Confidence-Based Pedestrian Tracking in Unstructured Environments Using 3D Laser Distance Measurements

Marcel Häselich, Benedikt Jöbgen, Nicolai Wojke, Jens Hedrich, Dietrich Paulus
Active Vision Group, University of Koblenz-Landau
Universitätsstr. 1, 56070 Koblenz, Germany
{mhaeselich, bjoebgen, nwojke, jenshedrich, paulus}@uni-koblenz.de

***Abstract*— Detection and tracking of pedestrians is an essential task for autonomous outdoor robots. Modern 3D laser range finders provide a rich and detailed 360 degree picture of the environment. Unstructured environments pose a difficult scenario where a variety of objects with similar shape to a human like shrubs or small trees occur. Especially in combination with partial occlusions, sensor noise, and concussions from traversing rough terrain.**

In this paper, we address the problem of tracking multiple pedestrians in unstructured 3D point clouds by extracting and discarding additional candidates from vegetation and other structures. Our approach works exclusively on LIDAR-based features and uses split and merge steps that create additional candidates which are classified with a Support Vector Machine. Tracking is performed by using a particle filter where particles require a recurring re-detection with a high confidence value from the classifier. Thus, we enforce a rapid disposal of particles without adequate re-activation.

Our evaluation is performed in differently populated and vegetated scenarios as well as on publicly available datasets. Experiments revealed high recall values of the tracking approach, a reduced detection precision in inhomogeneous environments, and the capability of robustly tracking pedestrians in difficult scenarios without the usage of additional sensors.

I. INTRODUCTION

Autonomous robots nowadays conquer increasingly difficult outdoor environments. Safe navigation requires robust obstacle avoidance and is a key challenge within this domain. Pedestrians are challenging obstacles due to the either stationary or moving state, the highly varying appearance, and the complex movement characteristics. Compared to a vehicle, e.g., a pedestrian can move over difficult terrain surfaces and walk in arbitrary directions whereas cars are limited to the driving direction. Furthermore, in groups, pedestrians might be walking very closely and occlude each other. In this approach, we present a pedestrian detection and tracking system using a *Velodyne HDL-64E S2*. Classification of these sensor data in unstructured environments is challenging, as pedestrians can easily be confused with other objects like trees or bushes. Occasional false classifications are unavoidable, considering sensor noise, occlusions, and similar objects in an unknown and highly inhomogeneous area.

In this work, the challenging task of accurately tracking pedestrians without tracking false positives over time is approached. We describe an interaction-based approach between a detection algorithm using a Support Vector Machine (SVM) and a tracking algorithm based on particle filters. Besides the suppression of re-detections in regions where a particle filter is already tracking a pedestrian, the combination allows a rejection of tracked targets if there are insufficient re-detections by the SVM at the tracking position. Further, a direct connection of the SVM to the particle filter grants the opportunity to raise the weights of particles that are close to a re-detection of the SVM. The resulting tracks from the particle filters are stored and can be used for other applications, e.g. sidewalk or dirt track detection.

In a nutshell, our own contribution consists of a new segmentation strategy using the dp-means algorithm [12] and the sophisticated interaction between the SVM confidence and the extinction of particles which is directly coupled to SVM re-detections in our approach. Note that motion estimation of the ego vehicle is not part of this contribution.

This paper is organized as follows. In Sec. II we present a discussion of the state of the art. The detection algorithm is outlined in Sec. III with a description of the tracking algorithm in Sec. IV. Finally, in Sec. V we present the experimental results followed by a discussion in Sec. VI and a conclusion in Sec. VII.

II. RELATED WORK

Pedestrian detection is an active research topic and approaches for several sensor modalities exist. Camera-based approaches ([6] for an overview) mostly address the problem by detecting and tracking pedestrians in single images or in video streams. In addition, several Laser-based approaches emerged over the past years using the data of 2D or 3D laser range finders (LRFs).

Breitenstein et al. [2] present a tracking-by-detection approach for multiple persons using particle filters in color image series. Their implementation allows to track large numbers of moving pedestrians in 2D space without camera calibration or knowledge of the ground plane.

