

Feature Extraction from Dense 2D Range Images for Service Robotics

Peter Einramhof, Robert Schwarz and Markus Vincze
Automation and Control Institute (ACIN), Vienna University of Technology
Gusshausstrasse 27-29 / E376, A-1040 Vienna, Austria
pe,rs,vm@acin.tuwien.ac.at

Abstract. *In this paper we present a fast approach to extracting features from range images. These features are intended to serve as input to the perception system of a domestic service robot. In the first step, after initial noise filtering, a range image pyramid is computed and used together with a measure for local planarity to determine regions that require and allow further smoothing. Mixed pixels at depth discontinuities are identified. This is followed by extracting step and roof edges. Planar patches are detected with a focus on horizontal and vertical planar structures. Finally, a "label map" is created by assigning a label to each pixel indicating its membership to one of the feature types. We show feature extraction results of our approach on real range images recorded in a home-like office environment using a tilting laser range finder, and give examples for their use in typical service robotics tasks.*

1. Introduction

In classic service robotics the focus lies on transporting objects in controlled and well-structured environments like factories, hospitals or offices. A simple perception system is sufficient to provide data for the tasks of the robot, namely self localisation and obstacle avoidance. A popular sensor is the 2D laser range finder scanning parallel to the ground because the environment can be modelled by 2D grid or feature maps due to an abundance of unobstructed vertical structures such as walls or the faces of file drawers and closets. Processing the sensor data, for instance 361 range measurements for one 180° scan, requires only very little computational power to work in real-time, which is a necessity since the robot is moving in a dynamic environment.

In domestic service robotics the demands to the perception system are higher in comparison. Home

environments are challenging since they are cluttered and less structured. There are amorphous surfaces like curtains or other home textiles, and protruding surfaces such as table tops. The three-dimensional nature of this environment can no longer be neglected. To be useful for technically non-trained users, the offered services must be more than just safely navigating from A to B. The perception system has to support grasping and object detection, for instance. The latter also contributes to more natural human-robot interaction since it is closer to human perception: the user would rather like to call the robot to "the sofa in the living room" than to a location " x, y, θ ".

When changing over from 2D to 3D data (more exactly: 2.5D), the demands to the computational part of the perception system rise drastically due to the increased amount and complexity of the data. Nevertheless, the real-time requirement still stands. Since mobile robots run on batteries and a long time of autonomy is desired, the onboard computational power cannot be increased arbitrarily. To achieve real-time responsiveness despite this restriction, fast segmentation algorithms are required.

In this paper we address range image segmentation under real-time constraints as well as detection and removal of mixed pixels, outliers that occur at discontinuities within the range image. The features extracted in the course of our segmentation approach are intended to be useful for various tasks of indoor service robotics. Horizontal planar structures provide information for safe navigation and about support planes such as table tops. Vertical planar structures can be used for map building and self localisation, or represent doors. Step and roof edges define object boundaries and the transition between object and support plane or between object parts, which can serve as additional information to object detection

and eventually grasp point computation.

The remainder of this paper is structured as follows: Section 2 gives an overview of related work. Section 3 provides the motivation for the features we have selected as target result and describes our approach in detail. Section 4 provides experimental results on data recorded with a tilting laser range finder. Finally, Section 5 concludes with a summary and an outlook.

2. Related Work

When compared to intensity or colour images, which provide information about the surface properties of the objects observed by a sensor, range images encode the three-dimensional structure of the observed scene. The purpose of range image segmentation is to divide the image into features or regions that are meaningful with respect to a given task.

Comparison of quality and performance of different segmentation methods is difficult due to the lack of sound experimental evaluation. An exception is the field dealing specifically with the segmentation of objects with planar faces, published in [8] together with experimental data. A detailed overview of literature in that field is available in [4]. Segmentation methods can be roughly divided into edge-based and region-based approaches.

