EXTRACTION OF SEMANTIC INFORMATION FROM THE 3D LASER RANGE FINDER

Michal Gnatowski[‡] and Barbara Siemiątkowska[‡] and Jacek Szklarski[‡]

[‡] Institute of Fundamental Technological Research, Polish Academy of Sciences, Warsaw, Poland

Abstract In this paper a system for extracting semantic information in indoor and outdoor environment from 3D laser scanner is presented. The largest objects (like walls, floor, ceiling, etc.) are recognized by constructing an RGB image based on normal vectors and applying a simple rule-based system. More sophisticated techniques are used to detect the remaining ones: Haar-like features – to classify small and irregular objects, and Cellular Neural Networks – to distinguish between different types of ground on which the robot is able to operate.

1 Introduction

Classical navigation system consists of the following parts: mapping, localization and collision-free path-planning. A comprehensive overview of the existing navigation techniques can be found in (Thrun et al., 2005). Many environment representations have been proposed. One of the most popular is 2D representation (Elfes, 1987; Thrun et al., 2005) which has many limitations, for example the height of obstacles is not taken into account. In the last decade 3D sensors are being more popular which makes possible 3D map building. The following representations are used most often:

- point clouds (Triebel et al., 2006), (Vosselman et al., 2004)
- grid-based representation (Sakas and Hartig, 1992), (Siemiątkowska et al., 2007), (Gnatowski et al., 2008) which can be divided into: elevation maps (standard or extended) and multi-level surface maps.

In case of human-robot interaction the map has to contain not only metric but also semantic information (Rusu et al., 2008; Mozos et al., 2007). The main goal of our current project is to build a semantic map of unknown environment based on data obtained from a 3D laser scanner. In this paper we presented that semantic information can be extracted from the data obtained by a laser range sensor. This extraction is performed by a combination of image analysis techniques and background knowledge.

2 Data Acquisition

The experiments have been done on a mobile robot equipped with a head module comprising a 3-dimensional scanning laser rangefinder used for navigation and creating 3-dimensional representation of the environment. The module consists of a SICK LMS 200 scanning laser rangefinder installed on a rotating support, which can rotate the scanner around the horizontal axis within the angular range from -15 to +90. The scanning laser enables to measure the distance to the obstacle within 180 with resolution of 0.5. The data is subsequently transmitted to the control unit by means of an RS 422 bus. The module is powered by a DC planetary gear motor. The power is then transferred by means of a toothed belt transmission. Two rotational encoders measure the scanning velocity and angle. The first encoder installed on the motor shaft is used for regulating the position whereas the other is responsible for measuring the rotation angle directly on the rotation axis of the scanner. The two measuring systems allow precise steering and positioning of the sensor. The unit controlling the head enables both continuous as well as step-by-step modes of the head. PID control algorithms were used for positioning and controlling the drive unit. Communication with the main control unit is achieved by means of an RS 422 bus. Figure 1 (left) presents pure data obtained from the described device.

3 Texture-based object detection

In our approach polar coordinates obtained from the laser range finder are transformed into Cartesian coordinates and then they are represented as a set of normal vectors n. A color RGB image is constructed by assigning values of the coordinates n_x, n_y, n_z as colors red, green and blue accordingly. The normal vectors give very important information about surfaces which are observed - when two points belong to the same plane the normal vectors computed in this points are equal. This allows us to perform segmentation of data obtained from the laser range finder in office-like environment. The purpose of the first step is to perform a fast segmentation of the gathered data into areas, each one representing a flat polygon in the real scene. The most important areas are, of course, ceiling, floor, walls, doors, etc. Besides the list of polygons, some numbers characterizing physical properties of a polygon can be extracted as well, which can be used later for better object classification. Simple rule-based system allows to classify the surfaces like:

walls, ceiling, floors, doors, etc.(Siemiątkowska et al., 2009). The ramaing objects have to be classified using different methods. We performed the classification based on Haar-like features, which is described in the next section. Figure 1 presents the point cloud obtained from the laser range finder in the corridor and the result of classification. Gray area represents walls, dark blue - the floor and bright blue - the ceiling. The yellow area represents objects which are not classified.

