Registering, Integrating, and Building CAD Models from Range Data

Ruigang Yang and Peter K. Allen
Department of Computer Science, Columbia University, New York, NY 10027 *

Abstract

We introduce two methods for the registration of range images when a prior estimate of the transformation between views is not available and the overlap between images is relatively small. The methods are an extension to the work of [6] and [5] and consists of 2 stages. First, we find the initial estimated transformation by extracting and matching 3D space curves from different scans of the same object. If no salient features are available on the object we use fiducial marks to find the initial transformation. This allows us to alway find a satisfactory and even highly accurate transformation independent of the geometry of the object. Second, we apply a modified Iterative Closest Points algorithm (ICP) to improve the accuracy of registration. We define a weighted distance function based on surface curvature which can reduce the number of iterations and requires a less accurate initial estimate of the transformation.

1 Introduction

We are building a system to automatically recover a solid CAD model of an arbitrary object. This is an important problem with many applications including reverse engineering, generation of virtual reality and simulation models, and 3-D Fax transmission of object models. Automatic reconstruction of models of real objects involves 3 steps: 1) Data acquisition: In our case, range imagery from a number of different viewpoints is obtained. 2) Registration of different views into a common coordinate system. 3) Integration of views to form a complete (i.e. no open surface holes or dangling surfaces) watertight solid model.

In this system, an object is placed on a turntable and a solid CAD model is produced automatically. The data acquisition module includes a line scan range finder which is attached to an IBM 7575 robot (see figure 1). The object is scanned by moving the scanner along the robot's Z axis, and then the turntable is moved to obtain another scan of the object from a different viewpoint.

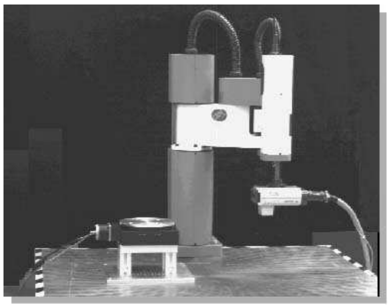

Figure 1: Robot, Laser rangefinder and turntable

We have previously described the method for building these models in [8, 9]. For each range scan, a mesh surface is formed and "swept" to create a solid volume model of both the imaged object surfaces and the occluded volume. This is done by applying an extrusion operator to each triangular mesh element, sweeping it along the vector of the rangefinder's sensing axis, until it comes in contact with a far bounding plane. The result is a 5-sided triangular prism. A regularized union operation is applied to the set of prisms, which produces a polyhedral solid consisting of three sets of surfaces: a mesh-like surface from the acquired range data, a number of lateral faces equal to the number of vertices on the boundary of the mesh derived from the sweeping operation, and a bounding surface that caps one end. Each of these surfaces are tagged as "seen" or "occlusion" for the sensor planning phase that follows.

Each successive sensing operation will result in new information that must be merged with the current model being built, called the composite model. The merging process itself starts by initializing the composite model to be the entire bounded space of our modeling system. The information determined by a newly acquired model from a single viewpoint is incorporated into the composite model by performing a regularized set intersection operation between the two. This paper focuses on the problem of integrating these models when the range scans come from arbitrary positions, necessitating a registration step to bring the different scans into alignment.

*This work was supported in part by an ONR/DARPA MURI award ONR N00014-95-1-0601, DARPA AASERT awards DAAHO4-93-G-0245 and DAAH04-95-1-0492, and NSF grants CDA-96-25374 and IRI-93-11877.

2 Overview of Registration Methods

Most registration methods can be divided into 2 main categories. The first avoids the registration problem altogether by relying on the precisely calibrated mechanical equipment to determine the motion transformation between views. These methods assume that the interview transformations provided by the data acquisition apparatus are sufficiently accurate to properly register the range views and do not need to be improved upon. These methods are relatively simple to implement. For a system using a turntable, the accuracy can be acceptable, although recalibration is an ongoing problem. However, a turntable approach will still usually leave the top and bottom (support) surfaces of an object unscanned.