Edge-based methods are inspired by human vision since humans have the principle that there is a discontinuity of some kind between two separable objects [17][12]. Edge pixels have a gradient assigned, that is, magnitude and direction of the greatest local change. Sappa and Devy [13] use an edge-based segmentation technique that consists of two stages: the first stage generates a binary edge map based on scan line approximation that considers only two orthogonal scan line directions. The second stage links the edge points by applying a graph strategy. In [1] Bellon and Silva present range image segmentation based on edge detection techniques with the aim of better preserving the object topology and shape in noisy range images. Their approach avoids fixed thresholds. Han et al. [5] propose a jump-diffusion method for segmenting a range image and its associated reflectance image in a Bayesian framework. Harati et al. [6] propose the metric "bearing angle", which is the incidence angle between the measurement beam and a surface. By thresholding the bearing angle and its first derivative, step and roof edges are detected. Since their target application is 3D in-

door SLAM they are rather interested in the remaining planar patches and thus remove all edges.

Region-based methods group pixels into regions using the criteria of proximity and homogeneity. These methods achieve grouping either by splitting the image into smaller regions [10], merging small regions into larger ones [8], or splitting and merging until all criteria are maximally satisfied ([2][9][7]. In more recent work, Gotardo et al. [3] present a robust estimator, derived from the RANSAC and MSAC estimators, whose optimization process is accelerated by a genetic algorithm. Their range image segmentation algorithm is based on planar surface extraction in preserving small regions and edge locations when processing noisy images. Similarly, Wang and Suter [15] propose a highly robust estimator (Maximum Density Power Estimator), which applies nonparametric density estimation and density gradient estimation techniques in parametric estimation ("model fitting"). According to the authors it can tolerate more than 85% outliers. Weingarten et al. [16] use probabilistic plane fitting to extract large planar surfaces from range images as input to mapping the environment for mobile robotics.

Similarly to Bellon and Silva [1], we also make use of standard image processing as much as possible for edge detection. To increase robustness, various methods are combined in a voting scheme, and also the metric "bearing angle" [6] is incorporated. The contribution of this paper is the proposal of a processing pipeline that addresses range image segmentation in the context of robotics holistically and whose computational complexity is linear in the number of pixels, capable of running in real-time.

3. Approach

In this section we discuss what features are extracted in the course of our segmentation approach as well as their relevance in the context of domestic service robotics. This is followed by a detailed discussion of the individual processing steps of our approach.

3.1. Target Features

The vertical axis is an important reference (direction of gravity). This information is incorporated into the sensor data and associated processing algorithms via the known geometry and kinematics of the setup consisting of sensor and robot, and optionally via inclinometers.

In (manmade) indoor environments horizontal and vertical planar structures are dominant; together they also define the room structure. Horizontal planar structures such as the ground, table tops or the seats of chairs play the role of support planes on which the robot moves and on which obstacles or objects of interest are located. Vertical planar structures represent walls, doors, the faces of closets or the bodies of objects. Especially walls define the boundaries of the indoor environment (together with ground and ceiling) and can serve as features for the robot's self localisation.

Step edges occur at object boundaries, more exactly at the transition between foreground and background. We consider step edges as part of foreground objects. Roof edges occur at the transition between parts of objects or between object and support plane, and also at top rims and high-curvature surfaces. Roof edges can be concave or convex.

3.2. Preprocessing

As preprocessing step a 3x3 median filter is applied to the range image I . This removes spurious outliers and reduces noise. Especially the former would otherwise affect neighbouring range measurements in the course of optional, subsequent smoothing. Due to the small kernel size edges are largely kept intact and the computational effort is low.

It is expected that the input range images are either dense, or, if they have holes, that they can be made dense by applying a fast depth diffusion approach [14].

Depending on the remaining noise level of the range image further smoothing steps might be necessary, repeating median filtering or additionally applying Gaussian smoothing. In the latter case we use a normalized 3x3 binomial kernel.

The final step of preprocessing is the generation of a Gaussian range image pyramid (Figure 1). To do so, the initial range image is convoluted with a normalized 5x5 binomial filter followed by subsampling the result by a factor two in each dimension. These two steps are repeated on the resulting images. For performance reasons we make use of the separability of the filter kernel and directly incorporate subsampling: a 1D horizontal kernel is applied to every second pixel column of the range image. This is followed by the application of the translated i.e. vertical 1D kernel to every second pixel row of the intermediary result.

