

3D Laser Scanning System and 3D Segmentation of Urban Scenes

L. C. Goron, L. Tamas, I. Reti, G. Lazea
Robotic Research Group, Technical University Cluj-Napoca
{lucian.goron, levente.tamas, istvan.reti, gheorghe.lazea}@aut.utcluj.ro

Keywords: 3D laser scanning system, 3D point cloud, urban segmentation, plane extraction

Abstract - This paper dwells upon the promising 3D technology for mobile robots and automation industry. The first part of the paper describes the design details of our own 3D Time of Flight (TOF) scanning system based on 2D laser range finder. The second part presents a specific segmentation technique for 3D outdoor urban environments by the common detection of plane models. In a few words, the technique separates the raw data into sparse and dense points, followed by the segmentation of the dense points into urban background and foreground objects. In the end we present some experimental results of real-world data-sets taken from the repository^{1,2} of the Leibniz University in Hannover, Germany.

I. INTRODUCTION

The paper includes two main parts. The first part is focused on the sensor design and its construction while the second part deals with the segmentation of urban scenes represented through point cloud data.

There are many possibilities to acquire 3D information from the surrounding environment. The measurement methods can be divided into three major categories based on applied sensor and sensing technology: stereo vision with two or more cameras [1], active triangulation [2] and time-of-flight (TOF) measurements. One of the most precise TOF measurement systems is based on laser scanners. Further on, this paper is focusing on this particularly type of sensor. The 3D laser range finder (LRF) is a relatively new and active remote sensing system which has been applied also in the mobile robotics domain in the last ten years. The application domain of these sensors is extending from the underground mine measurement to the urban environment and even aerial 3D image acquisition.

Exploring the outdoor urban environment is the next milestone of future robotic systems. Due to advanced scanning technologies the vision of widespread 3D data is becoming reality. Large scale urban scenes are now analyzed and processed using their 3D point cloud data representation. This is playing an important role in mobile robotics applications such as object or scenario detection, navigation, obstacle avoidance, surveillance and even service robots.

The segmentation of 3D data-sets [3] is a vital part of understanding urban scenarios. Besides the essential task of separating independent objects, segmentation can be helpful for localization, classification and feature extraction. Performing the task at hand is very difficult due to background vs. foreground related problems [4]. Furthermore, the real-world

data-sets are noisy and present an uneven sampling¹, thus being the result of ground-based scans² with point densities that dominate from the direction of which the scan is taken. In addition to this the low- or non- reflective surfaces are barely or not even represented (e.g. shiny objects or windows).

A vast amount of work has been done in this direction with encouraging results. But from the increasing scale and complexity of 3D data-sets, raises the need of simple processing techniques with well-known methods [5], especially from computational geometry. The problem of extracting 3D surfaces is widely spread in point cloud computation [6, 7] using high density data-sets. On the other hand, these approaches are not suitable for mobile robots applications. In opposition, the use of plane fitting combined with plane sweeping [8] is much relevant to the task at hand. In cluttered indoor scenes, e.g. in a kitchen or office, the environment is mostly composed out of squared furniture and walls. Meanwhile in an outdoor urban scene, e.g. university campuses or streets, we encounter buildings, walls, fences, trees, cars and of course humans.

Our approach presents a robust segmentation technique for separating background objects from foreground objects by fitting 3D plane models. At first, the noisy points are removed from the raw data, thus being referred to as sparse points. The remaining points, also called dense points, are then used for the segmentation. After correctly detecting the plane models, two different point clouds are obtained. One which represents the large scale objects (e.g. ground, buildings and walls) and the other one containing small scale objects (e.g. humans, cars, trees, fences). At the end, bounding boxes are attached for each foreground object by segmenting the foreground.

The paper has the following structure. In Section II we are discussing about state of the art and some related work. Section III presents the 3D laser scanning system. In Section IV the segmentation technique is detailed and in the last section we present the experimental results.

II. RELATED WORK

The TOF laser sensor is usually based on a transmitter-receiver laser diode pair which can give distance information from a few centimeters till hundreds of meters with a relative accuracy of less than 1 %. Commercial laser range finders like Sick, Leica, Riegl or Velodyne make use of a rotary mirror system through which the laser beam is swept along a surface in order to gain 2D or 3D information [9].

