Automatic generation of ground truth for the evaluation of obstacle detection and tracking techniques

Hatem Hajri*, Emmanuel Doucet*†, Marc Revilloud*, Lynda Halit*, Benoit Lusetti*,
Mohamed-Cherif Rahal*

*Automated Driving Research Team, Institut VEDECOM, Versailles, France
†InnoCoRe Team, Valeo, Bobigny, France

Abstract— As automated vehicles are getting closer to becoming a reality, it will become mandatory to be able to characterise the performance of their obstacle detection systems. This validation process requires large amounts of ground-truth data precisely describing the pose and kinematics of the obstacles surrounding the vehicle. The creation of such datasets currently requires the lengthy and arduous process of manually interpreting sensor data. In this paper, we propose a novel methodology to generate ground-truth kinematics datasets for specific objects in real-world scenes. Our procedure requires no annotation whatsoever, human intervention being limited to sensors calibration. We present the recording platform which was exploited to acquire the reference data, then fully describe our data generation process. A detailed and thorough analytic study of the propagation of errors in our procedure is also performed. This allows us to provide detailed precision metrics for each and every data item in our datasets.
The main contributions of this paper reside in the data-generation methodology, the analysis of the error propagation, and in the creation of a new dataset.

I. INTRODUCTION

Object detection and tracking both play a crucial role in autonomous driving. They are low-level functions upon which many other increasingly high-level functions are built. These functions include Intention prediction, Obstacle avoidance, Navigation and planning. Being depended on by so many functions, the task of obstacle detection and tracking must be performed with a high level of accuracy and be robust to varying environmental conditions. However, the generation of ground truth data to evaluate obstacle detection and tracking methods usually involves manual annotation, either of images, or of LIDAR point clouds [1][2][3].

This paper showcases a method which takes advantage of the multiplication of autonomous driving platform prototypes in research structures to generate precise and accurate obstacle ground truth data, without requiring the usual phase of painstaking manual labelling of raw data.

Firstly, the methodology applied to generate this data will be described in general terms, and some specific technical topics such as the sensors time-synchronisation method used to collect data will be presented. Then, a thorough analysis of errors propagation is performed. Finally some plots of the collected data are given together with potential applications.

II. GENERAL METHOD DESCRIPTION

The proposed method requires two or more vehicles to generate obstacles dynamics ground truth data. The first vehicle -the ego-vehicle- is equipped with a high-precision positioning system (for example GNSS with Real Time Kinematics (RTK) corrections coupled with an IMU), and various environment perception sensors (for example LIDARs, cameras or RADARs). The other vehicles -the target vehicles- only need to be equipped with a high-precision positioning system, similar to that of the ego-vehicle. By simultaneously recording the position and dynamics of all vehicles, it is possible to express the kinematics of all equipped vehicles present in the scene, in the ego-vehicle frame of reference, and therefore to generate reference data for these vehicles. This reference data can then be used to evaluate the performance of obstacle detection methods applied to the environmental sensors data collected on the ego-vehicle.

III. DATA COLLECTION SETUP

This section, presents in details the set of sensors which were available for the data collection, and details the method applied to ensure the synchronicity of the recording process taking place in different vehicles. Three vehicles were used during this data collection campaign : the ego-vehicle was a Renault Scenic equipped to acquire environment perception data with a high accuracy. The target vehicles were Renault ZOEs, modified to be used as autonomous driving research platforms, and therefore equipped with precise positioning systems.

A. Ego-vehicle perception sensors

To record perception data, the ego-vehicle is equipped with two PointGrey 23S6C-C colour cameras, a Velodyne VPL-16 3D laser scanner (16 beams, 10Hz, 100m range, 0.2° horizontal resolution), a cocoon of five Ibeo LUX laser scanners (4 beams, 25Hz, 200m range, 0.25° horizontal resolution) covering a field of view of 360° around the vehicle.

The cameras are positioned inside the car, behind the windscreen with a baseline of approximately 50cm. The Velodyne VLP-16 is positioned on the roof of the vehicle at a position and height that minimise the occlusion of the laser beams by the vehicle body. The Ibeo LUX are all mounted at the same height of approximately 50cm, two on each front wing (one pointing forward, one to the side), and one at the back, pointing backwards (see Figure 2).

