RANSAC-based DARCES: A New Approach for Fast Automatic Registration of Partially Overlapping Range Images

Chu-Song Chen[1], Yi-Ping Hung[1], Jen-Bo Cheng[1,2]

[1]Institute of Information Science, Academia Sinica, Nankang, Taipei, Taiwan

[2]Department of Computer Science and Information Engineering,
National Taiwan University, Taipei, Taiwan

Email: hung@iis.sinica.edu.tw

1 Introduction

Registration of two partially-overlapping range images taken from different views is an important task in 3D computer vision. In general, if there is no initial knowledge about the poses of these two views, the information used for solving the 3D registration problem is mainly the 3D shape of the common parts of the two partially-overlapping data sets.

To make the discussion more rigorous, we give a formal definition of the 3D registration problem to be solved in this paper as follows:

Given two data sets $D_M = D_{M_O} \cup D_{M_U}$ and $D_S = D_{S_O} \cup D_{S_U}$ satisfying the following conditions:

1. $D_{M_O} = \{p_i = [x_i, y_i, z_i]^t \mid i = 1, 2, ..., n_o\}$ and $D_{S_O} = \{Rp_i + t + N_i \mid i = 1, 2, ..., n_o\}$, where R is a rotation matrix, t is a translation vector, and N_i is a noise term. That is, D_{M_O} and D_{S_O} are two data sets which can be transformed to be exactly aligned

for the *noiseless case* (i.e., $N_i = 0$) through the rigid-motion (R, t). The overlapping regions of D_M and D_S are D_{M_O} and D_{S_O}, respectively.

2. D_{M_O} (or D_{S_O}) is the unique largest possible overlapping region whose *overlapping ratio* i.e. $n_o/|D_M|$ (or $n_o/|D_S|$ is larger than a threshold Ω, where $|D_M|$ and D_S are the cardinalities of D_M and D_S, respectively. The purpose of the assumption of this condition is to avoid some meaningless solutions. Without this assumption, even the case that the two data sets overlap at a single point is a legal solution.

The *3D registration problem* considered in this paper is to find the *correct* rigid-motion (R, t) which can make the overlapping ratio larger than a threshold Ω.

In the past, a popular type of approaches to solve the 3D registration problem is the *iterative approach* [5][10][14][26]. The 3D registration problem can be formulated as a nonlinear parameter-estimation problem, and an iterative approach minimizes the error function iteratively if an initial estimate of the rigid-motion parameters is given in advance. Iterative approaches have the advantages that they are fast and easy-to-implement. However, the drawbacks are that (i) they require a good initial estimation to prevent the iterative process to be trapped in a local minimum, and (ii) there is no guarantee of getting the correct solution even for the noiseless case. Another popular type of methods is the *feature-based approach*. Feature-based approaches extract invariant local features at first, and then find the correspondences of features for the estimation of the rigid transformation between two partially-overlapping 3D data sets [11][17][24]. Feature-based approaches have the advantages that they do not require initial estimates of the rigid-motion parameters. Their drawbacks are mainly that (i) they can not solve the problem in which the 3D data sets contain no prominent/salient local features, and (ii) a large percentage of computation time is usually spent on preprocessing, which include the extraction of invariant features [11][24] and the organization of the extracted feature-primitives (e.g., sorting [11], hashing[19], etc.).

In practice, when the 3D data contained in two range images are to be registered, we can treat one of them as the *model data set* and the other as the *scene data set*. From

this point of view, registration of two partially-overlapping 3D data sets is similar to the work of *3D object recognition* but in which there is only a single model object stored in the database for matching. However, in a recognition task, it needs not to align the overlapping regions of the scene and the model data sets very accurately. Coarse alignment is usually sufficient for the verification purpose in a recognition task. Hence, the methods which were developed for 3D object recognition can also be employed for solving the 3D registration problem in a coarse manner [2][9][22].