¹<http://kos.informatik.uni-osnabrueck.de/3Dscans/hannover1.tgz>

²<http://kos.informatik.uni-osnabrueck.de/3Dscans/hannover2.tgz>

In order to get information from the third dimension, the standard 2D laser scanners are often used with an auxiliary rotary mechanical system. The 2D laser is then mounted on that system, obtaining the third degree of freedom for the laser beam. Such an approach based on servo actuator system was used in [10, 11].

The main motivation for developing a custom 3D laser scanning system is the fact that the available commercial 3D laser are either not suitable for mobile robotics applications (they are too heavy and they need too much power for data acquisition) or they are too expensive compared to a standard 2D LRF [12]. The designed 3D LRF meets the requirements of the state of the art 3D laser sensors used in the mobile robotics community [13] and the cost of it is less than 10% of a commercial version.

Over the last decade, the task of processing point clouds data taken from urban environments has become more and more relevant. One popular application is the segmentation of urban models as described in [14] using the Hough transform. In opposition we have the state of the art [3, 4] where algorithms are developed to locate, segment, represent and classify most small objects in scanned point clouds of a city. Despite the great results, the system proposed in [3, 4] would be considered an overkill for simple urban tasks due to its large amount of resources needed.

Other approaches for segmentation of urban scenes are using systems that combine 3D laser scanning with vision [15, 16]. The authors of [15] describe a method for segmenting and detecting artificial objects by making use of structure and appearance information. The structure information detection is computed from 3D range data-sets and the appearance information from the image datasets. Unfortunately, the approach from [15] was not tested for real-life scenes and we cannot anticipate the results. On the other hand, in [16], we have a segmentation technique based on salient regions. It is a bottom-up process without any high-level priors, models or learning being robust against noise and outliers.

In our case, unlike images, we cannot use colors or textures as cues and unlike most computer graphics and CAD segmentation problems, the input is a noisy point cloud representing a scene rather than a clean surface model of an individual object. Our system does not need an objects' database for matching models or human supervision.

III. CONSTRUCTION OF 3D LRF

This section introduces the design and construction details regarding the 3D laser module. This module is based on a commercial Sick LMS200 2D laser product for which an auxiliary mechanical part was constructed in order to obtain 3D data sets.

A. 3D Sensor Design and Construction

The key component of the 3D sensor is the 2D LRF for which the rotary platform was designed. There are more possibilities to rotate the LRF, i.e. around the yaw, pitch and roll axes, thus achieving a *yawing*, *pitching* or *rolling* 3D sensor [13]. Each of these setups has its own advantage and disadvantage.

As for the mobile robots the most common approach is the pitching scan, this solution was adopted for the current design. The mechanical design shown in the Fig. 1 has two parts. One fixed part containing the driving servo motor (left) and the rotation encoder (right) and the other is the mobile rotary part on which is placed the Sick LMS200.

For the driving motor a Hitachi 12 [V] servo motor was chosen with a minimum rotation of 0.45° , while for the rotation sensor a varying resistor was considered. The motor control and the serial interface to the PC were solved using an AVR microcontroller based Cerebot2 type board. This type of board as well as the other mechanical and electrical components of the prototype are low cost products and available on the market.



Figure 1. The design (left) and the prototype (right) of the 3D sensor

The Sick LMS200 has a depth resolution of 2 [cm] and an angular resolution of 0.25° , 0.5° , or 1° depending on the configuration. The scanning cone of the device can be set either to 100° or 180° depending on the actual needs, while the maximum range of readings is up to 80 [m]. The scanning time is around 15 [ms] and additional time is required to send the data to the PC at 9600, 19200, 38400 or 500000 [kb/s]. Thus a complete 3D scan may require a few seconds.

B. Measurement Model and Data Acquisition

For a pitching type of scanner the third information about a point is from the pitch angle information. The coordinates of a 3D point result from the distance to the surface, the yaw measurement angle of the beam and the pitch angle moving of the mechanical part. Thus a scan point can be represented as a tuple of the form $(\rho_i; \theta_i, \gamma_i)$ where ρ represented the depth information from the LMS and θ, γ are the yaw and pitch measurement angles from the reading. The Cartesian coordinates of a point hence can be computed by the means of

$$p = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos \gamma & 0 & \sin \gamma \\ 0 & 1 & 0 \\ -\sin \gamma & 0 & \cos \gamma \end{pmatrix} + \begin{pmatrix} \rho \cos \theta \\ \rho \sin \theta \\ 0 \end{pmatrix} \quad (1)$$

In Equation (1) the displacement between the center of the robot and the 3D sensor was not taken into account. This could be introduced into the mathematical model by means of an additional translation term.