The second category involves methods that derive the registration transformation between range images from the information contained in the range image. Transformation parameters are gradually updated and refined until the range views are precisely registered. A feedback function measuring the quality of the registration is used (e.g [3]). This has been an active research area over the last few years; a recent paper [7] gives an overview of different methods in this category. These methods can break the hardware limit and enable us to register different views scanned from arbitrary view points. For an overview of previous work, the paper by Barequat and Sharir [1] contains an extensive bibliography along with the description of their method based on geometric hashing.

The Iterative Closest Points (ICP) algorithm[2], [5] is a well known method that is used to register images with significant overlap. The ICP algorithm starts from an initial configuration of two views and iteratively finds their best correspondence. This algorithm is proven to converge but it can get trapped in a local minimum, The convergence of the this algorithm towards the global minimum is known to depend largely on the initial configuration. In our problem, the views we are trying to register may be quite distant from each other, with little overlap, and this can cause problems in determining a good initial estimate of the transformation between views. Below, we describe 2 methods we have used to to find the initial estimate and a modification to the ICP algorithm that can help the system find a better registration.

3 Finding Initial Transformation

Our method is a 2-step feature-based registration algorithm. In step 1 we derive the initial transformation from the features in the range data, such as points, edges and curvatures. Then, in step 2, we apply our modified ICP algorithm based on [5]. We note that the second step is optional, since with objects with salient features, the output of step 1 may be accurate enough.

The method we use is based on the work of Feldman and Ayache [4] and Gueziec and Ayache [6]. This approach makes use of intrinsic features extracted from the range data such as space curves and that an intrinsic reference frame can be associated with each point along the curve.

Our work in automatically building CAD models has tried to reduce the number of views needed to capture an object's shape. Reducing the number of scans (and associated data set size) can cause a problem: the scans that need to be registered are scanned from relatively far away view points. Thus, many regions in one range scan are not visible in the other scans. So instead of picking up registration points randomly, we choose points that have a gradient over a threshold. Because points with a high gradient usually belong to the edges or contour of the object, they can form salient feature points. These points can then be joined by curves to get a highly structural description of the range data. Further, these characteristic curves are stable with respect to rigid transformations, and can tolerate partial occlusion due to their local nature - this is very important in our case.

Our approach involves the use of polygonal curves formed by locally joining distinguished point on the surface of a mesh. We first find edge points from the range data, and then filter out any isolated points. The remaining local points are grouped into curves, keeping only these points which form a curve of more than P points. In the experiments below, we chose $P = 15$.

Once we have a set of linked edge points, we can fit them to a 3-D space curve using a cubic polynomial approximation:

$$\begin{cases} y = ax^3 + bx^2 + cx + d \\ z = ex^3 + fx^2 + gx + h \end{cases} \tag{1}$$

In a parametric representation, the above equation can be written as

$$u = ti + (at^3 + bt^2 + ct + d)j + (et^3 + ft^2 + gt + h)k \tag{2}$$

Using this curve approximation, we can compute the curvature and torsion of each point $P_i(x, y, z)$ using the cubic polynomial approximation.

Denoting differentiation with respect to u by a dot ($\dot{u}$) the curvature κ, and torsion ψ of P_i are

$$\kappa = \left| \frac{\dot{u} \times \ddot{u}}{\dot{s}^3} \right|, \psi = \left| \frac{\dot{u} \cdot (\ddot{u} \times \dddot{u})}{\dot{s}^3 \cdot \dot{\kappa}^2} \right| \tag{3}$$

where $\dot{u} = i + (3at^2 + 2bt + c)j + (3et^2 + 2ft + g)k$, $\dot{s} = |\dot{u}|$.

Having represented a curve as a cubic polynomial approximation, we now wish to compute a rigid transformation which will match curves in 2 different views. Formally, our problem can be stated as follows: We are given a set of target(stationary) curves M in target view and a set of source curves S in the source view which needs to be transformed. We wish to find a rigid transformation of S which will minimize the distance between M and S.

