list of 2D geometric features describing road obstacles is given. This list is split in four categories : generic, vehicle, pedestrian, and, obstacle centered features. To express these features, the following magnitude are defined :

- O_S, the center of S.

Generic

Nr	Feature Name	Nr	Feature Name	Nr	Feature Name
f1	Number of points	*f2*	Width	*f3*	Length
f4	Standard deviation	*f5*	Mean deviation from median	*f6*	Linearity
f7	Circularity	*f8*	Radius	*f9*	Boundary length
f10	Boundary regularity	*f11*	Mean curvature	*f12*	Mean angular difference
f13-f14	Left & Right jump distance	*f15*	Speed	*f16*	Orientation
f82	Normalized number of points	*f83*	Mean intensity	*f84*	Max intensity
f85	Intensity standard deviation	*f47-f51*	*f2-f5*, *f16* Derivative		

Vehicle

Nr	Feature Name	Nr	Feature Name
f17	Angle between s_1 and s_2	*f18*	Length of s
f19-f20	Max & Min length in $\{s_1,s_1\}$	*f21*	Ratio between length of s_1 and s_2

Pedestrian

Nr	Feature Name	Nr	Feature Name
f22-f23	Min & Max distances in $\{O_SO_1,O_SO_2\}$	*f24*	Distance O_1O_2
f25	Ratio between distances O_SO_1 and O_SO_2	*f26-f39*	*f2-f8* Min & Max in $\{L_1,L_2\}$
f40-f42	*f8-f10* Ratio between L_1 and L_2	*f52-f59*	*f2-f5* Derivative in $\{L_1,L_2\}$
f60-f63	*f7-f8* in $\{L_1,L_2\}$		

Obstacle Centered

Nr	Feature Name
f43-f46	*f2-f5* in R_S
f64-f69	Lateral & Longitudinal standard deviation in $\{R_S, R_1, R_2\}$
f70-f73	Min & Max lateral & longitudinal standard deviation in $\{R_1, R_2\}$
f74-f75	Ratio between lateral & longitudinal standard deviation in R_1 and R_2
f76-f81	Lateral & Longitudinal standard deviation derivative in $\{R_G, R_1, R_2\}$

TABLE I: List of 2D geometric features.

- L_1 and L_2, the *sub-clusters* extracted by applying 2-means clustering on S.
- O_1 and O_2, the centroid positions of L_1 and L_2.
- R_S, R_1 and R_2, the sets S, L_1 and L_2 transformed in the obstacle coordinate system, i.e., centered in O_S and oriented towards its direction of motion.
- s, a segment fitted on *cluster* S.
- s_1 and s_2, the two segments fitted on S to describe two sides of a vehicle.

Generic features are composed with the features introduced in [6] to describe persons in 2D laser data, *cluster* length and object orientation (*f1-f16*). Besides, *f82* is dividing number of points in S by distance to the sensor, i.e, *f1* is normalized wrt distance, and, some descriptors are calculated from laser points intensity (*f83-f85*).

Vehicle features are defined by fitting one or two segments in the *cluster* and calculating diverse values on these segments (*f17-f21*).

Pedestrian features are defined from leg detection in obstacles *clusters*. Various values are calculated from leg *sub-clusters* such as distance between the two centroids O_1 and O_2, distance of these points to the *cluster* center O_S (*f22-f25*), and generic features applied to each *sub-cluster* (*f26-f42*, *f60-f63*).

Obstacle Centered are various descriptors extracted by transforming the measured *cluster* in the obstacle coordinate system (*f43-f46*, *f64-f81*).

Finally, association of consecutive measurements permits to calculate the derivative of some characteristics (*f47-f59*, *f76-f81*).

In conclusion, we define 85 features for road obstacle description by combining already existing descriptors with general characteristics concerning usual road obstacles. In the next section, we explain how a classifier can be constructed from all this set.

C. Adaboost based cluster classification

In this work, it has been chosen to apply supervised learning method to generate a classifier from the previously defined feature set. We briefly describe here the learning method introduced in [18] and successfully applied to laser data in [6] called adaboost.