Preprocessing is an important technique which enables the 3D object recognition problem to be solved in an efficient way. It is because that for a recognition task, there are multiple models stored in a database. If some primitives or attributes are pre-extracted for each of the model-object in an *off-line* process, the recognition speed can be increased. For example, to deal with a recognition request, the scene data set is preprocessed only once, and the thereby extracted primitives or attributes are used for matching with that of the models stored in the database. If there are M models in the database, the matching time required for a recognition request is $t_p + Mt_m$, where t_p is the average time for preprocessing the scene data set, M is the number of models contained in the database, and t_m is the average time required to match each model. If M is large, t_m dominates the computational speed for the total recognition task. Hence, to reduce t_m is a major concern to build an efficient 3D object recognition system. On the other hand, the time required for a 3D registration task is $2t_p + t_m$ because the preprocessing of both the model and the scene data sets are required but the matching procedure has to be performed only once. Hence, if a complex pre-processing strategy is used, t_p may dominate the time required for the entire registration task. In fact, many existing methods take a big percentage of time in preprocessing to solve a 3D registration problem [4][11][24][25] [1]. Our goal in this paper is to solve the 3D registration problems in a fast and reliable manner. We propose a new method – the *data-aligned rigidity-constrained exhaustive search* (DARCES),

[1]For example, many feature-based methods take about 20-30 minutes to extract and organize useful feature-primitives, while the time required for finding the correspondences of features is less than one minute [4][24][25].

which can check all possible data-alignments of two given 3D data sets in an efficient way. In our approach, very useful constraints which exploit the rigidity relation among some pre-selected control points are utilized to speed up the search of candidates. The use of these constraints allows our method to solve the 3D registration problem efficiently while requiring no preprocessing and no initial estimate of the 3D rigid-motion parameters.

To solve the partially-overlapping 3D registration problem. the *random sample consensus* (RANSAC) scheme [16] is integrated in our search procedure. RANSAC is a general robust estimation method for surface or model fitting. Robust estimation means that model fitting is not influenced by *outliers*. In this paper, we propose the RANSAC-based DARCES method to solve the 3D registration problem. The RANSAC-based DARCES approach treats the 3D registration problem as an optimization problem. However, unlike the approach using a global optimization technique without any constraints in the search space as introduced in [6], it uses the rigidity-constraints among the points contained in a 3D data set to greatly restrict the possible search space. Therefore, our method is much more efficient. Also, we can prove that the proposed RANSAC-based DARCES approach can solve the partial matching problem of two 3D data sets with few random trials (see Section 4).

This paper is organized as follows. In Section 2, we review some previous related work. Section 3 introduces the DARCES method proposed in this paper. Section 4 introduces the RANSAC-based DARCES approach. In Section 5, the RANSAC-based DARCES method is integrated into a coarse-to-fine structure to speed up the 3D registration. Section 6 shows some experimental results. Finally, some conclusions and discussion are given in Section 7.

2 Priori Related Work

The existing methods for 3D registration can be classified into four classes: *feature-based approaches*, *pre-modeling approaches*, *iterative approaches*, and *optimization approaches*.

We briefly review them in the following.

2.1 Feature-Based Approaches

In a feature-based approach, local feature-primitives have to be extracted in advance. Registration of range images is achieved through the correspondence of feature-primitives that are invariant to rigid-motion.

In [24], Stein and Medioni proposed the *splash structure* to be used for 3D object recognition and registration. A splash is a local map describing the distributions of surface normals along a geodesic circle. They compute a normal by approximating the environment of a normal with triangular patches of small sizes. Every triangle votes for a triangle normal. The average of the three closest triangle normals is the surface normal. Along a given location p, they determine the surface normal $\vec{n}$, and call this normal the *reference normal* of a splash. A circular slice around $\vec{n}$ with the geodesic radius r is computed. Starting at an arbitrary point on this surface circle, a surface normal is determined at every point on the circle. Consider the normals along a geodesic circle indexed with a circular angle θ. Draw a mapping for the pan angle and the tile angle with respect to θ results a 3D curve. These 3D curves are then encoded with a polygon approximation. The curvature and torsion angle of the 3D curves are stored in a hash table for efficient matching.

In the Chua and Jarvis approach [11], 3D data sets are first pre-processed by fitting over the $N \times N$ neighborhood patch of each 3D point using a least square method and examining the residual fit error of that patch. The purpose of the fitting procedure is to allow the differential-geometry-based parameters (such as principal curvatures and Darboux frames) to be estimated more reliably. Some dispersed points with small fit errors are then selected as *seeds*. The principal curvatures and Darboux frames are computed from the local regions of each seed. By applying the assumption that the angle between the two viewing directions must not exceed a given threshold under orthogonal projection, they proposed a view-hemisphere analysis method to select a minimal non-redundant set

of seeds such that at least one of them will lie on the overlapping region. After selecting the first seed, its corresponding seeds with similar feature attributes in the model data set are found, and the Darboux frames of them are used to estimate a set of rough rigid-transformations. Another two seeds are then selected to be lying on the overlapping region based on these transformations via the view-hemisphere analyses. After selecting three seeds in the scene data set, their principal curvatures and Darboux frames are served as local invariances to find many matching candidates on the model data set. For each matching candidate, a hypothesized rigid-motion is generated. Finally, all hypotheses are verified through a transformation/alignment process. In this method, the success of the algorithm requires that the Darboux frames to be computed sufficiently accurate.