Fig. 1.   The perception vehicle used : two ibeo LUX, the VLP16, GNSS antenna and the precision odometer are visible
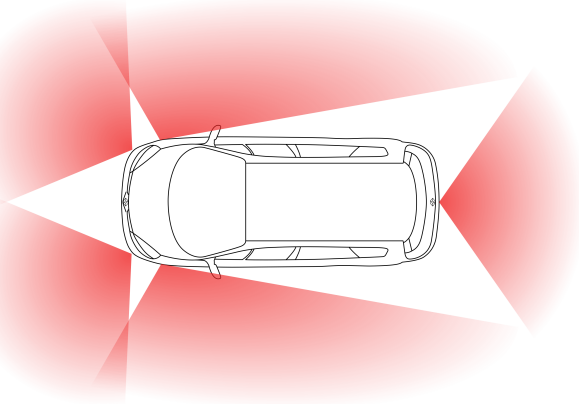

Fig. 2.   Ibeo LUX cocoon setup

B. Precise localisation system

The accuracy of the data generated using our method relies entirely on the accuracy of the vehicles' positioning systems. Therefore, each vehicle was equipped with state of the art positioning sensors : a choke-ring GNSS antenna feeding a GNSS-RTK receiver coupled with a high-grade, fibre optic gyroscopes-based iXblue Inertial Measurement Unit. Additionally, the perception vehicle is equipped with a high-accuracy odometer mounted on the rear-left wheel, while the target vehicles are equipped with a Correvit® high-accuracy, contact-less optical odometer. The data emanating from these sensors is fused using a Kalman Filter-based robust observer which jointly estimates the IMU and external sensors biases. The performance of this system can be further improved in post-processing by employing accurate ephemeris data and smoothing techniques. Table I provides an overview of the combined performance of our positioning system and of the the aforementioned post-processing.

C. Time synchronisation

One of the biggest challenges of performing data collection distributed across multiple vehicles is to precisely synchronise the clocks used for time-stamping the data in each platform. This is especially true when acquiring data in high-speed scenarios. Highway scenarios, in which the absolute

TABLE I

POSITIONING SYSTEM PERFORMANCE

	Heading (deg)	Roll/Pitch (deg)	Position X,Y (m)	Position Z (m)
Nominal GNSS signal	0.01	0.005	0.02	0.05
60 sec GNSS outage	0.01	0.005	0.10	0.07
300 sec GNSS outage	0.01	0.005	0.60	0.40

value of the relative velocity of vehicles may reach 70m/s, require the synchronisation of the vehicle clocks to be at least accurate to the millisecond. This inaccuracy induces an incompressible positioning error, which adds to that of our positioning system (see Subsection VI-A).

To achieve such a precise synchronisation, a Network Time Protocol (NTP) server fed with the pulse-per-second (PPS) signal provided by our GNSS receivers was installed in each vehicle to synchronise the on-board computers in charge of recording all the data. Prior to any recording session, the NTP servers and computer clocks were allowed 12 hours to converge to a common time.

IV. SENSORS CALIBRATION

Generating ground truth data requires very accurate sensors calibration. Given the difficulty of calibrating LIDAR sensors relatively to cameras[1], we propose the following calibration pipeline : first, the positioning system is calibrated, then the cameras are calibrated intrinsically, and finally the rigid transformations relating cameras and LIDARs to the vehicle frame are estimated. To take into account any potential effect of the sensors positioning on their calibration, all sensors intrinsic parameters estimation processes take place once the sensors are installed in the vehicle.

A. Positioning system calibration

The calibration of the positioning system consists in calculating the position and attitude of all positioning sensors (GNSS antenna, optical odometer) in the frame of reference of the IMU. After a phase of initialisation during which the vehicle remains static to allow the estimation of all sensors biases and the convergence of the RTK, the vehicle is manoeuvred. By comparing the motion information emanating from each sensor and comparing it to that of the IMU, one is then able to determine the rigid transformation between the said sensor and the IMU.

B. Cameras calibration

To estimate the intrinsic parameters of our cameras, and the relative pose of our stereo pair, we sweep the space in front of the cameras at three different depths, making sure to cover the whole field of view of each camera. We then use a mixture of Geiger's checkerboard pattern detector [4] and of the sub-pixellic corner detector from openCV to extract the corners of the checkerboard. These are then fed to openCV's `stereoCalibrate` function to jointly estimate

the intrinsic parameters of each camera and their relative pose. The set of parameters thus obtained typically yields re-projection errors of less than 0.3 pixel.