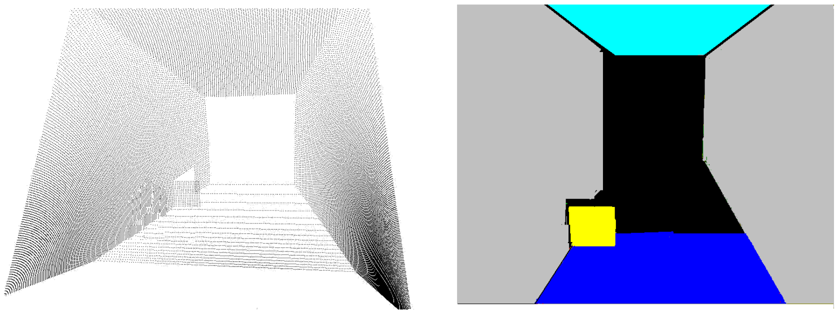

Figure 1. *Left:* Point cloud; *Right:* Detected walls, the ceiling and the floor. Yellow area represents unrecognized objects.

We performed the similar experiments in outdoor environment (Figure 2). We classify the ground into three classes: flat surface (pavement), grass and obstacles. The ground is recognized based on metric information and Cartesian coordinates, but in order to distinguished pavement from the grass the Cellular Neural Network (CNN) is used.

4 Haar-like features

Detecting regular and large areas like floor or ceiling in indoor environment in the way described above is a fairly easy task. However, for more complex objects the approach for the detection process should be different. Treating laser scanner data as an image makes it possible to apply well known methods for recognition from image analysis. Here we propose a simple classifier based on Haar features and rules related to geometrical features of objects. Algorithms which are applied here are available in OpenCV library and they implement methods proposed by (Viola and Michael, 2001) for basic set of Haar-like features and by (Lienhart and Maydt, 2002) for rotated features which enhance the basic set. Similar work, however, based on reflectance images and depth images has been proposed by, e.g., (Nüchter et al., 2005,

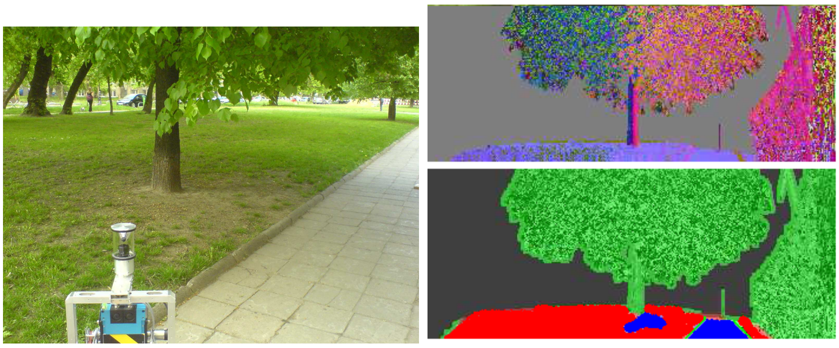

Figure 2. *Left:* The scene where the experiment was performed; *Right upper:* Normal vectors on RGB image; *Right lower:* Recognized objects: grass(red), pavement(blue) and unrecognized (green);

2004).

In the first step we define set of classes $\mathcal{C}$ of objects, like stairs, windows, doors, wastebaskets, and any objects which a mobile robot should be able to recognize. Then for each class a specific Haar-like classifier is created. This is done by training the classifier with use of a large set of positive and negative sample images representing objects for the given class and part of background respectively. All of the images should be scaled to the same size, we use 20x20. As the set of negative examples we use large number of arbitrary images representing a scene without any object from the class of interest.

The first stage of the classification process is to search an image constructed from laser data for objects belonging to any class $\mathcal{C}$. The classifier can be applied to any region of an image giving *true* if the region is likely to contain pattern similar to one of those from the positive samples set, *false* otherwise. Analysis is very fast so one can try many different regions with varied sizes from all parts of the image. By doing this in a loop one can search entire image for any object belonging to any class from set $\mathcal{C}$. Figure 3 shows a result of such analysis when searching with classifier trained for detection of a wastebasket.