Also the error induced by the misalignment between the rotation axes of the laser mirror and the pitching axis is not discussed. This introduces a systematic error which can be detected by tests and eliminated by considering a constant term

in the above equation. A more detail discussion regarding the error budget can be found in [17].

A typical indoor scan for an area of interest (AOI) between -30 and +60 pitch angle is presented in Fig. 2. As it can be seen in this figure, the objects being closer to the right and left edges of the AOI are represented with a higher density of 3D points, while in the center the density is lower. This is due to the distribution properties of the pitching scan acquisition mode. In case that the AOI would be the central region, than the yawing or rolling scan method should be chosen.

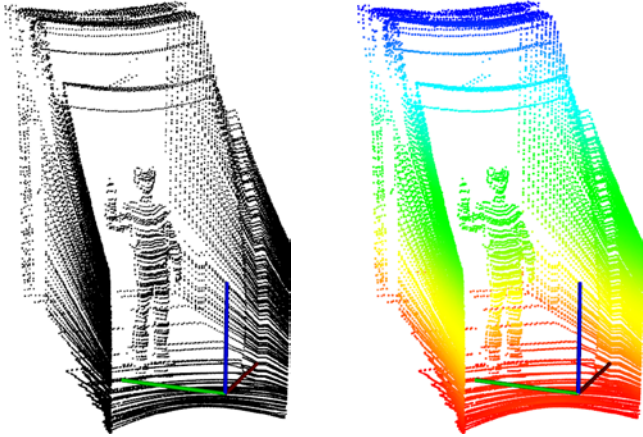


Figure 2. Simple indoor 3D laser scan. Raw point cloud data (left) and elevation map (right)

The above mentioned scanning system is still in development and being fine-tuned. The point cloud presented in Fig. 2 is just a simple indoor scan for demonstration. We want to have complete and complex point clouds and our goal is outdoor urban 3D scanning.

IV. SEGMENTATION OF POINT CLOUDS

Our method takes the raw data and cleans it using a statistical analysis of point densities. Points which are sparse are not taken into consideration for the following steps. The entire scene is then fitted with 3D plane models. Meanwhile our system is filtering models which are not suitable. Thus being very small/large planes or planes with scattered inliers. Next step will be to separate the background objects from the foreground objects. All the models' inliers we will refer to as background and the rest will be the foreground points. In the end we will compute bounding boxes for each independent object in the foreground point cloud. In some cases we compute the principle component analysis (PCA) for a better positioning of the bounding box.

In the next subsections we present the aforementioned steps in detail with the use of visual aids.

A. Removing Noise

Fitting models to a noisy point cloud can be a difficult task. This is the reason why noise removal is very important. The data-sets on which we tested our algorithm are taken with ground-based scanners. These will return point densities that dominate from the direction of which the scan was taken, thus resulting in an uneven sampling of points.

For overcoming this problem we propose an analysis of point densities. The algorithm is constructing first a kd-tree with all the points of the raw data. The number of approximate nearest neighbors³ (ANN) for every point is then computed and compared with a median threshold. The points which present a larger neighborhood as the threshold are considered dense points and the other are considered sparse points.

