Model-based segmentation and recognition from range data

Jan Boehm
Institute for Photogrammetry
Universität Stuttgart
Germany

Keywords: range image, segmentation, object recognition, CAD

ABSTRACT

This paper aims at developing a model-based system for the object recognition of three-dimensional objects with curved surfaces using range images. The model data is represented using a CAD-model, providing a mathematical precise and reliable description of arbitrary shapes. The proposed method is based on model-based range image segmentation, using curvature as invariant features. By integrating model information into the segmentation stage, the segmentation process is guided to provide a partitioning corresponding to that of the CAD-model. The work provides a way to detect objects in arbitrary positions and derive the transformation onto a CAD-model. Thereby it contributes to the development of automated systems in the areas of inspection, manufacturing and robotics.

1. INTRODUCTION

The availability of highly capable three-dimensional sensor systems creates the need for adequate systems for automated processing of three-dimensional data. Contrary to conventional intensity images, which contain information on the intensity of an objects surface, range images store the distance of each point on a surface to the sensor. In order to process these data on a high level of abstraction, new concepts need to be developed, exceeding two-dimensional image analysis. In the context of developing automated three-dimensional computer vision systems, this paper brings two aspects into focus: partitioning the image data and establishing a correspondence in-between image- and model-data.

The stage of partitioning the sensed data, known as segmentation is crucial to the image understanding process. We employ a model-based strategy, where curvature information extracted from the model is used as a priori information for a classification of input pixels, which are grouped into regions.

The features extracted during segmentation stage are brought into correspondence with the features of the model using a tree-search strategy. The number of possible pairings is reduced due to the fact that model information has been utilized during segmentation. After successful pairing, the three-dimensional transformation of the model onto the object in the scene is computed, completing the object recognition process. An overview of the complete process is given in figure 1.

Our own range image data in the examples of this work were acquired with a triangulation sensor using a coded light approach. The system is described in Brenner et al. (1999).[1] However the application of the proposed method is not limited to this sensor type but can be adapted to any range acquisition device or measurement system, such as described in Boehm et al. (2001).[2]

2. SEGMENTATION

Segmentation is the process of partitioning an image into meaningful entities. What exactly meaningful means, depends on the specific application. In the context of object recognition segmentation is typically used as feature extraction. It shall therefore be robust and derive a partitioning sufficiently similar to that of the model. When range images are used it is common to compute a grouping of the pixels into regions, where each region represents a surface of the object.

Many segmentation algorithms have been proposed in the literature, which work solely data-driven, see for example Hoover et al. (1996)[4] and Powell et al. (1998).[5] These approaches do not need any a priori knowledge

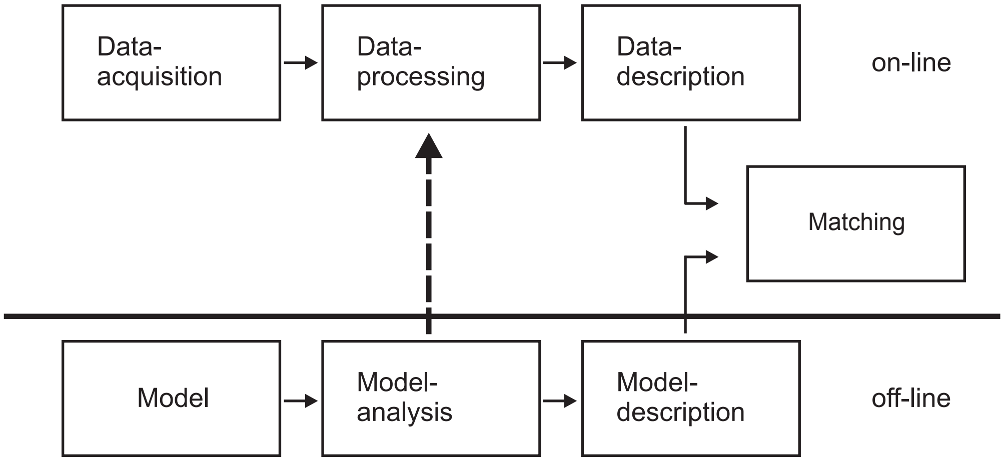


Figure 1. Overview of the object recognition process in reference to Arman and Aggarwal (1993).[3] Information of the model analysis is incorporated into the data processing.

of the objects shape. These approaches can be successful in cases where the partitioning of the boundary representation of an object is unique. This is particularly true for polyhedral objects. Also for objects which have sharp edges at the surface boundaries or for objects with a limited variety of surface types, such as scenes solely composed of cylinders, data driven approaches can be employed.

