Pedestrian Detection and Tracking in Urban Environment using a Multilayer Laserscanner

Samuel Gidel, Paul Checchin, Christophe Blanc, Thierry Chateau, and Laurent Trassoudaine, *Lasmea, France*

Abstract—Pedestrians are the most vulnerable participants to urban traffic. The first step toward protecting pedestrians is to reliably detect them in a real time framework. In this paper, a new approach is presented for pedestrian detection, in urban traffic conditions, using a multilayer laser sensor mounted on board a vehicle. This sensor, placed on the front of a vehicle collects information about distance distributed according to 4 planes. Like a vehicle, a pedestrian constitutes in the vehicle environment an obstacle which must be detected, located, then identified and tracked if necessary. In order to improve the robustness of pedestrian detection using a single laser sensor, a detection system based on the fusion of information located in the 4 laser planes is proposed. The method uses a non-parametric kernel density based estimation of pedestrian position of each laser plane. Resulting pedestrian estimations are then sent to a decentralized fusion according to the 4 planes.

Temporal filtering of each object is finally achieved within a stochastic recursive Bayesian framework (Particle Filter), allowing a closer observation of pedestrian random movement dynamics. Many experimental results are given and validate the relevance of our pedestrian detection algorithm in regard to a method using only a single-row laser-range scanner.

Index Terms—Pedestrian detection, LIDAR, intelligent vehicle, fusion, SIR PF, Parzen kernel method.

1 INTRODUCTION

The pedestrian detection is an essential functionality for intelligent vehicles, since avoiding crashes with pedestrians is a requisite for aiding the driver in urban environments. Currently, in France, more than 535 pedestrians die in road accidents every year, while several hundred thousands are injured. Most accidents (70%) take place in urban areas where serious or fatal injuries often happen at relatively low speeds. So, it is important to develop a pedestrian detection system. These issues take place in the context of the LOVe Project (Software for vulnerables observation) which aims at improving road safety, mainly focusing on pedestrian security [1].

In this paper, a system for pedestrian detection based on an approach using a multilayer laser sensor mounted on board a vehicle is presented. The system is composed of a single multilayer laser sensor mounted on board a vehicle with a variable scan area limited to 150° . It is designed to work in a particularly challenging urban scenario, in which traditional pedestrian detection approaches would yield non-optimal results. Because it must detect all the pedestrians (moving or static), no motion model of persons proposed in the litterature is convenient. Using a laser system in this way presents many difficulties including occlusions, non-rigid targets, obvious limitations of this sensor (no information about shape, contour, texture, color of objects) and varying atmospheric conditions (rain and fog).

For a broad review of the various sensors used for pedestrian detection, one can consult [2] where piezoelectric, radar, ultrasound, laser range scanner sensors and cameras operating in the visible or in the infrared are described. Using video sensors to solve the problems of detection and identification seems natural at first, given the capacity of this type of sensor to detect/analyze the size, the shape and the texture of a pedestrian. Many methods to detect human beings were developed in computer vision based on monocular or stereoscopic images [3]-[6]. However, the strong sensitivity to atmospheric conditions, the wide variability of human appearance, the limited aperture of this sensor and the impossibility to obtain direct and accurate information concerning depth have, among other reasons, given rise to an interest for the development of

a detection method starting from an active sensor like a radar or a laser sensor. In this article, we have chosen to focus on the latter type of sensor. Thus, we are interested in the development of a pedestrian detection technique using only data from a 4-layer laser sensor such as the one developed by [8]. This type of sensor, especially in its mono layer version, has already been used in a great number of practical mobile robotic applications such as SLAM (Simultaneous Localization and Mapping), navigation of robots, detection, localization and tracking of moving objects [9]–[19].

In real traffic conditions, the pitch of a vehicle in motion can cause the system to fail if a single-row laser range scanner is used. In fact, a small pitch movement ($< 1^{o}$) can move the laser plane 50 cm, 30 m away, which can change the information contained in the laser layer. We propose to use the information located in the 4 laser planes in order to solve this drawback and improve the robustness of pedestrian detection algorithm. After this detection stage, a Sampling Importance Resampling based Particle Filter (SIR PF) is used in order to track more easily the pedestrian random movement which can include abrupt trajectory changes.

This paper is organized as follows: in Section 2, a review of articles related to our research interests is carried out in order to position our work in relation to existing methods. In Section 3, our approach, the system and the sensor are described. In the LOVe project framework, the Renault manufacturer uses the IBEO ALASCA XT on board an experimental vehicle. The first part of Section 4 is dedicated to the segmentation of the laser image. In the second part, the method developed to isolate pedestrian objects and to merge the 4 laser layers is described. Finally, in Section 5, the SIR PF is used to track pedestrians. The results obtained on real data from several scenarios are presented in Section 6.

2 RELATED WORK

Many research works have been carried out over the last years concerning pedestrian detection from a laser sensor. The pedestrian recognition application on-board vehicles is particularly challenging with a laser sensor due to the wide range of possible pedestrian appearances, occlusions and the cluttered (uncontrolled) background that are involved. The articles related to this research work can be divided into two main approaches:

- detection and/or tracking pedestrian in a dynamic mode, in other words these methods detect only a moving pedestrian;
- detection and/or tracking pedestrian in static and dynamic mode, in other words these methods detect both static and moving pedestrians.

