Autonomous Person Following with 3D LIDAR in Outdoor Environments

Karsten Bohlmann, Andreas Beck-Greinwald, Sebastian Buck, Henrik Marks, Andreas Zell

Abstract:

The capability of a robot to follow autonomously a person highly enhances its usability when humans and robots collaborate. In this paper we present a system for autonomous following of a walking person in outdoor environments while avoiding static and dynamic obstacles. The principal sensor is a 3D LIDAR with a resolution of 59x29 points. We present a combination of 3D features, motion detection and tracking with a sampling Bayesian filter which results in reliable person detection for a low-resolution 3D-LIDAR. The method is implemented on an outdoor robot with car-like steering, which incorporates the target's path into its own path planning around local obstacles. Experiments in outdoor areas validate the approach.

Keywords: 3D perception, person detection, person following, car-like steering

1. Introduction

In the future we will likely see more and more robotic assistants in outdoor work environments, e.g. in agriculture, construction or forestry. In such fields humans and robots will closely work together, and robots, which won't be fully autonomous in the near future, will still require human guidance. A basic, but very helpful, capability for such a robotic helper is the ability to follow on command a human worker to a goal. Such a task is challenging for the robot, as it has to detect a moving target, predict its movements and follow with constant distance while avoiding previously unknown static and dynamic obstacles.

In the presented system we use a 3D laser ranging sensor, which has compared to vision-based systems the advantage that reliable obstacle detection and person tracking can be executed simultaneously with a single sensor. Our particular sensor is based on a resonating MEMS mirror and is certified for use in full sunlight (<=100000Lux), a major feature for deployment in outdoor areas, in opposite to indoor systems like the Microsoft Kinect. Compared to other outdoor-capable 3D-scanning systems, e.g. manufactured by Velodyne with a pricing in the high five-figure range, the sensor is still within a lower price range. The challenge, and the main contribution of this work, is to detect a person reliably in low-resolution sets of 3D point data from a moving platform. For the person detection we focus on using legs and feet, as it can be safely assumed that in outdoor working environments human workers do not wear floor-length coats or skirts. Legs as main features offer the advantage, that they are also detectable for ranging sensors with a limited field of view in vertical direction, as all the while the same sensor is needed to detect obstacles on the ground in the driving path of the robot.

1.1. Related work on person tracking and robot following

The detection of persons in range data has a long history in robotics literature. Among others the approaches differ in the choice of the principal sensor, the feature extraction techniques, the methods for tracking and predicting moving targets and the targeted environment. [8] presents an approach using an occupancy grid to detect pedestrians with a 6-layer LIDAR. Stacked 2D features with a 3D LIDAR are employed in [9]. The method works reliably in near ranges, using an Velodyne 64E S2 with about 20 times the resolution of the sensor used in this work. However, recognition rates drop to 63% when persons are represented by less than 200 points in the scan image. The same sensor is employed in [10], where a variant of GentleBoost together with temporal and static descriptors is used for recognition of pedestrians and three other object classes. A similar classifier is used together with contour features in [6]. Other types of 3D-ranging sensors have also been used for person detection, e.g. [2] uses a time-of-flight-sensor with a Kalman filter for tracking and in [11] a system based on a RGB-D sensor is presented with a Reversible Jump Markov Chain particle filter for fusing a number of different detection algorithms.

To our knowledge most of the publications about people-following robots focus on indoor environments. In [4] an indoor robot follows the exact path which the target person took. This behavior is compared with a direct-to-goal behavior with respect to human acceptance. A hybrid strategy is suggested but not elaborated. A contribution of this work is a heuristic implementation of this behavior. A few works deal with robots and ranging sensors in outdoor areas [1, 5]. [3] shows a vision-only based approach for an outdoor robot in unstructured environments.

2. Person Detection and Tracking

The flowchart in Fig. 1 shows the main steps of the presented person detection approach. After preprocessing the LIDAR data and estimation of the ground plane candidate hypotheses are generated by first extracting suitable segments from each horizontal scan line of the range data and then grouping segments to segment blocks. Segment blocks are classified and then tracked using a particle filter.