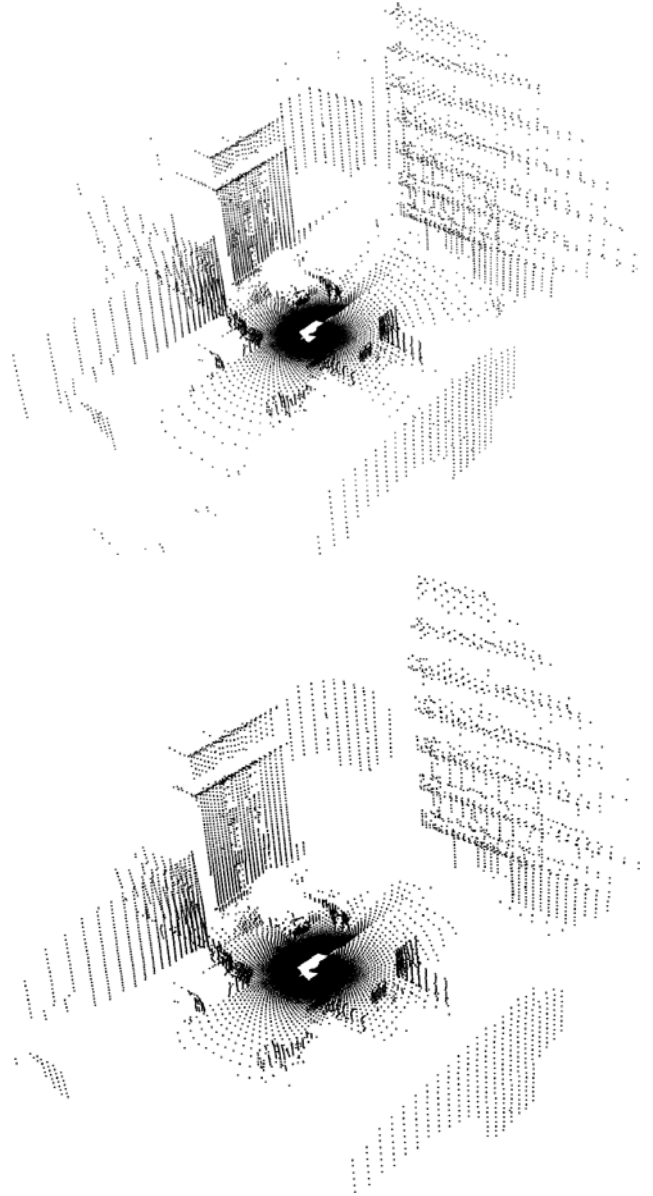


Figure 3. Point cloud data with (up) and without (down) noise

The low density points are considered noise and will be removed, while the high density points will be used for fitting plane models. The results of the noise removal procedure can be seen in Fig. 3.

B. Fitting 3D Plane Models

This is the most important part of our method. We believe that a reasonable segmentation of the urban scene background can be performed by fitting just plane models. For that we will

³<http://www.cs.umd.edu/~mount/ANN/>

use the RANSAC algorithm [18] combined with a filtering procedure for selecting valid models. While fitting one model, the algorithm will compare the number of its inliers to a model threshold for avoiding parasite plane fittings such as models with few inliers. This does not mean that models which contain a high number of inliers are automatically validated. There is still a possibility that RANSAC could return rogue planes. These models may seem to be broken into different regions, thus inliers being scattered in 3D space.