Several methods suggest using a sensor in a dynamic mode only. Zhao et al. [16] seek and track among the laser data binomials of clusters of points (equivalent to the feet of a pedestrian) which have a known periodic movement. They use a Kalman filter. In other methods the systems is mounted on the mobile platform. Prassler et al. [9], Lindström et al. [12] and Elfes [7] detect pedestrians by assimilating them to particular moving areas depending on their size through a temporal analysis of an occupation grid. Schulz et al. [14] use each local minimum to compute a set of two-dimensional position probability grids, each containing all the probabilities that a pedestrian's legs are at position $\langle x, y \rangle$ relative to the robot. They use a SJPDAF in order to track people.

In the second category, they detect and/or track pedestrians in a static and dynamic mode. Fod et al. [17] propose a method which subtracts a background model to aggregate laser measurements into big blobs and then matches the previous blobs with the new ones for each scan. Their tracking algorithm is based on a Kalman filter. In other methods, the system is mounted on the mobile platform. Szarvas et al. [15] clustered the laser data points before classifying them with a Convolutional Neural Network classifier. Montemerlo et al. compute probabilities based on disparities in x-y space. The probabilities of each point are computed from the Euclidean distance between these points and the closest object, be it a person or an occupied map cell. Fuerstenberg et al. [8] achieve pedestrian recognition by classifying objects thanks to preestablished criteria (vehicles, vulnerable items, etc.), then track pedestrians using a Kalman filter.

All these methods presented above do not use the 4 laser layers from a multilayer laserscanner before making a final decision and furthermore use the traditional Kalman filter which does not seem to take into account the right assumptions in our view (linear evolution model and Gaussian noise).

3 OVERVIEW

3.1 Proposed approach

In order to find a method enabling outdoor pedestrian detection from a 4-plane laser sensor as the one developed by the IBEO company, different works listed in the bibliography guided us in this research. Without prior knowledge of the number of obstacles in the observed scene, a segmentation, and classification algorithm [8] which clusters together the laser measurements and classifies them in different geometrical classes to keep only the objects having a "pedestrian" shape have been chosen. Our approach introduces a new idea which consists in using the information located in the 4 laser planes, before making a final decision. This algorithm uses the Parzen method suggested in Cui's article [22] to extract beforehand the pedestrian objects in each plane before merging them. Contrary to Cui's et al., Parzen methods are used to detect pedestrians without focusing on parts of the body (legs for instance). Moreover the use of the Gaussian kernel containing the geometrical information related to a pedestrian is original. In fact, to improve the performances of pedestrian detection algorithms based on a single laser layer sensor, the basic idea is that missing information at a given time t in some layers can be compensated with the other layers and the wrong information located in one or two layers can be rejected by using the others; and thus the rate of correct detection is increased. In order to track pedestrians from a moving vehicle, a SIR PF [27] is used because it proved an efficient way to track a varying number of targets when a priori knowledge or assumption about the movement of a pedestrian was not available. An overview of this algorithm is presented in Fig. 2.

3.2 The IBEO Laserscanner

In the LOVe project framework, the Renault manufacturer uses the IBEO ALASCA XT. The IBEO laserscanner (see Fig. 1) has a variable scan area up to 270° but limited here to 150° for our experiments. The laserscanner is mounted in the center of the frontal area of Renault test vehicle. From this position, the sensor can detect all relevant objects in front of the vehicle. The manufacturer indicates that the IBEO sensor has a range measurement up to 128 m with a accuracy of +/- 5 cm. The angle of resolution varies with scan frequency (at 20 Hz the resolution angle is 0.5°) thus providing 300 measurements per channel and scan. These scan planes have a total opening angle of approx. 3.2° . Two SMAL video cameras mounted on the top of the vehicle (see Fig. 1) simultaneously record the scene.



Fig. 1: The IBEO ALASCA XT Laserscanner and the Renault test vehicle.

4 ALGORITHM PROPOSED FOR PEDES-TRIAN DETECTION

In this section, the different modules of the object detection algorithm (see Fig. 2) are presented. In the context of the LOVe project, the algorithm has to detect all the moving and not moving pedestrians who are located in front of the vehicle. The first step of the algorithm is a segmentation phase. Then a presentation of the uncommon use of the 4 laser layers and the adopted method in order to extract pedestrian objects is carried out, where, two complementary goals are desired: filtering false detections due to information in one or two layers by using the others; increasing the rate of correct detections which can appear in a single layer by seeking confirmation of detection in the other three.

4.1 Segmentation

Extracting observation from sensor data is the first fundamental stage of any object detection algorithm. For that purpose, the process starts by grouping all the measures of a scan into several clusters, according to the distance between two consecutive points P_i and P_{i+1} , followed by line fitting of the points in each cluster. To extract segments or clusters, the algorithm chosen for our application is part of the algorithms presented and evaluated in [21]. The selected technique minimizes the orthogonal distance between the points of measurement and the



Fig. 2: Pedestrian detection algorithm using a multilayer laser sensor.

estimated line. The process continues by incorporating into clusters (called beacons) all the points that have not been approximated as segments and which represent a large group of points (points close to each other).

Without a priori knowledge of the number of obstacles in the observed scene, our segmentation technique gathers the points in various geometrical classes with the aim of extracting from the laser image the characteristics relating to the walls, cars or panels and to keep only the objects having a "pedestrian" signature (see § 4.2). All the points gathered in the geometrical classes are eliminated from the initial laser layer before looking at the "pedestrian" signature.