2.1. Preprocessing of range data and ground plane detection

The current ground plane is estimated from range using the RANSAC algorithm. The plane detection is speeded up by taking the first two points from fixed spots

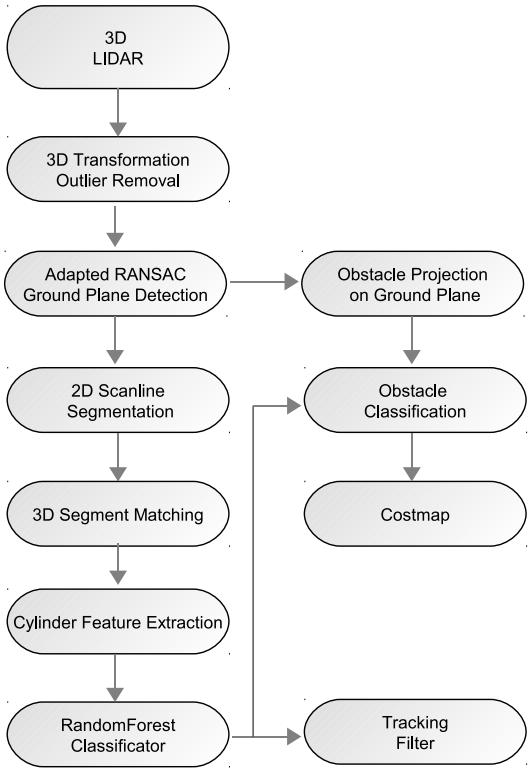


Figure 1: Structure of Person Detection

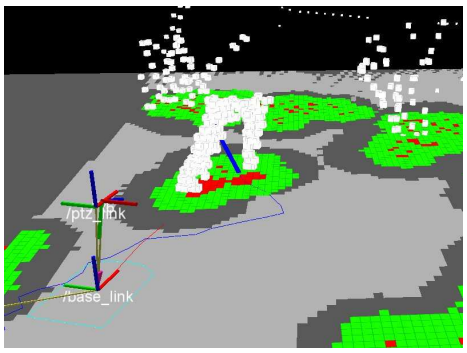


Figure 2: Local map with inflated obstacles and LIDAR scan points above (white boxes). Points classified as ground filtered out.Screenshot taken from rviz

in front of the robot and only sampling the third point. Using the detected plane all points belonging to it are removed from further computations. To obtain a local map of the surroundings all other 3D points are projected on this plane. Points with a height above the plane greater than the height of the robot or person are excluded from obstacle detection.

2.2. Segment-block-based feature detection

Legs, or actually leg parts, are modelled as cylinders. The extraction of cylinders from the point cloud is performed in three stages: In the first stage each 2D scanline is separated into sets of segments using a jump detection algorithm. Segments with lengths outside an interval are purged. For each remaining segment a circle is then fitted to the points. Here, it has to be taken into account that, due to noisy range measurements, low resolution of the scanner, disadvantageous viewing angle and loose-fitting clothes, the circle-estimation might yield very imprecise data. In the next stage horizontal segments are combined to segment blocks. Two segments are grouped if they overlap each other in vertical direction with at least 50%. For each segment block the following features values are calculated:

- Average width w_S of segments as euclidean distance between first and last segment point
- Standard deviation of width σ_{wS}
- Total number of segments n_S combined in segment block
- Height h_S of segment block above estimated ground plane
- Number c_s of segments within block, which can be fitted a circle with parameters reasonable for a leg
- Average diameter d_s of circular segments
- Standard deviation of circle diameters σ_{dS}

Segment blocks are assigned a score w_b using a random forest classifier which assigns for each feature a value describing how much this value is adequate for a human leg.

2.3. Classification and tracking of multiple hypotheses

The tracking of person pose hypotheses is implemented using a sequential importance resampling (SIR) filter. The probability density of the target person's position is approximated with a set of m particles. Each particle represents a hypothesis s^i for position $\mathbf{x}$, velocity $\mathbf{v}$ person with an associated importance weight w.

The ego-motion of the target-following robot is compensated using odometry data therefore at the beginning of each sensor cycle all samples are shifted using the robots displacement and rotation and their estimated own velocity:

$$\mathbf{x}_k^i = {}^kO_{k-1}(\mathbf{x}_{k-1}^i + T \cdot v_{k-1}) \quad (1)$$

with ${}^kO_{k-1}$ the transformation from previous to new robot frame and T the time between two sensor updates. In the correction step the current sensor measurement is used to calculate weights for each sample. In this application, sensor measurements are affected with non-gaussian noise. The targeted person can be occluded and should not be mismatched with another person. We propose a set of three techniques to address this problem.