The position and orientation of our cameras are obtained in a semi-automatic fashion. It involves minimising the reprojection error of a known pattern in the images. Using laser tracers, the position of the cameras in the vehicle, and the position and orientation of the vehicle relative to the ground pattern are precisely measured. The orientation of the cameras in the vehicle frame is then estimated by minimising the reprojection error of the pattern characteristic points in the image using a Levenberg-Marquardt minimisation method.

C. LIDARs calibration

1) Velodyne VLP-16 calibration: The objective of the Velodyne to IMU calibration process is to determine the rigid transformation $T_{Vel \rightarrow IMU}$ between the Velodyne reference frame and that of the IMU. Our Velodyne to IMU calibration process is fully automated. Using Iterative Closest Point, we match the 3D point clouds acquired during the calibration manoeuvre one to another to generate a trajectory. Then, the pose of the IMU is re-sampled to the timestamps of the Velodyne scans. The estimation of $T_{Vel \rightarrow IMU}$ knowing both trajectories is similar to the *hand-eye calibration problem*. This estimation is performed using a non-linear optimisation process applied to 1000 pose samples [5]. This rigid transformation estimate is then refined by repeating the process, using $T_{Vel \rightarrow IMU}$ and the linear and angular velocities of the vehicle to correct the motion-induced distorsion of the Velodyne point clouds. This process usually converges in just one iteration.

2) Ibeo LUX: The calibration of an Ibeo LUX cocoon is slightly more complicated than that of a single Velodyne, as it involves simultaneously calibrating all the sensors. Indeed, calibrating each LUX separately will almost certainly result in a poorly consistent point cloud when aggregating clouds from all sensors. Likewise, precisely calibrating one sensor, and then calibrating all sensors relative to their neighbour will also lead to such poor results. A simple way to ensure the global coherence of the cocoon calibration is to use the point cloud from a calibrated Velodyne as a reference, and to calibrate all Ibeo LUX sensors relative to this spatially coherent reference.

V. GROUND-TRUTH DATA GENERATION

The objective of the ground-truth generation process is, at every instant, to precisely and accurately describe the position, orientation and kinematics of an observed obstacle in the frame of reference of the ego-vehicle. This data can then be used to estimate the accuracy and precision of an obstacle detection method or sensor.

A. Notations

Let $X_i^k = (x, y, v_x, v_y, \psi)_i^k$ be the state of vehicle i, with $(x, y)_i^k$ the position of its reference point, $(v_x, v_y)_i^k$ its velocity vector, and ψ_i^k its yaw angle, all expressed in the frame of reference k.

In the rest of the paper, $\cdot_e$ and $\cdot_t$ respectively denote state variables of the ego and of the target vehicle, and $\cdot^{UTM}$ and $\cdot^{ego}$ respectively denote a variable expressed in the Universal Transverse Mercator and in the ego-vehicle frame of reference.

B. Processing

Generating a set of obstacle ground truth data from high accuracy positioning recordings is a two step process :

- generate the relative position and dynamics of the obstacles relative to the ego-vehicle,
- generate data carrying obstacle semantics from the previously generated data.

For each sensor recording, a whole set of ground truth data is generated, so as to provide ground truth synchronised with the sensor data. At each sensor data timestamp, the position and dynamics of each vehicle are estimated from the positioning system recording using temporal splines interpolation.

From this, simple kinematics and velocity composition formulae allow the computation of the relative position and dynamics of the target vehicles in the ego-vehicle frame of reference :

$$\begin{bmatrix} x \\ y \end{bmatrix}_t^{ego} = R(-\psi_e^{UTM}) \begin{bmatrix} x_t - x_e \\ y_t - y_e \end{bmatrix}^{UTM} \quad (1)$$

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix}_t^{ego} = R(-\psi_e^{UTM}) \begin{bmatrix} v_{x_t} - v_{x_e} + \dot{\psi}_e(y_t - y_e) \\ v_{y_t} - v_{y_e} - \dot{\psi}_e(x_t - x_e) \end{bmatrix}^{UTM} \quad (2)$$

$$\psi_t^{ego} = \psi_t^{UTM} - \psi_e^{UTM} \quad (3)$$

Where $R(\alpha)$ denotes a rotation in SO_2 of angle α

C. Exploitation

The reference relative positioning data thus obtained can then be used to generate ground truth perception data. One can for example generate the bounding box of the target vehicle in the ego-vehicle frame of reference to evaluate the performance of LIDAR or image-based object detection and tracking algorithms. Another possibility is to use 3D models of the target vehicles and to project them in the camera images to automatically generate a partial image segmentation.