However in the general case the partitioning of the boundary representation of an object is not unique (see for example Mortenson (1997)[6]). For such a case a purely data driven approach is less suitable. Especially when segmentation is used as a feature extraction procedure for object recognition purposes, it is crucial that the partitioning of the range data coincides with that of the model. The information of the specific partitioning, which has been chosen for the model, is stored in its CAD model. It is therefore desirable to incorporate this information into the segmentation process.

Since at the segmentation stage no relation is established in-between the coordinate system of the model and that of the sensed data, no information of the model can be used which depends on the chosen coordinate system. Features invariant to orientation and translation have to be used. This leads to the use of features based on differential geometry.

We follow the work of Besl (1988)[7] for using mean and Gaussian curvature to describe the local properties of a surface. We compute the mean curvature as

$$H = \frac{(1 + f_y^2) f_{xx} + (1 + f_x^2) f_{yy} - 2 f_x f_y f_{xy}}{2(f_x^2 + f_y^2 + 1)^{\frac{3}{2}}}$$

and the Gaussian curvature as

$$K = \frac{f_{xx} f_{yy} - f_{xy}^2}{(f_x^2 + f_y^2 + 1)^2}$$

from the first and second derivatives of the surface function f. The surface function is a second order polynomial which is fit to the range data in a least squares estimation process.

While Besl proposed to use the curvature information to classify each pixel of a range image into one of eight fundamental surface types, we extend the scheme to classify the pixels according to the model information. To extract this information we use the CAD system Pro/Engineer to extract the curvature information for each surface patch of the CAD model. This information is stored in a table and is used to define the classes for the classification process. The interface to the CAD system and the output for the displayed part are given in figure 2.

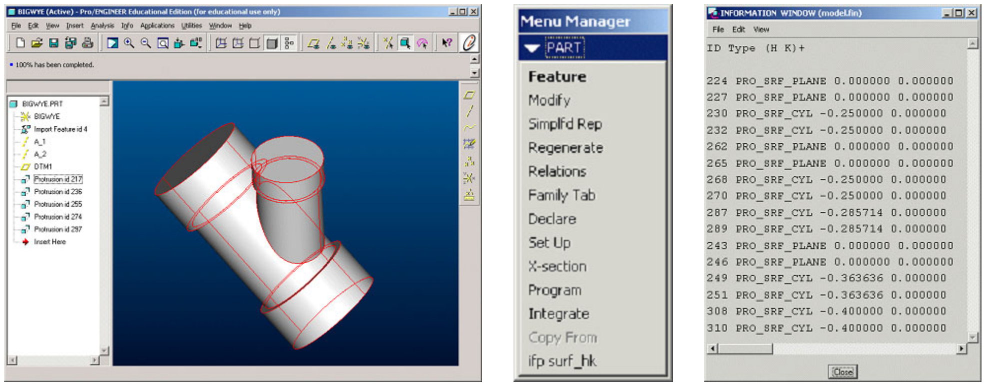


Figure 2. The interface to the Pro/Engineer CAD system and the corresponding curvature output.

Surface type	H	K
Plane	−0.0000	0.0000
Cylinder	−0.0200	0.0000
Sphere	−0.0200	0.0004

Table 1. The table that holds the curvature information of the test scene.

We use a simple nearest neighbor classification scheme in order to assign each pixel to its corresponding class. Pixels which have no corresponding class are assigned the background label. After initial classification each pixel is either labeled according to its corresponding surface or remains unlabeled.

2.1. Test Scene

The proposed algorithm is tested on a simple test scene. The scene consists of a plane, a cylinder and a sphere, three basic surface types commonly encountered. An image of the test scene is given in figure 3. The model curvature information is given in the table 1. The information of this table graphed in the two-dimensional feature space spanned by mean and Gaussian curvature corresponds well to the curvature histogram of the acquired range data as is shown in figure 4.