Using the best current target estimation to weight observed segment blocks A pedestrian walks typically with a speed of less than 10km/h. This implies that the target person moves only a small distance between two sensor readings, and position candidates at greater distance to the previous position are less probable than closer estimations. To account for this we calculate a weight for each segment block b^j using the distance to the last best target position estimate $\mathbf{x}^*$and Gaussians distribution:

$$w_d^j = (2\pi T)^{-\frac{1}{2}} \exp(-\frac{1}{2T}(\mathbf{b}^j - \mathbf{x}^*)^\intercal(\mathbf{b}^j - \mathbf{x}^*)) \quad (2)$$

The factor T in the equation takes account of the fact, that the uncertainty about the targets's position grows with increasing time between sensor updates. The target position is extracted using the robust mean where the position

is calculated as the weighted mean of the samples in a window sized ϵ around the best particle $s^{\max}$, the one with the maximum weight:

$$\mathbf{x}^* = \sum_{i=1}^{m} \mathbf{x}^i w^i : |\mathbf{x}^i - \mathbf{x}^{\max}| \leq \epsilon \quad (3)$$

Assuming noise-free sensor readings, this method would calculate the goal position as the center between the observed pair of legs.

Identifying the best fitting observation for each sample In a noisy natural environment there are often numerous observations, which are possible target candidates. However, a target hypothesis s^i should not get a stronger importance weight if there are multiple observations nearby, but should be associated with a single observation. Therefore for each sample we calculate the distance to each segment block weighted with the quality of the segment block w_b^j

$$w_s^{i,j} = (2\pi)^{-\frac{1}{2}} \exp(-(\mathbf{b}^j - \mathbf{x}^i)^\intercal (\mathbf{b}^j - \mathbf{x}^i)) \cdot w_b^j \quad (4)$$

and select the block with the maximum resulting weight as the associated observation. Then, the predicted weight $\overline{w}^i$ is

$$\overline{w}^i = \max_j (w_s^{i,j}) \cdot w_d^{jmax} \quad (5)$$

Accounting for occlusions and sensor noise A disadvantage of approximating probability distributions with a sampling importance filter is particle degeneration, i.e. there are no particles left at the actual position of the goal. A typical cause in our case are temporary occlusions of the target or sensor noise. To account for this we calculate the final weight for each sample in the current time step k using a decay factor ϵ_n:

$$w_k^i = (\epsilon_n + (1 - \epsilon_n)\overline{w}^i) \cdot w_{k-1}^i \quad (6)$$

The value ϵ_n is increased if there is an segment block between the position of the particle and the sensor.

3. Motion planning and obstacle avoidance for person following

A walking person is most of the time capable to leave behind a wheeled non-holonomic robot in uneven outdoor terrain. Therefore it is assumed that the person moves cooperatively by walking on paths where the robot is able to follow. We also assume that the environment is not known in advance and features dynamic obstacles, like for example other persons. Thus, the robot can only plan its movement within the range of the LIDAR and must implement a reactive behavior to avoid dynamic obstacles.

After extracting the ground plane as described in 2.1 a local 2D occupancy grid map is generated from LIDAR data by projecting all obstacle points on the ground plane. The map module is set up as a rolling windows which always keeps a list of detected obstacles with the robot in its center and thus serves as memory of the robot for passed obstacles. The area in the sensor's field of view is constantly updated while objects behind the robot fall out of the map with the robot's movement. The pose transformation of obstacles is executed using odometry data. For path planning all obstacles are inflated in two levels: First by the radius of the in-circle of the robots footprint and second with a safety margin (Green and dark grey areas in Fig. 2, obstacles in red).

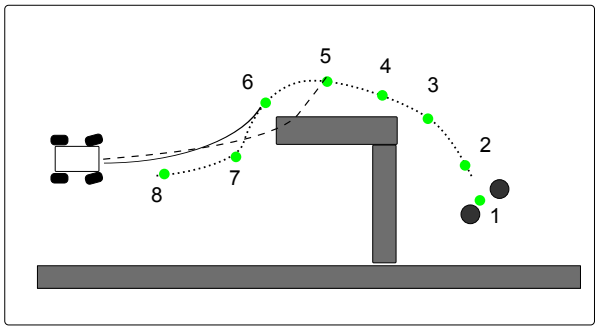


Figure 3: Motion planning using recorded path of target person and locally detected obstacles. The robot drives towards the most distant recorded waypoint which is directly approachable.

