Construction of Articulated Models from Range Data

A. P. Ashbrook, R. B. Fisher, N. Werghi and C. Robertson
The Division of Informatics
The Institute of Perception, Action and Behaviour
The University of Edinburgh
5, Forrest Hill, Edinburgh, EH1 2QL
anthonya@dai.ed.ac.uk

Abstract

In this paper we present an algorithm for automatically building models of articulated objects from range data. These models not only describe the surface shape of the object but also describe the kinematics that constrain the movement of one object component in relation to another. This is more difficult than building models of rigid objects because the association of surface measurements to object components must be determined. The algorithm is demonstrated on a difficult object with free-form surfaces.

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. Users of Computer-Aided Design technology would like to be able to make improvements to a manufactured part and then update their CAD model to reflect this. This provides a very effective design cycle. In an industrial production setting it is useful to compare the geometry of manufactured parts with models of the design so that flaws can be detected.

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. The problem then is to determine a rigid transformation for each viewpoint that maps all of the measurements from that viewpoint into a common coordinate frame. This is commonly known as the *registration* problem. Finally, the measurements are used to construct a surface representation. This might be a CAD model for an industrial application or a polyhedral mesh for a graphics application.

A limitation of this approach is that it assumes that the object does not change shape between views. This is fine for rigid objects but presents a problem for

BMVC 1999 doi:10.5244/C.13.19

modelling more complex, deformable objects. In this paper we consider the more complex problem of modelling both the geometry *and* the kinematics of *articulated* objects from example range images. Articulated objects are the simplest objects in the class of deformable objects. For clarity we define articulated objects as those objects consisting of a number of rigid parts that are connected by non-rigid joints [1]. In this paper we refer to a rigid part of an articulated object as a *component*.

Most of the related work on modelling articulated objects utilises manually constructed models for either motion analysis or recognition of articulated objects, for example [2] and [3]. Recently Kakadiaris *et al* [4] presented a system for reconstructing the shape and kinematics of people from video images but specific knowledge of the human form was exploited. Also, this approach utilises a long sequence of images through which features can be tracked. Our objective is to construct arbitrary models from small numbers of example range images in which the objects are observed in different pose configurations.

The problem of automatically building models of articulated objects from range data is complex because data points in each range image that belong to the same rigid component of the articulated object must be brought into registration but the association between surface measurements and object components is unknown. Without knowing this association registration is difficult and without knowing the registration determining this association is difficult. In previous work we have considered the problem of identifying rigid subsets of data from pairs of range images of articulating objects [5, 6, 7].

Our approach to this problem is to search for local registration solutions by matching local surface shape features between pairs of surfaces. Surface measurements brought into registration are then labelled as being part of a rigid surface and are associated with a component of the articulated model. Results for each pair of range images are then merged to produce a complete geometric model of the object. Finally the relationship between components of the model seen in different views is used to estimate the position of joint axes.

This pairwise approach does not provide an optimal solution to the problem although the results are very satisfactory. An optimal solution should simultaneously register all surface measurements whilst also taking account of the constraints imposed by the joints. In the future we intend to use the solution given by our current approach as an initial solution in an iterative, optimal algorithm.

2 The Model Construction Algorithm

Given a number of range images of an articulated object we would like to build a model that describes the object's surface geometry and kinematics. This suggests the following three objectives:

1. To determine the association of surface measurements in each range image to components of the articulated object.

2. To establish the transformation that registers surfaces measurements belonging to the same object component between viewpoints.

3. To estimate the position of the joints that connect different object components.

In this paper we simplify the problem by only considering rotational joints with a single degree of freedom. This is not a fundamental limitation of our approach however.

Here we present an overview of the model construction algorithm that we have developed. The details of each stage of the algorithm are presented in the following subsections.

1. A surface representation of triangular facets is constructed for each of the example range images.
2. For each pair of surfaces, the rigid transformations that bring a substantial proportion of the surfaces into registration are determined.
3. Surface facets in good registration, for each rigid transformation, are used to form *partial* component models.
4. Overlapping partial component models are grouped to form *complete* component models.
5. The relationship between pairs of components is used to estimate the relative position of the model joints.