Boosting has the ability to create a strong classifier from N weak classifiers. The only requirement for the weak classifiers is to perform slightly better than random guess. Therefore, a common approach is to train a simple classifier for each feature f_i. This classifier is described in eq. (2) where h_i is the weak classifier for f_i, θ_i the threshold between the two classes, and p_i is either $+1$ or -1 modifying the inequality direction.

$$h_i(S) = \begin{cases} +1, & \text{if } p_i f_i(S) < p_i \theta_i \\ -1, & \text{otherwise} \end{cases} \tag{2}$$

During the training phase, the weak classifier with the highest classification score is repeatedly selected at each step $n = 1, ..., N$. As all features are used to train a weak classifier, a feature f_i can be selected to generate a classifier at different steps, e.g., n_1 and n_2, but, with different threshold values θ_{n_1} and θ_{n_2}. Finally, the strong classifier is given in eq. (3) as a combination of the weak classifiers h_n weighted by α_n. Usually, classification decision is given by the sign of $d(.)$.

$$d(S) = \sum_{n=1}^{N} \alpha_n h_n(S) \tag{3}$$

To create a multi-class classifier, a one-versus-all strategy, i.e., for each class $C \in \mathscr{C}$ a classifier is constructed by calculating the boundary for class C versus the other classes, has been chosen. This approach is interesting to obtain a measure of how likely a pattern belongs to a certain class.

In our approach, decision values are calculated for every class, and no decision is taken at this level. Class assignment to each *track* is described in the next section.

D. Track class estimation

At the final stage, a class $C \in \mathscr{C} \cup C_0$ is assigned to each *track* T_k, where C_0 corresponds to the 'No Object' class, based on *cluster* classification.

Let us consider that T_k is composed of k *clusters* $\{S^1, ..., S^k\}$. The multiclass classification of S^k returns c decision values $\{d_i^k\}_{i=1,...,c}$ from eq. (3). Hence, the vector $\mathbf{D}^k = [d_1^k, ..., d_c^k]^T$ can be considered as an observation of the class of T_k to estimate the *posterior* distribution $p(C_i|\mathbf{D}^k, ..., \mathbf{D}^1)$ for each class C_i. At each time k, thanks to Bayes rule, $p(C_i|\mathbf{D}^k, ..., \mathbf{D}^1)$ can be updated from the previous time $k-1$, and, the likelihood of the observation given the class C_i, $p(\mathbf{D}^k|C_i)$, as described in eq. (4).

$$p(C_i|\mathbf{D}^k, ..., \mathbf{D}^1) = \frac{p(C_i|\mathbf{D}^{k-1}, ..., \mathbf{D}^1)p(\mathbf{D}^k|C_i)}{\sum_{i=0}^{c} p(C_i|\mathbf{D}^{k-1}, ..., \mathbf{D}^1)p(\mathbf{D}^k|C_i)} \quad (4)$$

Furthermore, the likelihood function $p(\mathbf{D}^k|C_i)$ is defined in eq. (5) where $\mathrm{sigm}(.)$ is the sigmoid function [1] often used with neural networks. This function takes its values in $\mathbb{R}$ and returns a real value in the interval $]0, 1[$. Thus, it can measure how likely a *cluster* belongs to each class. Then, likelihood for class C_0 is defined as the minimum of all the likelihoods of not belonging to a class, i.e., how likely a *cluster* does not belong to any of the predefined classes. This likelihood is not a probability distribution function (pdf), however, the denominator in eq. (4) is a normalizing term which ensures the *posterior* distribution sums to 1.

$$p(\mathbf{D}^k|C_i) = \begin{cases} \min\limits_{j=1,...,c} (1 - \mathrm{sigm}(d_j^k)), & \text{if } i = 0 \\ \mathrm{sigm}(d_i^k), & \text{if } i \neq 0 \end{cases} \quad (5)$$

Finally, the estimated class $\hat{C}$ of T_k is defined in eq. (6) by the Maximum a Posteriori (MAP).

$$\hat{C} = \underset{C_i \in \mathscr{C}}{\mathrm{argmax}}(p(C_i|\mathbf{D}^k, ..., \mathbf{D}^1)) \quad (6)$$

IV. Performance evaluation

A. Acquisition of laser data

The proposed solution is evaluated with Kitti data set [19] which gives a large number of laser data in many road environments (city, campus, residential and road). The available laser sensor is a Velodyne HDL-64e providing data with 360° horizontal field of view on 64 layers, which makes about 125000 points, every 100 ms. As our approach is designed for 2D laser data, we simulated a single layer laser from Velodyne data 60 cm above the ground. Moreover, Velodyne data are not annotated in Kitti data set. Therefore, we labeled seven sequences with different classes of frequent road obstacles such as : pedestrians, cyclists, vehicles, trees and posts. To this end, Table II shows the number of *clusters* and *tracks* contained in the 2D data set.