An intrinsic reference frame can be associated with each point along the curve (the Frenet frame). Given a pair of points(m_{ip}, s_{jq}), where point m_{ip} belongs curve m_i, and point s_{jq} belongs to curve s_j, a unique rigid transformation $D = (R, T)$ can be defined. Because we can associate Frenet Frames (t, n, b) with each point, the rotation R that brings the two frames aligned is given simply by the outer product of the 2 frames orthogonal basis vectors: $R = (t, n, b)(t\prime, n\prime, b\prime)^T$. For the translational component, suppose that o is the origin in the global reference frame, and that os_{jq} denotes the vector from o to s_{jq} while om_{ip} is the vector from o to m_{ip}, then $T = om_{ip} - R(os_{jq})$.

We define the $NearestPoint(CurveSet, p)$ as a function where $CurveSet$ is a set of 3D space curves, and p is a point. It will return a point in $CurveSet$ which is closest to p. Then, we start the following matching process.

1. A subset of k significant points on the curve s_j is selected according to a stable criterion. In our case, points of curvature extrema are used.

2. Choose one point s_{jq} from the k significant points, find a point m_{ip} in the target curve set M which has a similar curvature and torsion values.

3. Compute the rigid transformation D for (m_{ip}, s_{jq}) point pair, and apply D to all points in curve s_j.

4. Compute the distance d_j between s_j and M where

   $$d_j = \sum_q |NearestPoint(M, s_{jq}) - Ds_{jq}|^2 / |s_j|, \quad (4)$$

   where $s_{jq} \in s_j$ and $|s_j|$ is the number of points in curve s_j.

   If the distance d_j is below a threshold, we consider this transformation is a candidate transformation and put it into the candidate bin.

5. go to step (1) until all significant points in all curves in S have been tested.

6. For every transformation matrix D_k in the candidate bin, apply it to every curve s_j in S , compute the global distance g_k

   $$g_k = \sum_{j,q} |NearestPoint(M, s_{jq}) - D_k s_{jq}|^2 / N \quad (5)$$

   where N is the number of points in the source curve set S. The final transformation D_{final} is the one which results in the least global average distance between M and S.

In the experiments using this method, we found that the accuracy of this method is, like most curvature based methods, object dependent. For objects without salient features, we can use the method described in the next section.

4 Registration Using Fiducial Marks and Modified ICP Algorithm

The curvature-based method is not guaranteed to converge to a global minimum if there are no salient features on the surface of the object. In these cases, we can place fiducial marks on the object. These marks provide very good visual checkpoints to see how good the registration is. We use long thin pins as the registration points. We plant these pins on the object in a way so that they can be seen from many viewpoints, and they are distributed asymmetrically.

Because the pins are long and thin, we can easily distinguish them by using an edge detector on the range data. Having identified each pin in 2 different images, we can match them uniquely based upon their asymmetrical distance measures. This sets up a 1-to-1 correspondence between the registration pins from 2 different views, and from this we can compute the transformation D.

Once we find an initial transformation as described above, we can use the ICP algorithm to improve the fit.

We first give an overview of Turk and Levoy's ICP algorithm, which was used in their program called Zipper [5]. It can quickly register pairs of triangular meshes representing the range images. Formally, given a pair of meshes, A and B, Zipper finds the nearest position a_i on Mesh A to each vertex of Mesh B. To achieve a higher accuracy, position a_i is not limited to the position of a vertex, it may be anywhere on a C^0 continuous surface - that means a_i may be a vertex of a triangular mesh element, may be a point within a triangular mesh element or it may lie on a triangular mesh edge. Zipper calculates the rigid transformation that minimizes the distance between the pairs of points. The procedure will iterate until convergence.

The algorithm adds a distance threshold to the basic ICP method described by Besl and McKay[2] to avoid matching any point P of one scan to a remote part of another scan, which is not likely to correspond to P. In practice, the distance threshold is set to twice the spacing s between range points. So for some high-resolution scans, it requires the error of the initial configuration to be within 1 mm. We tested several greater distance thresholds (4s, 6s, etc.), all with unsatisfactory results. First, it took longer for every iteration, because it tried to find out the nearest points in a larger space. Second and more important, it tended to erroneously achieve excessive overlaps between the target and source scan.