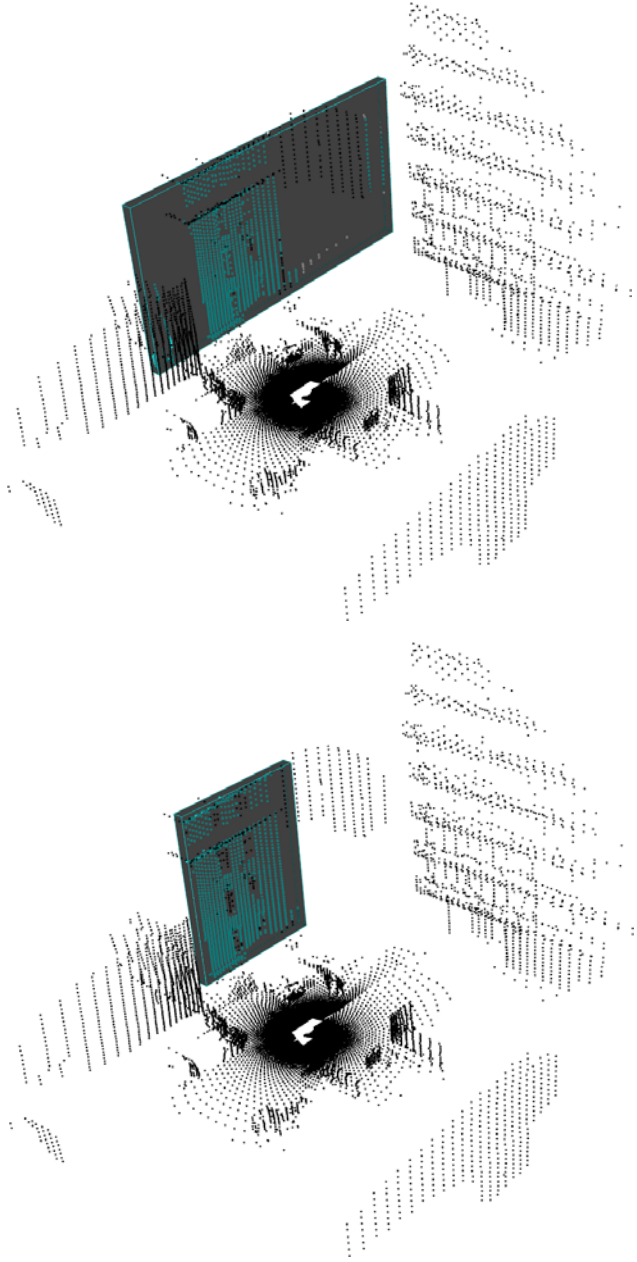


Figure 4. Rogue plane (up) and valid plane (down)

We can solve this impediment by computing a clustering procedure inside the model’s inliers. For this we would need the kd-tree mentioned in Subsection A. The idea behind this is to start with a query point and search for nearest neighbors within a given radius. Then, repetitively, search the nearest

neighbors of the last iteration’s neighbors within the same radius until there are no more. In the end, we will keep the cluster with the largest number of inliers, thus being the validated model.

As it can be seen in Fig. 4, the rogue plane has its inliers, colored with cyan, scattered with large spatial gaps between them. In opposition, we also show, in the same figure, the correctly fitted plane as an output of our method.

The fitting procedure will stop when the numbers of points drop below a minimum, out of which no reasonable plane can be fitted any more.

C. Separating Background from Foreground

In this subsection we will discuss the easiest part of our approach. As mentioned in Section I, our goal is to separate the large scale objects (e.g. ground, buildings and walls) from the small scale objects (e.g. humans, cars, fences and trees).

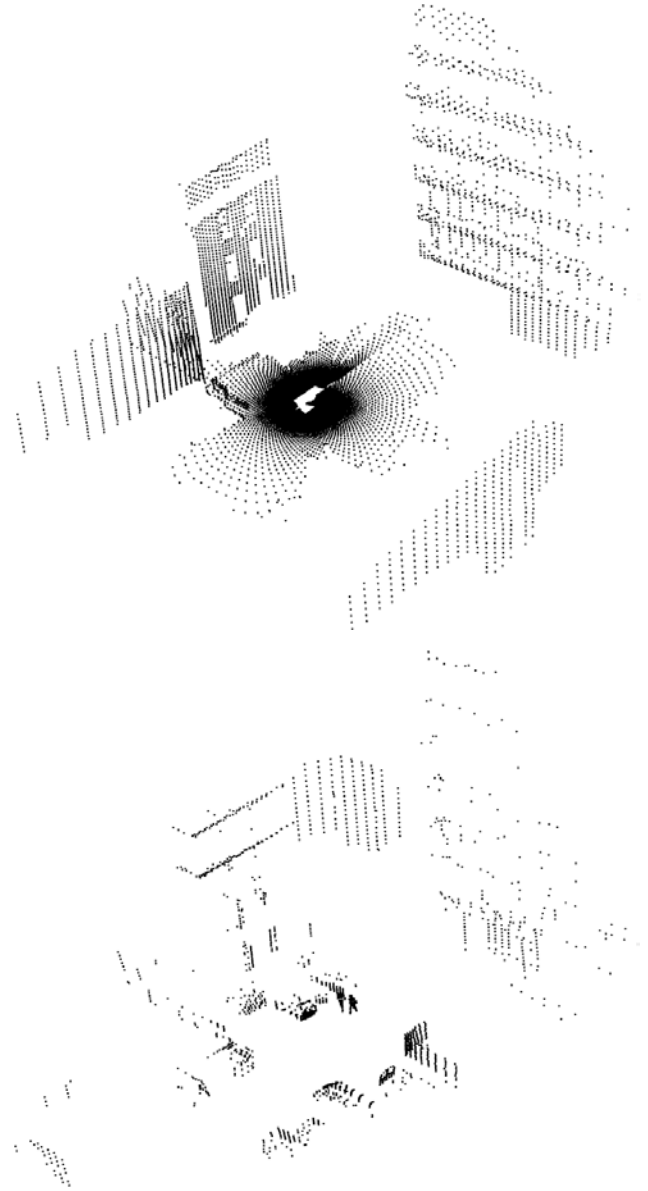


Figure 5. Segmenting background (up) from foreground (down)

After computing the plane models we will save all their inliers into one single point cloud data and this will be the background. The remaining points will represent the foreground. In Fig. 5 we have the results of the proposed segmentation. As it can be seen the foreground point cloud contains a large amount of noise.