VI. UNCERTAINTY PROPAGATION ANALYSIS

In this section, a sensitivity analysis of the proposed ground truth generation process is performed, to characterise the accuracy and precision of the generated data, depending on the performance of our positioning systems, and the clock shift between the vehicles.

The inputs of the generation process are made of position, velocity and heading estimates provided by a GNSS-INS fusion

system. These can be modelled as independent, Gaussian random variables[6].

Therefore, the position, velocity and yaw angle are treated separately, so as to limit the calculation hurdle.

A. Position

Equation (1) yields :

$$\begin{bmatrix} x \\ y \end{bmatrix}_t^{ego} = F(dx^{UTM}, dy^{UTM}, \psi_e^{UTM})$$

with :

$$F(dx, dy, \psi_e) = \begin{bmatrix} dx \cos \psi_e + dy \sin \psi_e \\ dy \cos \psi_e - dx \sin \psi_e \end{bmatrix}.$$

Lemma 1: Let Ω be a Gaussian random variable such that $\Omega \sim \mathcal{N}(m_\Omega, \sigma_\Omega^2)$. Then:

$$\mathbb{E}(\cos(\Omega)) = \cos(m_\Omega) \, e^{-\sigma_\Omega^2/2}$$
$$\mathbb{E}(\sin(\Omega)) = \sin(m_\Omega) \, e^{-\sigma_\Omega^2/2}$$

Proof: Using the explicit expression of the characteristic function of a Gaussian variable yields:

$$\mathbb{E}(\cos(\Omega) + i \sin(\Omega)) = \mathbb{E}(e^{i\Omega}) = e^{i \, m_\Omega - \sigma_\Omega^2/2}$$

The real and imaginary part of this expression yield the desired result. ∎

Under the assumption that $\mathrm{Var}(dx) = \mathrm{Var}(dy) = \sigma_{dx}^2$, and noting $\mathrm{Var}(\psi_e) = \sigma_\psi^2$:

$$\mathrm{Cov}(F(dx, dy, \psi_e)) = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$$

With :

$$a = \sigma_{dx}^2 + \mathbb{E}(dx)^2 \mathrm{Var}(\cos \psi_e) + \mathbb{E}(dy)^2 \mathrm{Var}(\sin \psi_e)$$
$$- \mathbb{E}(dx)\mathbb{E}(dy)\sin(2\mathbb{E}(\psi_e))e^{-\sigma_\psi^2}(1 - e^{-\sigma_\psi^2})$$

$$b = \sigma_{dx}^2 + \mathbb{E}(dx)^2 \mathrm{Var}(\sin \psi_e) + \mathbb{E}(dy)^2 \mathrm{Var}(\cos \psi_e)$$
$$+ \mathbb{E}(dx)\mathbb{E}(dy)\sin(2\mathbb{E}(\psi_e))e^{-\sigma_\psi^2}(1 - e^{-\sigma_\psi^2})$$

$$c = \frac{1}{2}\sin(2\mathbb{E}(\psi_e))e^{-\sigma_\psi^2}(1 - e^{-\sigma_\psi^2})(\mathbb{E}(dx)^2 - \mathbb{E}(dy)^2)$$
$$- \mathbb{E}(dx)\mathbb{E}(dy)\cos(2\mathbb{E}(\psi_e))e^{-\sigma_\psi^2}(1 - e^{-\sigma_\psi^2})$$