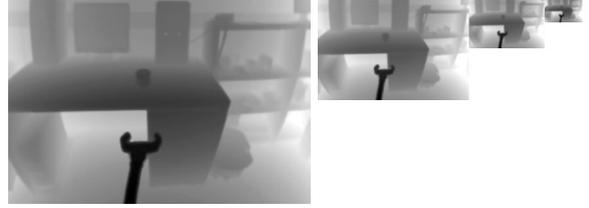


Figure 1. Example of an image pyramid of a range image showing a table scene with a robotic gripper. Subsampled images were only created for as long as both dimensions were multiples of two.

3.3. An Indicator for Local Planarity

We consider the nine range values $r_{i+m,j+n}$ (i and j are the column and row indices) from any 3x3 neighbourhood of an approximately locally equally-spaced range image. Tilting laser scanners, stereo and time-of-flight cameras fulfil this requirement. If these range values can be explained by the linear equation

$$r_{i+m,j+n} = k_i m + k_j n + r_{ij}, \quad m, n \in [-1; 1] \quad (1)$$

then they are part of a (close to) planar patch (k_i and k_j are constants). Please note that the coordinate system defined by (m, n, r) is not exactly Euclidean which can be neglected in practice. If Equation 1 is perfectly satisfied, there is an interesting relationship between gradient magnitude g_{ij} and the population standard deviation σ_{ij} of these values of the 3x3 neighbourhood around r_{ij} :

$$\sqrt{\frac{2}{3}} \frac{g_{ij}}{\sigma_{ij}} = 1.0 \quad (2)$$

The gradient magnitude (and direction θ_{ij}) is calculated using the Prewitt operator:

$$g_{ij} = \sqrt{g_{x,ij}^2 + g_{y,ij}^2}, \quad \theta_{ij} = \arctan\left(\frac{g_{y,ij}}{g_{x,ij}}\right) \quad (3)$$

$$g_{x,ij} = \frac{1}{6} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} * I \quad (4)$$

$$g_{y,ij} = \frac{1}{6} \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} * I \quad (5)$$

and the standard deviation is calculated as

$$\sigma_{ij} = \sqrt{\frac{1}{9} \sum_{m=-1}^1 \sum_{n=-1}^1 (r_{i+m,j+n} - \mu_{ij})^2} \quad (6)$$

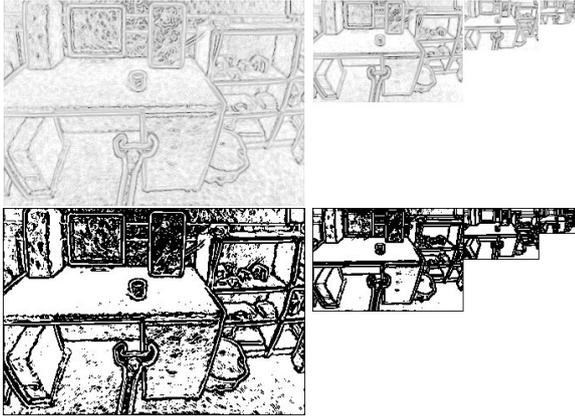


Figure 2. *Top*: Ratio of the local gradient magnitudes and the standard deviations. *Bottom*: Ratio thresholded by the empirically found value 0.985. Black represents the value 0.0, white represents 1.0.

$$\mu_{ij} = \frac{1}{9} \sum_{m=-1}^1 \sum_{n=-1}^1 r_{i+m, j+n} \quad (7)$$

The special case that all nine values are equal and as a result the standard deviation is zero must be intercepted – the result is manually set to 1.0 since Equation 1 is perfectly satisfied. In noisy regions, at depth discontinuities (step edges) or in regions with high curvature (roof edges) the result is smaller than 1.0.

Figure 2 shows the result of computing the ratio from the image pyramid in Figure 1 (top) and thresholding it (bottom). For planar patches that are viewed by the sensor under an acute angle, the ratio is close to 1.0, but for planar patches that are perpendicular to the sensor’s viewing direction the impact of measurement noise is clearly visible. At smaller resolutions the influence of noise on planar patches disappears while edges remain stable. Also, the curvature i.e. the non-planarity of rounded objects becomes more prominent at lower resolutions. The ratio at full resolution and its behaviour across scales is used to determine what regions require and allow further smoothing (planar but noisy). Even if noise remains, the identified planar regions serve as good seed regions for region growing.