2.1 Surface Reconstruction

The data used to build the articulated model comprises a number of range images of the articulated object in different poses. Each range image is a set of 3-dimensional measurements of the underlying object surface. For each range image a representation, S_i, of the object's surface is determined. In this work the surface representation used is a mesh of triangular facets.

$$S_i = \{{t_1}^i, \cdots, t^i_{N_i}\} \tag{1}$$

where t_i is a triangular facet of the mesh with vertices $\mathbf{u}_i$, $\mathbf{v}_i$ and $\mathbf{w}_i$.

$$t_i = [\mathbf{u}_i \mathbf{v}_i \mathbf{w}_i] \tag{2}$$

A number of algorithms have been proposed for reconstructing a triangular faceted mesh from a set of points. In the work presented here an initial, regular mesh was constructed by forming a scalar field from the sampled point data and then using the *Marching Cubes* algorithm to find the iso-surface [8]. The resulting regular mesh was then refined to minimise the number of facets whilst maintaining most of the surface shape using a mesh simplification algorithm by Garland and Heckbert [9].

2.2 Surface Registration

Given two surface meshes, each representing an articulated object in a different pose, a number of rigid transformations exist that bring parts of the surfaces into mutual alignment. Each of these transformations, T^a, brings surface measurements belonging to the a^{th} object component into registration. The objective of this part of the algorithm is to determine all of the rigid transformations that register rigid subsets of the two surfaces.

Many of the registration algorithms in the literature are based on the *Iterated Closest Point* algorithm of Besl and McKay [10]. These algorithms are generally unsuitable for the registration of articulated data because they adopt a global methodology. The solution to the registration of articulated data is a number of local transformations each of which brings only one of any number of rigid components into alignment.

Here we perform surface registration using a technique we have developed previously [11, 12]. This method uses a novel representation of local surface shape to find local surface correspondences. A RANSAC [13] algorithm is then used to estimate the registration transformations that bring significant areas of the two surfaces into alignment.

2.3 Partial Component Models

If a rigid transformation can be found that brings a substantial proportion of two surfaces into registration then it is concluded that the overlapping sections of the surfaces represent some part of a rigid component of the articulated object. These overlapping sections of the surfaces are then used to build a partial model of the object component. Non-overlapping sections of either surface might also represent the same component but this cannot be determined without reference to other example surfaces and are not included in the partial component model. For two surfaces, S_i and S_j, brought into registration by a rigid transformation, T^a_{ij}, the partial component model is represented by the following graph.

$$\boxed{S^a_{ij}} \xrightarrow{T^a_{ij}} \boxed{S^a_{ji}}$$

where S^a_{ij} and S^a_{ji} are the subsets of the surfaces S_i and S_j that are brought into mutual alignment by the rigid transformation. The superscript is used to indicate that these surface patches represent part of object component a. The surface patches S^a_{ij} and S^a_{ji} are defined formally as follows.

$$S^a_{ij} = \left\{ t \in S_i : d(T^a_{ij} t, S_j) < \tau \right\} \quad \text{and} \quad S^a_{ji} = \left\{ t \in S_j : d(T^a_{ji} t, S_i) < \tau \right\} \qquad (3)$$

where the function $d(Tt, S)$ provides a measure of how well a triangle, t, is aligned with a surface S when it is transformed by T. We use the integral of the squared distance from each point on the triangle to the nearest point on the complete surface. Note that in our convention $T_{ij} = T^{-1}_{ji}$. τ is a distance tolerance.