Therefore, under the assumption that the variance of the position error is similar from one vehicle to another, and along North and East axes ($\sigma_x^2 = \sigma_y^2 = \sigma_{pos}^2 = \frac{1}{2}\sigma_{dx}^2$), and that the maximal distance between the ego-vehicle and an obstacle is d_{max}, we propose the following upper bound for the position error covariance matrix:

$$a \leq 2\sigma_{pos}^2 + 2d_{max}^2(1 - e^{-\sigma_\psi^2}),$$
$$b \leq 2\sigma_{pos}^2 + 2d_{max}^2(1 - e^{-\sigma_\psi^2}),$$
$$c \leq \frac{3}{2}d_{max}^2(1 - e^{-\sigma_\psi^2/2}).$$

B. Velocity

Equation (2) yields :

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix}_t^{ego} = G\left((dx, dy, dv_x, dv_y, \psi_e, \dot{\psi}_e)^{UTM}\right)$$

with $G(dx, dy, dv_x, dv_y, \psi_e, \dot{\psi}_e)$ equal to :

$$\begin{bmatrix} \cos \psi_e(dv_x + \dot{\psi}_e dy) + \sin \psi_e(dv_y - \dot{\psi}_e dx) \\ \cos \psi_e(dv_y - \dot{\psi}_e dx) - \sin \psi_e(dv_x + \dot{\psi}_e dy) \end{bmatrix}$$

Using the independence of the input variables, it is possible to express the covariance matrix of $G(dx, dy, dv_x, dv_y, \psi_e, \dot{\psi}_e)$ as a function of the first and second order moments of the input variables. Due to space limitations, we only give bounds for the elements of this matrix. Similarly to Section (VI-A), these bounds tend to 0 as the variances of the input variables tend to 0.

Lemma 2: Let X and Y be two independent random variables with means m_X and m_Y and variances σ_X^2 and σ_Y^2. Let Ω be a Gaussian random variable $\mathcal{N}(m_\Omega, \sigma_\Omega^2)$ independent of X, Y. Let $Z = \cos(\Omega)X + \sin(\Omega)Y$. Then :

$$\mathrm{Var}\,(Z) \leq \sigma_X^2 + \sigma_Y^2 + (1 - e^{-\sigma_\Omega^2})(|m_X| + |m_Y|)^2$$

Proof: Using Lemma 1, we have the following bound:

$$\mathrm{Var}(Z) \leq \sigma_X^2 + \sigma_Y^2 + m_X^2 \mathbb{E}(\cos^2(\Omega)) + m_Y^2 \mathbb{E}(\sin^2(\Omega))$$
$$- e^{-\sigma_\Omega^2}(m_X^2 \cos^2(m_\Omega) + m_Y^2 \sin^2(m_\Omega))$$
$$- m_X m_Y \sin(2m_\Omega)e^{-\sigma_\Omega^2}(1 - e^{-\sigma_\Omega^2})$$

Using the identity $\cos^2(\Omega) = \frac{1 + 2\cos(\Omega)}{2}$ and again Lemma 1, we get :

$$m_X^2 \mathbb{E}(\cos^2(\Omega)) - e^{-\sigma_\Omega^2}m_X^2 \cos^2(m_\Omega) \leq m_X^2(1 - e^{-\sigma_\Omega^2})$$

The same identity holds with X replaced with Y and cos replaced with sin. This gives the desired bound. ∎

Let us denote the mean and variance of a Gaussian variable Z respectively by m_Z and σ_Z^2. Under the assumption that $\mathrm{Var}(dv_x) = \mathrm{Var}(dv_y) = 2\sigma_{vel}^2$ and $\mathrm{Var}(dx) = \mathrm{Var}(dy) = \sigma_{dx}^2$, we have :

$$\mathrm{Var}(dv_x + \dot{\psi}dy) = 2\sigma_{vel}^2 + \sigma_{dx}^2\sigma_{\dot{\psi}}^2 + m_{\dot{\psi}}^2\sigma_{dx}^2 + m_{dy}^2\sigma_{\dot{\psi}}^2$$

$$\mathrm{Var}(dv_y - \dot{\psi}dx) = \mathrm{Var}(dv_x + \dot{\psi}dy)$$

Let $\begin{bmatrix} a & c \\ c & b \end{bmatrix}$ be the covariance matrix of $\begin{bmatrix} v_x \\ v_y \end{bmatrix}_t^{ego}$. Using the previous results, the upper bounds for a, b and c are:

$$a, b \leq 4(\sigma_{vel}^2 + \sigma_{pos}^2\sigma_{\dot{\psi}}^2 + \dot{\psi}_{max}^2\sigma_{pos}^2) + 2d_{max}^2\sigma_{\dot{\psi}}^2,$$
$$+ 4(1 - e^{-\sigma_\psi^2})(v_{max} + d_{max}\dot{\psi}_{max})^2$$

$$c \leq \sqrt{a^2}\sqrt{b^2},$$

where v_{max} is an upper bound for dv_x and dv_y and $\dot{\psi}_{max}$ is an upper bound for $\dot{\psi}$.