3.4. Mixed Pixels

In range images (or equivalently: depth images) from tilting laser scanners, time-of-flight cameras and also stereovision, there are range measurements at depth discontinuities that do not correspond to any physical structure. These mixed pixels have values that lie somewhere between the valid foreground

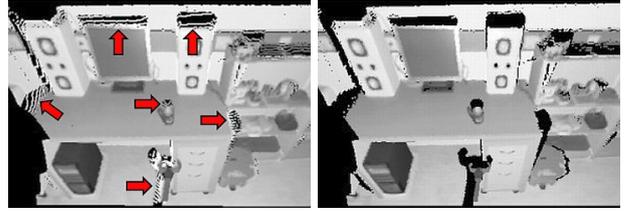


Figure 3. 3D data computed of our initial range image. At depth discontinuities occur mixed pixels (marked by arrows, left) that must be removed (right).

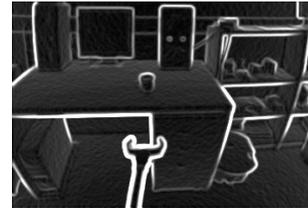


Figure 4. Standard deviation for each pixel of the range image within a 3x3 neighbourhood. Maximum deviation was clipped to 0.05m for better visibility of low deviation regions.

and background range measurements (Figure 3, left). They are problematic because they seemingly connect foreground and background points to one contiguous object and thus have to be removed (Figure 3, right).

The first step in detecting mixed pixels is to find depth discontinuities, that is, abrupt changes in the values of neighbouring pixels of the range image. But what extent of change is ”abrupt”? Clearly, this depends on the measurement noise of the used sensor. To determine the noise level, we use the standard deviation of the range values for each pixel within a 3x3 neighbourhood that has already been computed in the previous section. Figure 4 shows that the standard deviation is high at depth discontinuities. The standard deviation computed for each pixel votes for a bin of a histogram with a bin width of 0.1mm. We determine the location of the (first) peak of the histogram and consider it as ”sigma” of the noise. Three times this ”sigma” serves as threshold to determine depth discontinuities within the standard deviation image. While this threshold works well, there is a problem when neighbouring pixels’ beams intersect with planar structures at larger distances and at acute angles. In such cases the local change in measured range is well above the threshold and would thus register as discontinuity.

In [6] the authors propose the metric ”bearing angle” (Figure 5), which is the incidence angle $\beta_{ij,h}$ between the measurement beam and a surface. It is

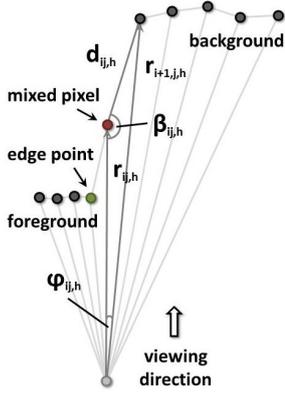


Figure 5. Horizontal bearing angle $\beta_{ij,h}$ (viewed from above). $r_{ij,h}$ and $r_{i+1,j,h}$ are two neighbouring range measurements of the row j of the range image, enclosing an angular increment $\phi_{ij,h}$. $d_{ij,h}$ is the distance between the intersection points of the measurement beams $r_{ij,h}$ and $r_{i+1,j,h}$ with the surface.

computed as

$$\beta_{ij,h} = \arccos\left(\frac{r_{ij,h} - r_{i+1,j,h}\cos(\phi_{ij,h})}{d_{ij,h}}\right) \quad (8)$$

$$d_{ij,h} = \sqrt{r_{ij,h}^2 + r_{i+1,j,h}^2 - 2r_{ij,h}r_{i+1,j,h}\cos(\phi_{ij,h})} \quad (9)$$

At real depth discontinuities this angle takes on values close to 0° or 180° , i.e. the beam would be close to parallel to the (imaginary) surface. Although Harati et al. use solely this metric, it is problematic at short range i.e. objects close to the sensor, like in the case of a table scene. Due to the in general small angular increment $\phi_{ij,h}$ between two neighbouring measurement beams, $r_{ij,h}$ and $r_{i+1,j,h}$, and the resulting small lateral distance between them at close range, the bearing angle $\beta_{ij,h}$ rather reflects the measurement noise than the geometry of the scanned object in such cases.