4.2 Pedestrian extraction in the 4 laser planes

This section presents the detection system based on the fusion of information located in the 4 horizontal planes; this system allows to improve pedestrian detections in comparison with a method using only a single-row laser-range scanner. The main idea is to reduce false detections which may appear in a single layer by seeking confirmation of detection in the other three. Part of the occlusion problem of an object by a smaller one is also solved thanks to the angular variation of ~ 1.07° between the 4 laser layers. After the segmentation step and the four object classifications in each layer, we propose, a nonparametric method based on a discrete modeling of the probability density of each laser reading using a kernel density estimator [22].

Initially and for each layer, all the points which were not filtered by the segmentation module are used in the Parzen density estimator for calculating the pedestrian probability densities located in the observed laser plane. In the fusion module, all the "pedestrian objects" detected in each laser layer are projected onto a plane parallel to the ground plane. Then, the Parzen density estimator is also used in order to compute the pedestrian probability density located in the observed scene by the 4 laser planes.

4.3 Pedestrian Detection

Using a laser scan, the system must deduce from the position of laser points, a non-parametric representation of the likelihood. Thus the complexity of this model will depend directly on the assumptions made from the pedestrian shape located in a laser image.

• How many raw observations?

• How are these laser measurements dispatched? To answer these questions, we propose an original approach to build a non-parametric model based on kernel functions, allowing a smart selection of the most pertinent 2D laser points from a likelihood analysis function. A likelihood discriminating function permits the classification of each laser point as the pedestrian gravity center or not. This method is not supervised, so no prior knowledge is required to process a laser scan.

Let $\mathbb{Z} \doteq {\{\mathbf{z}_k\}}_{k=1,...,N_s}$ denote the vector composed by 2D laser readings which have not been filtered by the segmentation module in a laser plane. We defined a Bernouilli random variable $w_k \in {\{w_1, w_2\}}$ given by $w_k = w_1$ if the associated event is classified as pedestrian gravity center or $w_k = w_2$ in all other cases.

The likelihood function $p(\mathbf{Z}|w_k)$ can compute the probability that a laser point belongs to the pedestrian gravity center. We propose to illustrate the likelihood $p(\mathbf{Z}|w_k)$ by a non-parametric model using

an estimation based on kernel functions (Parzen window model).

$$p(\mathbf{Z}|w_k) = \frac{1}{Nbpts} \sum_{i=1}^{Ns} \varphi(\mathbf{z}_k, \mathbf{z}_i)$$
(1)

where N_s represents the total number of points present in the image, *Nbpts* represents the number of theoretical points that a pedestrian should send back into a laser layer according to distance (*D*). Finally $\varphi(\mathbf{z}_k, \mathbf{z}_i)$ is the kernel function which allows to modify the zone of influence of a point with its neighbours, it is defined by:

$$\boldsymbol{\varphi}(\mathbf{z}_k, \mathbf{z}_i) = \exp[-\lambda_c \cdot d_c(\mathbf{z}_k, \mathbf{z}_i)]$$
(2)

The λ_c parameter permits to adjust the weight calculation. The d_c distance used is a Mahalanobis distance defined by:

$$d_c(\mathbf{z}_k, \mathbf{z}_i) = (\mathbf{z}_k - \mathbf{z}_i) \Sigma_{\boldsymbol{\varphi}}^{-1} (\mathbf{z}_k - \mathbf{z}_i)^T$$
(3)

with Σ_{φ} , the covariance matrix associated to both pedestrian geometrical components (width and thickness) in a laser image.

$$\Sigma_{\varphi} = \begin{bmatrix} \sigma_{width}^2 & 0\\ 0 & \sigma_{thickness}^2 \end{bmatrix}$$
(4)

Finally the function that weighs the likelihood function $p(\mathbf{Z}|w_k)$ depending on the sensor characteristics, the sought objects and on the detection distance is defined by:

$$Nbpts = \frac{W}{D \cdot \tan \theta} \tag{5}$$

This expression takes into account the pedestrian's dimensions (width: W), the angular resolution (θ) of the sensor according to the distance (D) separating the obstacle from the vehicle.

For each plane, the 2D laser points having the highest probability $\mathbf{z}_k \in w_1$ are chosen by the maximum likelihood estimator as the pedestrians' positions:

$$\mathbf{Z}_{k}|k = \arg\max_{k}(p(\mathbf{Z}|w_{k} \in w_{1}))$$
(6)

Once the pedestrian gravity center is defined, the next step consists in searching all the points belonging to the points group of the gravity center. These points will be eliminated if their distance $d_c(\mathbf{z}_k, \mathbf{z}_i)$

is lower than the threshold α . Thus the point list L^i below is eliminated:

$$L^{i} = \{d_{c}(\mathbf{z}_{k}, \mathbf{z}_{i}) < \alpha\}, \alpha = MaxPedestrianWidth/2$$
(7)

This algorithm is reiterated while the likelihood maximum estimator contains a value which is higher than the trust threshold $\delta \in [0, 1]$ (see Algorithm 1).