C. Yaw angle

The variance of the relative yaw estimate ψ_t^{ego} is trivially derived from equation (3). Considering that the heading error covariance is similar in the target and ego-vehicle ($\sigma^2_{\psi_t^{UTM}} = \sigma^2_{\psi_e^{UTM}} = \sigma^2_\psi$), we get:

$$\mathrm{Var}(\psi_t^{ego}) = 2\sigma^2_\psi$$

D. Results

Using the analytic upper bounds for the covariance matrices of the generated ground truth position, speed and yaw angle data calculated in the previous sections, we can estimate the precision of the data generated using the process described in section V.

The following values will be used :

σ_{pos}	0.02	m
σ_{vel}	0.02	m.s^{-1}
σ_ψ	$1.75.10^{-3}$	rad
d_{max}	50	m
v_{max}	36	m.s^{-1}
$\dot{\psi}_{max}$	1	rad.s^{-1}

They are representative of the performance of our positioning system, of the distance at which obstacles become truly observable, and of the maximal dynamics of the vehicle in typical use cases.

These values, when injected in the upper bounds of the position and velocity covariance matrices yield the following results :

$$\|\mathrm{Cov}((x,y)_t^{ego})\|_F^{1/2} \leq 0.12 \ m$$
$$\|\mathrm{Cov}((v_x,v_y)_t^{ego})\|_F^{1/2} \leq 0.30 \ m.s^{-1}$$

These values represent the root mean square error on the position and velocity information yielded by our ground truth generation process. We stress the fact that they are obtained by taking extreme values for all variables describing the dynamics of both vehicles, which can never actually be encountered simultaneously in real life situations.

VII. DATA VISUALISATION

In this section some figures of the acquired data are given. These data were collected during a tracking scenario on a track located in Versailles France. The track has approximative length of 3.2 km and is represented in Figure 3 together with the starting and arriving points of the vehicles.

The following figures show temporal variations of the relative positions x, y, relative velocities vx, vy and orientation ψ of the tracked vehicle, as estimated by our LIDAR's firmware (in the ego vehicle frame), and the corresponding ground truth values deduced from RTK sensors as explained in the previous section. All these quantities are in the International System of Units.
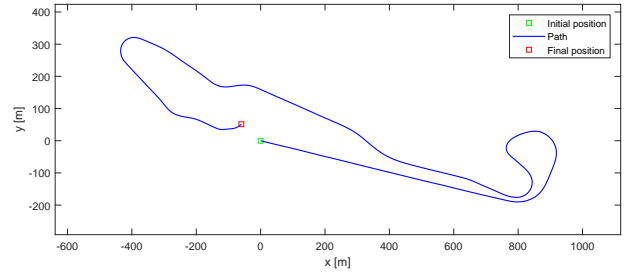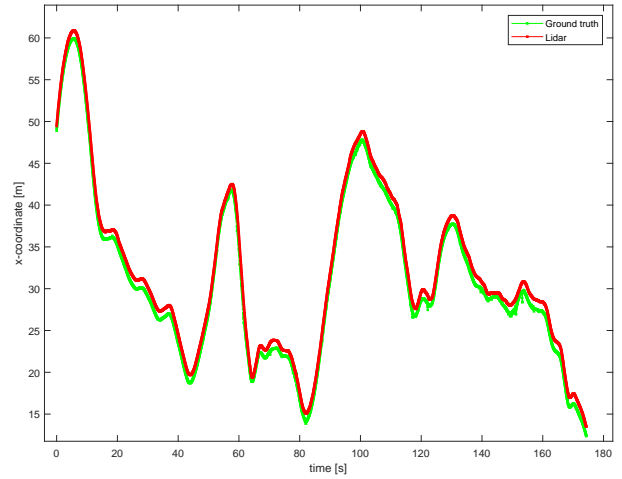