Since thresholding the standard deviation image yields wrong depth discontinuities at greater distances and thresholding the bearing angle yields wrong depth discontinuities at close distances, we multiply both thresholded results, which leaves only discontinuities where both methods agree (figure 6, top). The threshold value for the bearing angle depends on the sensor; for our setup we experimentally found angles of smaller or equal 5° or greater or equal 175° ¹. According to the principle sketched

¹The computationally costly arccos can be skipped - we can directly compare against the cosine of the extremal angle.

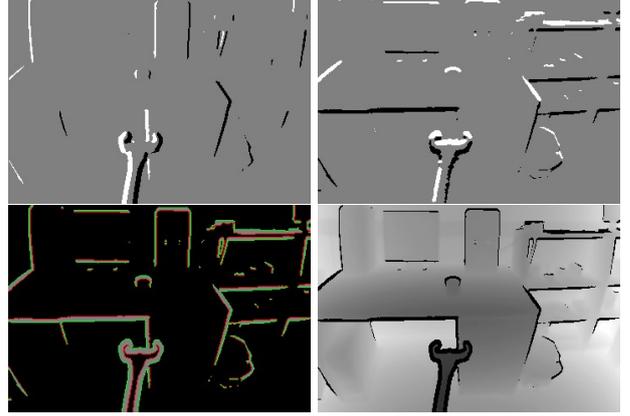


Figure 6. *Top*: Sign of the thresholded horizontal and vertical bearing angles. *Bottom left*: Mixed pixels (gray) that are enclosed by step edges i.e. valid foreground pixels (red) and valid background pixels (green). *Bottom, right*: Range image without mixed pixels.

in Figure 5, mixed pixels are determined from both bearing angle images and combined (Figure 6, bottom left).

3.5. Step Edges

In our definition step edges are pixels of the range image at depth discontinuities that (1) are valid pixels and (2) belong to the foreground. Thus, they represent boundaries of foreground objects. Since the first derivative of range images produces undesired strong gradient magnitudes at planar areas that are further away and that are intersected by the measurement beams or visual lines of the sensor at acute angles, we use the second derivative. The latter is sensitive to measurement noise, thus, the range image is first further smoothed, but only those areas for which our previously determined indicator (Section 3.3) allows it. In the next step, the filtered range image is convoluted with a 3×3 mask that has "-8" as central element and ones as 8-neighbours (3×3 8-neighbour Laplace kernel). The result is an image that has positive values at the edges of foreground objects. At smooth parts of the range image, there is only a small response due to noise. This step is followed by the detection of zero crossings, which we guide by the local gradient directions that have already been computed in a previous section (Equation 3). For each pixel with a positive value of the second derivative the two neighbours along the local gradient direction are checked. If one of them is zero or negative, the positive pixel is kept as edge pixel. Pixels with negative values of the second derivative are replaced by zero and a histogram of the positive values is created.

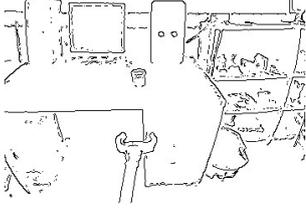


Figure 7. Step edges.

As described in the previous section, the location of the first peak is detected and three times its value is used as threshold. Finally, single-standing pixels that have no further pixel in its 8-neighbourhood are removed since they stem from noise (Figure 7).

