Segmentation of Range Data into Rigid Subsets using Surface Patches

A. P. Ashbrook, R. B. Fisher, C. Robertson and N. Werghi
Department of Artificial Intelligence
The University of Edinburgh
5, Forrest Hill, Edinburgh, EH1 2QL
anthonya@dai.ed.ac.uk

Abstract

In this paper we consider one aspect of the problem of automatically constructing geometric models of articulated objects from multiple range images. Automatic model construction has been investigated for rigid objects, but the techniques used do not extend easily to the articulated case. The problem arises because of the need to register surface measurements taken from different viewpoints into a common reference frame. Registration algorithms generally assume that an object does not change shape from one view to the next, but when automatically building a model of an articulated object, it is necessary for the modes of articulation to be present in the example data. To avoid this problem we propose that raw surface data of articulated objects is first segmented into rigid subsets, corresponding to rigid subcomponents of the object. This allows a model of each subcomponent to be constructed using the conventional approaches and a final, articulated model to be constructed by assembling each of the subcomponent models. We describe an algorithm developed to segment range data into rigid subsets based on surface patch correspondences and present some results for the planar patch case.

1 Introduction

The ability to automatically acquire geometric models from example objects is useful in a growing number of application areas. In the field of computer graphics, the need for improvements in realism requires more complex models, but manual model construction is time-consuming and difficult. In industrial settings it is useful to capture the geometry of existing parts either for the purpose of inspection or to enable exact replicas to be manufactured automatically. For reasons which will be clarified shortly, current techniques are generally limited to constructing models of single rigid objects. In this paper we suggest how these algorithms might be augmented to allow the automatic construction of articulated objects, increasing the scope of this technology. For clarity we define articulated objects as those objects consisting of a number of rigid parts that are connected by non-rigid joints [1].

The established approach for automatic model construction begins by taking surface measurements from a number of viewpoints so that all of the object's surface is captured. Typically, this will be done with a range finder such as a laser striper or stereo vision system. Either of two different approaches can then be taken. In the first approach surface primitives are fitted to the raw data in each of the views of the object, and then the different views are *registered* by aligning similar primitives [2]. The second approach registers the raw data initially, using local surface shape (see for example [3, 4]) and then surface primitives are fitted directly to the registered surface data [5]. The second approach is favoured because it makes maximum use of the raw data when surface fitting and avoids the problem of having to piece together possibly fragmented surface patches from different viewpoints that are not perfectly aligned.

Whichever approach is taken, the registration process assumes that the shape of the object does not change as the surface data is acquired. If, however, we wish to automatically capture the geometry and kinematics of an articulated object, then the object's shape must change from example to example. This means that current registration algorithms cannot be used directly. Rather than developing new registration algorithms we propose here that the raw measurement data is segmented into *rigid subsets* each corresponding to a rigid subcomponent of the object. This will enable models of each subcomponent to be constructed independently using existing technology and a final, articulated model to be formed by assembling each of the subcomponents.

The algorithm we have developed processes a pair of range images at a time and segments each of them into N sub-images, where N is the number of independently moving, rigid subcomponents which are present in the data. This processing is carried out in two distinct stages. In the first stage, the rigid transformation that aligns most of the data in the first image with corresponding data in the second is estimated. This is done by segmenting the range data into surface patches and then finding the transformation that aligns most of the patches in the first image with potential correspondents in the second. In the second stage, the movement of each surface patch between the two images is compared to the estimated transformation and labelled as belonging to the current rigid subcomponent if it is in agreement. These two stages are then iterated until no surface data remains unlabelled and the required segmentation is obtained.

We have already published a rigid part segmentation algorithm which utilises point-to-point as opposed to surface patch-to-surface patch correspondences between two range images [6, 7]. In this approach the transformation that aligns object subcomponents are estimated by aligning rigid coordinate frames defined at corresponding surface points. This approach is, however, limited to non-developable surfaces on which an invariant, rigid coordinate frame can be uniquely defined. This limitation has motivated the development of the surface patch based algorithm presented here but it is our intention to eventually combine the two approaches to produce a general purpose, rigid subcomponent segmentation algorithm.

In the next section a detailed description of our rigid part segmentation algorithm is presented. This is followed by a demonstration of the algorithm applied to some real range image data. Finally we summarise the contribution made by this work and briefly discuss the aims of ongoing and future work.