A way to improve this algorithm is to use the local curvature value as a guide in selecting matching points as suggested in [4]. Thus, we can increase the distance threshold without the risk of possible excessive overlaps in the result, since we are using curvature as well as distance to measure match fidelity.

This improvement can also reduce the number of iterations for the algorithm to converge because the distance threshold is increased, it can find out relatively-remote matching points in one step which may take several steps if a smaller distance threshold is used. To avoid possible oscillation around the convergence point we use a decreasing distance threshold during every iteration,

To use this method, it is necessary to measure the curvature values of every point in the scan, not just high gradient points along space curves. We estimate first and second derivatives using discrete operators and from these estimates, we can compute the fundamental forms of the surface, which describe the local curvature at each point on the surface. From the fundamental forms we can compute the principal curvatures at each point on the surface. In practice, we only use the maximum principal curvature as the curvature value of the point.

Now that every point is associated with a curvature value, we can use the following weighted distance function to define the closest point:

$$Dis = \left| \frac{\kappa_p - \kappa_q}{\kappa_q} \right| \cdot |P - Q| \qquad (6)$$

here, P, Q are 2 points on 2 different scans respectively, and κ_p, κ_q are their curvature values.

In practice, because one point P's closest point is a 3D space point on a C^0 continuous surface, its curvature value is the interpolation of the values of its neighborhood vertexes. We pre-compute the curvature value at every vertex on the target surface and interpolate the curvature for any interior point on the triangular mesh element.

Having a set of correctly matched points, which is the most challenging task in the ICP, we can use a minimization function to find the rigid transformation which will minimize the least-squared distance between the points pairs.

5 Experimental Results

An eye-in-hand system (shown in Fig 1) was used for the acquisition of the range images. It consists of an rangefinder attached to a IBM 7575 robot. The rangefinder can move along the vertical axis while scanning. A precision turntable was used to accurately rotate the object so that sampling from different view points can be achieved.

In the experiments below, it is important to separate out resolution issues of the scanner itself, the set inter-

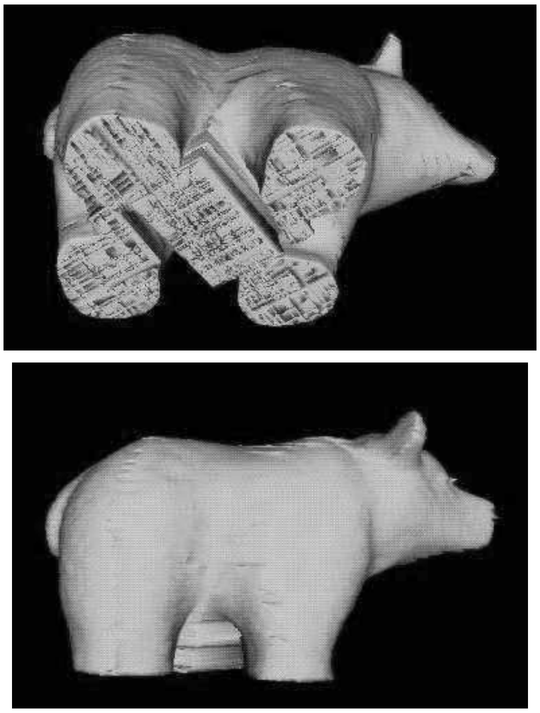

Figure 2: Initial composite model of toy Bear, integrating 4 turntable views, as seen from 2 distinct viewpoints, Note the incorrect abdomen area due to occlusion of the range finder by each of the 4 legs of the bear.

section method of building solid models, and the registration procedure which is the focus of this paper. While the registration procedure works very well, there are still artifacts in the models introduced by the scanner, and removing these is a separate problem.

