

EXTENDING GENERALIZED HOUGH TRANSFORM TO DETECT 3D OBJECTS IN LASER RANGE DATA

Kourosh Khoshelham

Optical and Laser Remote Sensing Research Group, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands
k.khoshelham@tudelft.nl

KEY WORDS: Laser scanning, Point cloud, Object Recognition, 3D Generalized Hough Transform, Automation

ABSTRACT:

Automated detection and 3D modelling of objects in laser range data is of great importance in many applications. Existing approaches to object detection in range data are limited to either 2.5D data (e.g. range images) or simple objects with a parametric form (e.g. spheres). This paper describes a new approach to the detection of 3D objects with arbitrary shapes in a point cloud. We present an extension of the generalized Hough transform to 3D data, which can be used to detect instances of an object model in laser range data, independent of the scale and orientation of the object. We also discuss the computational complexity of the method and provide cost-reduction strategies that can be employed to improve the efficiency of the method.

1. INTRODUCTION

Automated extraction of objects from laser range data is of great importance in a wide range of applications. Reverse engineering, 3D visualisation, industrial design monitoring and environmental planning are a few examples of the applications that require 3D models of objects extracted from images or laser range data. A 3D model provides an abstract description of the object, which can be processed and visualised more easily and efficiently. The process of object extraction consists of two main tasks. The first task is detection, in which the presence of an object in the data is verified, and its approximate location is found (usually by labeling the data points that belong to the object). The second task is modeling, where the detected object is represented with a 3D geometric model that is most adequate in terms of such criteria as accuracy, compactness, the domain of the object and the application requirements. The detection step plays a key role in the successful modeling of the object. If the object is properly detected in the data, the modeling can be carried out more reliably and accurately.

Existing approaches to the detection of objects in range data can be divided into two major categories: data-driven approaches and model-driven approaches. Data-driven approaches are mainly based on segmentation (Khoshelham, 2006; Rottensteiner and Briese, 2003; Sithole, 2005), clustering (Filin, 2002; Vosselman, 1999) and classification (Forlani et al., 2006; Oude Elberink and Maas, 2000). While these methods have been commonly applied to the laser range data of 2.5D surfaces, their application to more complex 3D scenes is not always possible. For instance, in laser range data of industrial installations many objects are partially occluded and data-driven methods fail to correctly detect these objects in the data. Model-driven approaches, on the contrary, are more robust in the presence of partial occlusion, since they incorporate some form of knowledge about the shape of the object. The object model can be represented, among other

representations, as a set of voxel templates (Greenspan and Boulanger, 1999) or spin images (Johnson and Hebert, 1999), which are matched against the data or as a set of parameters that mathematically define the object. In the latter case, Hough transform (Duda and Hart, 1972; Hough, 1962) has been used to determine the model parameters as well as the data points that belong to the object (Olson, 2001).

The application of Hough transform is restricted to simple objects that can be represented with few parameters, such as planes, spheres and cylinders. Vosselman et al., (2004) describe a Hough-based method for the detection of planes and spheres in a point cloud. Rabbani (2006) developed an extension of this method that can be used for the detection of cylinders. Figure 1 demonstrates the application of Hough transform to the detection of cylinders in a point cloud. As can be seen, the curved parts joining the cylinders have not been extracted because these parts cannot be expressed in parametric forms with few parameters.

This paper concentrates on the detection of 3D objects with arbitrary shapes in a point cloud. The objective of this paper is to develop a new extension of Hough transform, which can be used to detect instances of a complex object model in laser range data, independent of the scale and orientation of the object.

The paper has five sections. Section 2 provides an overview of the standard and generalized Hough transform as applied to 2D images. In section 3, the principles of the 3D generalized Hough transform is described. A discussion on the computational complexity of the method is presented in section 4. Conclusions appear in section 5.

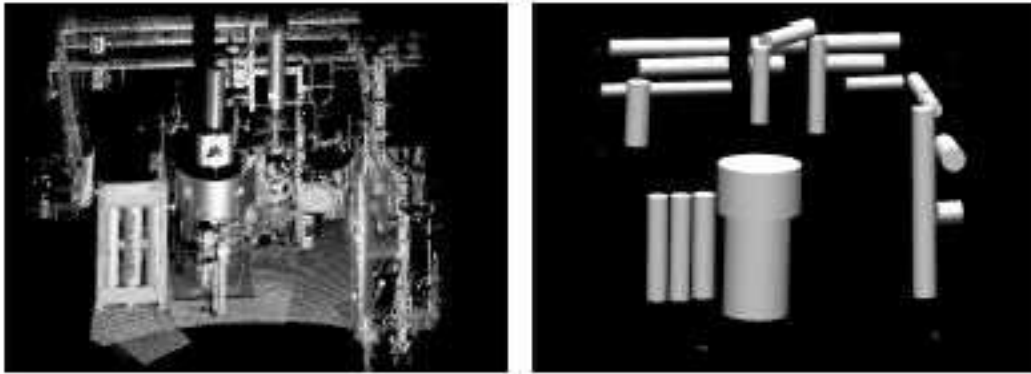


Figure 1. Detection of cylinders in a point cloud using Hough transform (from Rabbani (2006)). The curved parts joining the cylinders cannot be extracted using this method.