Modern 3D LRFs like the Velodyne HDL-64E enable detection and tracking within a 360 degree field of view around the robot and yield up to 120,000 measurement per rotation. Recently, a number of approaches used the data of LRFs to detect and track moving objects in outdoor scenarios. Within the context of the DARPA Urban Challenge, Petrovskaya and Thrun [16] present an approach to detect and track vehicle with a particle filter in 3D data. Their generated model uses multiple rectangles for precise motion estimation of detected cars.

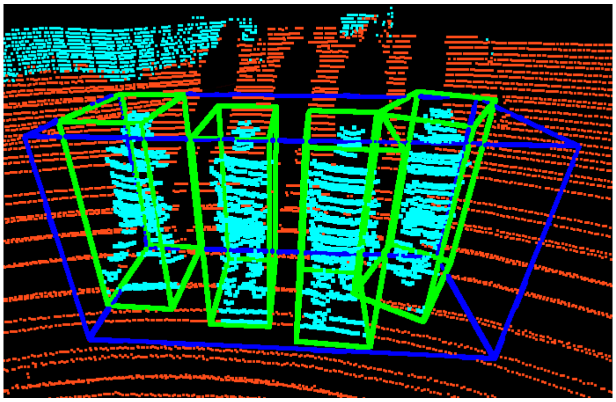


Fig. 1. Accepted pedestrians candidates (green frames) extracted from a large segmented group (blue frame). The clustering algorithm separates large objects into smaller ones.

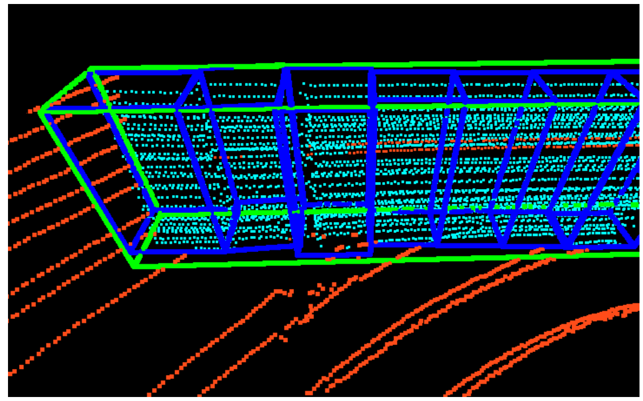


Fig. 2. Rejected pedestrian candidates (blue frames) from a wall (green frame). The clustering algorithm re-merged the smaller boxes. The large object is also rejected because it is now too large, again.

Scholer et al. [19] use a Velodyne HDL-64E LRF to detect and track people in 3D point clouds. Their approach allows to track partially and fully occluded persons over a certain amount of time in indoor areas.

Spinello et al. [21] present a combined bottom-up top-down detector for pedestrians in Velodyne HDL-64E data in urban outdoor environments. Their approach is independent from ground plane assumptions and the described detector uses a layered person model. For tracking, a multi-target multi-hypothesis approach is used and their system achieves high equal error rates within a limited range of 20 m.

Navarro-Serment et al. [15] use geometric and motion features to detect and track pedestrians while driving in outdoor regions. Their algorithm detects objects using a virtual 2D slice and then classifies each object using statistical pattern recognition techniques. Kidono et al. [11] extend features of [15] by a slice feature and by reflection intensities of their Velodyne HDL-64E. Their approach classifies pedestrians with a SVM in road environments and is able to deal with low spatial resolutions targets.

Thornton et al. [23] present a multi-sensor approach including a 3D LRF for human detection and tracking in cluttered environments.

A probabilistic person detector on multiple layers of 2D laser range scans classified using AdaBoost [8] is presented by Mozos et al. [14]. Each layer detects a different body part and the conducted experiments reveal robust detection rates in cluttered environments.

Premebida et al. [17] present a laser-based pedestrian detection system and focus on information extraction from LRFs. In their work, the authors explore and describe the potential of LRFs in pedestrian classification and present results with different classification techniques using automotive and industrial LRFs.

Carballo et al. [3] fuse multiple 2D LRFs on two layers to detect pedestrians in uncluttered indoor environments. Gidel et al. [9] present a pedestrian detection and tracking approach using an LRF system on multiple layers, too. The proposed detector fuses the information from four layers and tracking is performed with a particle filter. Sato et al. [18] use an LRF with six scanning layers to track pedestrians with a Kalman filter in urban environments. The approach maps objects with a certain height to a grid map and uses global nearest neighboring based data association.