In [25], Thirion proposed a new type of feature points, namely, the *extremal points* of 3D surfaces, based on geometric invariants. The relative positions of those 3D points are invariant to 3D rigid-transformations. In their method, some invariant attributes include the types, the principal curvatures, and the distances and the orientations for pairs and triplets of extremal points are used to generate matching hypotheses for 3D registration. To verify the hypotheses, they create a 3D hashing table with the 3D spatial coordinates of the points of the model data set. The hypothesized rigid-transformations are then verified using an iterative procedure similar to the one proposed in [5].

In the above, local features around a 3D point (referred to as *point features*) are used for matching and registration. In [17], 3D invariant curves are extracted and used as *curve features*. In the preprocessing stage of their work, three rotation and three translation parameters are served as invariances between the Frenet frame at a given basis point and the Frenet frame at every other point on the 3D curves. This is repeated for every point along the 3D curves. During recognition, an arbitrary basis point is selected on the 3D curves extracted from the scene data set. For this basis point, the rigid-motion invariants to all the other points are computed, and votes for the model curve and basis point which has already been computed in the preprocessing stage. The rigid-motion with the highest vote indicates the optimal transformation.

In principle, the existence of features is a necessary condition for a feature-based approach to perform the 3D registration task. Hence, feature-based approaches has a limitation that they are not suitable for the case that the data sets contain few prominent/salient features.

2.2 Pre-Modeling Approaches

In a pre-modeling approach, 3D data sets are first represented by a surface model. Usually, the surface model used for registration is the *mesh* model. Then, the registration is performed directly on mesh models rather than on the orignal data sets.

For example, the data set are first converted into a triangulated mesh by using a hierarchical triangulation method [4]. The hierarchical triangulation method can compute from a range image a description of the object surface at different levels of resolutions. After that, acceptable triangles from both views are selected. Triangles are acceptable if they satisfy some constraints, e.g., they are larger enough, at least has a neighbor on three of their edges, and the directions of their normals are approximately parallel to the viewing directions, etc. Then, hypotheses of matches of the two views are generated for those triangle pairs having similar shape, normal orientations, and concavity/convexity. Rigid-motion parameters can then be found by verifying these hypotheses.

In [18], registration starts with a predefined *reference mesh* (e.g., a sphere). The initial reference mesh is deformed to a *surface mesh* using an iterative fitting algorithm until the mesh fits the object surface. The one-to-one correspondences of the nodes on the surface mesh to those on the reference mesh are recorded during the deformation. A spherical image are built by mapping the nodes of the surface meshes to the nodes of a reference mesh on the sphere. On each node of the surface mesh, an index called the *simplex angle* reflecting the local surface curvature is defined and thereby form a *simplex angle image* (SAI). The search for an optimum alignment between the scene SAI and the model SAI is implemented by pre-computing all possible relations of the tessellations on the SAI.

In general, the pre-modeling procedure is usually very time-consuming, and thus be-

comes a bottleneck for the entire 3D registration task. Hence, this type of methods may not suitable when requiring fast registration.

2.3 Iterative Approaches

Some researchers have proposed the registration methods requiring neither feature extraction nor pre-modeling. Besl and McKay proposed the iterative closest point (ICP) algorithm which can reduce the registration error iteratively [5]. Chen and Medioni [10] proposed a similar method to reduce the registration error via iterative refinements. In both methods, a set of uniformly sampled data points (referred to as *reference points*) can be used instead of the extracted feature-primitives. Given an initial estimate of the rigid-transformation, each iteration of their methods contains two steps: (1) establishing point correspondences and (2) computing the rigid motion by minimizing a least-square error measure. These methods are quite simple and effective for the refinement of an existing rigid-motion estimation, and thus are recently very popular for fine registration. However, a limitation of these methods is that they can be applied only for the case that the initial estimate of the rigid-transformation is good enough.