In the first experiment, we took 5 scans of a toy polar bear. The first 4 scans were obtained by rotating the bear on a precision turntable by 90 degree intervals. We integrated the first 4 scans to generate the initial model shown in figure 2 using our swept-mesh/set-intersection method of building closed volumes [9]. However the abdomen of the bear was not modeled correctly since it was occluded in all 4 of the 90 degree scans. The fifth scan was taken after we manually flipped the bear on the turntable so the bottom of the bear was facing the rangefinder, as shown in figure 3.

Figure 4a shows the registration points computed from the bottom view, and figure 4b shows the approximated space curves after filtering. Fig 5 shows the final model in both mesh and volume representations. After registration via our method, the bottom view is integrated with the initial model (shown in Fig 2) to produce that model. As we see, the abdomen is modeled accurately in this model, which greatly improves the quality of the model.

Our next experiment object is a mug (shown in Fig 6), this is a challenging object because of the thin wall

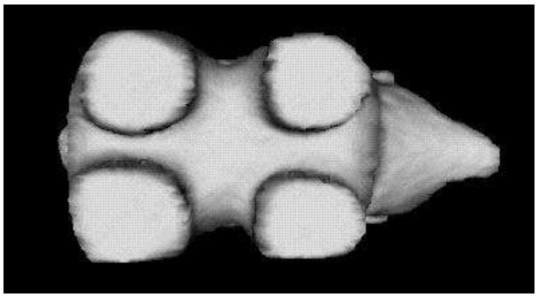

Figure 3: Bottom view of toy Bear

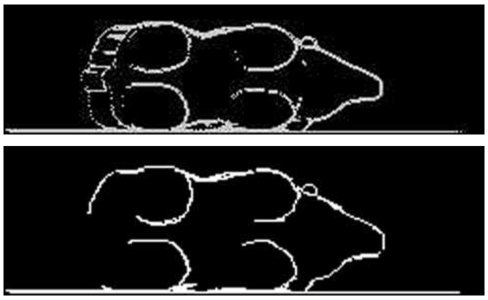

Figure 4: a)Registration points derived from curvature measures and b) space curve generated by linking curvature points.

in the upper part of the mug, and the small diameter of the handle which can make misregistration very obvious. We took five scans including one top view. We first used the feature-based method described in section 2 to derive the initial transformation, then applied our modified ICP algorithm to improve the accuracy. Fig 7 shows the final integrated model. The thin wall of the mug is modeled accurately, as is the handle.

In the next experiment the object to be registered is a tape dispenser, with smooth surfaces and a lack of feature points to allow curvature based registration. We placed 3 long pins on the tape dispenser as fiducial markers. Every pin is about 15 mm in length with a diameter of 0.7mm. 5 scans were taken, including one top view. Fig 8 shows the initial configuration of the 2 views, the top view (darker one) should be transformed to be aligned with the front view (lighter one). The transformation applied to the top view was derived from the method described in Section 4 and Fig 9 shows the final model which is the integration of the 5 views. The long thin sticks with a diameter less than 1 mm are accurately modeled and can still be seen in the model.

6 Conclusion

In this paper, we presented 2 methods for the problem of registering multiple range data scans in order to accurately build a solid 3D CAD model. Our experiments