There are different strategies for a robot to follow a person. A simple but efficient method is the greedy strategy where the robot attempts to drive the shortest path towards a target point behind or next to the person. An alternative approach is to constantly record the path the target person is walking, and to follow this path as closely as possible. This incorporates the idea that the human has walked the optimal path around encountered obstacles and lets the robot benefit from human intelligence. However, this approach may lead to unnecessary motions of the robot and erratic behavior in case the target is momentarily lost. In [7] a hybrid approach as the combination of these strategies is recommended, with a heuristic switching between both behaviors.

The basic idea for the approach implemented in this work is outlined in Fig. 3. The position of the target is constantly recorded, so the robot always has a list of waypoints and thus the walked path of the person to follow. In the given example the position of the target person is not directly reachable. In this case, the history of waypoints is searched backwards, starting with the persons position, to find a suitable point to drive to. As the robot has Ackermann-type car-like steering, it drives along circle segments with a minimal possible turning radius determined by the wheelbase and the maximum possible steering angle. The cost of a path to a waypoint is determined by summing up costs of the cells in the cost map.

For local obstacle avoidance a potential field approach is used. The speed of the robot is controlled with a PID controller.

4. Employed hardware

4.1. Robot platform

We conducted experiments with an outdoor-robot developed and built at our department (Fig 4). The plat-

Figure 4: Robot with mounted 3D laser ranging sensor

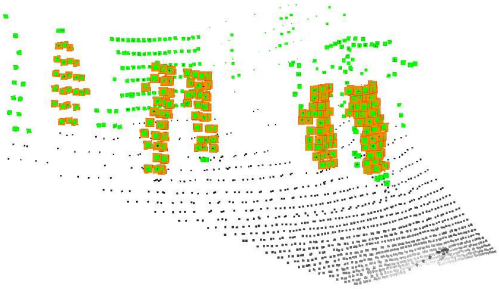


Figure 5: Scene with two persons in the field of view recorded with the FX-6-LIDAR. Ground is displayed as small black dots, detected segment blocks as orange squares and unclassified obstacles as green dots.

form is based on a RC-model truck in 1:8 scale with Ackermann-type steering which was outfitted with a 32-bit microcontroller for real-time control and an embedded dual-core PC for high-level tasks. The odometry system is implemented using an encoder inside the wheels for counting wheel revolutions and hall sensors for measuring the steering angle. To improve the quality of the positional estimation the odometry values are fusioned with an IMU containing a magnetic compass using a Kalman filter. With single-axis steering the robots have a minimal turning radius of 0.87m. The robots are capable of driving in rough terrain with slopes up to 35°. The robot was equipped with an U-Blox-6-GPS for logging and analyzing the test drives.

4.2. 3D LIDAR Sensor

The employed 3D laser ranging sensor is an FX-6 shown manufactured by Nippon Signal Co. It is based on an MEMS chip with a resonating micro-mirror. It scans a pyramid-shaped area with a vertical angle of 50° and a horizontal angle of 60°. The sensor operates with a frame rate of 16Hz and provides a resolution of 59x29 dots. The usable range is 5m with a resolution of 1cm and an accuracy of 5%. Fig. 5 displays the typical output of the sensor.

5. Experimental Results

We tested our person-following robot system with a number of test drives in the area surrounding the faculty building shown in aerial view in Fig. 6. The environment is typical for semi-urban areas, with paved roads, meadows, but also gravel walks and wood-like areas with slopes up

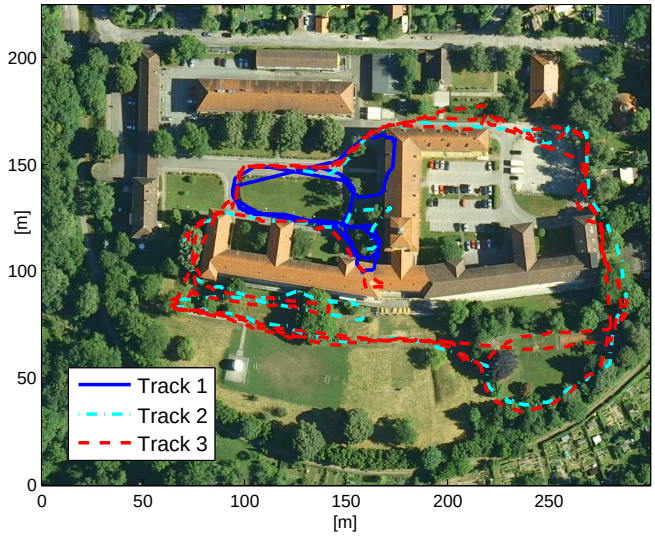


Figure 6: Trajectories of the robot following a person, red path three times around the faculty building with a total length of 2083.4m

to 20%. Obstacles encountered by the robot include trees, bushes, and lamp posts. At the final stage of our experiments we conducted two test drives within different environments (Tracks 1-2 in Fig. 6) and as final experiment a long drive with a length of 2083.3m three times around the building (Track 3). In all three test drives together the robot followed a person for a total distance of 3.5km.