Algorithm 1 Pedestrian detection algorithm with a non-parametric estimator

Enter: set of points composed by 2D laser measurements which have not been filtered by the segmentation module $\mathbf{Z} = {\mathbf{z}_i}_{i=1,...,N_s}$

Compute the likelihood function: $p(\mathbf{Z}|w_k)$

initialization: m = 0 and $\mathbf{Z}_0 = \mathbf{Z}$

repeat

m = m + 1

Extraction of maximal likelihood point

$$\hat{\mathbf{z}}_m = \mathbf{Z}_k | k = \arg\max_k (p(\mathbf{Z}_m | w_k \in w_1))$$

Compute the associated points set

$$\mathbf{L}^m = \{\mathbf{Z}_i \in \mathbf{Z} | d_c(\hat{\mathbf{z}}_m, \mathbf{z}_i) < \alpha\}$$

Update input set

$$\mathbf{Z}_{m+1} = \overline{\mathbf{Z}_m \cap \mathbf{L}^m}$$

while: reached stopping criteria: $p(\mathbf{Z}_m | w_k \in w_1) < \delta$ M = m

return The set of points selected: $\hat{\mathbf{Z}} = \{\hat{\mathbf{z}}_1, \hat{\mathbf{z}}_2, ..., \hat{\mathbf{z}}_M\}$

4.4 False detection filtering

What is a false detection? A false detection is an object classified as "pedestrian" when it is not. Indeed, in the complex urban environment, a lot of detections are unfortunately wrong because many objects are not totally filtered such as cars, trucks, buses, poles, trees, crash barriers, etc. So, this module checks the size and the orientation angle of the segments labelled for the Parzen algorithm as "pedestrian" in order to filter all the segments with a size lower than 30 cm or greater than 18° or lower than -18° in the laser reference frame.

4.5 Fusion of the 4 layers

Once all the positions $\hat{\mathbf{Z}} = {\{\hat{\mathbf{z}}_k\}_{k=1,...,M}}$ resulting from the 2D raw observations representing the gravity center of the pedestrians located in each laser plane are known, one must verify if the 4 laser planes confirm the same information: that is known as the fusion stage. First, the fusion of the information located in the 4 laser planes consists in projecting all the points $\hat{\mathbf{Z}}$ in the same plane. Then the fusion is carried out by a similar method to pedestrian detection (see Algorithm 1) with Σ_{ψ} , the covariance matrix associated to the laser sensor inaccuracy in the two dimensions (*x* and *y*).

$$\Sigma_{\psi} = \begin{bmatrix} \sigma_x^2 & 0\\ 0 & \sigma_y^2 \end{bmatrix}$$
(8)

The function that weighs the likelihood function $p(\hat{\mathbf{Z}}|w_k)$ depending on the sensor characteristics, on the sought objects and on the detection distance is defined by:

$$N_l = \sum_{i=1}^{4} N_l(i)$$
 (9)

with
$$N_l(i) = 1$$
 if $(0 < (hc + D \cdot \tan(\phi_i)) < H)$ (10)

These expressions take into account the pedestrian's dimensions (height: *H*), angular spacing (ϕ_i) between layers according to the distance (*D*) separating the obstacle from the vehicle and the height of the sensor (*hc*) in relation to the ground. The 2D prominent laser points $\hat{\mathbf{z}}_k \in w_1$ are selected as the pedestrian position by the maximum likelihood estimator (see equation 6).

Once the pedestrian gravity center is defined, the next step consists in searching all the points belonging to the points group of the gravity center. These points will be eliminated if their distance $d_c(\hat{\mathbf{z}}_k, \hat{\mathbf{z}}_i)$ is lower than β . Thus the point list \mathbf{L}^i below is eliminated:

$$\mathbf{L}^{i} = \{d_{c}(\hat{\mathbf{z}}_{k}, \hat{\mathbf{z}}_{i}) < \beta\}, \beta = SensorInaccuracy (11)$$

This algorithm is reiterated while the likelihood maximum estimator contains a value higher than the trust threshold $v \in [0, 1]$.

6

5 ALGORITHM PROPOSED FOR PEDES-TRIAN TRACKING

The choice of the tracking algorithm depends directly on the application. In the case of pedestrian tracking, no prior knowledge or assumption about their 2D motion which can be very uncertain is assumed. The most commonly used framework for tracking is based on a Bayesian sequential estimation.

Under such assumptions (stochastic state equation and/or non linear state and/or non Gaussian noises), particle filters are particularly well adapted.

The SIR PF and the derived Auxiliary and Regular particle filters, as proposed by Gordon et *al.* [27], are the most popular particle filters to estimate non-Gaussian density probability or a non-linear evolution model.

5.1 SIR PF

In the following section the theory of the sequential Monte Carlo methods in the framework of multiple object tracking is briefly reminded. For more details, the reader can refer to Gordon's work [27]. Let us consider a discrete dynamic system:

$$\mathbf{X}_k = f(\mathbf{X}_{k-1}) + \mathbf{W}_k \tag{12}$$

$$\mathbf{Z}_k = h(\mathbf{X}_k) + \mathbf{V}_k \tag{13}$$

where \mathbf{X}_k represents the state vector and \mathbf{Z}_k the measurement vector at instant *k*.

Particle filters provide an approximate Bayesian solution to discrete time recursive problems by updating a rough description of the posterior filtering density $p(\mathbf{x}_k | \mathbf{z}_{1:k})$. This a posteriori belief represents the state in which the objects are.

The main purpose of particle filters is to approximate the prior distribution of the recursive Bayesian filter $p(\mathbf{x}_k | \mathbf{z}_{1:k-1})$ as a set of N_s samples, using the following equation:

$$p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) = \frac{1}{N_s} \sum_{i=1}^{N_s} \delta(\mathbf{x}_k - \mathbf{x}_k^i)$$
(14)

where δ is the discrete Dirac function. Then the posterior distribution $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ can be estimated by:

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = p(\mathbf{z}_k|\mathbf{x}_k) \sum_{i=1}^{N_s} p(\mathbf{x}_k|\mathbf{x}_{k-1}^i) \qquad (15)$$

This approach can be implemented by a bootstrap filter or a SIR PF.