D. Foreground Segmentation

Our final goal is the robust segmentation of self-independent objects of the foreground point cloud. Dealing with a noisy data set the method will use the analysis of point densities as mentioned in Subsection A. The result is shown in Fig. 6.

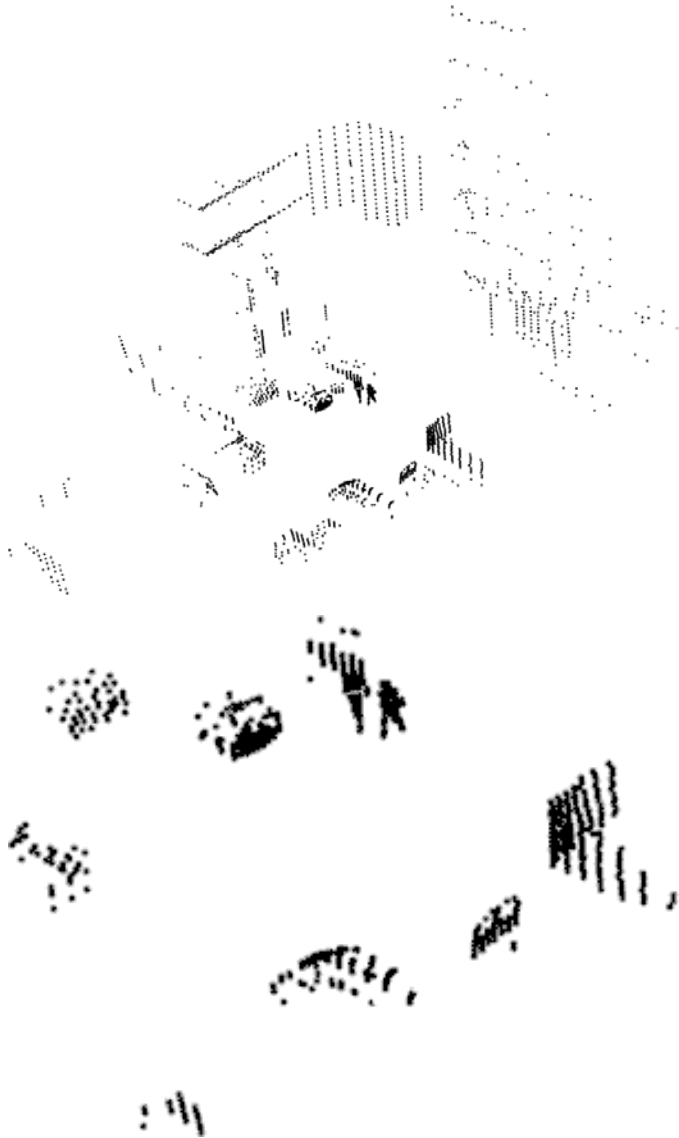


Figure 6. Original (up) and cleaned (down) foreground point cloud

In the end we make use of the clustering procedure presented in Subsection B to segment the cleaned foreground into regions which make up the self-independent urban objects. For each independent region the algorithm will compute a bounding box to improve visualization.

V. EXPERIMENTAL RESULTS

We applied our method to a couple of urban point clouds scenes taken from the repository of the Leibniz University in Hannover, Germany. The reason for this is that we are not able to perform 3D laser scans of urban environments at the current time. Meanwhile the 3D laser scanning system at hand is still under development.

The point clouds used for experimental results have different orientations containing common urban objects such as buildings, walls, fences, cars, trees and humans. The described algorithm returned fairly robust and straightforward results, as it can be seen in Fig. 7. The small variations in the results for the same dataset are due to the random element in the sample consensus approach, but the method gives consistent approximations.

An issue would be the under-segmentation of closely positioned objects, like e.g. the two humans in the upper-right corner of Fig. 7. For yet another example of under-segmentation related problems please take a look at Fig. 7, but this time in the lower-right-corner, where two cars were put in the same bounding box.

There is one small inconvenience regarding the visualization of the fitted models. In the second row of Fig. 7 we see that the bounding boxes of the planes are not correctly aligned.