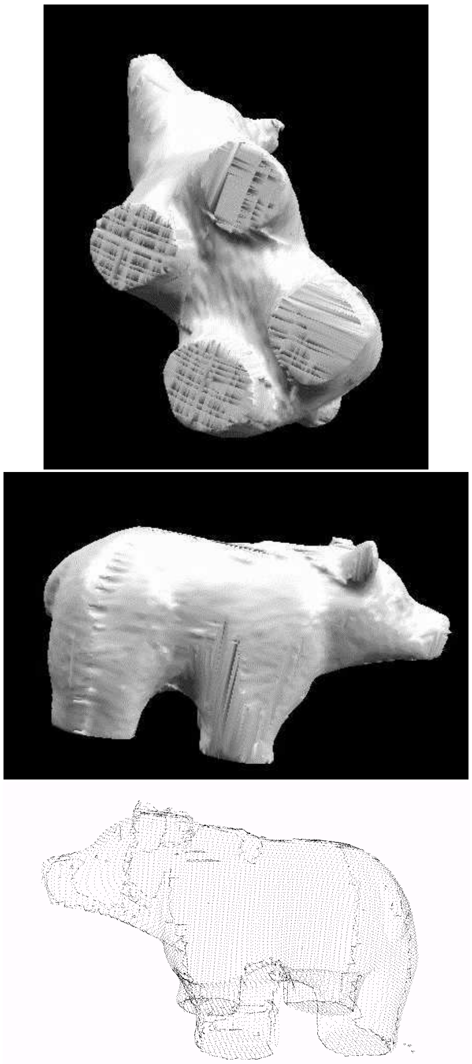

Figure 5: Final model with integration of bottom view (solid and mesh representations).

Figure 6: A picture of the Mug

showed that both methods achieved a satisfactory accuracy. Depending upon the availability of salient features in the object, we may apply the appropriate method. This has become an important component in our automatic 3D modeling system. The system allows a user to build a 3D CAD model without the need for user registration of images.

Future work on the project includes experiments with reducing noise in range data, because curvature values are known to be sensitive to noise. In addition, we are experimenting with a form of super-resolution for range image scanning to remove small sensor artifacts in each scan.

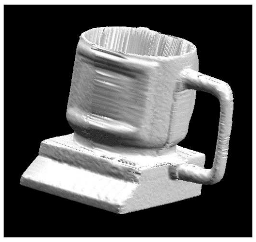

Figure 7: Integrated Model of the Mug

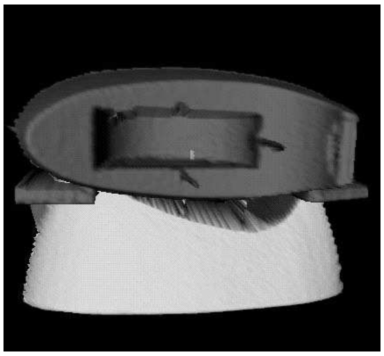

Figure 8: Two views of tape dispenser model before registration, overlaid on each other

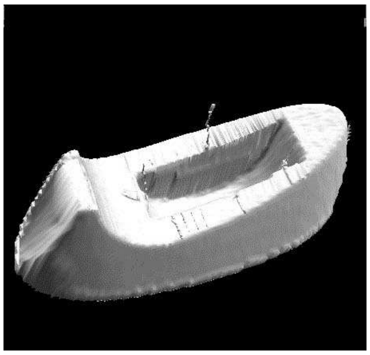

Figure 9: Final model after integration of all 5 views

References

[1] G. Barequet and M. Sharir. Partial surface and volume matching in three dimensions. *IEEE PAMI*, 19(9):929–948, September 1997.

[2] P. J. Besl and N. McKay. A method of registration of 3D shapes. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(2), 1992.

[3] Y. Chen and G. Medioni. Object modeling by registration of multiple range images". In *Proceeding of the 1991 IEEE International Conference on Robotics and Automation*, 1991.

[4] J. Feldmar and N. Ayache. Rigid, affine and locally deformable registration of free-form surface. Technical Report 2200, Inria, 1994.

[5] G.Turk and M. Levoy. Zippered polygon meshes from range images. In *Proceeding SIGGRAPH '94*, pages 311–318, July 24-29 1994.

[6] A. Gueziec and N. Ayache. Smoothing and matching of 3d space curves. *International Journal of Computer Vision*, 12(1):79–104, 1994.

[7] H. Hugli and C. Schutz. Geometric matching of 3D objects: Assessing the range of successful initial configurations. In *Int. Conf. on Advances in 3D Digital Imaging and Modeling*, May 1997.

[8] M. Reed and P. Allen. A robotic system for 3-D model acquisition from multiple range images. In *IEEE Conf. on Robot. and Autom.*, April 22-25 1997.

[9] M. Reed, P. Allen, and I. Stamos. Automated model acquisition from range images with view planning. In *Comp. Vision and Pat. Rec. Conf.*, June 16-20 1997.