III. PEDESTRIAN DETECTION

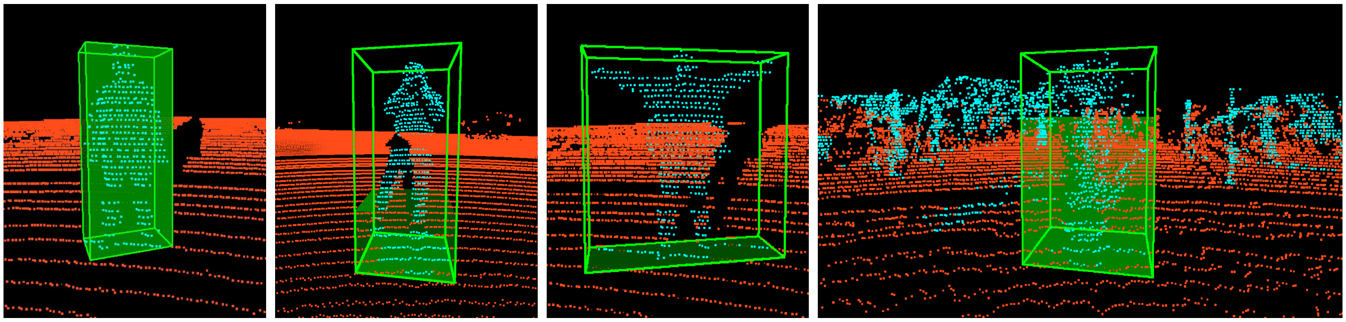
The following part describes our detection approach and is divided by descriptions of our ground removal, clustering, feature extraction, and classification steps.

A. Ground Removal

The Task of ground removal is to separate obstacle points from ground points (cf. [11], [15]) in order to reduce the computational load and to perform a first selection of candidates. In unstructured environments, an assumption of a planar ground surface with a fixed ground height is likely to fail. Planar ground regions are often interrupted by hills, ditches, and other surface variations. In our approach, we first insert the 3D data into a three-dimensional grid with a resolution of $0.1 \times 0.1 \times 0.1\,\mathrm{m}^3$ per cell. Afterwards, the cells are sorted according to their height, computed from the highest and the lowest point in each cell, from high to low. Each of those candidate cell is not only classified wrt. the contained points, but also wrt. the neighboring cells. A sliding window with a height threshold is used to calculate the absolute difference between the highest and the lowest point among all point of the current cell and in addition all points of neighboring cell. Based on the resulting values, a cell is classified as *Empty*, *Ground* or *Obstacle*.

B. Clustering

The ground removal algorithm yields groups of cells which contain obstacles. Afterwards, those groups need to be prepared for feature extraction and classification. On the one hand, the clustering algorithm needs to cluster obstacle cells to groups of sufficiently large 3D distance measurements which represent pedestrian candidates. On the other hand, the algorithm needs to split up large groups of 3D points in order to separate pedestrians from other obstacles which are close to them, including trees, building and other pedestrians.



(a) (b) (c) (d)

Fig. 3. Selected detection results with confidence displayed as green filling. A standing pedestrian with a SVM-confidence of $> 99\,\%$ is shown in (a) and a running pedestrian with a confidence of $\approx 35\,\%$ in (b). The pedestrian with a confidence of $< 1\,\%$ in (c) outstretches his arms while wearing an open jacket and is a good example for the limitations of the detector. In (d), a tree in a forest region falsely classified as pedestrian with a confidence of $\approx 65\,\%$.

Unstructured outdoor scenes are more complex than urban environments considering different elevation levels in rough terrain, diverse vegetation, and sensor noise.

Our approach consists of two steps and aims to find a high number of pedestrian candidates even if they are located close to other objects. We explicitly choose not to discard anything in the slightest similarity to a human. Thus, our next step is to further analyze clusters that are too large to represent a single candidate, which in turn yields increased runtime. In order to keep the overall runtime low, several algorithms were inspected. Approaches like k-means [13] exhibit fast runtimes but the original algorithm requires knowledge of the number of clusters k in advance and an imprecise initial configuration of the medians may drastically increase the runtime. The k-means++ extension [1] improves the initial distribution and provides faster runtimes. The problem of an unknown k is solved by the dp-means algorithm [12], which we use and that starts with $k = 1$ and increments it, if a cluster grows too large, and in addition exhibits fast runtimes. A resulting separated cluster is exemplary shown in Fig. 1 where the replaced cluster is framed in blue and the new clusters are framed in green.