The results show a considerable amount of points which were misclassified, i.e. were assigned a wrong label or were not assigned a label but should have. These misclassifications are caused by false curvature estimation. Especially on the cylinder it becomes evident that not all of the pixels could be classified correctly. These classification errors have to be removed in a refined processing step. The most dominant regions, i.e. regions above a certain size threshold, are selected as seed regions for a region growing process. Region growing is implemented as a morphological operation. A 3×3 mask is moved over the dataset. When a neighbor to the point of interest (the center of the mask) has a label assigned, the point of interest is checked for compatibility to that region. In case it is found to be compatible it is assigned the label of the corresponding region. If there are conflicting regions, i.e. there are different regions adjacent to the point of interest, the largest region is preferred. This is also the case if the center pixel is already labeled. The results of the pure classification and of the refined processing are shown in figure 5.

2.2. SAMPL Test Data

Additional tests of the classification algorithm were made based on data from the SAMPL database maintained by P. Flynn.[8] The database provides CAD models and range images for various simple parts. Most of them contain only cylindrical and planar surfaces and have few surfaces (less than ten). The data set we present here

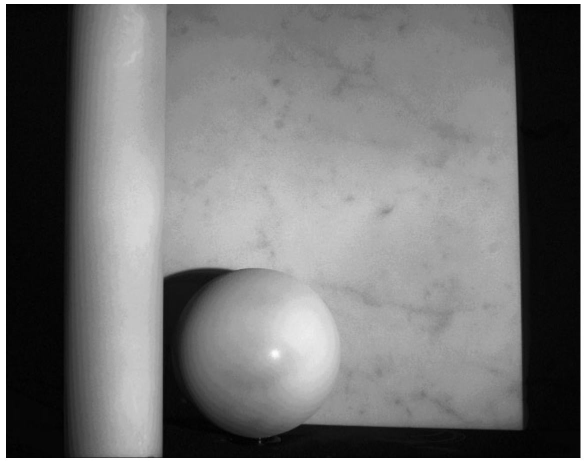
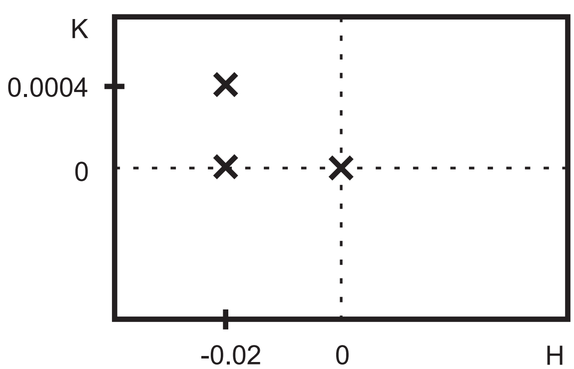
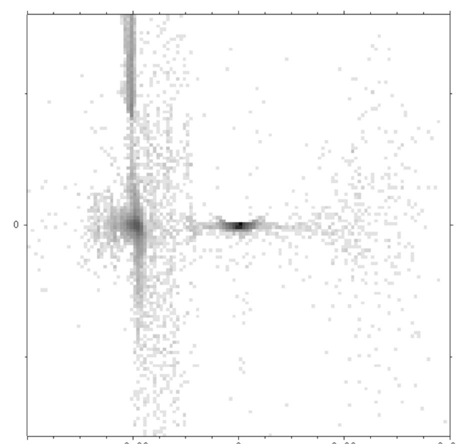


Figure 3. Intensity image of the test scene consisting of three basic surface types.