3.6. Roof Edges

For each range value r_{ij} of the previously smoothed range image the associated 3D point or position vector $\vec{r}_{ij} = (x_{ij} \ y_{ij} \ z_{ij})^T$ is computed using the sensor's projection matrix. The coordinates of each 3D point are stored in individual arrays of the same size as the initial range image and at the same array cell position as its associated range value. In this way the initial neighbourhood is maintained. For each 3D point we compute the surface normal (Figure 8, left) based on its 3x3 neighbourhood:

$$\vec{v}_{ij,1} = \vec{r}_{i+1,j+1} - \vec{r}_{i-1,j-1} \quad (10)$$

$$\vec{v}_{ij,2} = \vec{r}_{i+1,j-1} - \vec{r}_{i-1,j+1} \quad (11)$$

$$\vec{n}_{ij,1} = \vec{v}_1 \times \vec{v}_2 \quad (12)$$

$$\vec{n}_{ij} = \frac{\vec{n}_{ij,1}}{|\vec{n}_{ij,1}|} \quad (13)$$

Surface normal computation is done for all images of the pyramid. We decide if we use the surface normal from a higher resolution or propagate one from the next-lower resolution based on the values of the local planarity measure presented in Section 3.3. If the value for a pixel at a higher resolution is below the value of its associated pixel at the lower resolution, the surface normal is propagated. We start at the lowest resolution and proceed upwards. A similar approach was proposed by Oehler et al. in [11], however, with a focus on 3D point clouds.

In the next step the dot product is computed between horizontal and vertical pairs of neighbouring surface normals. The resulting two images are thresholded. We use 15° as minimum local change of the angle between neighbouring normals to be accepted as roof edge. The roof edges from the thresh-

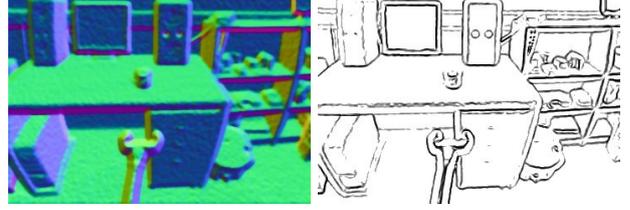


Figure 8. *Left:* Surface normals. *Right:* Roof edges.

olding results for the horizontal and vertical dot products are combined (Figure 8, right).

3.7. Planar Patches

The surface normals are multiplied (dot product) with the unit vector of the vertical axis. As stated earlier, the information about the vertical direction has to be supplied from outside, either from the known geometry of the setup or from inclination sensors. The result of the dot product is thresholded. We allow a deviation of 10° from the vertical axis (horizontal plane) for the surface normals of horizontal (vertical) planar structures.

3.8. Label Map

Mixed pixels, step edges, roof edges, vertical and horizontal planar patches have so far been stored in individual maps that have the same size as the initial range image from which they were derived. To each pixel of the range image we assign a label according to the local feature type. If more than one feature type has activation at a pixel location, a prioritization is applied: Mixed pixels, then step edges, roof edges, vertical and finally horizontal planar patches. An example for the co-occurrence of two features are roof and step edges; the former also occur at depth discontinuities. The final result is a label map (Figure 10, right column).

4. Experimental Results

The following two subsections describe the sensor used for data acquisition and the data itself as well as practical results achieved on that test data.

4.1. Test Data Acquisition

A tilting 2D laser range finder (Figure 9, left) built from a SICK LMS 100-10000 scanner and a SCHUNK PW 70 rotary tilt unit was used to capture test data. Each captured frame provides 360x500 range and intensity measurements. With an angular resolution of 0.25° horizontally and 0.125° vertically, the field of view is $90^\circ(\text{H}) \times 62.5^\circ(\text{V})$. The sen-



Figure 9. Tilting laser range scanner for capturing the test data.

Resolution	Point count	CPU time (ms)
360x500	180,000	83.3
360x250	90,000	41.5
250x160	40,000	20.5
180x125	22,500	10.1

Table 1. Computation times at different resolutions

scanner was mounted onto a mobile robot at a height of about 125cm with respect to the ground (Figure 9, right). The top of the vertical field of view is parallel to the ground plane, its bottom is tilted downwards by 62.5° . This configuration allows scanning table scenes as well as detecting obstacle directly in front of the robot and up to the robot’s height. One 3D scan takes about 20 seconds. In order to simulate a frame rate of about 10Hz, a stop-motion technique was applied. That is, after each scan the robot and dynamic objects in the scene were moved by a small distance or angle according to the simulated speed and frame rate. The data consists of 14 sequences with a total of 2,136 frames. The recorded sequences address robotic tasks such as obstacle detection, self localisation, object detection, and grasping.