The aforementioned step results in an over-clustering of large obstacles, which are now separated wrt. the extent of pedestrians. Hence, our next step is to re-merge nearby clusters without a gap between them. The distance between the centers of two clusters is divided into eight histogram bins and if the two bins in the middle contain less than 80% of the measurements compared to the average number of measurements in all bins, a gap is detected. An example of re-merged clusters is shown in Fig. 2. Here, the new smaller clusters cannot be separated adequately and the original group is maintained. This strategy has the advantage of finding pedestrian candidates close to any other obstacles at the expense of an increased runtime and the possibility of more false-positives.

C. Features

Our feature vector per cluster consists of 8 different features introduced in the literature, where f_1 and f_2 are presented by Premebida et al. [17] and describe the number of points included in a cluster and the minimum distance of the cluster to the sensor. Navarro-Serment et al. [15] apply a Principal Component Analysis (PCA) to the clusters, which represents f_3 to f_7. Those features are the 3D covariance matrix of a cluster, the normalized moment of inertia tensor, the 2D covariance matrix in different zones (cf. [15]), the normalized 2D histogram for the main plane and the normalized 2D histogram for the secondary plane. In another approach, Kidono et al. [11] introduce two additional features. The first one, the slice feature of a cluster, forms our last feature f_8 and aims to differentiate pedestrians from false positives in the shape of trees or poles. A cluster is partitioned into slices along the z-axis and for each slice the first and second largest eigenvalue is calculated. As the descriptive power of the slice feature decreases over longer distances, only a rough estimate remains in long distances. The other feature introduced by Kidono et al. considers the distribution of the reflection intensities in the cluster. Since our LRF is not calibrated wrt. the intensities, this feature could not be integrated.

D. Classification & Training

The task of the classifier is to perform a binary classification between pedestrian and non-pedestrian as precise as possible. As proposed by Kidono et al. [11], we use a SVM with radial basis function (RBF) Kernel [4], [5] together with the feature vector described in the previous subsection. For training the SVM we annotated pedestrians in different datasets from the campus Koblenz of the University of Koblenz-Landau, which contain vegetation, trees and buildings. Positive samples are extracted from annotations and negative samples are generated at random from clusters that have not been marked as pedestrians by a human annotator. For this work, we used the LIBSVM library from Chang and Lin [4] that computes a separating hyperplane which discriminates between the pedestrian and non-pedestrian. Our classifier returns a vector of confidence values in $[0, 1]$ for each class how likely a candidate belongs to the class. The probability of the highest rated class is used as input value for the tracking system (cf. Sec. IV-B) and can further be used to reject candidates with a low probability.

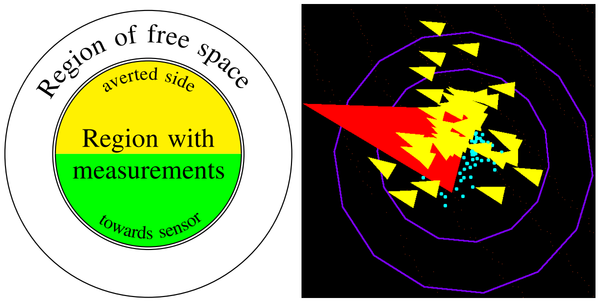


Fig. 4. Cylindrical measurement model of the particle filter. The left image outlines the different regions of the model. The right image shows the model applied to real sensor data with displayed particle hypothesis (cyan: range readings; yellow: particle; red: best particle).

IV. PEDESTRIAN TRACKING

Our initial idea for the tracking was to use virtual 2D scans and was inspired by Petrovskaya and Thrun [16]. Virtual 2D scan have successfully been used during the DARPA Urban Challenge to identify vehicles in urban scenarios. With an adopted measurement model we observed that, with a high resolution 2D scan, many pedestrians could be successfully tracked. Unlike, in urban scenarios varying inclination angles of the terrain and many occlusions, e.g. caused by vegetation, are problematic. Hence, we developed a new measurement model and decided to focus on an sophisticated interaction of the detection algorithm with the particle filter. This allows our approach to discard hypotheses early and enables it to deal efficiently with false detection that occur frequently in unstructured environments.