2 The Rigid Part Segmentation Algorithm

Given a pair of range images of an articulated object, $\mathcal{R}^a$ and $\mathcal{R}^b$, the goal is to segment the range data into subimages where each subimage represents a rigid subset of the data. We present an overview of our proposed segmentation algorithm here, followed by a more detailed description of the most important stages.

1. **Segment the range data into surface patches**

   Each of the range images, $\mathcal{R}^a$ and $\mathcal{R}^b$, is segmented into a set of surface patches, $\{\mathcal{S}^a\}$ and $\{\mathcal{S}^b\}$ respectively. Currently, only planar patches are used by later stages of the algorithm.

2. **For each surface patch in $\{\mathcal{S}^a\}$ find corresponding patches in $\{\mathcal{S}^b\}$**

   Geometric constraints may be used to form correspondences between surfaces patches. In the work presented here, however, we consider each planar patch in $\{\mathcal{S}^b\}$ to be a potential correspondence with each planar patch in $\{\mathcal{S}^a\}$.

3. **Determine constraints on the rigid transformations that align object subcomponents from surface patch correspondences**

   Each correct surface patch correspondence provides a constraint on the rigid transformation that aligns the object subcomponent that the surfaces belong to. Incorrect surface correspondences will provide erroneous constraints but these will be rejected later.

4. **Vote for each rigid transformation constraint in a discretized parameter space**

   A Hough transform is used to accumulate evidence for the transformations that align each of the object subcomponents. Surface patches belonging to the same subcomponent will provide constraints which intersect at the correct alignment transformation parameters, producing a peak in the Hough space. To account for the significance of larger patches, votes are weighted by the area of the smallest patch in each pair. The 6-parameter rigid, transformation space is partitioned into a 3-parameter rotation space and a 3-parameter translations space for efficiency.

5. **Surface patches are labelled, based on their contribution to the highest peak in the Hough transform**

   The highest peak in the Hough transform is isolated and then the transformation constraints contributing to the peak are identified. The surface patches associated with those constraints are then labelled as either *definitely*, or in some special cases (see Section 2.4) *possibly*, belonging to the current subcomponent.

6. **Steps 2 through to 5 are repeated until no more peaks can be found**

   The Hough transform is reconstructed using the surfaces without *definite* labels to find the next most significant subcomponent. This is repeated until no more peaks are found in the Hough transform.

7. **Surface patches labelled as *possibly* belonging to a subcomponent are confirmed**

   A verification stage is used to confirm whether surface patches labelled as *possibly* belonging to a particular subcomponent do in fact belong to that part. This is done by aligning each of these surfaces with the corresponding range image and looking for good agreement.

2.1 Surface Patch Segmentation

We have used a surface patch segmentation algorithm which segments the surface data into quadric patches using the local surface curvature. The full details of this algorithm can be found in the literature [8, 9, 10] but can be summarised as follows:

1. The range data is smoothed using edge preserving, diffusion smoothing.
2. The mean and Gaussian curvatures are calculated at each data point.
3. Seed surface patches are formed by grouping neighbouring pixels whose mean and Gaussian curvature have the same sign.
4. Seed patches are either eliminated or cleaned up using a morphological operator.
5. Quadric surfaces are fitted to the seed patches and pixels are regrouped according to the surface patch they are closest to. This is repeated until a stable segmentation is obtained.

In the future we intend to use generic quadric patches to segment range data into rigid subsets but we currently only use planar patches.

2.2 Calculating the Transformation Constraints

Each pair of corresponding surface patches provides a constraint, $\mathbf{T}_i$, on the parameters of a rigid transformation. For planar surface patches the constraint fixes 3 of the 6 degrees of freedom and can be defined as:

$$\mathbf{T}_i(\alpha, du, dv) = \mathbf{f}(\hat{\mathbf{n}}_m, \hat{\mathbf{n}}_n, d_m, d_n; \alpha, du, dv) \quad (1)$$

where $\hat{\mathbf{n}}_m$ and $\hat{\mathbf{n}}_n$ are the surface normals of the two patches and d_m and d_n are the corresponding perpendicular distances from the surfaces to the origin. The details of the function $\mathbf{f}$ are presented in Appendix A. The parameters α, du and dv represent the remaining degrees of freedom.