Fig. 3: Pedestrian tracking algorithm.

With regard to the dynamics of the pedestrians' movements (see equation 12), we suppose that no prior information on their trajectory (change in pace, direction, sudden stop, etc.) is available. In order to predict all these trajectory modifications as well as possible, an evolution model with a circular motion [29] is used. The heading angle is used as a disturbance of the predicted trajectory. The model with circular motion applied to each particle is defined below.

$$F_{k} = \begin{bmatrix} 1 & \frac{\Delta T \sin(\theta_{k})}{\theta_{k}} & 0 & -\frac{\Delta T}{\theta_{k}} \cdot (1 - \cos(\theta_{k})) \\ 0 & \cos(\theta_{k}) & 0 & -\sin(\theta_{k}) \\ 0 & \frac{\Delta T}{\theta_{k}} \cdot (1 - \cos(\theta_{k})) & 1 & -\frac{\Delta T \sin(\theta_{k})}{\theta_{k}} \\ 0 & \sin(\theta_{k}) & 0 & \cos(\theta_{k}) \end{bmatrix}$$
(16)

and $\theta_{k+1} = \theta_k + b_g$ with $b_g \sim \mathcal{N}(0, \sigma_g)$. σ_g is the standard deviation of the heading angle concerning the pedestrian's trajectory.

The state vector used summarizes all the information observed in the scene, i.e. the number of observed pedestrians and their characteristics:

$$\mathbf{X}_k = (O_k, \mathbf{x}_{1,k}, \dots, \mathbf{x}_{N,k}) \tag{17}$$

with O_k a discrete random variable representing the number of pedestrians present in the scene and $\mathbf{x}_{N,k} = (\mathbf{p}_{N,k}, \mathbf{I}_{N,k})$ the state vector associated to the object *N*. The 2D positions and speed characteristics are given by $\mathbf{p}_{N,k}$. $\mathbf{I}_{N,k}$ gives identification, age, and the number of points that a pedestrian sends back into a laser layer. According to equation 13

$$\mathbf{Z}_k = [I_{2*2}]\mathbf{x}_{k+1} + \mathbf{V}_k \tag{18}$$

where \mathbf{x}_k represents the object position. Finally, noise, \mathbf{V}_k and \mathbf{W}_k are assumed to be a Gaussian

function, of zero-mean and of respective covariances:

$$Q_k = \begin{bmatrix} \sigma_x^2 & 0\\ 0 & \sigma_y^2 \end{bmatrix}, \quad R_k = \begin{bmatrix} \sigma_{cx}^2 & 0\\ 0 & \sigma_{cy}^2 \end{bmatrix}$$
(19)

The variances of the added noises depend on the maximum movement amplitude possible for a pedestrian i.e. $\sigma_x = \sigma_y = 2$ m and maximum errors of sensor measurements are $\sigma_{cx} = 0.2$ m and $\sigma_{cy} = 0.2$ m.

5.2 Track management module

To allow modification to the number of objects, Khan et *al.* [26] introduce the RJMCMC (Reversible Jump Markov Chain Monte Carlo) methods. Indeed, as the number of visible objects may change, the state space dimension may also change following the set of RJMCMC moves defined by the user. For example, the move set can be included: {update, birth, death, merge, split,...}.

In order to change the number of tracked objects, a track management module is used. Its definition is summarized below:

- If an observation cannot be associated with the set assumption, then the track management module proposes a new assumption.
- If an assumption does not find out any observation over 500 ms, the track management module proposes to suppress the assumption. In this case, of course, an evolution model helps to guide state space exploration of the SIR PF algorithm with a prediction of the state.

The limitation of the exploration state is given by the maximum displacement speed of a pedestrian (≈ 2 m/s).

5.3 Limitation of exploration state and data association

Multiple objects tracking with a particle filter generally uses a data association step, in which each target is mapped to an object hypothesis. Conventional methods such as Nearest Neighbor Standard Filter (NNSF), Joint Probabilistic Data Association Filter (JPDAF) [28] and also the Multi Hypothesis Tracking (MHT) [25] calculate a region delimiting the space where future observation are likely to occur [28]. Such a region is called *validation gate* or *gate*. Selecting a too small gate size may lead to miss the target originated measurement, whereas selecting a too large size is computationally expensive and increases the probability of selecting false observations.

In our framework, the validation gate G_k can be approximated by an ellipsoidal region given by a Gaussian density which is given by $p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \mathcal{N}(\hat{\mathbf{x}}_k, \mathbf{P}_k)$ with:

$$\hat{\mathbf{x}}_k = \sum_{i=1}^{N_s} w_k^i \mathbf{x}_k^i \tag{20}$$

$$\mathbf{P}_{k} = \sum_{i=1}^{N_{s}} w_{k}^{i} (\mathbf{x}_{k}^{i} - \hat{\mathbf{x}}_{k}) (\mathbf{x}_{k}^{i} - \hat{\mathbf{x}}_{k})^{T}$$
(21)

where $\hat{\mathbf{x}}_k$ and \mathbf{P}_k are the first two moments of the predicted Gaussian density. In this case, the validation window is the ellipsoid of size N_z (dimension of measurement vector) defined such as:

$$G_k = \{ \mathbf{z}_k : (\mathbf{z}_k - \hat{\mathbf{z}}_k) \mathbf{S}_k^{-1} (\mathbf{z}_k - \hat{\mathbf{z}}_k)^T \le \gamma \}$$
(22)

Where $\mathbf{S}_k = \mathbf{H} \cdot \mathbf{P}_k \cdot \mathbf{H}^t + \mathbf{R}$ is the covariance of the innovation corresponding to the true measurement. The threshold γ is obtained from the Chi-square tables for N_z degrees of freedom and represents the probability that the (true) measurement will fall in the gate.