A. Measurement Model

Our measurement model approximates pedestrian geometry as cylindrical shape of non-zero depth (cf. Fig. 4). The likelihood of range readings is modeled according to three different regions omitting the height value. A region of free space in form of an outer cylinder is modeled around an inner cylinder with one half facing the sensor and the other half facing away. The majority of range readings are expected to fall in the region facing the sensor (green cylinder-half). A minority is expected on the other side (yellow cylinder-half) as humans represent solid objects and laser rays passing a human occur infrequently, e.g. during limb movement. The region around the pedestrian is expected to contain few to none points. Hence, the measurement models separates neighboring obstacles adequately while taking pedestrians very close to other objects into account, too.

B. Particle Filter & Particle Extinction

Tracking is performed using a Rao-Blackwellized particle filter [7] with 40 particles for each target where we estimate target extent separately for each positional hypothesis. For the measurement model we follow the derivations of [16]. Each pedestrian hypothesis consists of a 2D position, an orientation, a rotation angle wrt. the sensor, a velocity, and a circular extent. The velocity compensates for pedestrian movement by applying a model of constant velocity due to the small possible change within one rotation of the LRF. In case of a false detection, e.g., caused by a tree or a bush, a particle filter would be initiated and remain on the target until it disappears from view. Since false detections occur inevitably in the target domain, we sought a solution to handle them. Confidence-based approaches (in images [2], [22]) additionally grade system estimations to reinforce correct inferences. In our approach, a particle filter is discarded if either no re-detection occurs for a predefined time or if re-detections occur but the confidence is insufficiently low to maintain the target. The first criteria allows continuous tracking in case of occlusions for a short period of time and the second criteria ensures that no particle filter remains on invalid targets for a longer period of time. In other words, the idea is that many low-confidence detections or few high-confidence detections are both able to maintain a target, even if the tracker yields inaccurate results.

V. EXPERIMENTS

Our evaluation was performed on a laptop with an Intel(R) Core(TM) i7 QM with 1.73 GHz and 8 GB RAM. Training of the SVM was performed on datasets gathered in vegetated areas on the University campus in Koblenz and the training datasets are completely disjoint from all evaluation datasets. During the experiments, all parameters were fixed and pedestrians were annotated within a range of up to 20 m. The first two datasets were recorded in Koblenz. *Gravel terrain* represents a gravel parking lot on rough terrain with a number of puddles, some bushes and an overgrown lamp post. The *Outdoor area* dataset was recorded on a road turn next to vegetated hillside with two buildings and a tree. Two Velodyne HDL-64E datasets with pedestrian ground truth were published by Spinello et al. [20], [21]. The first dataset *Polyterrasse* contains pedestrians and bicycles in an urban area. *Tannenstrasse*, the second dataset, was recorded in downtown area and contains additional traffic participants such as trams and cars. Tables I and II show the detection respectively the tracking performance on the datasets. The true positives (TP) denote the number of correctly detected/tracked pedestrians, the false positives (FP) denote the number of incorrectly detected/tracked objects (false alarms), and the false negatives (FN) denote the missed targets. Precision is calculated as $\frac{TP}{TP+FP}$ and Recall as $\frac{TP}{TP+FN}$. Runtimes of the algorithm are summarized in Table III. Results are analyzed and discussed in the next section.