2.3 Representing *Noisy* Constraints

In the absence of measurement error, each of the constraints on the transformation of a particular subcomponent will intersect at the appropriate point in the parameter space. In practice, measurement errors are inevitable and unless these are accounted for in an appropriate manner, the robustness of the segmentation algorithm will suffer.

To ensure robust performance, measurement errors are propagated through the algorithm to determine the error on each of the constraints. If the errors on all of the plane parameters are represented by the covariance matrix Σ_{plane} (these errors are themselves calculated by error propagation) then the error on each constraint $\mathbf{T}_i$ can be approximated as:

$$\Sigma_i(\alpha, du, dv) = \nabla\mathbf{f}^T \Sigma_{plane} \nabla\mathbf{f} \quad (2)$$

where $\nabla\mathbf{f}$ is the matrix of partial derivatives of $\mathbf{f}$ with respect to the plane parameters, commonly known as the Jacobian matrix.

Having established the error on each constraint we can define a probability density function, $P_i(\mathbf{T}; \alpha, du, dv)$, to integrate into the Hough transform instead of making a single vote. In fact, it is the logarithm of P_i which is integrated into the Hough transform so as to turn a product of probabilities into a sum of independent terms. For convenience we make the simplifying assumption that the errors are Gaussian so that:

$$P_i(\mathbf{T}; \alpha, du, dv) \propto e^{-\chi^2} \quad (3)$$

where:

$$\chi^2 = [\mathbf{T}_i - \mathbf{T}]^T \Sigma_i [\mathbf{T}_i - \mathbf{T}] \quad (4)$$

This provides a probability for each cell in the Hough space, defined by $\mathbf{T}$, which is a function of the parameters α, du and dv. To evaluate this probability we choose parameter values which maximise P_i. The result is that each constraint is now *smeared* in the parameter space according to its sensitivity to measurement errors. It is worth noting the relationship between the constraint error and the size of the planar patches used to derive the constraint. For small patches the error on each of the plane parameters is relatively large which results in a more *smeared* constraint entry in the Hough transform. For large patches the constraint entry becomes much sharper as the constraint error is small.

2.4 Surface Patch Labelling

The highest peak in the Hough transform space provides an estimate, $\mathbf{U}_j$, of the transformation that aligns most of the surface data between the range images. Each of the constraints, $\mathbf{T}_i$, is considered to have contributed to the peak under the condition that:

$$\left([\mathbf{U}_j - \mathbf{T}_i]^T \Sigma_i [\mathbf{U}_j - \mathbf{T}_i]\right)^{\frac{1}{2}} \leq 3.0 \quad (5)$$

In other words, the constraint is within 3 standard deviations of the peak given the estimated error. Surface patches are then labelled as either *possibly* or *definitely* belonging to the current subcomponent if they are associated with any of the constraints contributing to the peak. The *possibly* label is used if the surface patch is perpendicular to the axis of rotation. This is a degenerate case and can only be resolved later by verification.

3 Experiments

Rigid part segmentation results are presented in this section for 2 sets of range data acquired using a laser striper. For both experiments the same algorithm parameters have been used. These are detailed in Table 1.

3.1 Experiment 1

Figures 1(a) and (b) present a pair of range images each of which contains 2 rigid subcomponents. The subcomponents are simple metallic blocks with planar surfaces. Figure 1(a) depicts the components in their original position and Figure 1(b) depicts the components after they have undergone a relative transformation. The dimensions of these range images are 165x182 pixels and 152x175 pixels respectively.