VI. CONCLUSIONS AND FUTURE WORK

In this paper we described a 3D laser scanning system and a simple and robust approach for segmenting urban scenes represented by point cloud data sets.

While still being in development, the scanning system is showing promising results, as seen in Fig. 2. In the near future we will be able to perform 3D urban scans and process our own point clouds.

As for the segmentation approach we can say that it accomplished its goals by providing a simple and efficient way of separating self-independent urban objects and solving the background vs. foreground problems. For the future work we propose a procedure for merging overlapping planes and one for splitting intersecting models.

We are also intending to implement an algorithm which combines the fitting of plane models with the one of model lines for improved segmentation. The reason behind this is that line models can approximate better at a detail scale.

Another direction could be the reconstruction of urban environments without the dynamic objects like e.g. cars or humans. For this we would need to implement a principle component analysis procedure to compute the orientation of models and to solve the alignment problem.

REFERENCES

- [1] S. A. Lacroix, D. Mallet, G. Bonnafous, S. Bauzil, "Autonomous Rover Navigation on Uneven Terrains: Functions". *International Journal of Robotics*, pp. 917-942, 2002.
- [2] Perceptron. ScanWorks 3D Brochure. www.perceptron.com. 2009.

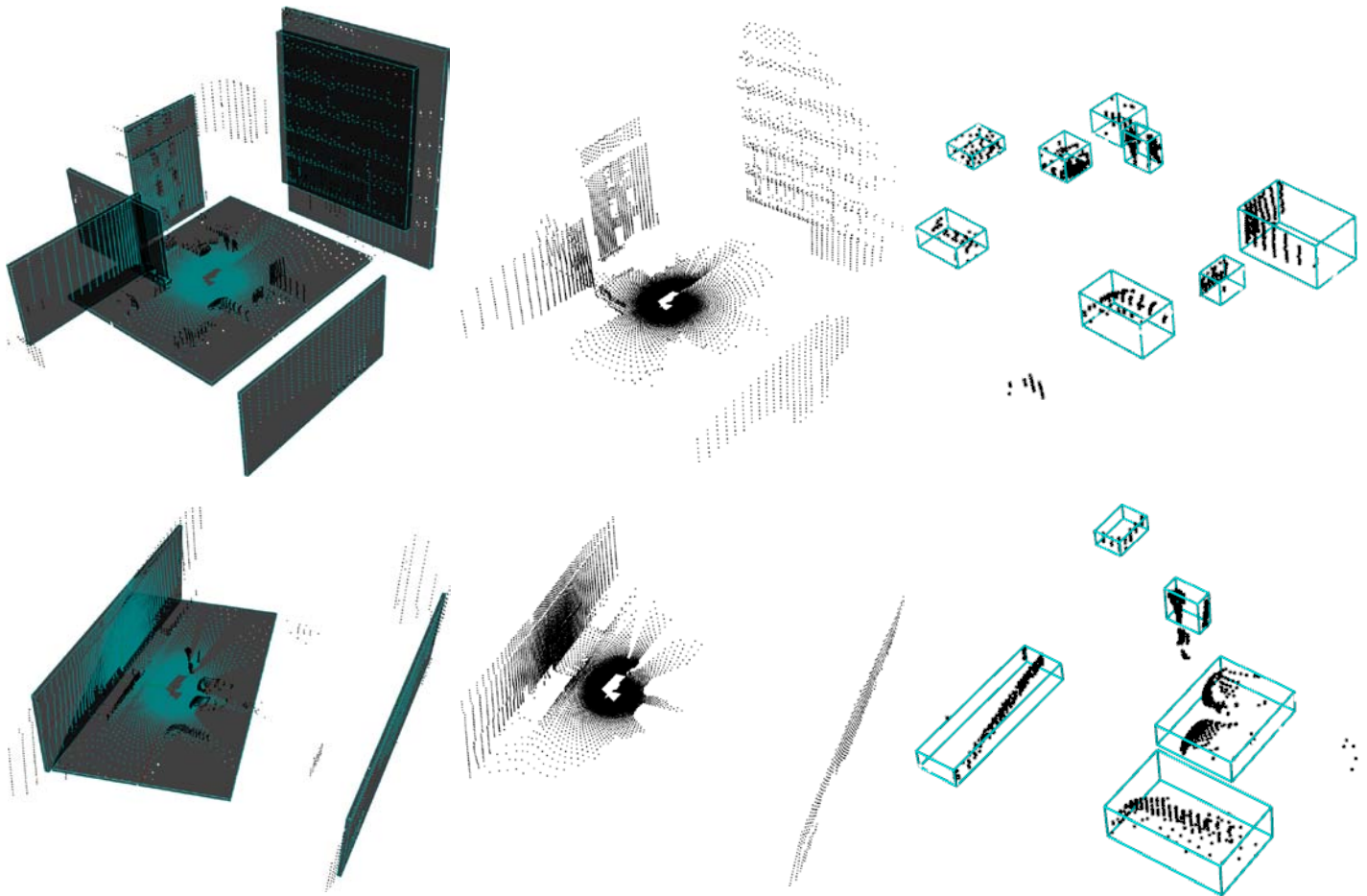


Figure 7. Two urban scenes were selected for visualizing the major steps of the implemented method. From left to right: (a) fitted 3D plane models, (b) background of scene, (c) segmented foreground

- [3] A. Golovinskiy, V. G. Kim, and T. Funkhouser. "Shape-based recognition of 3D point clouds in urban environments". *ICCV*, 2009.
- [4] A. Golovinskiy and T. Funkhouser. "Min-Cut Based Segmentation of Point Clouds". *IEEE Workshop on Search in 3D and Video (S3DV) @ ICCV*, 2009.
- [5] J. Poppinga, N. Vaskevicius, A. Birk, and K. Pathak. "Fast Plane Detection and Polygonalization in Noisy 3D Range Images". *International Conference on Intelligent Robots and Systems (IROS)*, 2008.
