Efficient 3D Vertex Detection in Range Images Acquired with a Laser Sensor

Dimitrios Katsoulas and Lothar Bergen

Institute for Pattern Recognition and Image Processing,
Computer Science Department,
Albert-Ludwigs-Universität Freiburg, Germany
dkats@informatik.uni-freiburg.de,
http://lmb.informatik.uni-freiburg.de

Abstract. In many branches of industry, piled box-like objects have to be recognized, grasped and transferred. Unfortunately, existing systems only deal with the most simple configurations (i.e. neatly placed boxes) effectively. It is known that the detection of 3D-vertices is a crucial step towards the solution of the problem, since they reveal essential information about the location of the boxes in space. In this paper we present a technique based on edge detection and robust line fitting, which efficiently detects 3D-vertices. Combining this technique with the advantages of a time of flight laser sensor for data acquisition, we obtain a fast system which can operate in adverse environments independently of lighting conditions.

1 Introduction

This paper addresses the depalletizing problem (or bin picking problem) in the context of which a number of objects of arbitrary dimensions, texture and type must be automatically located, grasped and transferred from a pallet (a rectangular platform), on which they reside, to a specific point defined by the user. The need for a robust and generic automated depalletizing system stems primarily from the car and food industries. An automated system for depalletizing is of great importance because it undertakes a task that is very monotonous, strenuous and sometimes quite dangerous for humans. More specifically, we address the construction of a depalletizer dealing with cluttered configurations of boxes as shown in Fig. 1.

Existing systems can be classified as follows: systems incorporating no vision at all and systems incorporating vision. The latter group can be further divided into systems based on intensity or range data. For an in depth discussion of existing systems the reader is referred to [7], where the superiority of systems employing range imagery is discussed as well.

One of the fastest and conceptually closest systems to the one we propose, is the one of Chen and Kak [2]. A structured light range sensor is used for data acquisition. A region based technique is used to segment the range image into surfaces. Since a completely visible vertex in the scene provides the strongest

B. Radig and S. Florczyk (Eds.): DAGM 2001, LNCS 2191, pp. 116–123, 2001.

Fig. 1. Intensity image

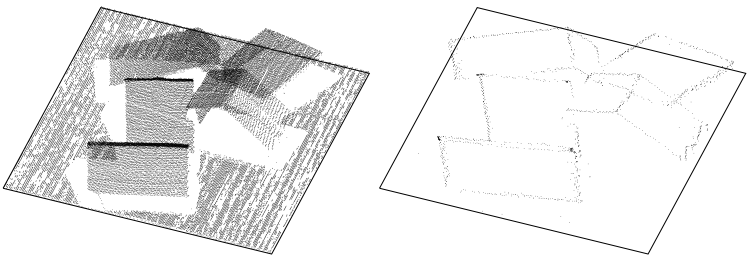
constraints for calculating hypotheses about the position of an object in the scene, vertices are detected by intersecting surfaces acquired from the segmentation. The main disadvantage of this approach is that since a vertex is computed as the intersection of three surfaces, the objects should expose three surfaces to the imaging source. In many cases this is not true, which results in a small number of localized objects per scan. In the extreme case of neatly placed objects of the same height, for example, no vertices can be detected at all. A second disadvantage originates from the fact that, according to [5], the region based segmentation methods of range data are time consuming operations. A third drawback, common to almost all the systems in the literature, is the usage of structured light techniques for range data acquisition. Since it is desirable for the projected light to be the only illumination source for the scene, they certainly perform better than camera based systems, but they still can not be used in uncontrolled industrial environments. Furthermore, they require time-consuming calibration.

This paper, in which a sub part of a depalletizing system is described, is organized as follows.

Initially, we describe the scene data acquisition module, which is capable of operating in a variety of environmental conditions, even in complete darkness. The description of a fast and accurate edge detector for object border extraction based on scan line approximation follows. The presentation of an algorithm for detecting vertices in a fast and robust manner is then given. This technique requires only two lines for vertex detection (as opposed to three surfaces) and therefore provides richer information about the objects and, as will be seen, at a low computational cost. Finally, a paragraph summarizing the system's advantages and outlining future work concludes.