2. AN OVERVIEW OF THE STANDARD AND GENERALIZED HOUGH TRANSFORM

Hough transform is a well known method for the detection of objects in 2D intensity images. The standard Hough transform is applicable to objects with an analytical shape such as straight lines, circles and ellipses; whereas, with the generalized Hough transform any arbitrary curve can be detected in a 2D image. The following sections briefly describe the standard and generalized Hough transform.

2.1 The standard Hough transform

The idea of Hough transform for detecting straight lines in images was first introduced by Hough (1962). In the original Hough transform, a straight line is parameterized as $y = mx + b$ with two parameters m and b . According to the number of parameters, a 2D parameter space is formed in which every point in the image space corresponds to a line $b = -xm + y$. A set of image points that lie on a same line $y = mx + b$ in image space correspond to a number of lines in the parameter space, which intersect at point (m, b) . Finding this intersection point is, therefore, the basis for line detection in Hough transform. The parameter space is realized in the form of a discrete accumulator array consisting of a number of bins that receive votes from edge pixels in the image space. The intersection point is determined by finding the bin that receives a maximum number of votes.

In addition to straight lines, Hough transform has been used to detect also other analytical shapes, such as circles and ellipses, in 2D images. The underlying principle for the detection of other analytical shapes is the same as for the straight line detection, and is based on constructing a duality between edge pixels in the image and object parameters in the parameter space. The dimensions of the parameter space, however, vary with respect to the parameterization of the object.

2.2 The generalized Hough transform

Ballard (1981) proposed a generalization of Hough transform to detect non-parametric objects with arbitrary shapes in 2D intensity images. In the generalized Hough transform, the object model is stored in a so-called R-table format. An arbitrary reference point is selected for the object, and for every pixel on the object boundary the gradient direction as well as the length and direction of a vector connecting the boundary pixel to the reference point are computed (Figure 2). The gradient directions, φ , serve as indices in the R-table to look up the

length, r , and direction, β , of the connecting vectors. Table 1 illustrates a general form of an R-table.

Table 1: R-table

Point	φ	\mathbf{r}
0	0	$(r, \beta)_{01} - (r, \beta)_{02} - (r, \beta)_{03} - \dots$
1	$\Delta\varphi$	$(r, \beta)_{11} - (r, \beta)_{12} - (r, \beta)_{13} - \dots$
2	$2\Delta\varphi$	$(r, \beta)_{21} - (r, \beta)_{22} - (r, \beta)_{23} - \dots$
...

The reconstruction of the object model from the R-table is straightforward:

$$\begin{cases} x_p = x_c - r \cdot \cos(\beta) \\ y_p = y_c - r \cdot \sin(\beta) \end{cases} \quad (1)$$

where (x_c, y_c) and (x_p, y_p) are respectively the coordinates of the reference point and a point on the boundary of the object. For the detection of the object model in the image, however, the coordinates of the reference point are not known. A 2D accumulator array is, therefore, constructed with the two parameters of the reference point as the axes. At every image edge pixel the gradient direction is obtained and then looked up in the R-table. The corresponding sets of r and β values are used to evaluate Equation 1, and the resulting x_c and y_c values indicate the accumulator array bins that should receive a vote. Once this process is complete for all edge pixels, the bin with the maximum vote indicates the reference point, and the edge pixels that cast vote for this bin belong to an instance of the object in the image.

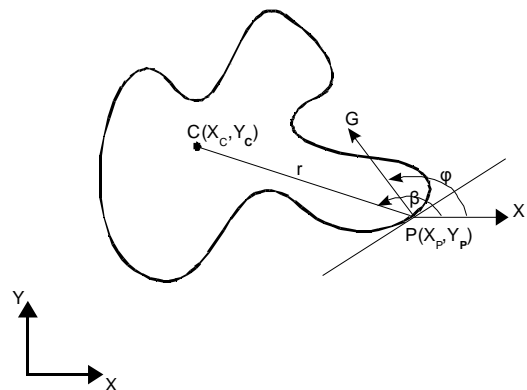


Figure 2: Parameters involved in the generalized Hough transform.