The ICP algorithm may trap in the local minima. To diminish the chances of occurrence of this problem, a stochastic approach combining the least median of squares estimator (LeMedS) and the ICP algorithm is developed [20]. In this approach, the error is measured using the median of the squares (MS) of the distances between two data sets. In each iteration, a set of reference points are first selected randomly and are used to execute the ICP algorithm. Finally, the rigid-transformation with minimal MS error for all iterations is then served as the result of 3D registration.

In the Dorai et. al approach [12] and [13], the method proposed by Chen and Medioni [10] is modified. First, an initial rigid-transformation is estimated by aligning the principal axes of the two data sets. Then, the initial rigid-transformation is iteratively refined using the Chen and Medioni approach. To execute the Chen and Medioni approach in a more reliable way, the distances of each pair of the reference points are used to delete the

mismatch pair of correspondences in advance.

2.4 Optimization Approaches

Because it is not easy to solve the 3D registration problem in general, Blais and Levine have expressed the 3D registration task as an optimization problem [6]. The error function in their work is computed by the sum of Euclidean distances between a set of control points on one of the surfaces and their corresponding points on the other. A stochastic optimization technique, very fast simulated reannealing (VFSR), is used to minimize the error function.

Our goal in this paper is to solve the 3D registration problems in a fast and reliable manner. We propose a new method – the RANSAC-based DARCES approach, which can solve the 3D registration problem without feature-extraction, pre-modeling, or initial transformation. In our approach, very useful constraints which exploit the rigidity-relation among some pre-selected control points are utilized to considerably reduce the search space. The RANSAC-based DARCES approach can be classified as an *optimization* approach. However, the proposed DARCES approach has the advantage that it can be implemented very efficiently. More details are given in Sections 3 and 4.

3 Data-Aligned Rigidity-Constrained Exhaustive Search (DARCES)

Given two data sets, namely, the *scene data set* and the *model data set*. In this section, we only consider a simpler 3D registration problem that the shape of the scene data set is *completely contained in* the shape of the model data set. We call such a problem the *fully contained 3D registration* (FC3DR) problem. After solving the FC3DR problem, two main modifications are considered. First, to solve the general partially-overlapping 3D registration problem. the DARCES approach is integrated into a *random sample consensus* (RANSAC) scheme [16], and thus form a *RANSAC-based DARCES approach* as

introduced in Section 4. Second, the RANSAC-based DARCES approach can be speeded up by integrating it into a coarse-to-fine structure as introduced in Section 5.

3.1 Overview of The DARCES Algorithm

Let us treat the 3D data contained in a dense range image as a 3D surface. In the first place, we have to select some *reference points* from the scene surface, as shown in Fig. 1(a). For example, we can perform a uniform sampling from the indexing grids of the range images to select the reference points, or we can use all the data points contained in the scene data set as the reference points (but will be less efficient). In the subsequent processing, a set of (at least three) *control points* are selected from these reference points, as shown in Fig. 1(b). The operation of the DARCES algorithm is somewhat like *sliding* the control points (of the scene surface) on the model surface. During the sliding, all the selected reference points can be brought to some positions through a rigid-motion determined by these control points. In our approach, if the distance between a 3D reference point and the model surface is smaller than a threshold, then the point is regarded as being *successfully aligned* with the model surface. For each sliding position, we count the number of successfully aligned reference points. Finally, the rigid-motion associated with the sliding position which has the highest count is considered as the solution of our registration task.

A rigid-transformation has to be computed in the DARCES method when the control points move to new poses. Therefore, at least three control points are needed because the determination of a unique rigid-transformation requires at least three point correspondences. In this paper, given a set of point correspondences, the Arun method [3] is used to compute the LMS transformation of them.

Basically, there are $C_{n_c}^{n_s}$ combinations of control points can be selected, where n_s is the number of data points contained in the scene data set, and n_c is the number of the control points. However, it is not necessary to try all of these combinations. In fact, only one of them is sufficient for finding the correct solution in the DARCES approach. How to choose