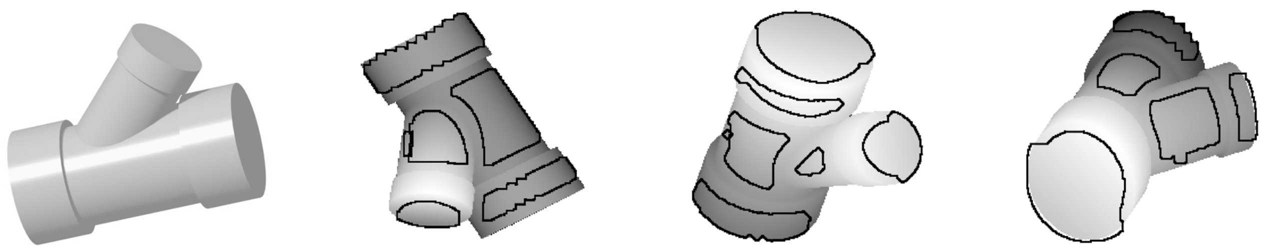




(a)

Figure 4. Left: the model curvature information as a two-dimensional plot. Right: the curvature histogram of the corresponding range image. The points of accumulation in the histogram correspond well to the points in the model's feature space.

Figure 5. The results of the classification (left) and of the refined processing (right).



(a) (b) (c) (d)

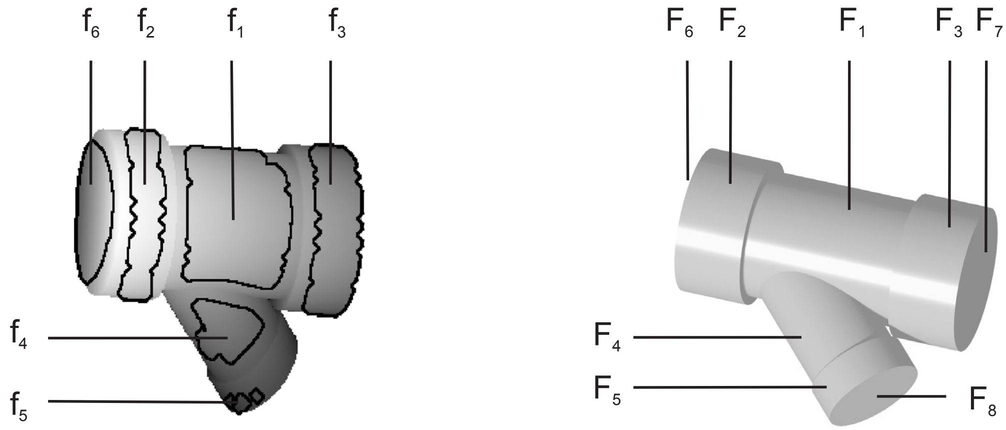
Figure 6. (a) The CAD model of "bigwye". (b)-(d) Segmentation results on the dataset. Segmentation of the surfaces is independent of the pose of the object.

contains five cylinders and the according planar surfaces. Figure 6(a) shows this dataset called bigwye. Figure 6(b)-(d) show the result of classification. The overall results show that the surfaces can be reliably detected independent of the pose of the object.

3. MATCHING

Once the surfaces have been segmented from the range data the correspondence of surfaces in the scene to surfaces of the model can be established. For this matching the set Ω of possible correspondences for each feature / region f_i of the range image is established. We note that since we know the curvature characteristics of each of the surfaces, we can eliminate those correspondences which connect surfaces of different curvature characteristics. Therefore the set Ω is typically smaller than the set of all features / surfaces F_i of the model. In figure 7 we show the results for one range image of the "bigwye" dataset. The size of the set of possible pairings ranges from one to three. This limits the size of the search space.

The final search for correspondences is carried out by a tree search algorithm following the principles of Grimson.[9] The tree search results in several hypotheses for the pose of the object in the scene. The final verification of the match is performed by computing the rigid body transformation of the centroid p_i of each region of the range image to the centroid P_i of the corresponding model surface. The coordinates of all centroids



$\Omega(f_1) = \{F_1\}$
$\Omega(f_2) = \{F_2, F_3\}$
$\Omega(f_3) = \{F_2, F_3\}$
$\Omega(f_4) = \{F_4\}$
$\Omega(f_5) = \{F_5\}$
$\Omega(f_6) = \{F_6, F_7, F_8\}$

Figure 7. Deriving the set of possible correspondences from the segmentation result. The segmentation of the dataset "bigwye" resulted in six regions f_1 to f_6. The CAD model consists of eight surfaces F_1 to F_8. The set of possible correspondences for each of the six regions are given.

are reduced by their mean value, which gives the centroids $p_i^{'}$ and $P_i^{'}$ respectively. The computation of the translation and rotation follows the approach of Horn.[10] The matrix

$$\mathbf{M} = \sum_{i=1}^{N} \mathbf{p}_i^{'} \mathbf{P}_i^{'T} = \begin{bmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{bmatrix}$$

gives the elements $S_{xx} = \sum_{i=1}^{N} p_{i_x}^{'} P_{i_x}^{'}$, $S_{xy} = \sum_{i=1}^{N} p_{i_x}^{'} P_{i_y}^{'}$ and so on.