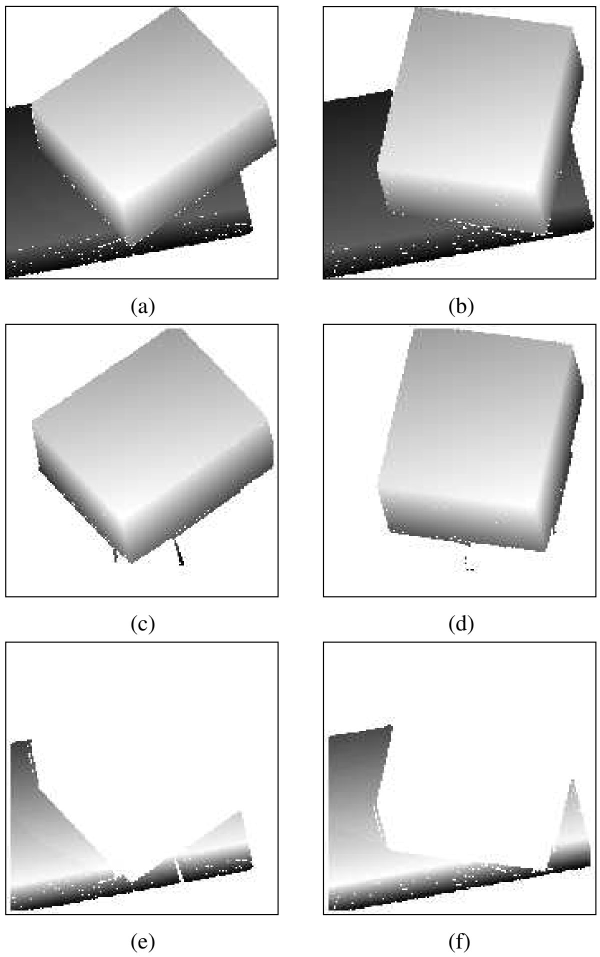

Figure 1. The range images used in experiment 1 and the segmentation results

The first pass of the rigid part segmentation algorithm took 213 seconds on a 200MHz Ultra Sparc for this data set. Figure 2 presents details of the rotation Hough transform. The lower of the 2 charts on the left presents the heights of each of the peaks detected in the rotation space in order of importance. The upper of these charts presents the positions in the rotation parameter space of the largest two peaks. The segmentation of the range data based on the largest of these peaks is presented in Figures 1(c) and (d).

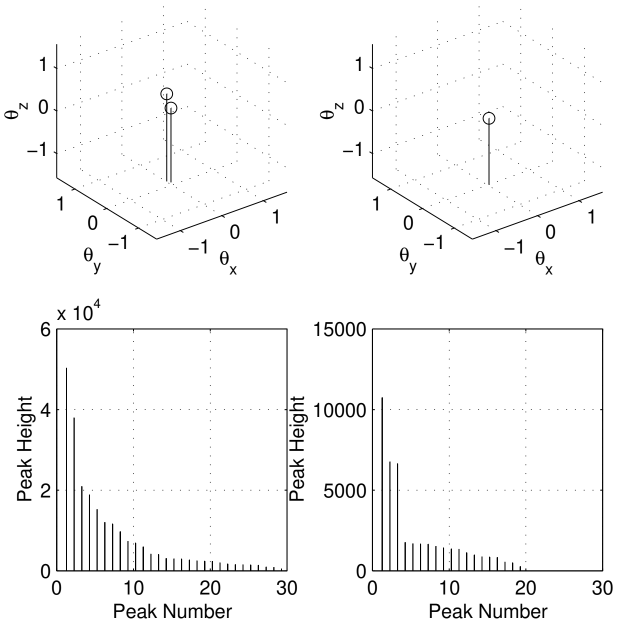

Figure 2. Details of the rotation hough transform constructed for the range data in Figure 1. The graphs on the left show the position of the first two peaks (top) and the height of all of the peaks found in order of significance (bottom) after the first pass of the algorithm. The graphs on the right show the peaks details after the second pass.

Having labelled patches belonging to the first rigid subcomponent the segmentation algorithm was iterated again, this time taking 74 seconds. The details of the rotation Hough transform peak for this pass are also presented in Figure 2. The height of this peak is now significantly smaller than in the first pass as much of the spurious Hough transform voting has been removed. The position of this peak has also moved, and is likely to give a better estimate of the second alignment transformation. The segmentation of the range data based on the largest of the new peaks is presented in Figures 1(e) and (f). No surface patches are left at this stage so the segmentation is terminated.

3.2 Experiment 2

Figures 3(a) and (b) presents a pair of range images each of which contains 2 rigid subcomponents. Figure 3(a) depicts the components in their original position and Figure 3(b) depicts the components after they have undergone a relative transformation. The dimensions of these range images are 198x114 pixels and 196x152 pixels respectively.