In this paper, the NNSF was chosen in order to match the different measurements with the different assumptions. The NNSF is the most popular and widely used algorithm for target tracking due to its computational simplicity and its low computation time. Because the measurements (or observations) sent by pedestrian detection algorithm (cf. § 4) are located in a few cluttered environment, the NNSF can be used with good performance [28].

6 EXPERIMENTS

This section presents the experiments which have allowed to validate the algorithm of pedestrian detection and tracking [20] [23].

6.1 Detection results

In order to evaluate the pedestrian detection algorithm, we have tried to answer at the following question:

How can we objectively evaluate the performance of pedestrian detection?

In order to define a framework for detection evaluation, it is important to understand what qualities are essential to a good laser-based detection method. To do so, it can be helpful to consider what constitutes a "golden" pedestrian detection algorithm by means of a laser sensor. One could argue that a good laser detection, in a real-life situation, should:

- detect all the moving and not moving pedestrians who are located in front of the vehicle;
- 2) get a false alarm rate equal to zero;
- detect objects in all weather conditions (rain, fog, etc.);
- 4) accurately estimate pedestrian position;
- 5) be fast (real time pedestrian detection).

So, this evaluation method focuses on the more generic tasks above. From this list of qualities, a ground truth has been created from one of our scenarios in order to evaluate our detection algorithm. This ground truth allows to know for each laser scan all the pedestrian objects located in the scene with their exact position. Few studies are proposed in the literature concerning the evaluation of pedestrian detection methods by means of a laser sensor. The ground truth proposed in our framework takes into account the sensor incapacity to detect the pedestrian occluded by other objects in order not to bias the pedestrian detection rate. In fact, occluded objects are a special case that can cause spurious errors to appear when evaluating the configuration.

So, "Ground Truth" scenario from our experiments allow us to evaluate the pedestrian detection algorithm with real data. All experiments are based on real laser scan sequences.

The "Ground Truth" scenario takes place in an urban environment. This experiment is obtained with the Renault test vehicle moving at real condition traffic. It includes several pedestrians (> 5) who appear or who disappear in the sensor area. The sensor resolution angle is 0.25° .

During experiments, a complete laser statement is memorized approximately every 140 ms. This algorithm is implemented in Matlab and C/C++ [30]. All the results are obtained using the same single set of parameters. This algorithm was tested in different situations such as an urban scene, a semi-urban scene or a car park. For each scan, the number of false detections is obtained by calculating the ratio:

rate_of_false_detections =
$$\frac{N_T - N_P}{N_T}$$
 (23)

with N_T the total number of detections and N_P the number of detected pedestrians. The rate of pedestrian detection is given by calculating the ratio:

rate_of_pedestrian_detection =
$$\frac{N_P}{N_{P-VT}}$$
 (24)

with N_{P_VT} the number of pedestrians who are effectively in the sensor area.

6.2 Advantage of the fusion of the 4-plane laser method

Table 1 show the advantage of the use of the 4 laser layers in order to significantly decrease the number of false detections. It can also be noticed that the rate of pedestrian detections is higher when using the 4 laser layers on our "Ground Truth" scenarios presented in last paragraph.

TABLE 1: Rate of false and correct detections according to the number of layers used for the scenarios presented in the article.

	One layer		4 layers	
	false	pedestrian	false	pedestrian
	detec-	detec-	detec-	detec-
	tion	tion	tion	tion
	rate	rate	rate	rate
dynamic scenario obtained with the Renault vehicle (50 s)	0.424	0.705	0.342	0.916

Fig. 4 presents some results of the ROC (Receiver Operating Characteristic) curve obtained with the detection algorithm. In fact, ROC curves are a standard way to display the performance of a set of binary classifiers for all feasible ratios of the costs associated with false positives and false negatives. The aim of ROC analysis is to display in a single graph the performance of classifiers for all possible costs of misclassification. In this paper, laser pedestrian detection is considered as a classifier parameterized by a threshold $\delta \in [0, 1]$ (see Algorithm 1), threshold which determines the proportion of false or true positives.



Fig. 4: Receiver Operating Characteristic (ROC) curve obtained from two algorithms (4 layers vs one layer).

All the results that are presented in the different figures, were obtained only from the detection method using 4 laser layers. Taking into account the results of the ROC curve, we have chosen to use a pedestrian acceptation threshold of 0.6 allowing to obtain a good compromise between false detection rate and pedestrian detection rate. To illustrate the detection obtained in an external environment, the detected pedestrians' estimated positions are projected in the video image. These experiments propose to deal with a great number of urban situations which allow to test the robustness of our method. It is interesting to notice that pedestrian detection is correct at a distance up to 20 m, which is difficult to achieve with an angle resolution of 0.25° .