1. $\mathrm{sigm}(x) = \frac{1}{1+e^{-x}}$

	Number of *clusters*	**Number of *tracks***
Pedestrians	2550	29
Vehicles	2557	36
Bicyclists	723	10
Static objects (Trees and posts)	2089	38

TABLE II: Description of the data set.

Laser based classification is evaluated with two different classification systems :

- *Arras :* state of the art for 2D pedestrian classification introduced in [6].
- *RoadFeat+Adaboost :* the extended set of road features presented in Sec.III-B with adaboost training.

During evaluation, each sequence among the seven annotated ones is alternately selected as **validation set**. The remaining sequences are used to train a classifier.

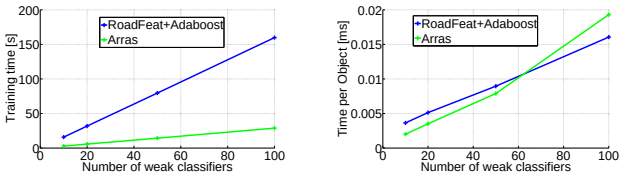
In the following sections, we first investigate which are the most relevant features used with adaboost for both *Arras* set and our *RoadFeat* set in road environment. Performances of these two feature sets are compared for every class, and, we focus on the performances of our classifier versus laser height. Finally, track classification results are shown.

B. Feature selection

In this section, the two classifiers are trained on our labeled data. Table III displays the five best weak classifiers corresponding to five features that are selected for each class. As mentioned in [6], *f8* which is the radius of the circle fitted to the *cluster* is the best feature to classify pedestrians with Arras features, but also to classify cyclists with both techniques. However, with our proposed feature set, *f32* corresponding to the *Maximum Mean Deviation from Median* between the two *sub-clusters* is the most important feature. Concerning vehicles, *f15* that is their measured *speed* is the best feature for both methods. In addition, we notice that the second best features for pedestrians, vehicles and cyclists are *f83* and *f85* which are calculated from laser impacts intensity. This shows how this value presents an interest for obstacle classification.

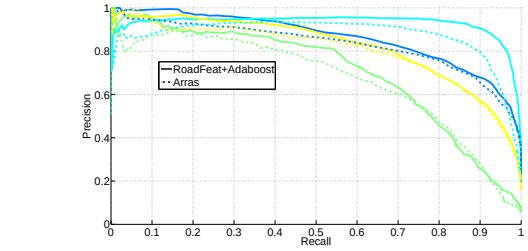
Class \ Selected feature		1st	2nd	3rd	4th	5th
Pedestrian	*Arras*	*f8*	*f12*	*f8*	*f15*	*f11*
	RoadFeat	*f32*	*f83*	*f39*	*f16*	*f31*
Vehicle	*Arras*	*f15*	*f7*	*f5*	*f7*	*f9*
	RoadFeat	*f15*	*f85*	*f7*	*f32*	*f40*
Cyclist	*Arras*	*f8*	*f6*	*f1*	*f9*	*f11*
	RoadFeat	*f8*	*f83*	*f6*	*f39*	*f1*
Static Obstacle	*Arras*	*f15*	*f1*	*f13*	*f5*	*f5*
	RoadFeat	*f32*	*f15*	*f13*	*f1*	*f5*

TABLE III: Best feature selected by adaboost weak classifier.



(a) (b)

Fig. 2: **Processing Time** - (a) Training - (b) *Cluster* classification.



Pedestrian Vehicle Cyclist Static Obstacle

Fig. 3: Precision Recall.

Our solution is implemented using matlab and is not yet designed for embedded systems, however, we can already bring some insights about computation time. Fig. 2 displays the time needed for training the classifiers and for classifying a *cluster* versus the number of weak classifiers. As features number is increased with our solution, the training time is much more important and requires more than $150s$ with 100 weak classifiers. However, we notice on Fig. 2.(b) that classification time is only dependent on the number of classifiers and is less than $0.02ms$ with 100 weak classifiers. In addition, due to increase of the whole set and calculation of more complex functions such as k-means, and segment fitting, time for calculating feature vector is augmented from $0.15ms$ with *Arras* features to $1.14ms$ with our *RoadFeat* features for every object.

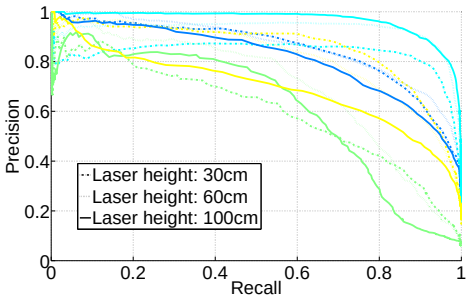
In conclusion, specific road features are selected in priority at the cost of a higher computation time. For system optimization, only the most efficient ones could be calculated.