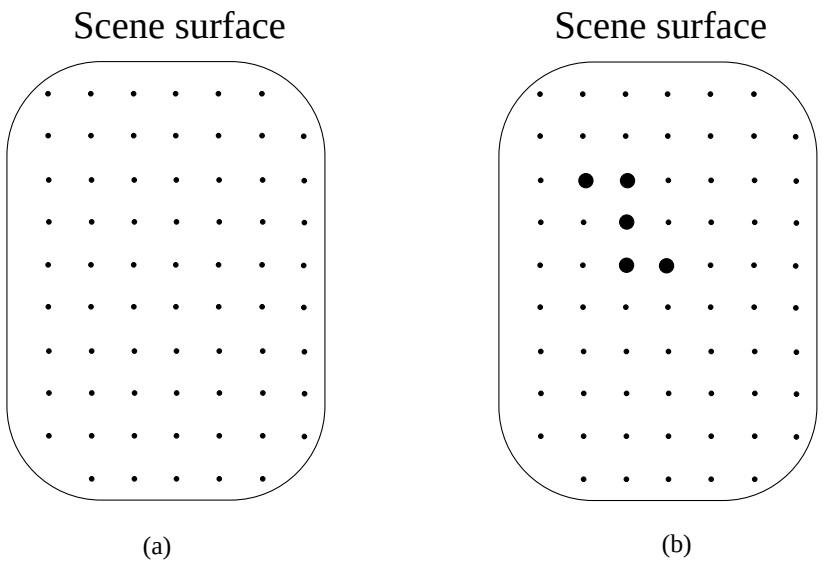


Figure 1: (a) Selection of the *reference points* in the scene surface. (b) Selection of a set of *control points* from these reference points.

an appropriate set of control points, and how many number of control points are required to be used for solving the FC3DR problem, will respectively be discussed in Sections 3.2.4 and 3.3. Another important issue in the DARCES approach is that once a set of control points is selected, how the rigidity-relation among them can be used to facilitate the search process. In the following, we first introduce the case using *three* control points in Section 3.2. The case using more than three control points will be considered in Section 3.3.

3.2 Using Three Control Points

In this section, we only consider the case that three control points are used. In the following, we call the three selected control points in the scene data set the *primary point* S_p, the *secondary point* S_s, and the *auxiliary point* S_a, respectively.

3.2.1 Search Range Reduction by Using the Rigidity Constraints Generated by the Control Points

First, in the model data set, consider the possible corresponding positions of the primary point S_p. Without using feature attributes, every 3D point contained in the model data set can be the possible correspondence of the primary point. Hence, the primary point will be hypothesized to correspond to each of the n_m points in the model data set, where n_m is the number of model points. It is obvious that our method can be easily extended to use some feature attributes associated with each 3D data point, e.g., 3D curvature or image luminance.

Suppose S_p is hypothesized to be corresponding to a model point M_p during the slide. Then, in the model data set, we try to find some candidate points corresponding to the secondary point S_s using the rigidity constraint. Assume that the distance between S_p and S_s is d_{ps}. The corresponding model point of S_s must lie on the surface of a sphere C_s whose center is M_p and radius is d_{ps}. That is, $C_s = \{\mathbf{p} = (x, y, z) \mid ||\mathbf{p} - M_p|| = d_{ps}\}$. In other words, once a corresponding model point of the primary point S_p was hypothesized, the search for M_s, the candidate model point corresponding to the secondary point S_s, can be limited to a small range which is the surface of a sphere with radius d_{ps}, as shown in Fig. 2.

After a corresponding model point of the secondary point S_s has also been hypothesized, we can then consider the constrained search range of the auxiliary point S_a. Assume that S_p and S_s respectively corresponds to the model points M_p and M_s now. The candidates of M_a, the corresponding model pont of the auxiliary scene data point S_a, can be found within a limited search range determined by M_p, M_s and d_{qa}, where d_{qa} is the distance between S_q and S_a, and S_q is the orthogonal projection of S_a to the line segment $\overline{S_pS_s}$, as shown in Fig. 3(a). It is easy to see that the candidates of M_a have to lie on the circle C_a centered at M_q with radius d_{qa}, where M_q is the 3D position corresponding to S_q. That is, $C_a = \{\mathbf{p} = (x, y, z) \mid ||\mathbf{p} - M_q|| = d_{qa}$, and $\overline{pM_q}$ is perpendicular to $\overline{M_pM_s}\}$.