These elements are used to form the symmetric 4×4 matrix

$$\mathbf{N} = \begin{bmatrix} (S_{xx} + S_{yy} + S_{zz}) & S_{yz} - S_{zy} & S_{zx} - S_{xz} & S_{xy} - S_{yx} \\ S_{yz} - S_{zy} & (S_{xx} - S_{yy} - S_{zz}) & S_{xy} + S_{yx} & S_{zx} + S_{xz} \\ S_{zx} - S_{xz} & S_{xy} + S_{yx} & (-S_{xx} + S_{yy} - S_{zz}) & S_{yz} + S_{zy} \\ S_{xy} - S_{yx} & S_{zx} + S_{xz} & S_{yz} + S_{zy} & (-S_{xx} - S_{yy} + S_{zz}) \end{bmatrix}$$

The Eigenvector of this matrix corresponding to the largest Eigenvalue can be interpreted as a quaternion which contains the desired rotation $\mathbf{R}$. The translation $\mathbf{t}$ is simply computed as

$$\mathbf{t} = \bar{\mathbf{p}} - \mathbf{R}\bar{\mathbf{P}}$$

The RMS of this estimation process is used as criterion to select the best match.

Figure 8 shows four instances of the "bigwye" dataset in different poses. For each pose the three best matches sorted by the RMS of the computed transformation are shown. We observe that the transformation computed with the smallest RMS corresponds to the correct pose for each of the cases.