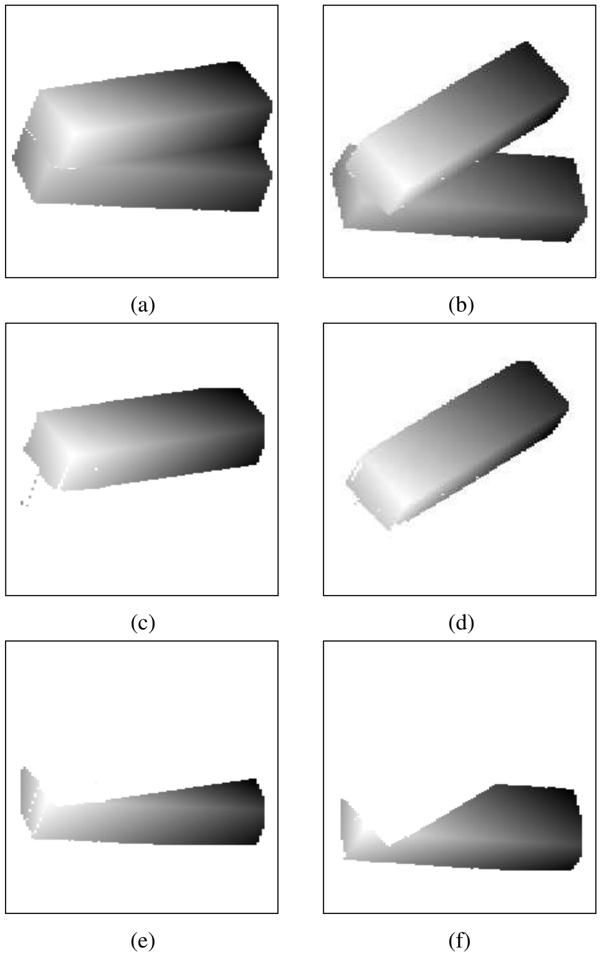

Figure 3. The range images used in experiment 2 and the segmentation results

The first pass of the rigid part segmentation algorithm took 242 seconds on a 200MHz Ultra Sparc for this data set. The segmentation of the range data based on the largest of these peaks is presented in Figures 3(c) and (e). Having removed the first rigid subcomponent the segmentation algorithm was iterated again, this time taking 74 seconds. The segmentation of the range data based on the largest of the new peaks is presented in Figure 3(d) and (f). One small surface patch is left at this stage so the segmentation is continued but no peaks are found and the process terminates.

Parameter	No. Bins	Full Range
Rotation H.T.	40x40x40	2πx2πx2π
Translation H.T.	20x20x20	100mmx100mmx100mm

Table 1. Algorithm parameters used in the two experiments

4 Conclusions and Future Work

As a first step in the process of automatically building geometric models of articulated objects from multiple range images, we have developed an algorithm that segments range images into rigid subsets. Evidence for the presence of rigid object subcomponents can be determined by measuring the transformations that align surface points and patches from one range image to another. Here only our new work on surface patches is reported but we have previously reported successful segmentation using point correspondences. Surface patches belonging to the same subcomponent can be identified as a set of surface pairings which undergo a similar transformation. The algorithm has been tested on real range data and shown to work effectively.

One advantage of this approach is that, in addition to performing the required rigid part segmentation, an estimate of the transformation that aligns each subcomponent between images is determined. This will provide a valuable initial estimate for performing accurate surface registration later on.

In this paper, the algorithm has only been tested on planar surface patches, but the approach can be extended to many different classes of surface. We are currently experimenting with cylindrical surfaces and intend to extend the algorithm to other degenerate surfaces, for example spheres. We have already published results on a rigid part segmentation algorithm which works on non-developable surfaces and intend to integrate these algorithms together to provide a general purpose tool.

Acknowledgements

The work presented in this paper was funded by a UK EPSRC grant GR/H86905.

A The Planar Patch Transformation Constraint

The alignment of a pair of planar patches, $\mathcal{S}_m^a$ and $\mathcal{S}_n^b$, provides 3 constraints on the 6 parameter rigid transformation, $\mathbf{T}_i$, where:

$$\mathbf{T}_i(\alpha, du, dv) = \mathbf{f}(\hat{\mathbf{n}}_m, \hat{\mathbf{n}}_n, d_m, d_n; \alpha, du, dv) \qquad (6)$$

In homogeneous coordinates this can be expressed as the product of a rotation, $\mathbf{r}_i$, and a translation, $\mathbf{t}_i$.

$$\mathbf{T}_i(\alpha, du, dv) = \mathbf{t}_i(du, dv)\mathbf{r}_i(\alpha) \qquad (7)$$