Track 1 took place on mostly even asphalt ground, with the exception of a few curbstones which needed to be crossed. In test drive 2 the robot was led downwards in a park-like area with grassy underground. The track continued on a gravel walk near a few lamp posts which resemble in diameter and height human legs. Near the first long curve the track continued on an ascending slope with a maximal slope of about 20% towards the building. After passing some trees and bushes the track ended on the paved road. During track 3 the robot followed the person continuously for about 30minutes and lost its target three times. In two cases the robot continued autonomously after the person returned to the robot, only in one case, at a sharp u-turn, the robot stopped in front of a wall and manual steering with remote control was necessary to continue the track. However, during track 2 the same spot was passed by the robot without any problems. Table 1 summarizes some key figures of the drives. As local map building and motion planning rely heavily on the quality of the odometry system, the long-term accuracy of the odometry was evaluated as well. Fig. 7 displays the trajectory of the robot on track 2 calculated from odometry data fusioned with the 3D-magnetic compass, which shows, compared to the cyan-colored GPS-trajectory in Fig. 6, the general stability of the odometry in outdoor areas. The odometry distance error is, as Table 1 shows, about 3.3% on asphalt, and approximately 9.5% for tracks on other ground types. The system also showed itself to be robust to other persons crossing the path between the target person and the robot (Fig. 9).

No.	Total track length (GPS) s[m]	Total track length (Odometry) s[m]	Avg. target distance d_{avg}[m]	Max. target distance $d_{\max}$[m]	Avg. target speed v_{avg}[m/s]	Max. target speed $v_{\max}$[m/s]	Avg. robot speed r_{avg}[m/s]	Max robot speed $r_{\max}$[m/s]	Max. slope [%]	Ground type
1	552.73	534.64	3.1	5.5	1.05	1.58	1.04	1.69	2	mostly asphalt
2	881.74	798.86	2.7	5.8	0.95	1.77	0.92	1.89	21.9	grass, asphalt, gravel
3	2083.26	1883.35	2.5	4.97	1.07	1.73	1.08	2.21	21.9	grass, asphalt, gravel

Table 1: Results of three person-following drives

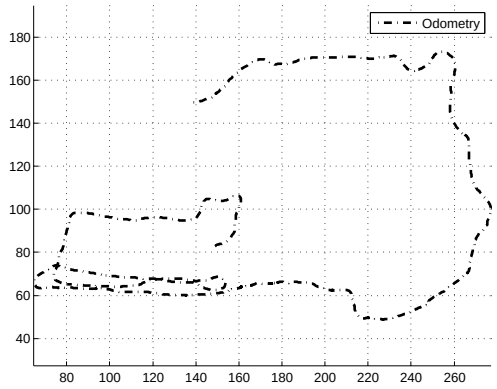
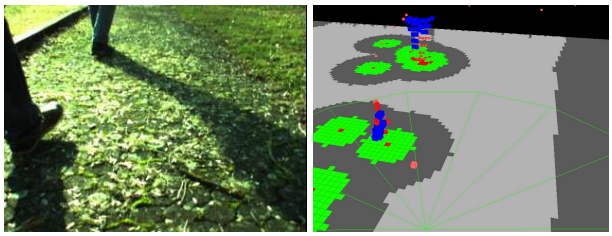


Figure 7: Compass-stabilized odometry trajectory of track 2.

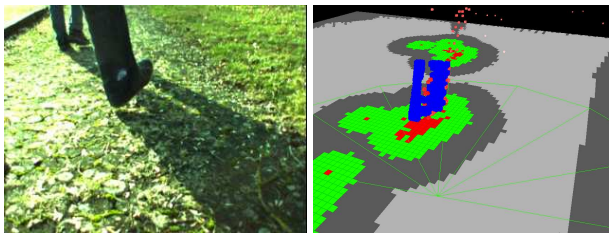
Figure 8: A part of track 2 and 3 (gravel walk) with ascending slope along a curved path.