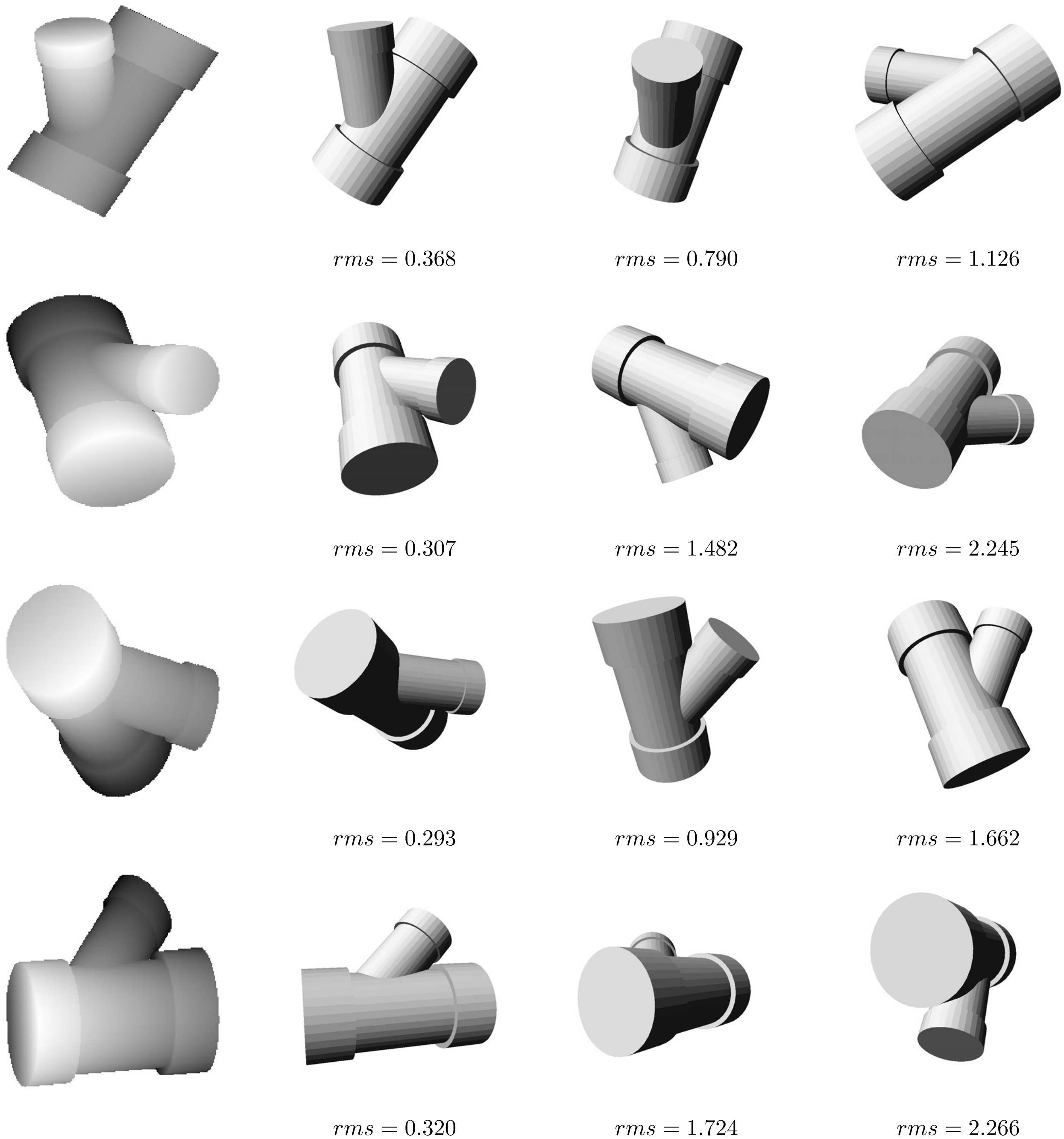


Figure 8. Result of four different range images of the "bigwye" dataset. Each row shows one pose of the object. The first column shows the range image. The second, third and forth column show the CAD model in the computed transformation. The transformations are sorted by their RMS value. The best transformation is in the second column.

4. SUMMARY

We have presented a method to detect objects in arbitrary positions and derive the transformation onto a corresponding CAD-model. The method employs an efficient technique for the model-based segmentation of range images. The proposed method is aimed at inspection and measurement task of industrial objects, but also has potential for a wide variety of applications in the areas of inspection, manufacturing and robotics. The topic of curvature estimation has not been mentioned within this paper, however it is still is a crucial part of the process and remains a topic of research.

REFERENCES

1. C. Brenner, J. Böhm, and J. Gühring, "Photogrammetric calibration and accuracy evaluation of a cross-pattern stripe projector," *Proc. SPIE Videometrics VI* **3641**, pp. 164–172, January 1999.
2. J. Böhm, J. Gühring, and C. Brenner, "Implementation and calibration of a multi sensor measuring system," in *Videometrics and Optical Methods for 3D Shape Measurement*, S. El-Hakim and A. Grün, eds., *Proceedings of SPIE* **4309**, pp. 154–161, SPIE, 2001.
3. F. Arman and J. K. Aggarwal, "Model-based object recognition in dense range images," *ACM Computing Surveys* **25**, pp. 5–43, March 1993.
4. 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**, pp. 673–689, July 1996.
5. M. W. Powell, K. W. Bowyer, X. Jiang, and H. Bunke, "Comparing curved-surface range image segmenters," in *International Conference on Computer Vision 1998*, pp. 286–291, 1998.
6. M. E. Mortenson, *Geometric Modeling*, John Wiley & Sons, 1997.
7. P. J. Besl, *Surfaces in Range Image Understanding*, Springer series in perception engineering, Springer-Verlag, 1988. Revised ed. of Ph.D. thesis, University of Michigan.
8. R. J. Campbell and P. J. Flynn, "A www-accessible database for 3d vision research," in *Proc. IEEE Workshop on Empirical Evaluation Methods in Computer Vision*, pp. 148–152, 1998.
9. W. E. L. Grimson, *Object Recognition by Computer*, MIT Press, 1990.
10. B. K. P. Horn, "Closed-form solution of absolut orientation using unit quaternions," *Journal of the Optical Society of America A* **4**, pp. 629–642, April 1987.