After all the three control points are successfully aligned on the model surface, a unique

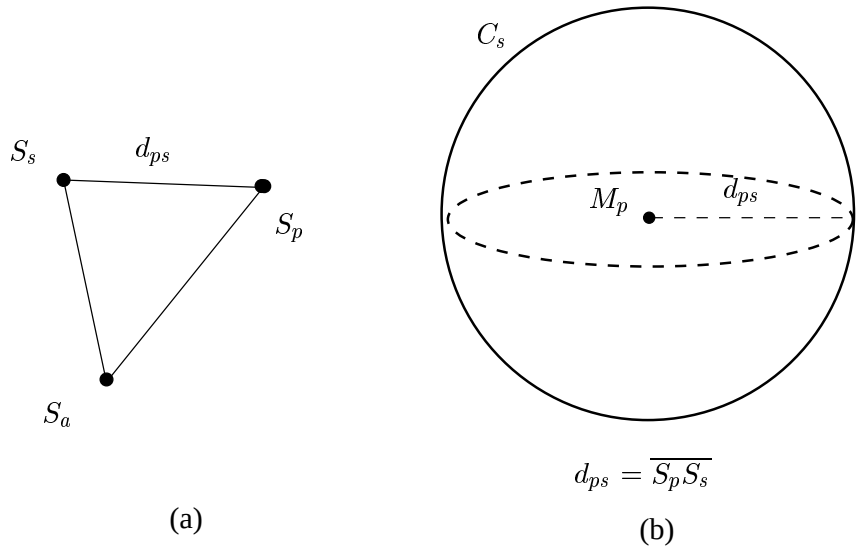


Figure 2: (a) The triangle formed by the three control points selected from the scene data set. (b) The search region of the secondary control point in the model data set is restricted to the surface of a sphere.

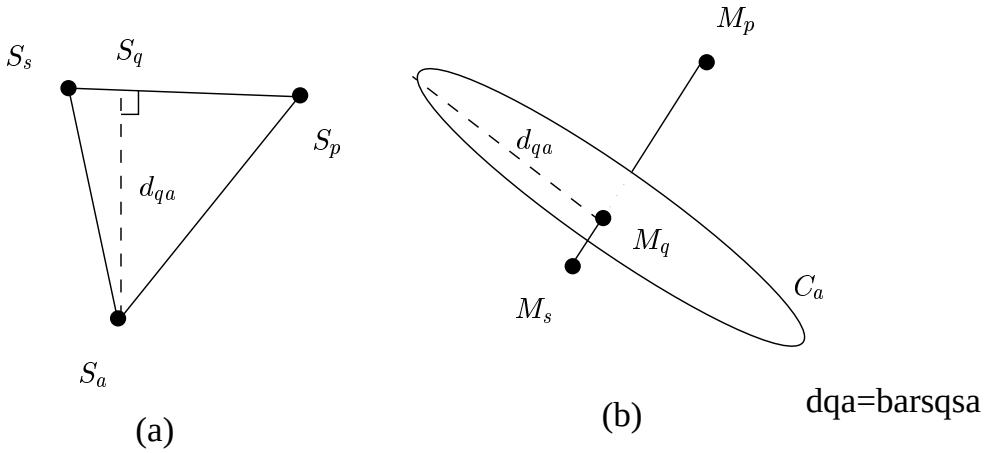


Figure 3: The search region of the auxiliary control point in the model data set is restricted to the contour of a 3D circle.

rigid-transformation, namely T_c, can be determined by using the three pairs of point correspondence: (S_p, M_p), (S_s, M_s) and (S_a, M_a). We then verify T_c by using all the reference points. With the rigid-transformation of T_c, all the reference points, $S_{r_1}, S_{r_2}, ..., S_{r_{n_r}}$, can be brought to new positions $S'_{r_1}, S'_{r_2}, ..., S'_{r_{n_r}}$. We count the number of occurrences, namely, N_o, that S'_{r_i} is successfully aligned on the model surface (i.e., the distance between S'_{r_i} and the model surface is smaller than a threshold) for all i, $i = 1, 2, ..., n_r$. Here, N_o is called the *overlapping number* of the transformation T_c, which is the LMS transformation determined by the three-point correspondence.

For each possible three-point correspondence, an overlapping number can be computed. Finally, the rigid-transformation with the largest overlapping number is selected as the solution of our registration task. [2]

In general, the DARCES method can provide the registration result to be accurate to some extent. To make the registration more accurate, some standard fine-registration procedure [5][10] can be used for refinement. In our method, an ICP-based method [26] is used to perform the fine registration.