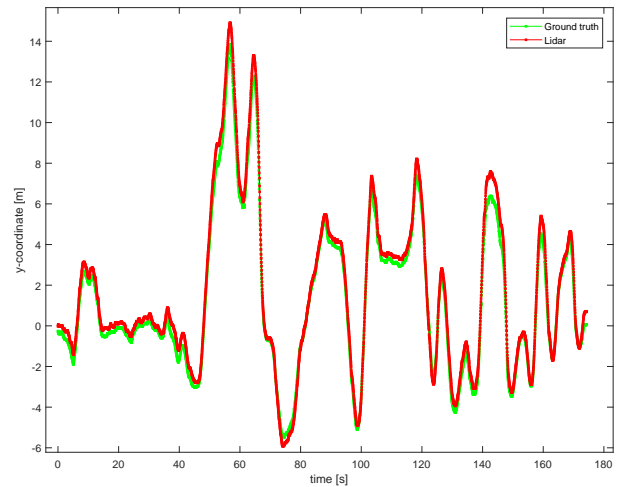Fig. 3.   Test track

Fig. 4.   Variations of x

Fig. 5.   Variations of y

VIII. APPLICATION TO THE EVALUATION OF PERCEPTION ALGORITHMS

This section gives some potential applications of ground truth data to the conception and evaluation of perception algorithms.
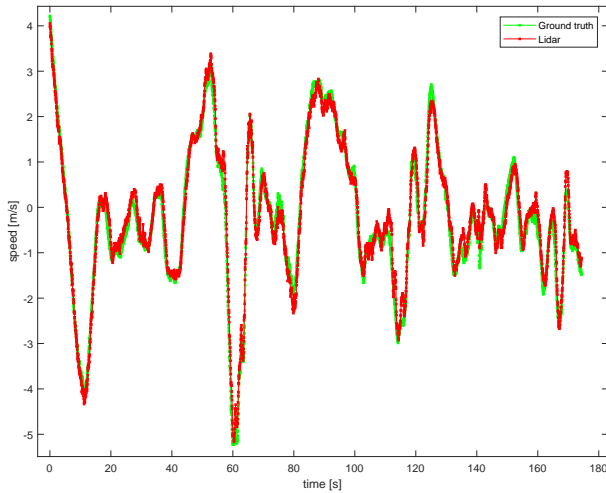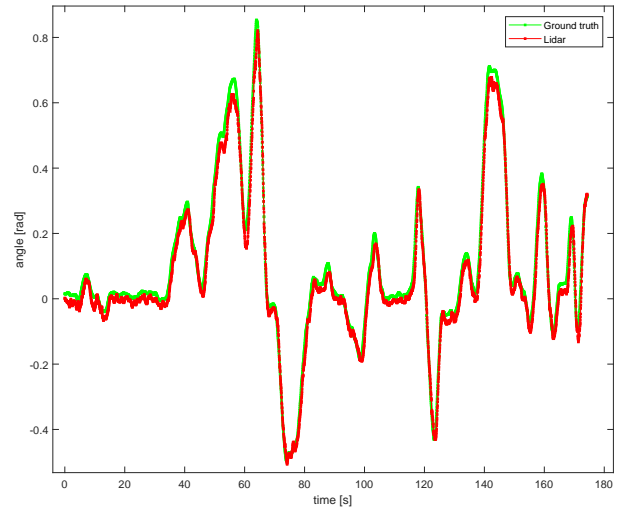
Fig. 6. Variations of vx
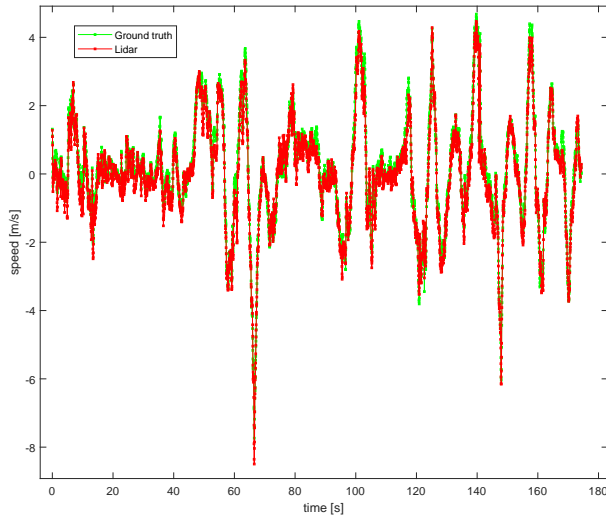