In the next step false-positives are removed. This is done under assumption that any object from class $\mathcal{C}$ represents a ,,continuous structure" in the real space 3D. First, a gray-scale image representing distance to obstacles is divided into N areas A_i. A flood fill algorithm is run on the image representing distance with pixel (0,0) as the seed point. The threshold for the

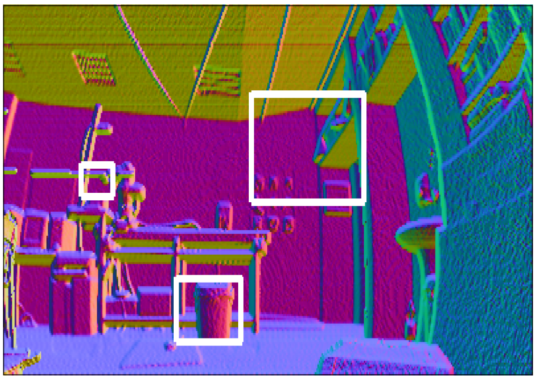

Figure 3. An RGB image constructed from laser measurements. White rectangles represent results of a Haar classifier trained for detection objects like a wastebasket. Note, that besides the wastebasket, there are two recognized regions which clearly should not be labeled as ,,wastebasket" (false-positives).

algorithm is constant and it corresponds to about 5 cm (difference between neighbor pixels is considered when flooding). If the resulting area is large enough, i.e., has total size grater than 15 pixels is marked with number 1 and then the same procedure is applied for a next pixel which has not been assigned to the area. In this way we obtain area number 2, and the procedure is repeated until all the pixels are assigned to one of the n areas. All areas which are to small to be classified are marked with number 0 and are not considered in any later stage of the process.

In order to check if any area H_i detected by the Haar classifier represents indeed a ,,continuous" object, it is checked if there exists i, so that $(A_{\mathrm{H}} \cap A_i)/A_i > 0.9$, where $A_{\mathrm{H}} \cap A_i$ is size of the common area between area detected by the classifier and the area A_i. In other words there should exist a continuous area A_i which covers most of the region recognized by the Haar-classifier (and A_i can not be much larger than this region). In this way all false-positives are rejected. Figure 4 shows 30 largest areas A_i together with rectangle which denotes the region recognized by the Haar classifier.

Naturally, each pixel of an image constructed from laser measurements corresponds to a 3D point from the associated point cloud. Therefore immediately there is available information about real size of any considered area A_i. For example, for the area labeled with white rectangle from Fig. 4

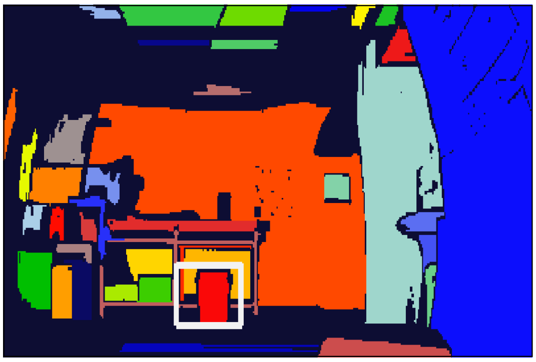

Figure 4. Areas A_i resembling continuous objects in real space (for each area, the real distance between neighbouring pixels is smaller than 5cm). The white rectangle denotes the region which has been recognized by the Haar classifier as a wastebasket.

all points can be placed into a box with size 27cm x 12cm x 31cm (width, depth, height). This information serves as an input for another rule-based classification scheme which filters out potentially bad recognized objects. For example, for the wastebasket we assume that its height must be between 20cm and 60cm and width between 20cm and 40cm.

5 Conclusions

In the article we discussed methods of extracting semantic information from 3D laser range finder. In contrary to visual cameras, using such 3D scanners gives data which are independent of lightening. Moreover, there is geometrical information available which can be used in the rule-based classification procedure. The goal of the research is to give the mobile robot an ability to recognise objects of certain classes. This will make possible to give commands like, *"Go to the* ***kitchen*** *and bring* ***a bottle of milk*** *from the* ***refrigerator****"*.