By using the above strategy, the search ranges depend on the radii of the 3D sphere and 3D circle. In principle, the three control points form a triangle. Therefore, if a smaller triangle is employed when selecting the three control points, a faster search speed can be achieved. However, if the triangle is selected too small, the computed rigid-transformation will be very sensitive to noise. Hence, how to determine an appropriate size for the above triangle is an important issue in our method. In Section 3.2.4, we will give a simple strategy about how to find an acceptable minimal triangle for the DARCES algorithm.

[2] In this paper, that the largest number of matching points was chosen to be the error measure of accepting the match is only because it is easy to implement and works well in our experiments. In general, there are no restrictions of the selections of the error measures in the DARCES method. Hence, different error measures such as the weighted averages of the distances of the large-enough overlapping regions (as introduced in [6]) can be used as well.

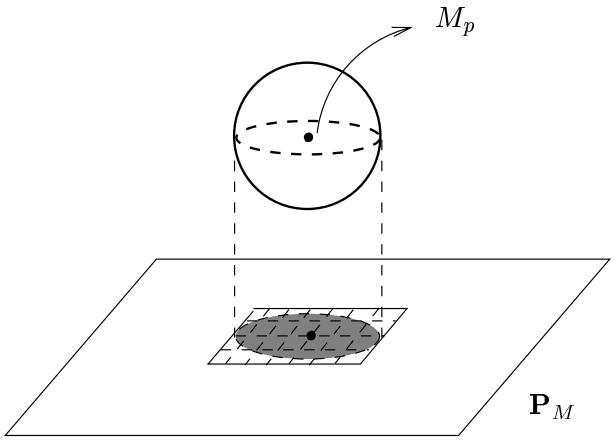


Figure 4: (a) The index plane of a set of 3D data points contained in a range image. (b)The 2D circular region of the projection of the 3D sphere onto the index plane, and the search of M_s is performed in a squared search region containing this circle.

3.2.2 Practical Way for the Implementation of the DARCES Method

For the implementation purpose, direct searches in the 3D space on the surface of a sphere or on the boundary of a circle may not be trivial because some complex data structures (e.g., voxels, tessellation of the sphere, etc) may be required. To implement the DARCES approach efficiently, we use a simple strategy which exploits the fact that a range image can be treated as the projection of the 3D points onto an *index plane*. In the following, assume that the index planes of the scene data set and the model data set are $\mathbf{P}_S$ and $\mathbf{P}_M$, respectively.

To search M_s in the model data set, the 3D sphere C_s is projected onto $\mathbf{P}_M$ and thus forms a 2D circular region on $\mathbf{P}_M$, as shown in Fig. 4. To simplify the implementation, a square search region is used instead of using the 2D circular region in our work. For each 3D point corresponding to the tessellation of the square search region, its distance to M_p is computed. Those 3D points whose distances are approximate to be d_{ps} are then recorded as the matching candidates of the second control point.

To search M_a, one method is to project the 3D circle C_a onto $\mathbf{P}_M$ too, and thus form a 2D ellipse on $\mathbf{P}_M$, as shown in Figure 5(a). The search can then be restricted on the 3D

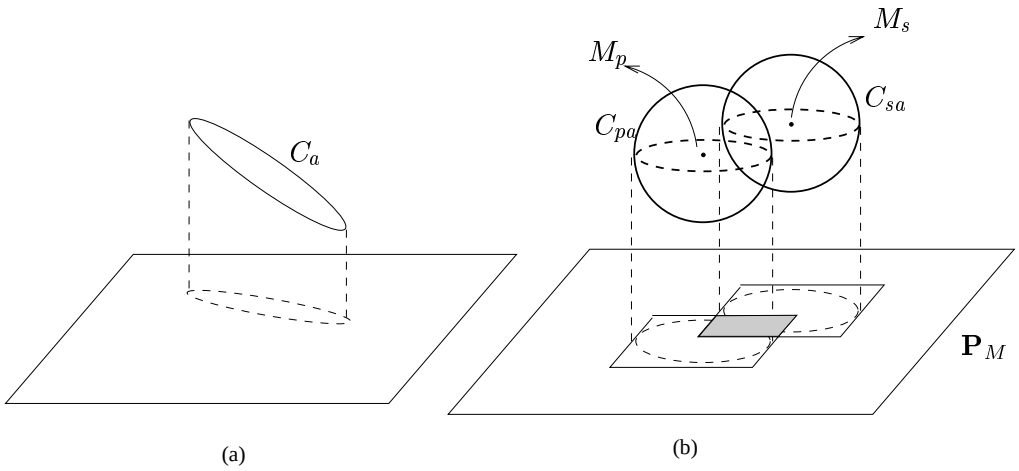


Figure 5: (a) The 2D ellipse of the projection of the 3D circle onto the index plane. (b) The search is performed in the intersection of the two squared regions.