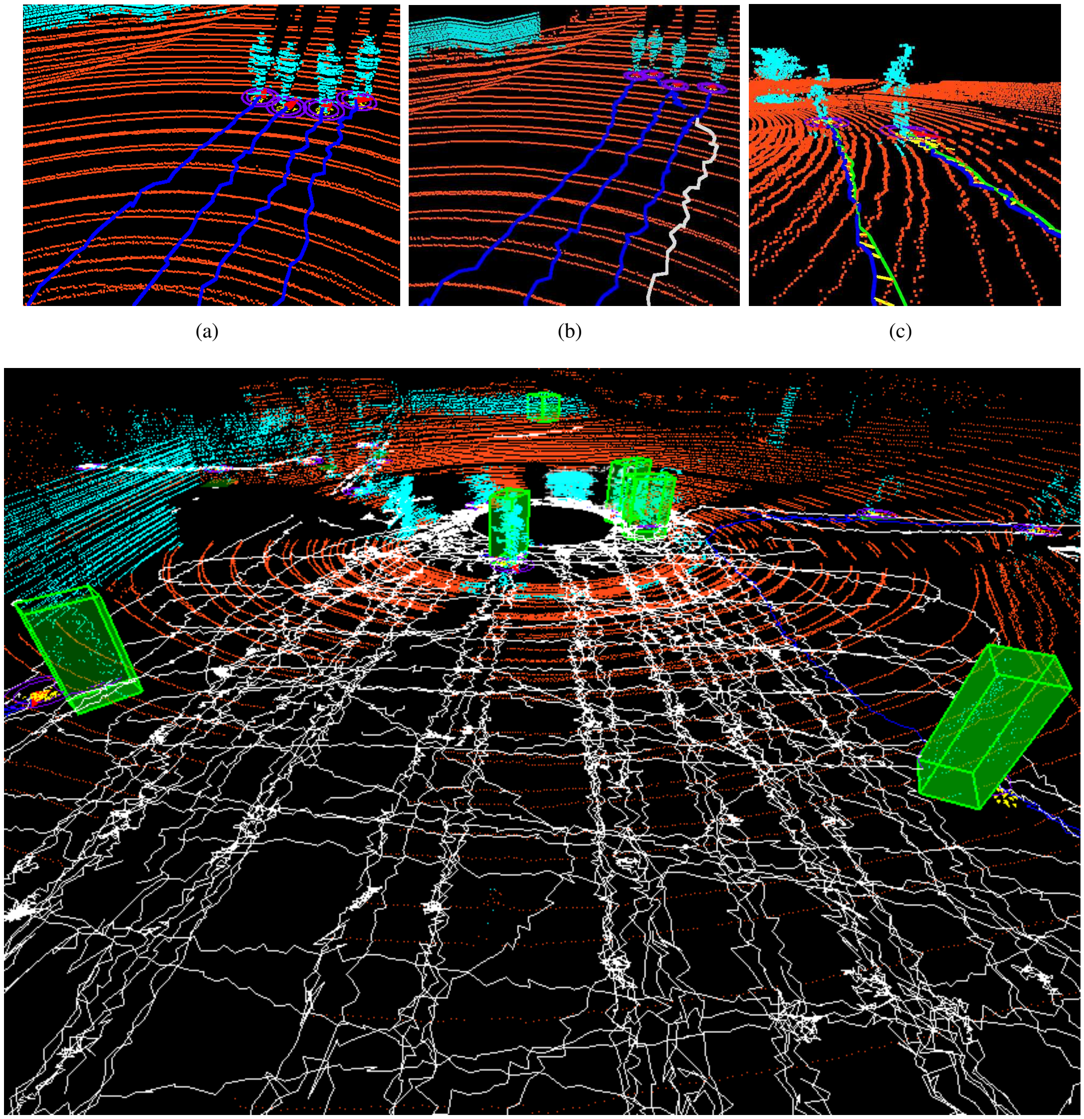
TABLE I

DETECTION RESULTS IN DIFFERENT ENVIRONMENTS.

Dataset	TP [No.]	FP [No.]	FN [No.]	Precision [%]	Recall [%]
Gravel terrain (12 Hz.)	2358	37	221	98.46	91.43
Outdoor area (12 Hz.)	3149	1434	1633	68.71	65.85
Polyterrasse (5 Hz.)	1849	846	886	68.61	67.61
Tannenstrasse (5 Hz.)	1348	2501	1116	35.02	54.71

VI. DISCUSSIONS

The presented approach yields interesting and likewise unexpected results. Firstly and as expected, the algorithm



(a) (b) (c)

(d)

Fig. 5. Tracking results in selected situations. Image (a) shows four pedestrians walking in the same direction and blue lines on the ground visualize active tracks. In (b), one of the pedestrians changed his direction and crossed path with the others. Since he became entirely occluded once, he had to be re-detected and the previous track is stored and displayed in white. The third image (c) shows a comparison of tracked persons (blue tracks) with given ground truth (green tracks). Correspondences with tracking result and ground truth are shown as yellow lines. The last image (d) depicts tracking results acquired over 10 min with a lot of stored tracks from previously tracked pedestrians in white. These tracks can be useful for other algorithms, e.g. terrain classification to identify sidewalks or, more general, inference on regions that were recently favored by other traffic participants.