6.3 Tracking evaluation

The behavior of the SIR PF is carried out on different sequences on board the vehicle (see Fig. 1) which served to validate our pedestrian detection algorithm. These sequences acquired in an urban scene (see Fig. 9), show one or more people walking alone across the scene, passing each other, meeting



Fig. 5: Result of pedestrian tracking on depth and theta positions. Measurements are always represented in gray circles.



Fig. 6: Result of pedestrian tracking on depth and theta velocity.

Several experiments are now presented demonstrating the performance of SIR PF algorithm. SIR PF algorithm was tested on real data. The presented scenario (see Fig. 10) include several pedestrians (> 5). In urban scene, the pedestrians move in all directions. The vehicle moves at a speed ranging from 0 to 50 km/h, which allows to test the robustness of this method.

7 DISCUSSION AND CONCLU-SION

In the future, vehicles are expected to become more intelligent and responsive, managing information



Fig. 7: Result of pedestrian tracking on depth and theta position error.



Fig. 8: Result of pedestrian tracking on depth and theta velocity error.

delivery in the context of driver's situation. Pedestrian protection is one method of accomplishing this goal. The study presented in this paper is related to the capability to detect pedestrians using only a laser sensor mounted on the front of a vehicle.

This work takes place in the LOVe Project which aims at improving road safety, mainly focusing on pedestrian security. The purpose is to design safe and reliable software for the observation of "vulnerables". A lot of research work has been carried out over the last years concerning pedestrian detection using a laser sensor. However, important issues still remain concerning reliability and mainly self-diagnosis algorithms. Intelligent systems still have to be developed before being integrated into mass-produced cars. Indeed, these systems must





Fig. 9: Video and laser screenshots in an urban environment. Below, the detected pedestrian in the laser image and above their respective projection in the video image.

not deliver any false information concerning the observed scene. The scientific community is highly aware to this necessity and this is one of LOVe's objectives.

In most methods, a single row laser range scanner version is used but it is unusual to take advantage of the complementarity of planes provided by a multilayer laser sensor. Furthermore, few papers [8] present complex urban environments in real traffic conditions. And, few authors propose a pedestrian detection and tracking system which allows to localize as accurately as possible all the pedestrians present in the scene, either at a standstill, or in motion. In this paper, we introduce a new scheme which meets these requirements.

This paper has presented a new algorithm to increase the safety and the possibly to avoid collisions with vulnerable road users. The goal of this work is to obtain a robust pedestrian detection algorithm allowing real time detection, location, identification and tracking. We first gave an account of our work concerning pedestrian detection using only a laser sensor that enabled us to attest the originality of the chosen approach concerning the sensor as well as

Fig. 10: Video and laser screenshots in a urban environment. Below, the detected pedestrians in the laser image and above their respective projections in the video image.

the algorithmic solution. We proposed a fusion of the 4 laser planes method based on the Parzen kernel method. This work allows to show that judicious use of 4 laser planes improves pedestrian detection and significantly decreases the number of false alarms. Moreover, a SIR PF allows to track the pedestrians; that enables to increase the robustness of the pedestrian detection algorithm and to manage the occlusion of a pedestrian by another one. At this stage of the study, we consider that Parzen's method allows after a decentralized fusion of the 4 planes an effective selection of the laser observation clusters having the geometrical characteristics of a pedestrian.

Currently, the results show that more than 90% (see Table 1) of collisions between pedestrians and vehicles could be detected if vehicles were equipped with our pedestrian collision avoidance system based on a laser sensor. However, frequent occlusions between objects, the obvious limitations of this sensor (no information about shape, contour, texture, color of objects), its sensibility to atmospheric conditions such as rain and fog, require to devise a method of laser/camera fusion to improve

a pedestrian collision avoidance system. Indeed, our pedestrian detection algorithm still returns about 30% (see Table 1) of false alarms. Therefore, the next research step will consist in developing a new method of laser/camera fusion in order to improve these results.