2.4 Complete Component Models

After all of the partial component models have been constructed for all possible surface pairings, partial component models belonging to the same object component are merged to form *complete* component models. This is achieved by merging partial components representing surface data that moves rigidly together. As each partial component model represents a rigid subset of data, two partial component models that share some common surface data must be rigidly attached and can be merged. If two surface patches, S^a_{ji} and S^b_{jk} share one or more surface facets then it is deduced that they belong to the same model component and are merged.

$$\boxed{S^a_{ij}} \xrightarrow{T^a_{ij}} \boxed{S^a_{ji}} \quad , \quad \boxed{S^b_{jk}} \xrightarrow{T^b_{jk}} \boxed{S^b_{kj}} \quad = \quad \boxed{S^A_{ij}} \xrightarrow{T^A_{ij}} \boxed{S^A_{ji} \cup S^A_{jk}} \xrightarrow{T^A_{jk}} \boxed{S^A_{kj}}$$

where $S^a_{ji} \cap S^b_{jk} \neq \emptyset$. Notice that the superscripts have now been updated to indicate that the patches belong to a large model component A. The original superscripts represent an arbitrary assignment of a surface patch to an object component but it was not possible to maintain consistency between different pairs of views. This problem is resolved when merging partial component models and is indicated by using a capital superscript.

The partial component merging procedure begins with an arbitrary partial component model which acts as a seed. This model is then grown by merging with other partial component models that share surface facets. When merging is complete the object component is represented by a connected graph whose nodes represent a collection of surface patches and whose arcs describe the transformations that register those patches. This process is repeated for any remaining partial component models until a complete model of each object component has been constructed.

2.5 Estimating the Joint Axis

Once the shape of each of the model components has been determined the articulated model is completed by determining the relationship between each of the components.

If a pair of object components are both visible in two range images then the relative transformation from one component to the other can be determined. We define the joint transformation, J^{AB}_{ij}, as the transformation of component A with respect to component B, determined from the data present in surfaces S_i and S_j. The joint transformation is given by the expression:

$$J^{AB}_{ij} = T^B_{ji} T^A_{ij} = \left(T^B_{ij}\right)^{-1} T^A_{ij} \tag{4}$$

The relationship between these transformations is presented in Figure 1. The joint transformation is easily represented as a rotation angle, a rotation axis direction vector and an axis position using standard results.

For each pair of views in which two components appear a separate estimate of the joint axis direction and position can be determined. We combine these estimates to derive an improved estimate of the joint in the final model. The final axis direction

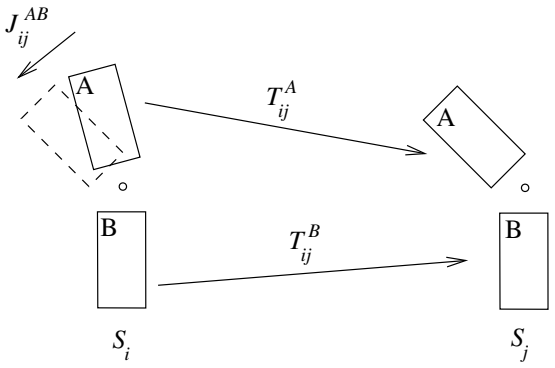

Figure 1: Determining the joint transformation J_{ij}^{AB} connecting components A and B from the two views S_i and S_j.

is calculated as the mean of the individual axis direction estimates. For the final axis position we determine the position of the point whose squared distance to each of the estimated axes is a minimum. The details of this calculation are presented in Appendix A.

3 Experiments

In this section the results of an experiment are presented in which a model of a doll's leg is constructed from eight range images. Figure 2 presents six of the eight surface meshes constructed from eight range images taken of the doll's leg in different poses. The range images were taken at approximately equal intervals around an axis parallel to the length of the leg. Each mesh comprises 1000 triangular facets and each face is represented by a geometric histogram which is 32 by 32 bins in size. A spherical window of radius 25mm around the centroid of each facet was used to define the local geometry represented by each histogram. The mean time to construct the 1000 histograms for each range image was 47 seconds.