C. Cluster classification

In this section, performances of adaboost based classification are established using the two feature sets defined in Section IV-A. Fig. 3 shows classification performances for each class. We can see on this figure that the defined *RoadFeat* feature set can outperform, or, at least, does not degrade, *Arras* feature set for every class. The major improvement is obtained with vehicle class with a precision of almost 90% for a recall of 90%.

For pedestrian class, we get a precision of 80% for a recall of 70% with the two methods. This is lower than the results introduced in [6] in an indoor environment, but, it is interesting considering the complexity of urban environment.

In Fig. 4, we investigated the influence of laser height on classification results for each class by training and evaluating



Pedestrian Vehicle Cyclist Static Obstacle

Fig. 4: Precision Recall versus laser height.

classifiers with three height values : low (30 cm), middle (60 cm), high (100 cm). For pedestrian class, the best results are obtained with middle and low height. Indeed, their legs may not be visible in laser data at higher height. Middle height is preferred to classify cyclists as their detection is better at this level. Finally, the best results for vehicle classification are obtained with high height. In conclusion, it would be interesting with 3D laser scanner to combine classification results of multiple layers with multipart fusion technique as it was proposed in [7].

D. Track classification

The integration of consecutive decision values is used to attribute a class to some *tracks*.

We selected an example for each moving target class : pedestrian, cyclist and vehicle. Results are shown in Fig. 5. For every sub-figure, left bottom figure represents the cluster likelihood at time k, right bottom figure the evolution of *posterior* probability, and, upper figure displays laser points in the bounding box of the *track* in the camera image. We can see on these examples how time integration helps to estimate *track* class. However, we can distinguish some errors at the beginning for the cyclist, Fig. 5.(b). Indeed, in this data, the cyclist is far away at the beginning, making detection very noisy and imprecise, and, leading to an incorrect classification. However, for road safety, this is maybe not a big issue because the class has to be correct when the obstacle comes closer to the vehicle which is the case in this example.

Fig. 6 shows *tracks* classification performances with a confusion matrix. For this evaluation, we considered only the *tracks* with more than ten *clusters* to create ground truth for every class. The class considered as outcome is the one assigned to each obstacle at the end of the experiment, i.e., when all the information is integrated. Accuracy of the classification method is about 87.4% which is very promising considering multiclass challenge. Besides, we notice only one misclassification and one false alarm for the vehicle class, and, three misclassifications and seven false alarms for pedestrian class. It is hard to give a conclusion concerning cyclists due to their small number. Finally, we can see in Fig. 6 that six static obstacles are classified as pedestrians which is the most important source of confusion.

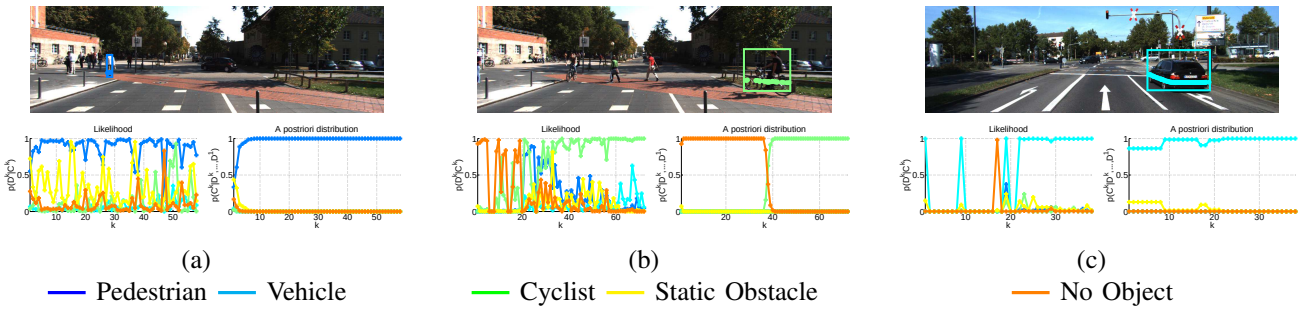


Fig. 5: **Examples of *track* classification** - (a) Pedestrian - (b) Bicyclist - (c) Vehicle

		Condition					
		No Object	Pedestrian	Vehicle	Cyclist	Static Obstacle	Precision
Test Outcome	No Object	36	2	0	2	5	80.0%
	Pedestrian	1	26	0	0	6	78.8%
	Vehicle	1	0	35	0	0	97.2%
	Cyclist	0	0	1	8	0	88.9%
	Static Obstacle	0	1	0	0	27	96.4%
	Recall	94.7%	89.7%	97.2%	80.0%	71.1%	Accuracy 87.4%

Fig. 6: *Track* confusion matrix

As a conclusion, the fusion of consecutive *cluster* classification results allow the estimation of the class for a full *track*.