2 Data Acquisition

The system comprises an industrial robot, namely the model KR 15/2 manufactured by KUKA GmbH, a square vacuum-gripper and a time of flight laser sensor (model LMS200, manufactured by SICK GmbH). The output of the laser sensor is a set of two-dimensional points, which are defined as the intersection of the objects with the sensor's scanning plane [4]. This set of planar points acquired from the laser sensor will be hereinafter referred to as scan line. The sensor is integrated on the gripper, and the latter is seamlessly attached to the robot's flange, as in [7]. In this way, we take full advantage of the flexibility for viewpoint selection made possible by a six degrees of freedom robot.



(a) Range image (b) Edge map

Fig. 2. Jumbled boxes

Due to the fact that we plan the robot to grasp objects in a top-to-bottom approach, we perform a linear scanning of the upper part of the pallet. The robot is programmed to execute a linear movement, the end points of which are the mid points of the two opposite sides of the rectangular pallet. The scanning plane of the sensor is always perpendicular to the direction of the movement and the sensor faces the upper side of the pallet. The set of scan-lines collected during the movement of the sensor is our 2.5D image.

One of the most salient problems of range imagery based on time of flight laser sensors, is the occurrence of noisy points caused by range or reflectance changes [1]. In order to discard these noisy points, before further processing, we apply the noise attenuation method described in [1] to the scan-lines of our range image. An acquired range image with attenuated noise is depicted in Fig. 2(a).

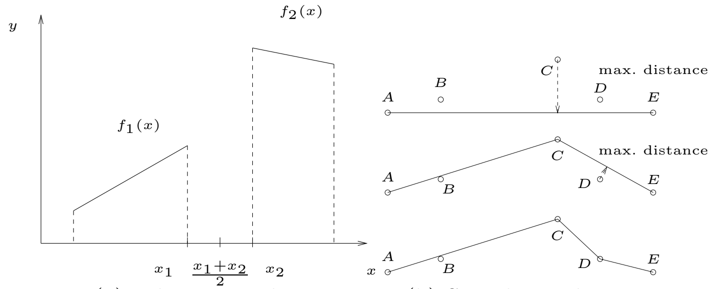
3 Edge Detection

The adopted scan line approximation algorithm, proposed in [6], splits a scan line into segments which are later approximated by model functions. With the help of these functions, edge strength values are computed for each pair of end-points of neighboring segments. To illustrate, the edge strength value definition for the two segments, approximated with the functions $f_1(x)$ and $f_2(x)$, which a linear in our case (Fig. 3(a)), but could be curves in general, is defined as

$$\text{JumpEdgeStrength} = |f_1(\overline{x}) - f_2(\overline{x})|,$$

$$\text{CreaseEdgeStrength} = \cos^{-1} \frac{(-f_1'(\overline{x}), 1)(-f_2'(\overline{x}), 1)^T}{\|(-f_1'(\overline{x}), 1)\| \|(-f_2'(\overline{x}), 1)\|},$$

where $\overline{x} = \frac{x_1+x_2}{2}$. The scan-line splitting technique [3] is illustrated in Fig. 3(b). If we suppose that our scan line comprises the points with labels A-E, a linear segment is initially estimated from the end points and the maximum distance of the line to every point of the scan line is calculated. If no point has a distance greater than a predetermined threshold T_{split} from the approximation curve, then the process stops for the particular scan line segment, since it is satisfactorily approximated. If the maximum distance is bigger than T_{split}, the whole process is repeated recursively (e.g. for the scan-lines AC and CE). This process is illustrated in Fig. 3(b). The approximating functions are defined on the segments originating from the splitting by means of a least square fitting.



(a) Edge strength (b) Scan line splitting

Fig. 3. Edge detection

The splitting algorithm is controlled by only one parameter T_{split}. Scan-line over-segmentation problems are solved when this threshold is increased. However, the arbitrary increase of T_{split} produces an under-segmentation phenomenon. In order not to lose edge information, we set a low value in the threshold and we applied the fast segment-merging technique suggested in [10]. This is where our detector differs from [6]. In each of the fitted segments, a significance measure (SM) is assigned as

$$SM = \frac{L}{e},$$

where L is the length of the segment and e is the approximation error of the least square fitting. The significance measure is based on a pseudo-psychological measure of perceptual significance, since it favors long primitives, provided they fit reasonably well. According to the merging procedure each segment is combined sequentially with the previous, the next and both the previous and next segments and for each combination a SM is computed. If one of the combinations results in a bigger SM than the one of the candidate segment, the corresponding segments are merged.