Fig. 7. Variations of vy

Fig. 8. Variations of ψ

A. Clustering Lidar raw data:

Robust obstacle detection using sensors such as Lidar is a key point for the development of autonomous vehicles. Lidar performance mainly depends on its low level processing (the clustering methods used for its raw data, estimations of the bounding boxes, velocities of the representative points of the clusters, etc.). The methodology presented in the paper can be used to evaluate the performance of the main outputs of a Lidar raw data clustering algorithm.

B. Lidar-Radar-Camera Fusion:

Lidars, Radars and Cameras are complementary sensors. Lidars are very accurate on obstacles positions and less accurate on their velocities. On the other hand, Radars are more precise on obstacles velocities and less precise on their positions. Cameras provide images and can be used to perform classification tasks. Fusion between these sensors aims at combining the advantages of each sensor to provide

permanent and more robust data (see for example [7], [8], [9], [10], [11]). In particular, the recent work [7] proposes a Lidar-Radar fusion algorithm with evaluation on the database generated in the present paper.

C. Dynamic models and prediction of obstacles motions:

Robust tracking of obstacles detected by sensors such as Lidars, Radars and Cameras is crucial for a good functioning of autonomous cars. Kalman filters are among the most used methods to track obstacles. These methods are usually coupled to preset models such as constant velocity, acceleration or curvature. Ground truth allows to learn true dynamic models of obstacles using statistics, neural networks etc. The learnt models are to be compared with the preset ones.

IX. SUMMARY AND FUTURE DEVELOPMENTS

In this paper, we have presented a method for automatically generating sets of ground truth data to support advances in the field of obstacle detection and tracking. We hope this methodology will help contributors of this area of research to challenge their approaches, and contribute to the development of robust and reliable algorithms. In the future, we intend to propose a complete benchmark to unify the usage of this dataset and the performance estimation of obstacle detection and tracking techniques.

REFERENCES

[1] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[2] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[3] H. Yin and C. Berger, "When to use what data set for your self-driving car algorithm: An overview of publicly available driving datasets," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Oct 2017, pp. 1–8.

[4] A. Geiger, F. Moosmann, O. Car, and B. Schuster, "Automatic calibration of range and camera sensors using a single shot," in *International Conference on Robotics and Automation (ICRA)*, 2012.

[5] F. Dornaika and R. Horaud, "Simultaneous robot-world and hand-eye calibration," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 4, pp. 617–622, Aug 1998.

[6] X. Niu, Q. Chen, Q. Zhang, H. Zhang, J. Niu, K. Chen, C. Shi, and J. Liu, "Using allan variance to analyze the error characteristics of gnss positioning," *GPS Solutions*, vol. 18, no. 2, pp. 231–242, Apr 2014. [Online]. Available: https://doi.org/10.1007/s10291-013-0324-x

[7] H. Hajri, L. Halit, B. Lusetti, and M.-C. Rahal, "Real time Lidar and Radar high-level fusion for obstacle detection and tracking," *submitted paper*, 2018.

[8] C. Blanc, B. T. Le, Y. Le, and G. R. Moreira, "Track to Track Fusion Method Applied to Road Obstacle Detection," *IEEE International Conference on Information Fusion*, 2004.

[9] M. S. T. G. D. Gohring, M. Wang, "Radar/lidar sensor fusion for car-following on highways," *IEEE Int. Conf. Autom. Robot. Appl.*, pp. 407–412, 2011.

[10] R. O. C. García and O. Aycard, "Multiple sensor fusion and classification for moving object detection and tracking," *IEEE Trans. Intelligent Transportation Systems*, vol. 17, no. 2, pp. 525–534, 2016. [Online]. Available: http://dx.doi.org/10.1109/TITS.2015.2479925

[11] A. Rangesh and M. M. Trivedi, "No Blind Spots: Full-Surround Multi-Object Tracking for Autonomous Vehicles using Cameras & Lidars," *https://arxiv.org/pdf/1802.08755.pdf*, 2018.