6. Conclusion

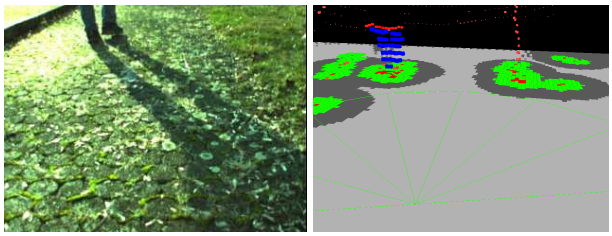
In this paper, we presented a robust robot system with a 3D LIDAR, which is able to track and follow a walking person in outdoor environments. Our experiments showed the method to work stable especially at ascending or descending slopes and when passing obstacles like trees and lamp posts. We have addressed the problem of using a low-resolution 3D laser scanner for the simultaneous tasks of obstacle detection and recognition of persons.



(a)



(b)



(c)

Figure 9: a) Another person starts crossing the path between the robot and the target person. b) The target is occluded. c) The crossing person is nearly out of view, the robot is still following the original target. Images were taken with a camera mounted on the LIDAR, where optical axes are aligned, but FOV is not identical.

AUTHORS

Karsten Bohlmann* – Computer Science Department, University of Tübingen,, Tübingen, Germany, e-mail: karsten.bohlmann@uni-tuebingen.de

Andreas Beck-Greinwald – Computer Science Department, University of Tübingen, Tübingen, Germany, Tübingen, Germany, e-mail: andreas.beck-greinwald@gmx.de

Sebastian Buck – Computer Science Department, University of Tübingen, Tübingen, Germany, Tübingen, Germany, e-mail: sebastian.buck@student.uni-tuebingen.de

Henrik Marks – Computer Science Department, University of Tübingen, Tübingen, Germany, Tübingen, Germany, e-mail: henrik.marks@student.uni-tuebingen.de
Andreas Zell – Computer Science Department, University of Tübingen, Tübingen, Germany, Tübingen, Germany, e-mail: andreas.zell@uni-tuebingen.de
* Corrresponding author

References

[1] Tirthankar Bandyopadhyay, Nan Rong, Marcelo Ang, David Hsu, and Wee Sun Lee. Motion planning for people tracking in uncertain and dynamic environments. In Proceedings of the IEEE ICRA 2009 Workshop on People Detection and Tracking Kobe, Japan, number May, pages 46–53, 2009.

[2] A. Bevilacqua, L. Di Stefano, and P. Azzari. People tracking using a time-of-flight depth sensor. In Video and Signal Based Surveillance, 2006. AVSS '06. IEEE International Conference on, page 89, nov. 2006.

[3] Jonathan Brookshire. Person following using histograms of oriented gradients. International Journal of Social Robotics, 2(2):137–146, 2010.

[4] Rachel Gockley. Natural person-following behavior for social robots. In in Proceeding of the ACM/IEEE international Conference on Human-Robot interaction HRI '07, 2007, pages 17–24. ACM Press, 2007.

[5] M. Kobilarov, G. Sukhatme, J. Hyams, and P. Batavia. People tracking and following with mobile robot using an omnidirectional camera and a laser. In Proceedings of the 2006 IEEE International Conference on Robotics and Automation, Orlando, Florida - May 2006, pages 557 –562, May 2006.

[6] Andreas Nüchter and Joachim Hertzberg. Towards semantic maps for mobile robots. Robot. Auton. Syst., 56(11):915–926, November 2008.

[7] R. Simmons R. Gockley, J. Forlizzi. Natural person-following behavior for social robots. In HRI 07: Proceedings of the ACM/IEEE international conference on Human-robot interaction, 2007.

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

[9] Luciano Spinello, Kai O. Arras, Rudolph Triebel, and Roland Siegwart. A layered approach to people detection in 3d range data. In Proc. 24th AAAI Conference on Artificial Intelligence, PGAI Track (AAAI'10), Atlanta, USA, 2010.

[10] Alex Teichman, Jesse Levinson, and Sebastian Thrun. Towards 3d object recognition via classification of arbitrary object tracks. In International Conference on Robotics and Automation, 2011.

[11] S. Savarese W. Choi, C. Pantofaru. Detecting and tracking people using an rgb-d camera via multiple detector fusion,. In Workshop on Challenges and Opportunities in Robot Perception (in conjunction with ICCV-11), 2011.