The first row of Figure 3 presents the best two solutions for the registration of surfaces 1 and 2. The two surface meshes have been rendered in two different shades of grey. It can be seen that each of the main components of the articulated object have been brought into approximate registration. The second row of Figure 3 presents the facets of the two surfaces in mutual alignment. These facets form partial component models. The mean time to find the registration transformations for each pair of surface meshes was 277 seconds.

The final, articulated model is presented in Figure 4. The first two rows present a number of views of each of the rigid components of the model. The overall shape of the components is good although the accumulation of registration errors due to the sequential nature of the algorithm can be seen in row 2, column 2. The lower part of the leg also has some faces which would be expected to belong to the thigh. These faces are almost orthogonal to the joint axis and overlap the thigh even when faces belonging to the lower leg are brought into registration. Some post processing is required to assign these faces to the correct component.

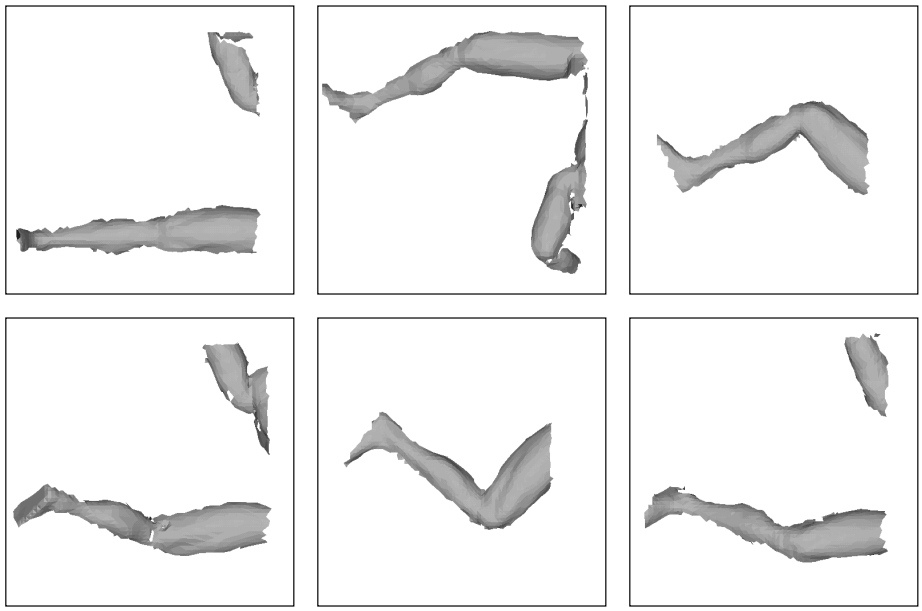

Figure 2: Six of the eight surface meshes constructed from the range data taken of the doll's leg in different poses.

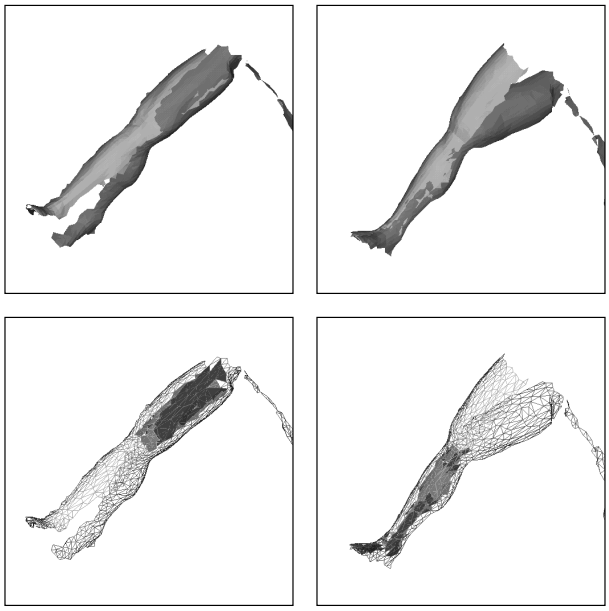

Figure 3: The best two solutions for the registration of surface meshes 1 and 2. The second row shows the surface facets in mutual alignment.