The generalized Hough transform can also be used to detect a rotated and scaled version of a model in an image. This is achieved by supplementing Equation 1 with a scale factor and a rotation angle, and the parameter space is expanded to a 4D accumulator array. The peak of the accumulator array determines the scale and rotation parameters in addition to the coordinates of the reference point, although at the price of a higher computational expense.

2.3 Modifications to Hough transform

Several modified variations of the Hough transform have been proposed to improve the performance of the method. Illingworth and Kittler (1988) provide a survey of these methods. Duda and Hart (1972) suggested a modification of the standard Hough transform by substituting the original slope-intercept parameterization of straight lines with a polar, angle-radius, parameterization. The polar parameterization leads to a bounded parameter space, unlike the original parameterization, and is, consequently, more computationally efficient. They also showed that standard Hough transform can be used to detect more general curves in an image. Gradient weighted Hough transform, as appears in Ballard's generalization, was first introduced by O'Gorman and Clowes (1976). The derivation of edge orientation information imposes very little computational cost, but greatly increases the efficiency of the method. Other methods that have been shown to improve the performance of Hough transform include Adaptive Hough transform (Illingworth and Kittler, 1987), Hierarchical Hough transform (Princen et al., 1990), and Randomized Hough transform (Xu et al., 1990).

3. EXTENSION OF GENERALIZED HOUGH TRANSFORM TO 3D DATA

In this section we present an extension of the generalized Hough transform to 3D data. The method will be referred to as 3D GHT in the subsequent parts of the paper. The 3D GHT follows the same principle as generalized Hough transform as outlined in Section 2.2. The main difference is that the gradient vector is replaced with a surface normal vector. The normal vectors can be obtained by triangulating the surface of the object or by fitting planar surfaces to small sets of points in a local neighbourhood. Vectors connecting each triangle to an arbitrary reference point are stored in the R-table as a function of the normal vector coordinates. A normal vector is constrained to be of unit length and is, therefore, defined by two orientation angles, φ and ψ , as depicted in Figure 3. A connecting vector is defined by two orientation angles, α and β , as well as its length r . These parameters can be derived from the coordinates of the reference point and the object boundary point:

$$\begin{aligned} r &= [(x_p - x_c)^2 + (y_p - y_c)^2 + (z_p - z_c)^2]^{1/2} \\ \alpha &= \arccos\left(\frac{z_c - z_p}{r}\right) \\ \beta &= \arccos\left(\frac{x_c - x_p}{r \sin(\alpha)}\right) \end{aligned} \quad (2)$$

This formulation results in a 2D R-table where all the connecting vectors, \mathbf{r} , are stored in cells whose coordinates are the orientation angles of the normal vectors. Figure 4 demonstrates how such a 2D R-table is constructed.

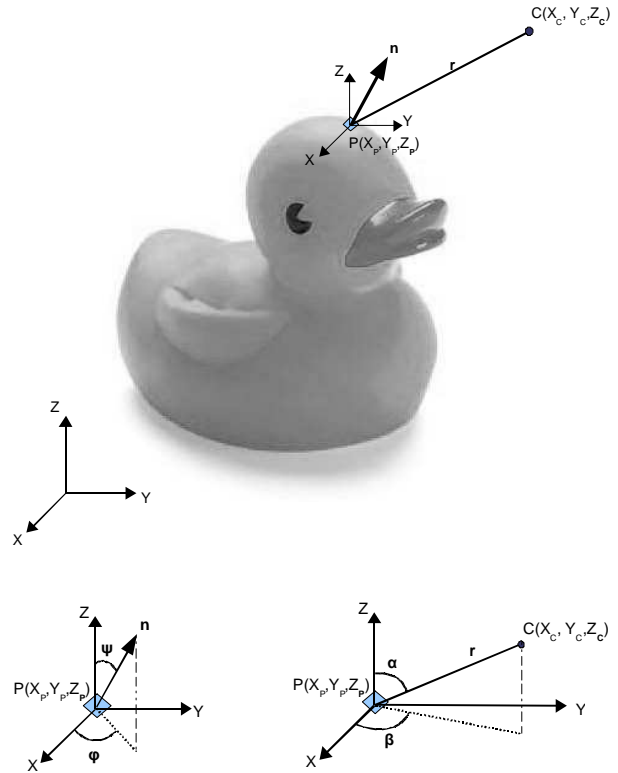


Figure 3: Parameters involved in the 3D GHT method.