In order to make use of some classical approaches from computer vision, we construct RGB images based on vectors normal to surface at each point provided by the range finder. These images are then used in a simple rule-based classification system, detection process based on Haar-like features, and Cellular Neural Networks which distinguish between different types of ground. Experiments indoor as well as outdoor were performed and the

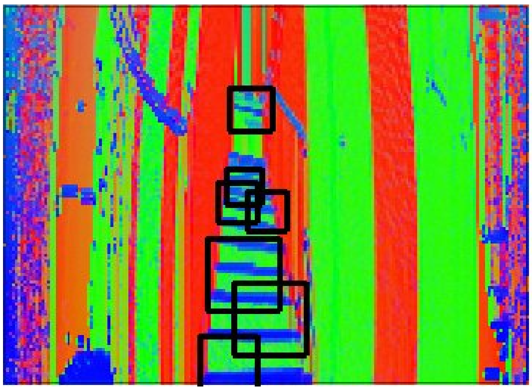

Figure 5. Another example of recognition using Haar-like features and geometrical information. Here the recognized areas are marked as „stairs" (for creating the image from normal vectors to RGB, absolute values $|\vec{n}_x|, |\vec{n}_y|, |\vec{n}_z|$ instead of $\vec{n}_x, \vec{n}_y, \vec{n}_z$ were used).

results are promising. In the future we will enlarge the list of identified objects.

6 Acknowledgment

This work was supported by the Polish Ministry of Science and Higher Education (grant 4311/B/T02/2007/33)

Bibliography

A. Elfes. Sonar-based real-world mapping and navigation. *IEEE Transactions on Robotics and Automation*, pages 249–265, 1987.

M. Gnatowski, B. Siemiątkowska, and R. Chojecki. Cellular neural networks in 3d laser data segmentation. In *9th WSEAS International Conference on NEURAL NETWORKS*, pages 84–88, 2008.

R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. *IEEE ICIP 2002*, 1:900–903, 2002.

O. M. Mozos, R. Triebel, P. Jensfelt, A. Rottman, and W. Burgard. Supervised semantic labeling of places using information extracted from sensor data. *Internationl Journal of Advanced Robotics Systems*, 2007.

Andreas Nüchter, Hartmut Surmann, and Joachim Hertzberg. Automatic classification of objects in 3d laser range scans. In *In Proc. 8th Conf. on Intelligent Autonomous Systems*, pages 963–970. IOS Press, 2004.

Andreas Nüchter, Kai Lingemann, Joachim Hertzberg, and Hartmut Surmann. Accurate object localization in 3d laser range scans. In *In Proceedings of the 12th IEEE International Conference on Advanced Robotics (ICAR '05)*, 2005.

R. Rusu, Z. Marton, N. Blodow, M. Dolha, and M. Beetz. Towords 3d point cloud based object maps for household environment. *Journal of Robotics and Autonomous Systems*, 56:927–941, 2008.

G. Sakas and J. Hartig. Interactive visualization of large scalar voxel fields. In *Vizualization*, pages 29–36, 1992.

B. Siemiątkowska, M. Gnatowski, and A. Zychewicz. Fast method of 3d map building based on laser range data. *Journal of Automation, Mobile Robotics and Intelligent systems*, 1(2):35–39, 2007.

B. Siemiątkowska, J. Szklarski, M. Gnatowski, and A. Borkowski. *Towards Semantic Navigation System, Recent Advances in Intelligent Information Systems.* Academic Publishing House EXIT, 2009.

S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics.* MIT-Press, Cambridge, Masachussets, USA, 2005.

R. Triebel, P. Pfaff, and W. Burgard. Multi-level surface maps for outdoor terrain mapping and loop closing. In *In Proc. of IROS*, pages 1–2, 2006.

P. Viola and J. Michael. Rapid object detection using a boosted cascade of simple features. *IEEE CVPR*, 2001.

G. Vosselman, B. Gorte, G. Sithole, and T. Rabbani. Recognizing structure in laser scanner point clouds. In *Proceedings of Conference on Laser scanners for Forest and Landscape assessment and instruments*, Freiburg, Germany, 2004.