V. Conclusion

Laser sensors have been proven to be a robust sensor for detection and tracking of moving objects. In this paper, a system classifying road obstacles for detected *clusters* and *tracks* in multiple classes is introduced. We believe that 2D laser scanner can be integrated in future vehicles. Therefore, we designed a system for these kind of sensors by defining a large amount of features. We build a multiclass *cluster* classifier showing how introducing a new feature set dedicated to road obstacles can bring good performances even in complex environments. It is shown that this fusion can disambiguate classification and give promising results. However, the extension of the feature set increases calculation time. Thus, in the final system, only features that are relevant to discriminate classes should be implemented.

Our future work includes the improvement of fusion rule to consider *cluster* detection and occlusion notes and fusion of laser based classification with classification produced by other sensors such as cameras.

Acknowledgment

This work benefits from a financial support of French funding "financement par le plan d'Avenir".

References

[1] WHO, "Global Status Report On Road Safety 2013 : supporting a decade of action," World Health Organization, Tech. Rep., 2013.

[2] T. Gandhi and M. Trivedi, "Pedestrian Protection Systems : Issues, Survey, and Challenges," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 3, pp. 413–430, Sept. 2007.

[3] B. Douillard, J. Underwood, N. Kuntz, V. Vlaskine, A. Quadros, P. Morton, and A. Frenkel, "On the segmentation of 3D LIDAR point clouds," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 2798–2805.

[4] G. Gate, A. Breheret, and F. Nashashibi, "Centralized fusion for fast people detection in dense environment," in *IEEE International Conference on Robotics and Automation*, 2009, pp. 76–81.

[5] S. Reuter, "Multi-object tracking using random finite sets," Ph.D. dissertation, Universität Ulm, 2014.

[6] K. O. Arras, O. M. Mozos, and W. Burgard, "Using Boosted Features for the Detection of People in 2D Range Data," in *IEEE International Conference on Robotics and Automation*, 2007, pp. 3402–3407.

[7] L. Spinello, K. Arras, R. Triebel, and R. Siegwart, "A Layered Approach to People Detection in 3D Range Data," in *AAAI*, 2010.

[8] A. Azim and O. Aycard, "Layer-based Supervised Classification of Moving Objects in Outdoor Dynamic Environment using 3D Laser Scanner," in *IEEE Intelligent Vehicles Symposium (IV)*, 2014.

[9] M. Himmelsbach and A. Müller, "LIDAR-based 3D object perception," in *Proceedings of 1st international workshop on cognition for technical systems*, 2008.

[10] L. Lamard, R. Chapuis, and J.-P. Boyer, "CPHD filter addressing occlusions with pedestrians and vehicles tracking," in *IEEE Intelligent Vehicles Symposium (IV)*, 2013, pp. 1125–1130.

[11] B. Pannetier, J. Dezert, and E. Pollard, "Improvement of multiple ground targets tracking with gmti sensor and fusion of identification attributes," in *IEEE Aerospace Conference*, 2008, pp. 1–13.

[12] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 433–449, 1999.

[13] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, USA, 2005, pp. 1–8.

[14] D. Schulz, W. Burgard, D. Fox, and A. B. Cremers, "People Tracking with a Mobile Robot Using Sample-based Joint Probabilistic Data Association Filters," *The International Journal of Robotics Research*, vol. 22, no. 2, pp. 99–116, 2003.

[15] L. Spinello and R. Siegwart, "Human detection using multimodal and multidimensional features," *IEEE International Conference on Robotics and Automation*, pp. 3264–3269, 2008.

[16] A. Teichman, J. Levinson, and S. Thrun, "Towards 3D object recognition via classification of arbitrary object tracks," *IEEE International Conference on Robotics and Automation*, pp. 4034–4041, 2011.

[17] A. Jain, R. Duin, and J. Mao, "Statistical pattern recognition : A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000.

[18] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, pp. 511–518.

[19] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving ? The KITTI vision benchmark suite," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3354–3361.