performs well on the Gravel terrain dataset which is less crowded than the other datasets. On the Outdoor area dataset, performance is affected by the additional candidates which reduce the precision. The decrease of the Velodyne HDL-64E frequency from 12 to 5 Hz has an unexpected but substantial influence on the SVM classification as the point density of a pedestrian drastically changes. Once detected, the particle filter tracks the pedestrians with 5 Hz, but the missing detections affect the tracking especially on the Tannenstrasse dataset. The recall rates in crowded datasets are expectably low due to the additional candidates required to separate pedestrians standing close to vegetation. In order to increase the recall values, we would have to exclude difficult candidates (which is antithetical to the whole idea of this approach) or exchange the classifier itself for another method. Considering the runtime results, the algorithms show an overall good performance. The bottleneck of our approach is the segmentation step, which is affected from the additional candidates created by our segmentation.

VII. CONCLUSIONS

We presented an approach for pedestrian tracking in unstructured environments that aims to identify pedestrians close to vegetation by using split and merge strategies on 3D clusters. While the algorithms separate pedestrians from other structures, a great number of newly created candidates affect precision and runtime. The proposed tracking approach performs well in all test environments, especially when

TABLE II

TRACKING RESULTS IN DIFFERENT ENVIRONMENTS. NOTICEABLE ARE IN PARTICULAR THE HIGH RECALL VALUES IN COMPARISON WITH THE AVERAGE DETECTION RECALL VALUES EXHIBITED IN TABLE I.

Dataset	TP [No.]	FP [No.]	FN [No.]	Precision [%]	Recall [%]
Gravel terrain (12 Hz.)	2545	165	34	93.91	98.68
Outdoor area (12 Hz.)	4556	6448	226	41.40	95.27
Polyterrasse (5 Hz.)	2476	1454	259	63.00	90.53
Tannenstrasse (5 Hz.)	1802	7218	662	19.98	73.13

TABLE III

SYSTEM RUNTIMES IN DIFFERENT ENVIRONMENTS

Dataset		Mean [ms]	Std. dev. [ms]	Min. runtime [ms]	Max. runtime [ms]
Gravel terrain 0 min 45 s 14 Hz	Segmentation	63.23	5.21	53.25	144.78
	Detection	9.76	4.74	0.00	25.43
	Tracking	1.48	0.96	0.00	4.31
	Overall	74.88	7.54	55.63	159.96
Outdoor area 12 min 46 s 14 Hz	Segmentation	141.03	57.08	96.89	349.27
	Detection	31.70	11.16	3.97	73.47
	Tracking	32.89	26.49	0.00	159.34
	Overall	206.56	64.66	115.84	582.29
Polyterrasse 2 min 59 s 5 Hz	Segmentation	63.86	5.05	48.70	87.85
	Detection	16.25	8.32	12.90	43.89
	Tracking	6.60	4.69	0.00	43.26
	Overall	87.53	12.32	57.54	137.17
Tannenstrasse 1 min 39 sec 5 Hz	Segmentation	551.63	199.05	76.65	2208.51
	Detection	44.47	14.15	7.68	100.10
	Tracking	96.53	74.86	0.00	591.14
	Overall	693.31	214.32	108.20	2390.22

regarding the low recall values of the SVM in cluttered environments. The SVM is unable to deal with a divergent update frequency of the Velodyne HDL-64E, thus lowering precision and recall on the publicly available datasets and creating a further challenge for the tracking system which proved reliable under these complicated circumstances.

Considering future work, an extension for crowded environments could complement the approach and reduce the maximum runtime when many candidates are present. It is also possible to exchange the SVM classification for a bank of classifiers on different height levels of the pedestrians as proposed in [20] in order to reduce the number of false positives. Further, a combination with a terrain classification algorithm [10] could yield more precise predictions of pedestrian movement directions according to surface conditions.

REFERENCES

[1] D. Arthur and S. Vassilvitskii. K-means++: The Advantages of Careful Seeding. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1027–1035, 2007.

[2] M. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. Gool. Robust Tracking-by-Detection using a Detector Confidence Particle Filter. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1515–1522, 2009.

[3] A. Carballo, A. Ohya, and S. Yuta. Fusion of Double Layered Multiple Laser Range Finders for People Detection from a Mobile Robot. In *Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 677–682, 2008.