REFERENCES

- [1] http://love.univ-bpclermont.fr/.
- [2] T. Gandhi and M.M. Trivedi, "Pedestrian Collision Avoidance Systems: a Survey of Computer Vision based Recent Studies", in *Procs. IEEE Intelligent Transportation Systems 2006*, Sept. 2006, pp. 976-981.
- [3] C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakas, and W. von-Seelen, "Walking Pedestrian Recognition", in *IEEE Trans. on Intelligent Transportation Systems*,1(3):155-163, Sept. 2000.
- [4] M. Bertozzi, A. Broggi, M. Del Rose, and M. Felisa, "A Symmetry-based Validator and Refinement System for Pedestrian Detection in Far Infrared Images", Seattle, USA, 2007.
- [5] A. Broggi, M. Bertozzi, M. Del Rose, M. Felisa, and A. Rakotomamonjy, "A Pedestrian Detector using Histograms of Oriented Gradients and a Support Vector Machine Classificator", in *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems* 2007, pages 144-148, Seattle, WA, USA, Sept. 2007.
- [6] A. Broggi, M. Bertozzi, R.I. Fedriga, C.H. Gomez, G. Vezzoni, and M. Del Rose, "Pedestrian Detection in Far Infrared Images based on the use of Probabilistic Templates", in *Procs. IEEE Intelligent Vehicles Symposium*, pages 327-332, Istambul, Turkey, June 2007.
- [7] A. Elfes, "Using Occupancy grids for Mobile Robot Perception and Navigation", Computer, vol. 22, no. 6, pp. 46-57, June, 1989.
- [8] K.C. Fuerstenberg, D.T. Linzmeier, and K.C.J. Dietmayer, "Pedestrian Recognition and Tracking of Vehicles using a Vehicle based Multilayer Laserscanner", in *10th World Congress on Intelligent Transport Systems (ITS)*, Madrid, Spain, November 2003.
- [9] E. Prassler, J. Scholz, and P. Fiorini, "Navigating a Robotic Wheelchair in a Railway Station during Rush Hour", Int. Journal on Robotics Research (IJRR), 18(7), pp. 760-772, 1999.
- [10] B. Kluge, C. Koehler, and E. Prassler, "Fast and Robust Tracking of Multiple Moving Objects with a Laser Range Finder", in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1683-1688, Seoul, Korea, 2001.
- [11] S. Wender, and K.C.J Dietmayer, "An adaptable Object Classification Framework", in Proc. of the IEEE Intelligent Vehicles Symposium (IV), Tokyo, Japan, 2006.
- [12] M. Lindström and J.-O. Eklundh, "Detecting and Tracking Moving Objects from a Mobile Platform using a laser Range Scanner", in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots* and Systems (IROS), pp. 1364-1369, 2001.
- [13] M. Montemerlo, S. Thrun, and W. Whittaker, "Conditional Particle Filters for Simultaneous Mobile Robot Localization and People-Tracking", in *Proc. of the IEEE International Conference* on Robotics and Automation (ICRA), Washington D.C, 2002.
- [14] D. Schulz, W. Burgard, D. Fox, and A.B. Cremers, "Tracking Multiple Moving Objects with a Mobile Robot", in *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.
- [15] M. Szarvas, U. Sakai, and J. Ogata, "Real-time Pedestrian Detection using LIDAR and Convolutional Neural Network", in Proc. of the IEEE Intelligent Vehicles Symposium (IV), Tokyo, Japan, 2006.

- [16] H. Zhao and R. Shibasaki, "A Novel System for Tracking Pedestrians using Multiple Single-Row Laser Range Scanners", IEEE Trans. on Systems, Man, and Cybernetics (SMC), Part A: Systems and Humans, Vol. 35, no. 2, pp. 283-291, march 2005.
- [17] A. Fod, A. Howard, and M. Mataric, "Laser-based People Tracking", in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, Washington D.C, 2002.
- [18] O. Frank, J. Nieto, J. Guivant, and S. Scheding, "Multiple Target Tracking using Sequential Monte Carlo Methods and Statistical Data Association", in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2003.
- [19] X. Shao, H. Zhao, K. Nakamura, K. Katabira, R. Shibasaki and Y. Nakawaga, "Detection and Tracking of Multiple Pedestrians by using Laser Range Scanner", in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2007.
- [20] S. Gidel, P. Checchin, C. Blanc, T. Chateau, and L. Trassoudaine, "Pedestrian Detection Method using a Multilayer Laserscanner: Application in Urban Environment", in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2008.
- [21] V. Nguyen, A. Martinelli, N. Tomatis, and R. Siegwart, "A Comparison of Line Extraction Algorithms using 2D Laser Rangefinder for Indoor Mobile Robotics", in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Edmonton, Canada, 2005
- [22] J. Cui, H. Zha, H. Zhao, and R. Shibasaki, "Laser-based Detection and Tracking of Multiple People in Crowds", in Computer Vision and Image Understanding (CVIU), Special Issue on "Advances in Vision Algorithms and Systems Beyond the Visible Spectrum", Volume 106, Issue 2-3, pp. 300-312, May 2007.
- [23] S. Gidel, P. Checchin, T. Chateau, C. Blanc, and L. Trassoudaine, "Parzen Method for Fusion of Laserscanner Data: Application to Pedestrian Detection", in Proc. of the IEEE Intelligent Vehicles Symposium (IV), Eindhoven, The Netherlands, 2008.
- [24] R.E. Kalman, "A New Approach to Linear Filtering and Prediction Problems", Trans. ASME J.Basic Eng, flight 82, pp.34-45, March 1960.
- [25] D.B. Reid, "Multiple an Algorithm for Alignment Targets", In Proc. CDC, pages 1202-1211, December 1978.
- [26] Z. Khan, T.R. Balch, and F. Dellaert, "An MCMC based Particle Filter for Tracking Multiple Interacting Targets." In ECCV, vol. 3024, pages 279-290, 2004.
- [27] N.J. Gordon, N.J. Salmond, and A.F.M. Smith, "Novel Approach to Nonlinear/non-Gaussian Bayesian State Estimate", IEEE Proc.-F, vol. 140, no.2, pp.107-113, 1993.
- [28] Y. Bar-Shalom and T.E. Fortmann, "Alignment and Data Association", New York: Academic, 1988.
- [29] X.R. Li and V.P. Jilkov, "A Survey of Maneuvering Target Alignment: Dynamic Models", in *6in Proc. of SPIE Conference one Signal and Data Processing of Small Target*, Orlando, FL, the USA, April 2000.
- [30] J. Falcou, J. Sérot, L. Pech and J.T. Lapresté, "Metaprogramming Applied to Automatic SMP Parallelization of Linear Algebra Code", *In Proceeding of EuroPar*, Las Palmas, Gran Canaria, August 2008.
- [31] F. Bardet and T. Chateau, "MCMC Particle Filter for Real-Time Visual Tracking of Vehicles", in Proc. of the IEEE Conference Intelligent Transportation System, Beijing, China, 2008.