4.2. Results

Our approach was implemented in C++ and tested on an Intel Core i5-430M notebook (2.24GHz, 4GB RAM) running 32bit OpenSUSE Linux 11.2. No optimizations such as SSE or multi-threading have been incorporating yet. The total amount of memory allocated for various buffers and lookup tables is slightly less than 3MB. The segmentation processing chain was applied to the recorded sequences at four different resolutions. Tab. 1 provides the respective average processing times per frame. The computational complexity is $O(n)$ with n being the number of pixels of the range image.

Figure 10 shows range images and associated label maps for three tasks a domestic service robots might have: grasping a cup on a table (top row), opening or closing a door (mid row), and detect-

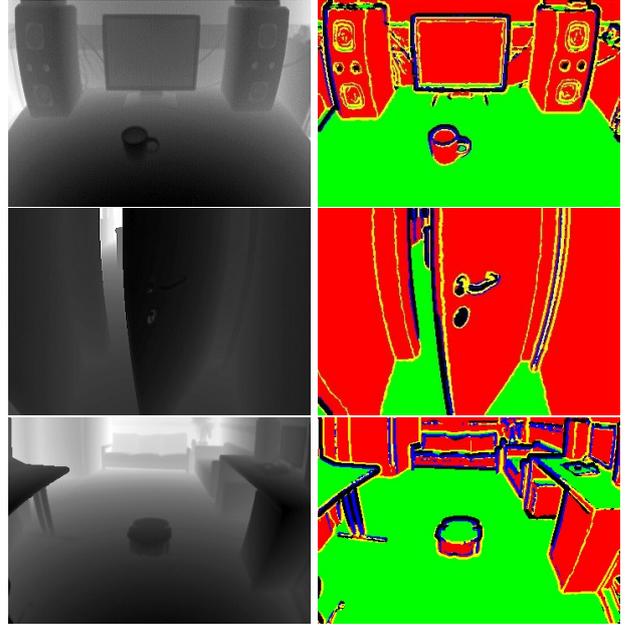


Figure 10. Range images (left column) and associated label maps for three scenes: objects on a table (top row), door handle (middle row), obstacles on the ground (bottom row). In each label map mixed pixels are black, step and roof edges are blue and yellow, and horizontal and vertical structures are green and red.

ing the closest obstacles within a relevant height region for obstacle avoidance (bottom row). Figure 11 shows applications of the extracted features in the context of service robotics.

5. Conclusion

In this paper we have presented a fast approach to segmentation of dense range images. In the first part of our approach, after noise reduction and the generation of a range image pyramid, mixed pixels at depth discontinuities are detected and masked out. The second part extracts step edges, roof edges as well as planar horizontal and vertical structures. Finally, the pixels of the range images are labelled according to the local feature type. We demonstrated that our approach is capable of real-time processing of range images on mainstream notebook hardware.

Since the test data used in this paper only stemmed from a tilting laser range finder, future work will encompass adapting the approach to range images from other sources, especially the recently introduced and highly popular Kinect. Apart from an adaption to the different sensor model, speedups will be necessary to still achieve real-time performance on higher resolution data (VGA resolution). Finally, a thorough evaluation of our approach against state-of-the-art meth-