[4] C.-C. Chang and C.-J. Lin. LIBSVM: A library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):27:1–27:27, 2011.

[5] C. Cortes and V. Vapnick. Support-Vector Networks. *Machine Learning*, 20(3):273–297, 1995.

[6] P. Dollar, C. Wojek, B. Schiele, and P. Perona. Pedestrian Detection: An Evaluation of the State of the Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(4):743–761, 2012.

[7] A. Doucet, N. de Freitas, K. Murphy, and S. Russell. Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks. In *Proceedings of the Sixteenth Conference Annual Conference on Uncertainty in Artificial Intelligence*, pages 176–183, 2000.

[8] Y. Freund and R. E. Schapire. A Decision-Theoretic Generalization of Online Learning and an Application to Boosting. In *Proceedings of the European Conference on Computational Learning Theory*, pages 23–37, 1995.

[9] S. Gidel, P. Checchin, C. Blanc, T. Chateau, and L. Trassoudaine. Pedestrian Detection and Tracking in an Urban Environment Using a Multilayer Laser Scanner. *IEEE Transactions on Intelligent Transportation Systems*, 11(3):579–588, 2010.

[10] M. Häselich, M. Arends, N. Wojke, F. Neuhaus, and D. Paulus. Probabilistic Terrain Classification in Unstructured Environments. *Journal of Robotics and Autonomous Systems*, 61(10):1051–1059, 2012.

[11] K. Kidono, T. Miyasaka, A. Watanabe, T. Naito, and J. Miura. Pedestrian recognition using high-definition LIDAR. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 405–410, 2011.

[12] B. Kulis and M. Jordan. Revisiting k-means: New Algorithms via Bayesian Nonparametrics. In *Proceedings of the International Conference on Machine Learning*, pages 513–520, 2012.

[13] J. MacQueen. Some Methods for Classification and Analysis of MultiVariate Observations. In *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297, 1967.

[14] O. Mozos, R. Kurazume, and T. Hasegawa. Multi-Part People Detection Using 2D Range Data. *International Journal of Social Robotics*, 2(1):31–40, 2010.

[15] L. Navarro-Serment, C. Mertz, and M. Hebert. Pedestrian Detection and Tracking Using Three-dimensional LADAR Data. *International Journal of Robotics Research, Special Issue: Seventh International Conference on Field and Service Robots*, 29(12):1516–1528, 2010.

[16] A. Petrovskaya and S. Thrun. Model Based Vehicle Detection and Tracking for Autonomous Urban Driving. *Autonomous Robots, Special Issue: Selected papers from Robotics: Science and Systems*, 26(2–3):123–139, 2009.

[17] C. Premebida, O. Ludwig, and U. Nunes. Exploiting LIDAR-based Features on Pedestrian Detection in Urban Scenarios. In *Proceedings of the International IEEE Conference on Intelligent Transportation Systems*, pages 405–410, 2009.

[18] S. Sato, M. Hashimoto, M. Takita, K. Takagi, and T. Ogawa. Multilayer lidar-based pedestrian tracking in urban environments . In *Intelligent Vehicles Symposium*, pages 849–854, 2010.

[19] F. Scholer, J. Behley, V. Steinhage, D. Schulz, and A.B. Cremers. Person Tracking in Three-Dimensional Laser Range Data with Explicit Occlusion Adaption. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1297–1303, 2011.

[20] L. Spinello, K. Arras, R. Triebel, and R. Siegwart. A Layered Approach to People Detection in 3D Range Data. In *Proceedings of the AAAI Conference on Artificial Intelligence: Physically Grounded AI Track*, 2010.

[21] L. Spinello, M. Luber, and K. Arras. Tracking People in 3D Using a Bottom-Up Top-Down Detector. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1304–1310, 2011.

[22] S. Stalder, H. Grabner, and L. Van Gool. Cascaded Confidence Filtering for Improved Tracking-by-Detection. In *Proceedings of the European Conference on Computer Vision*, pages 369–382, 2010.

[23] S. Thornton, M. Hoffelder, and D. Morris. Multi-sensor Detection and Tracking of Humans for Safe Operations with Unmanned Ground Vehicles. In *IEEE Workshop on Human Detection from Mobile Platforms*, pages 103–112, 2008.