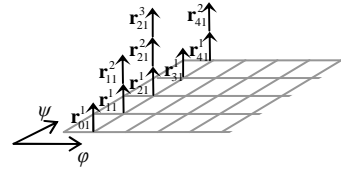


Figure 4: Storing \mathbf{r} vectors in a 2D R-table.

The reconstruction of the object model from the R-table is carried out by extending Equation 1 to 3D:

$$\begin{cases} x_p = x_c - r \cdot \sin(\alpha) \cos(\beta) \\ y_p = y_c - r \cdot \sin(\alpha) \sin(\beta) \\ z_p = z_c - r \cdot \cos(\alpha) \end{cases} \quad (3)$$

where α and β denote the orientation angles of the vector that connects a point p to the reference point c . For the detection of the 3D object model in a point cloud the three coordinates of the reference point are unknown parameters. Thus, the equations given in (3) are rearranged so as to express the unknown parameters as functions of the known variables:

$$\begin{cases} x_c = x_p + r \cdot \sin(\alpha) \cos(\beta) \\ y_c = y_p + r \cdot \sin(\alpha) \sin(\beta) \\ z_c = z_p + r \cdot \cos(\alpha) \end{cases} \quad (4)$$

Having obtained the object model in the form of the R-table, an algorithm for the detection of instances of this model in a point cloud can be outlined as follows:

1. Construct a 3D accumulator array with the three parameters of the reference point as the axes;

2. Compute the normal vector for every point in the point cloud and look up r vectors at coordinates (φ, ψ) of the 2D R-table;
3. Evaluate Equation (4) with the corresponding sets of r , α and β values to obtain x_c , y_c and z_c ;
4. Cast a vote (an increment) to the accumulator array bin corresponding to each set of x_c , y_c and z_c values;
5. Repeat the voting process for all the points in the point cloud;
6. The bin with the maximum vote indicates the reference point, and the 3D points that cast vote for this bin belong to an instance of the object in the point cloud.

In practice, the object appears in range data with an arbitrary rotation and scale. To account for the additional rotation and scale parameters, Equation (4) is modified as:

$$\mathbf{c} = \mathbf{p} + s\mathbf{M}_z \cdot \mathbf{M}_y \cdot \mathbf{M}_x \cdot \mathbf{r} \quad (5)$$

where $\mathbf{c} = (x_c, y_c, z_c)^T$, $\mathbf{p} = (x_p, y_p, z_p)^T$, $\mathbf{r} = (r \sin(\alpha) \cos(\beta), r \sin(\alpha) \sin(\beta), r \cos(\alpha))^T$ s is a scale factor and \mathbf{M}_x , \mathbf{M}_y and \mathbf{M}_z are rotation matrices around x , y and z axis respectively. The incorporation of a scale factor and three rotation parameters results in an expansion of the Hough space to seven dimensions. To evaluate Equation 5 and cast votes for the accumulator bins, a 4D space circumventing the entire range of scale factors and rotation angles must be exhausted. This implies that the crude application of the 3D GHT method to object detection can be very expensive. Therefore, cost-reduction strategies such as adaptive, hierarchical and randomized voting schemes are of great importance in the 3D GHT algorithm.

4. IMPLEMENTATION ASPECTS

The 3D GHT method as described in Section 3 is computationally expensive when the object appears in data with an arbitrary scale and rotation with respect to the model. The development of a cost-reduction strategy is thus the main challenge in the application of 3D GHT. In general, the execution time of Hough transform is more dominated by the voting process rather than by the search for a peak in the accumulator. In the absence of arbitrary scale and rotation, the number of required operations in the voting process is $O(M)$, where M is the number of points in the dataset. Thus, a desirable cost-reduction strategy must aim to reduce the number of points that are involved in the voting process. Randomized (Xu et al., 1990) and probabilistic (Kiryati et al., 1991) variations of the Hough transform work based on a random selection of a small number of data points, and are, therefore, suitable options for controlling the computational cost of the voting process..

In the presence of arbitrary scale and rotation, a 4D subset of the parameter space circumventing the entire range of scale factors and rotation angles is exhausted during the voting process. Consequently, the number of operations required in the voting process is $O(M \cdot N^4)$, where N is the number of intervals along each axis of the accumulator array. Clearly, a desirable cost-reduction strategy in this case must concentrate on the N^4 factor. The adaptive Hough transform (Illingworth and Kittler, 1987) reduces the number of intervals along axes since it begins with a coarse-resolution parameter space and increases the resolution only in the vicinity of the peak. The randomized Hough transform (Xu et al., 1990) also provides an efficient

strategy to reduce the number of bins that receive votes in the parameter space. In the randomized voting, instead of working with one point at a time, a number of points sufficient for the computation of all parameters are selected from the data. Once all the parameters are computed, only one bin in the accumulator array receives a vote. In the case of a 3D object with seven parameters, a set of three points must be selected from the data at each time. These points along with their respective \mathbf{r} vectors form nine equations of the form given in Equation 5, which can be solved for the seven parameters. Thus, for each randomly selected set only one vote is cast for a bin in the 7D accumulator array.