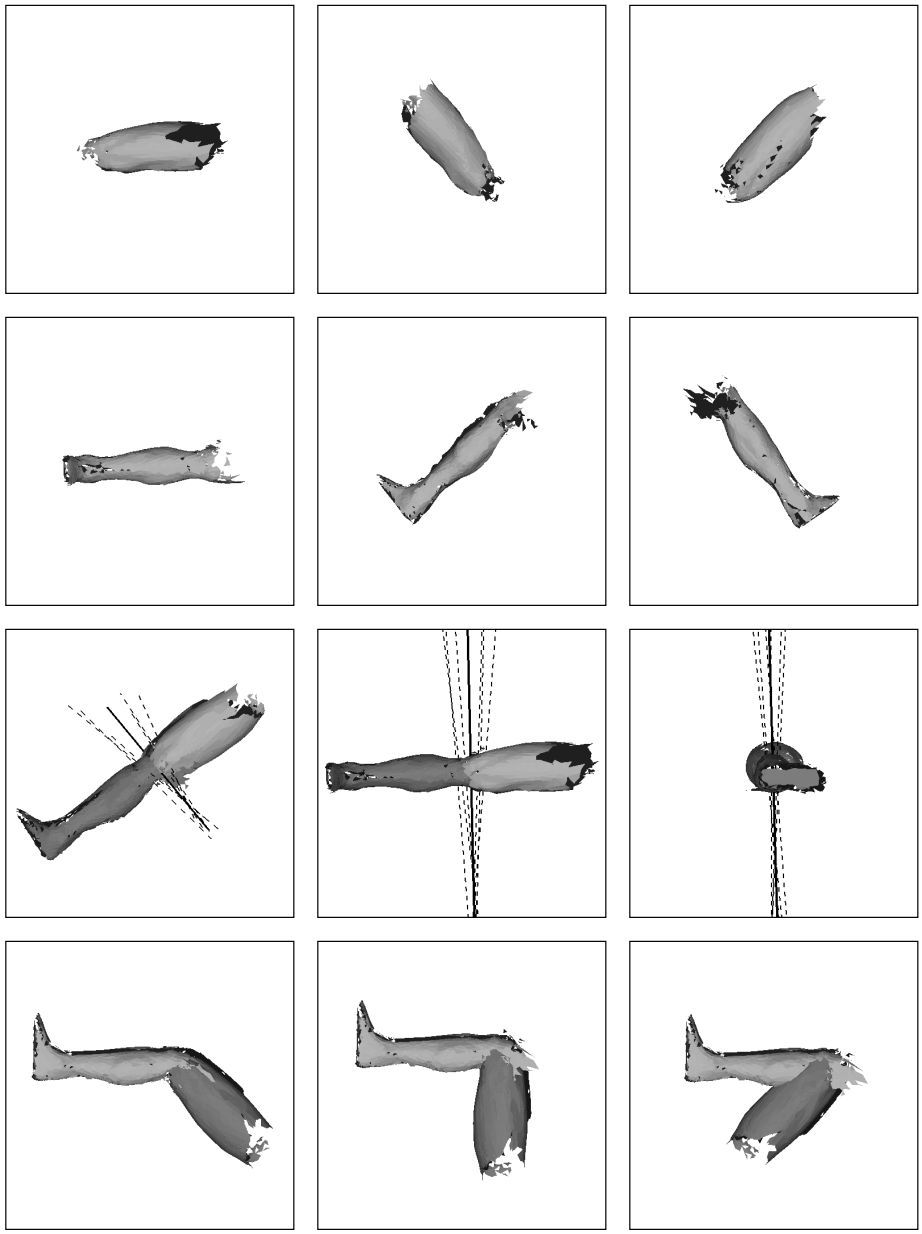

Figure 4: The first two rows present a number of views of each of the rigid components of the articulated model. The third row presents the estimated joint axes. The dotted lines are the estimates determined from each pair of views and the solid line is the combined estimate. The final row presents a view of the final model in a number of different poses.