- [6] J. Klein and G. Zachmann, "Nice and Fast Implicit Surfaces over Noisy Point Clouds", *In: SIGGRAPH*, Sketches and Applications, 2004.
- [7] N. J. Mitra and A. Nguyen, "Estimating Surface Normals in Noisy Point Cloud Data". *In Proceedings of the 19th Annual Symposium on Computational Geometry*, ACM Press, pp. 322–328, 2003.
- [8] D. Hähnel, W. Burgard, and S. Thrun, "Learning Compact 3D Models of Indoor Outdoor Environments with a Mobile Robot", *Robotics and Autonomous Systems*, vol. 44, no. 1, pp. 15–27, 2003.
- [9] J.G. Weingarten, G. Gruener, and R. Siegwart. "A State-of-the-Art 3D Sensor for Robot Navigation". *In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots (IROS)*, vol. 3, pp. 2155–2160, 2004.
- [10] H. K. Surmann, K. Lingemann, A. Nüchter, and J. Hertzberg. "A 3D Laser Range Finder for Autonomous Mobile Robots". *In Proceedings of the International Symposium on Robotics (ISR)*, 2001.
- [11] P. Jensfelt and S. Kristensen. "Active Global Localisation for a Mobile Robot Using Multiple". *IJCAI*, pp. 13-22, 1999.
- [12] F. Maurelli, D. Droschel, T. Wisspeintner, S. May, and H. Surmann. "A 3D Laser Scanner System for Autonomous Vehicle Navigation", *ICAR*, 2009.
- [13] O. Wulf and B. Wagner. "Fast 3D-Scanning Methods for Laser Measurement Systems". *In Proceedings of the International Conference on Control Systems and Computer Science*, 2003.
- [14] G. Vosselman and S. Dijkman. "3D Building Model Reconstruction from Point Clouds and Ground Plans". *In: International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV-3/W4, pp. 37-43, 2001.
- [15] L. Spinello and R. Siegwart, "Unsupervised Detection of Artificial Objects in Outdoor Environments", *In: 6th International Conference on Field and Service Robotics (FSR)*, 2007.
- [16] G. Kim, D. Huber, and M. Hebert, "Segmentation of Salient Regions in Outdoor Scenes using Imagery and 3D Data", *IEEE Workshop on Application of Computer Vision (WACV)*, 2008.
- [17] A. Zhang and S. Hu. "Fast Continuous 360 Degree Color 3D Laser Scanner". *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 409-415. 2008.
- [18] R. Schnabel, R. Wahl, and R. Klein. "Efficient RANSAC for Point-Cloud Shape Detection". *In Proceedings of the Computer Graphics Forum* vol. 26, no. 2, pp. 214–226. 2007.