An initial version of this algorithm applied on scan-lines is presented in [7], where a discussion concerning the algorithm's accuracy is also presented. The application of the algorithm to the range data shown in Fig. 2(a) is depicted in Fig. 2(b). The edge map was created by applying the detector to every row and column of the range image and by retaining the edge points with high jump edge strength values and crease edge strength values around 90 degrees.The time needed for the creation of the edge map in the input range image comprising 424 rows and 183 columns was about 2.8 seconds on an Intel Pentium III, 600MHz processor.

4 3D Vertex Detection

As discussed in the introduction, the use of lines as low-level features provides us with rich information about the scene at a low computational cost. In the following section we shall demonstrate how line segments can be fitted robustly to the edge data. We then point out how a-priori knowledge about the objects can be used to obtain vertices from the line segments.

4.1 Line Detection in 3D

For the line fitting we have chosen the Hough transform (HT) for it's robustness with respect to noisy, extraneous or incomplete data. Unfortunately, the computational complexity and memory requirements of the standard Hough transform (SHT) rise exponentially with the number of parameters under detection. Therefore, the SHT is unsuitable for 3D-line detection involving four parameters.

An attempt to reduce the number of parameters by applying the SHT to the projection of edge data to the planes defined by the coordinate axis proved unsuccessful due to the high line density.

The probabilistic Hough transforms (PHTs) (cf. [9]) offer another approach to reduce the computational complexity by calculating and accumulating feature parameters from sampled points. While this reduces the computational complexity of the accumulation to some extent, the complexity of the peak detection and the memory requirements remain unchanged.

Leavers developed the dynamic generalized Hough transform (DGHT) (cf. [8]), a technique which allows for the use of one-dimensional accumulators if a coarse segmentation of the objects can be obtained and if the features under detection can be suitably parameterized. This technique, which belongs to the group of PHTs, reduces the memory requirements from R^n to nR, where R is the resolution of the parameters space and n the number of parameters. Inspired by the DGHT, we have developed a technique for 3D-line fitting, which will be described in the remainder of this section.

Initially, the connected components (CCs) of the projection of the edges onto the ground plane are detected and sorted according to their size. Line fitting is performed by iterating three steps: coarse segmentation of the edge points using the biggest CC as seed, robust line parameter estimation and removal of edge points in the vicinity of the robustly estimated line. In the coarse segmentation step a 3D-line is fitted in the least square sense to the points of the CC. Points within a certain distance of this line are used as input to robust parameter estimation. In the robust parameter estimation step the parameters of the line consisting of the direction vector $\boldsymbol{d}$ and the starting point $\boldsymbol{p}$ are calculated. Pairs of points are sampled, their difference vectors normalized and the components of the normalized vectors accumulated in three one-dimensional accumulators. The maxima of the three accumulators yield the components of the direction vector $\boldsymbol{n}$. The data points are then projected on to the plain through the origin whose normal vector is $\boldsymbol{d}$. The 2D coordinates of the data points in the plane are then accumulated in two one-dimensional accumulators and provide us with the point $\boldsymbol{p}$.