5. CONCLUSIONS

In this paper we presented an extension of the generalized Hough transform to detect arbitrary 3D objects in laser range data. The procedure of storing a 3D model in a 2D R-table was demonstrated, and a method for the detection of instances of the model in a point cloud, based on a voting process, was described. It was discussed that the voting process can be computationally expensive in the case that the object appears in data with an arbitrary scale and rotation with respect to the model. The employment of a voting process based on the randomized Hough transform was, therefore, suggested to reduce the computational cost of the method.

REFERENCES

- Ballard, D.H., 1981. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2): 111-122.
- Duda, R.O. and Hart, P.E., 1972. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15: 11-15.
- Filin, S., 2002. Surface clustering from airborne laser scanning data, *Proceedings of the Photogrammetric Computer Vision, ISPRS Commission III Symposium*, Graz, Austria, pp. 119-124.
- Forlani, G., Nardinocchi, C., Scaioni, M. and Zingaretti, P., 2006. Complete classification of raw LIDAR data and 3D reconstruction of buildings. *Pattern Analysis & Applications*, 8(4): 357-374.
- Greenspan, M. and Boulanger, P., 1999. Efficient and reliable template set matching for 3D object recognition, *2nd International Conference on 3-D Digital Imaging and Modeling*, Ottawa, Canada, pp. 230-239.
- Hough, P.V.C., 1962. Methods and means for recognizing complex patterns, US patent 3069654.
- Illingworth, J. and Kittler, J., 1987. The adaptive Hough transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5): 690-698.
- Illingworth, J. and Kittler, J., 1988. A survey of the Hough transform. *Computer Vision, Graphics and Image Processing*, 44: 87-116.
- Johnson, A.E. and Hebert, M., 1999. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5): 433-449.
- Khoshelham, K., 2006. Automated 3D modelling of buildings in suburban areas based on integration of image and height data, *International Workshop on 3D Geoinformation (3DGeoInfo '06)*, Kuala Lumpur, Malaysia, pp. 381-393.
- Kiryati, N., Eldar, Y. and Bruckstein, A.M., 1991. A probabilistic Hough transform. *Pattern Recognition*, 24(4): 303-316.

- O'Gorman, F. and Clowes, M.B., 1976. Finding picture edges through collinearity of feature points. *IEEE Transactions on Computers*, 25(4): 449-456.
- Olson, C.F., 2001. A general method for geomtric feature matching and model extraction. *International Journal of Computer Vision*, 45(1): 39-54.
- Oude Elberink, S. and Maas, H.G., 2000. The use of anisotropic height texture measures for the segmentation of airborne laser scanner data, *International Archives of Photogrammetry and Remote Sensing*, 33 (Part B3/2), pp. 678-684.
- Princen, J., Illingworth, J. and Kittler, J., 1990. A hierarchical approach to line extraction based on the Hough transform. *Computer Vision, Graphics and Image Processing*, 52(1): 57-77.
- Rabbani Shah, T., 2006. Automatic reconstruction of industrial installations using point clouds and images PhD thesis Thesis, Delft University of Technology, Delft, 154 pp.
- Rottensteiner, F. and Briese, C., 2003. Automatic generation of building models from Lidar data and the integration of aerial images, *International Archives of Photogrammetry and Remote Sensing, Volume XXXIV/3-W13*, Dresden, pp. 170-184.
- Sithole, G., 2005. Segmentation and classification of airborne laser scanner data. PhD Thesis, Delft University of Technology, Delft, 185 pp.
- Vosselman, G., 1999. Building reconstruction using planar faces in very high density height data, *ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery*, Munich, pp. 87-92.
- Vosselman, G., Gorte, B.G., Sithole, G. and Rabbani Shah, T., 2004. Recognising structure in laser scanner point cloud, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 46, Part 8/W2, Freiburg, Germany, pp. 33-38.
- Xu, L., Oja, E. and Kultanen, P., 1990. A new curve detection method: randomized Hough transform (RHT). *Pattern Recognition Letters*, 11(5): 331-338.