The rotation constraint is the set of rotations that align the surface patch normals, $\hat{\mathbf{n}}_m$ with $\hat{\mathbf{n}}_n$. This is determined by finding an arbitrary rotation that aligns $\hat{\mathbf{n}}_m$ with $\hat{\mathbf{n}}_n$ and then rotating around $\hat{\mathbf{n}}_n$. We can align $\hat{\mathbf{n}}_m$ with $\hat{\mathbf{n}}_n$ by rotating by π radians about the bisector, $\hat{\mathbf{n}}_b$, to $\hat{\mathbf{n}}_m$ and $\hat{\mathbf{n}}_n$. This gives the value of the constraint at $\alpha = 0$.

$$\mathbf{r}_i(0) = \mathbf{rot}(\hat{\mathbf{n}}_b, \pi) \qquad (8)$$

where $\mathbf{rot}(\hat{\mathbf{n}}, \alpha)$ describes the rotation around an axis defined by $\hat{\mathbf{n}}$ by α radians. A good reference for this can be found in [11]. The constraint is then found by rotating around $\hat{\mathbf{n}}_n$.

$$\mathbf{r}_i(\alpha) = \mathbf{rot}(\hat{\mathbf{n}}_n, \alpha)\mathbf{rot}(\hat{\mathbf{n}}_b, \pi) \qquad (9)$$

The translation that aligns a pair of planes is only constrained in the direction of the plane normals. Note that the first plane has now been rotated so that it is parallel with the second.

$$\mathbf{t}_i(0, 0) = \hat{\mathbf{n}}_n(d_n - d_m) \qquad (10)$$

We now define a pair of orthogonal vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$, which are mutually orthogonal to the plane normals $\hat{\mathbf{n}}_n$, such that:

$$\hat{\mathbf{u}} \times \hat{\mathbf{v}} = \hat{\mathbf{n}}_n \qquad (11)$$

The translation constraint is then given by:

$$\mathbf{t}_i(du, dv) = \hat{\mathbf{n}}_n(d_n - d_m) + \hat{\mathbf{u}}du + \hat{\mathbf{v}}dv \qquad (12)$$

References

[1] Kambhamettue, C. and Goldgof, D. B., "Nonrigid Motion Analysis", Handbook of Pattern Recognition and Image Processing: Computer Vision, Ed. Young, T. Y., Chapter 11, 1994.

[2] Orr, M. J. L., Hallam, J., Fisher, R. B., "Fusion through Interpretation", Proc. 2nd European Conf. on Computer Vision, (ed) G. Sandini, pp 801–805, St. Margherita Ligure, Italy, 1992.

[3] Besl, P. J. and McKay, N. D., "A Method for Registration of 3-D Shapes", IEEE PAMI, 14(2), pp239-256, 1992.

[4] Eggert, D., Fitzgibbon, A. W. and Fisher, R. B., "Simultaneous Registration of Multiple Range Views for use in Reverse Engineering", Proc. ICPR, pp243-247, 1996.

[5] Fisher, R. B., Fitzgibbon, A. W. and Eggert, D., "Extracting Surface Patches from Complete Range Descriptions", Proc. International Conference on Recent Advances in 3-D Digital Imaging and Modeling , Ottawa, 1997.

[6] Ashbrook, A. P. and Fisher, R. B., "Constructing Models of Articulating Objects: Range Data Partitioning", The International Conference on Recent Advances in 3-D Digital Imaging and Modeling , Ottawa, 1997.

[7] Ashbrook, A. P. and Fisher, R. B., "Segmentation of Range Data for the Automatic Construction of Models of Articulated Objects", The IEEE Non-rigid and Articulated Motion Workshop, Puerto Rico, 1997.

[8] Besl, P. J. and Jain, R. C., "Segmentation through variable-order surface fitting", IEEE PAMI, 10(2), pp167-192, 1988.

[9] Fitzgibbon, A. W., Eggert, D. W. and Fisher, R. B., "High-level CAD Model Acquisition from Range Images", Computer-Aided Design, Vol 29(4), pp321–330, 1997.

[10] A. Hoover, G. Jean-Baptiste, X. Jiang, P. J. Flynn, H. Bunke, D. Goldgof, K. Bowyer, D. Eggert, A. Fitzgibbon, R. Fisher, "An Experimental Comparison of Range Segmentation Algorithms", IEEE Trans. Pat. Anal. and Mach. Intel., Vol 18(7), pp673–689, July 1996.

[11] Paul, R. P. "Robot Manipulators: Mathematics, Programming, and Control", MIT Press, 1981.