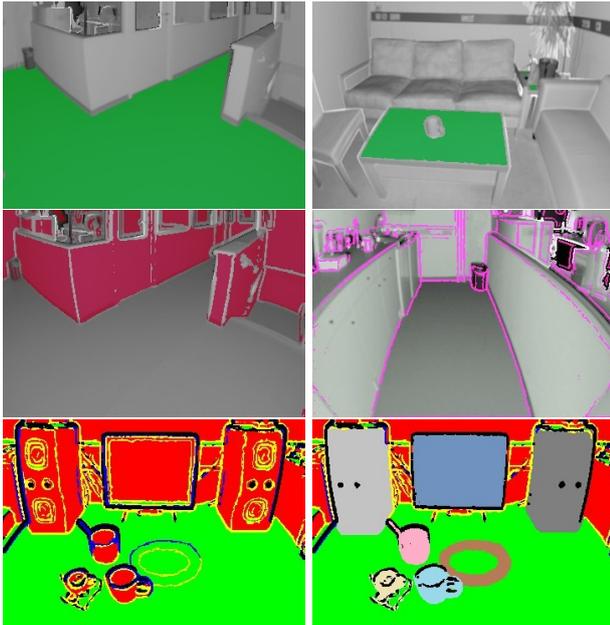


Figure 11. Examples for applications of the extracted features. Horizontal planar patches for obstacle detection (floor) and a table scene (table top), vertical planar patches (walls) for self localisation and edges reflecting transition between and boundaries of patches. Identifying object candidates by removing the table plane, projecting the resulting single-standing point clouds onto a grid in that plane and clustering them.

ods but also within the intended robotic use cases has to be conducted.

Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n^o248623.

References

- [1] O. Bellon and L. Silva. New improvements to range image segmentation by edge detection. *IEEE Signal Processing Letters*, 9(2):43–45, February 2002. 2
- [2] Y. Chang and X. Li. Adaptive image region-growing. *IEEE Transactions on Image Processing*, 3(6):868–872, November 1994. 2
- [3] P. Gotardo, O. Bellon, and L. Silva. Range image segmentation by surface extraction using an improved robust estimator. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings*, pages II: 33–38, 2003. 2
- [4] M. Haindl and P. Žid. Multimodal range image segmentation. In O. Goro and D. Ashish, editors, *Vision Systems, Segmentation and Pattern Recognition*. I-Tech Education and Publishing, Vienna, Austria, 2007. 2
- [5] F. Han, Z. Tu, and S. Zhu. Range image segmentation by an effective jump-diffusion method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1138–1153, September 2004. 2
- [6] A. Harati, S. Gaechter, and R. Siegwart. Fast range image segmentation for indoor 3d-slam. In *Proceedings of the 6th IFAC Symposium on Intelligent Autonomous Vehicles*, September 2007. 2, 4
- [7] S. Hojjatoleslami and J. Kittler. Region growing: A new approach. *IEEE Transactions on Image Processing*, 7(7):1079–1084, July 1998. 2
- [8] A. Hoover, G. Jean-Baptiste, X. Jiang, P. Flynn, H. Bunke, D. Goldgof, K. Bowyer, D. Eggert, A. Fitzgibbon, and R. Fisher. An experimental comparison of range image segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:673–689, July 1996. 2
- [9] X. Jiang and H. Bunke. Fast segmentation of range images into planar regions by scan line grouping. In *Machine Vision and Applications*, pages 115–122, 1994. 2
- [10] K.-M. Lee, P. Meer, and R.-H. Park. Robust adaptive segmentation of range images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:200–205, 1998. 2
- [11] B. Oehler, J. Stückler, J. Welle, D. Schulz, and S. Behnke. Efficient multi-resolution plane segmentation of 3d point clouds. In *4th International Conference on Intelligent Robotics and Applications*, pages 145–156, Aachen, Germany, December 2011. 6
- [12] P. Palmer, H. Dabis, and J. Kittler. A performance measure for boundary detection algorithms. *Computer Vision and Image Understanding*, 63(3):476–494, May 1996. 2
- [13] A. Sappa and M. Devy. Fast range image segmentation by an edge detection strategy. *International Conference on 3D Digital Imaging and Modeling*, 0:292–299, 2001. 2
- [14] K. Varadarajan and M. Vincze. Real-time depth diffusion for 3d surface reconstruction. In *IEEE International Conference on Image Processing*, pages 4149–4152, Hong Kong, September 2010. 3
- [15] H. Wang and D. Suter. Mdpe: A very robust estimator for model fitting and range image segmentation. *International Journal of Computer Vision*, 59:139–166, 2004. 2
- [16] J. Weingarten, G. Gruener, and A. Dorf. Probabilistic plane fitting in 3d and an application to robotic mapping. In *IEEE International Conference on Robotics and Automation*, pages 927–932, 2004. 2
- [17] X. Zhang and D. Zhao. Range image segmentation via edges and critical points. In *Proceedings of SPIE*, volume 2501, pages 1626–1637, 1995. 2