The third row presents the estimated joint axis positions. The dotted lines represent the axes estimated from each pair of images in which both object components have been successfully registered. The solid line presents the final, combined axis estimate. Although the individual estimates vary substantially the final estimate is very satisfactory. This is evident from the fourth row which shows the final model in different poses which have been generated using the estimated joint axis.

To provide a quantitative assessment of the algorithm the variation of each of the axis estimates, compared to the final combined estimate, is presented in Table 1.

Axis direction error (degrees)	Axis position error (mm)
5.174	3.026

Table 1: The variation (standard deviation) in the estimates of the direction and position of the joint axis.

4 Conclusions and Future Work

In this paper we have presented a technique for constructing an articulated model from a small number of example range images of an articulating object in different pose configurations. The final model captures both the surface shape of the object's components and the object's kinematics. The technique has been demonstrated on a difficult problem of building a model of an object with free-form surfaces.

Although the results of this technique are satisfactory, by adopting a pairwise approach to the registration of the surfaces the final solution is suboptimal. Also, the estimation of the joint axes is determined after the surface registration has been completed. We are currently investigating how we can simultaneously register all of the data points whilst imposing the constraints imposed by the joints' axes. This is a difficult, non-linear optimisation problem. The approach presented here will provide an initial estimate which will be invaluable when searching for the optimal solution.

Acknowledgements

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

A Closest Point to a Set of Lines

Given a number of lines we would like to determine the position of a point such that the sum of the squared distances from the point and each line is minimised. This problem is simplified if each line is replaced by any two orthogonal planes whose intersection is the line. The squared distance to the line can then be replaced by the sum of the squared distances to the planes.

If each plane is represented by a unit normal vector $\mathbf{n}_i$ and the perpendicular distance from the origin to the plane, d_i, then the point $\mathbf{x}$ for which the sum of the squared distances to the planes is a minimum is given by the expression:

$$\mathbf{x} = (NN^T)^{-1}ND \tag{5}$$

where $N = [\mathbf{n}_1\mathbf{n}_2\cdots]$ and $D = [d_1d_2\cdots]^T$.

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] Jung, S. and Wohn, K., "Tracking and Motion Estimation of the Articulated Object: A Hierarchical Kalman Filter Approach", Real-Time Imaging, 3(6), pp415-432, 1997.

[3] Grimson, W.E.L.,"On the Recognition of Parameterized 2-D Objects", IJCV, 2, pp353-372, 1989.

[4] Kakadiaris, I. A. and Metaxas, D., "Three-dimensional Human Body Model Acquisition from Multiple Views", IJCV, 30(3), pp191-218, 1998.

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

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

[7] Ashbrook, A.P., Fisher, R.B., Robertson, C. and Werghi, N., "Segmentation of Range Data into Rigid Subsets using Planar Surface Patches", Proc. ICCV98, pp201-206, 1998.

[8] Lorensen, W.E. and Cline, H.E.,"Marching Cubes: A High Resolution 3D Surface Construction Algorithm", Computer Graphics, 21(4), pp163-169, 1987.

[9] Garland, M. and Heckbert, P. S., "Surface Simplification using Quadric Error Metrics", SIGGRAPH97, pp209-216, 1997.

[10] Besl, P. J. and McKay, N. D., "A method for registration of 3-D shapes", IEEE PAMI, 14(2), pp 239-256, 1992.

[11] Ashbrook, A.P., Fisher, R.B., Robertson, C. and Werghi, N., "Finding Surface Correspondence for Object Recognition and Registration Using Pairwise Geometric Histograms", Proc. ECCV98, pp674-686, 1998.

[12] Ashbrook, A.P., Fisher, R.B., Werghi, N. and Robertson, C., "Aligning Arbitrary Surfaces using Pairwise Geometric Histograms", Proc. NMBIA98, pp103-108, 1998.

[13] Fischler, M. A. and Bolles, R. C., "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Communications of the ACM, Vol. 24, pp381, 1981.