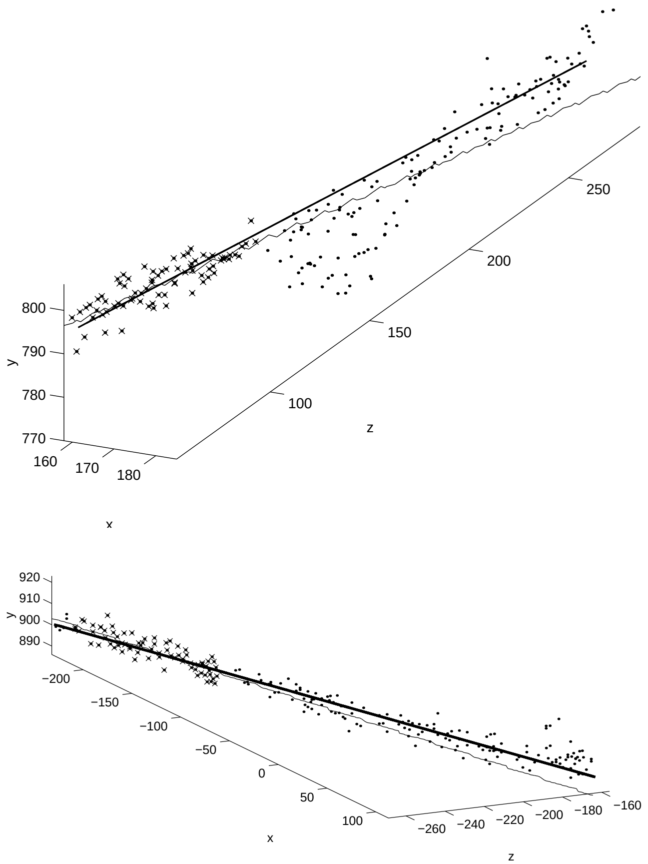


Fig. 4. 3D-line fitting

Figure 4 shows the results of our robust fitting technique for two segments: each CC is represented with crosses, the edge points with dots, the result of the least square fitting with a thin line and the final result with a thick line.

Due to the sparseness of the edge data in 3D-space, the precision of the lines fitted to the CC is largely sufficient for the segmentation of the edge data. The robustness with respect to outliers and the precision required for vertex detection is guaranteed by the parameter estimation with the one-dimensional accumulators and is clearly illustrated in Fig. 4.

4.2 Vertex Reconstruction

3D-vertices can now be obtained from close lines forming near 90 degree angles. The vertices can then be used in a hypothesis generation and verification step (cf. the system described in [2]) to accurately localize the boxes.

5 Conclusion

We have presented an efficient technique for the detection of 3D-vertices in range images. The joint use of edge detection and a technique inspired by the dynamic generalized Hough transform renders this technique fast and robust. This technique will be part of a real-time depalletizer for cardboard boxes. Due to the fact that vertices are obtained as the intersection of two lines (as opposed to the intersection of three planes), a rich set of hypotheses about the location of the boxes is generated. This increases the probability of grasping several objects per scan. Moreover, if the dimensions of the boxes are known it may be possible to achieve complete scene understanding. Another key feature of the depalletizer is the use of a laser sensor, which allows the system to be used in adverse environments and independent of the lighting conditions.

In the future, we plan to complete the construction of the system by adding a box location hypothesis generation and verification step.

Acknowledgments. This work was supported for the first author by the German ministry of economy and technology under the PRO INNO program, Grant No. KF 0185401KLF0. We also wish to thank Nikos Canterakis for the fruitful discussions.

References

1. M. D. Adams. Amplitude modulated optical range data analysis in mobile robotics. In Lisa Werner, Robert; O'Conner, editor, *Proceedings of the 1993 IEEE International Conference on Robotics and Automation: Volume 2*, pages 8–13, Atlanta, GE, May 1993. IEEE Computer Society Press.
2. C. H. Chen and A. C. Kak. A robot vision system for recognizing 3-D objects in low-order polynomial time. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1535–1563, November-December 1989.

3. Richard Duda and Peter Hart. *Pattern Recognition and Scene Analysis.* John Wiley and Sons, 1973.
4. SICK GmbH. *Measurement Software Tool (MST) technical description.*
5. A. Hoover. An experimental comparison of range image segmentation algorithm. *IEEE Trans. PAMI, 18(7):637–689, 1996.*, 1996.
6. Xiaoyi Jiang and Horst Bunke. Edge detection in range images based on scan line approximation. *Computer Vision and Image Understanding: CVIU*, 73(2):183–199, February 1999.
7. D.K. Katsoulas and D.I. Kosmopoulos. An efficient depalletizing system based on 2D range imagery. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA-2001), Seoul, Korea*, pages 305–312. IEEE Robotics and Automation Society, May 21–26 2001.
8. V. F. Leavers. *Shape Detection in Computer Vision Using the Hough Transform.* Springer-Verlag, London, 1992.
9. V. F. Leavers. Which Hough transform? *CVGIP: Image understanding*, 58(2):250–264, 1993.
10. P. Rosin and G. West. Nonparametric segmentation of curves into various representations. *IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 17, No. 12, 1995.*, 1995.