points corresponding to the boundary of the 2D ellipse on the indexing plane. However, to index the boundary of a 2D ellipse is also not an easy task. In our approach, to make the implementation more easier, we do not use the projection of the 3D circle C_a; instead, the projections of another two 3D spheres (C_{pa} and C_{sa}) are used, as shown in Fig. 5(b), where C_{pa} is the sphere whose center is M_p and radius is $\overline{S_pS_a}$, and C_{sa} is the sphere whose center is M_s and radius is $\overline{S_sS_a}$, respectively. The intersection of the two corresponding square regions of the two spheres C_{pa} and C_{sa} on the index plane is then used as the search region of the matching candidates of the auxiliary control point, as shown in Fig. 5(b).

The index plane is also used to facilitate the identification of successful alignments. Remember that if the distance between a 3D point and the model data set is smaller than a threshold, then the point is successfully aligned on the model data set. However, to compute the exact distance between a point and a data set requires to compute the distances of this point to all the points contained in the data set and then select from them the smallest one, which will be very time consuming. Hence, we identify the successful alignments by projecting this point onto the index plane of the model data set, and then compute the distances only in a neighbor region around the projecting position. If the minimal distance computed in the neighbor region is smaller than the threshold, then this

point is considered to be successfully aligned on the model data set.

3.2.3 Complexity Analysis: Case of Using Three Control Points

In the following, we briefly analyze the computational complexity of the DARCES approach if three control points are used. The analysis result can then be used as an important guideline for the development of the strategies to speed up the DARCES method (see Section 3.2.4 and Section 5).

[DARCES algorithm using three control points]

1. Select three control points S_p, S_s, and S_a.

2. for($M_p \leftarrow$ each data point contained in the model data set) {

3. for($M_s \leftarrow$ each data point whose projection is contained in the projected region of the sphere C_s, in the index plane) {

4. if(the distance between M_s and M_p is approximate to d_{ps}) {

5 for($M_a \leftarrow$ each data point whose projection is contained in the projected boundary of the 3D circle C_a), in the index plane) {

6. if(the distance between M_a and M_q is approximate to d_{qa}) {

7 compute the LMS rigid-transformation, namely, T_c, which transforms $\{S_p, S_s, S_a\}$ to $\{M_p, M_s, M_a\}$.

8 for($S_i \leftarrow$ each reference point contained in the scene data set, except S_p, S_s, and S_a) {

9 compute $M_i = T_c S_i$, if the distance between M_i and the model data set is smaller than a threshold, then M_i is labeled "successfully aligned".

}

10 $N_{T_c} \leftarrow$ the counts of the successful alignments with respect to T_c.

}

}

}

}

}

11 output T_c with the largest counts N_{T_c}.

To simplify our derivation, assume that the selected three control points have equal distances d from each other, i.e., they form a regular triangle as shown in Fig. 6. Since the first loop (steps 3-10) will be executed n_m times (where n_m is the number of the data points contained in the model data set), the time complexity is $O(n_m \cdot T_1)$ where T_1 is the time required for a single iteration of the first loop (i.e., steps 3-10). Let n_d be the equivalent number of pixels (in the index plane) for an edge segment of length d. Steps 3 and 4 take $O({n_d}^2)$ because the surface of a sphere whose radius is d roughly contains ${n_d}^2$ points in the index plane. Steps 5-10 will be executed $O(n_d)$ times because only the data points contained in the intersection of the sphere C_s and the model surface will pass the "*if*" statement in step 4, and the intersection is typically a 3D curve (or multiple 3D curves) containing roughly $O(n_d)$ points in the index plane. Hence, T_1 is equal to $O({n_d}^2 + n_d \cdot T_2)$ where T_2 is the time required for steps 5-10. Steps 5 and 6 takes $O(n_d)$ because a circle whose radius is d roughly contains n_d points in the index plane. Steps 7-10 will be executed $O(1)$ times because the intersections of the sphere C_a and the model surface are typically two points (or more isolated points). Hence, T_2 is equal to $O(n_d + 1 \cdot T_3)$ where T_3 is the time for the verification steps (i.e., steps 7-10). Since the verification takes $T_3 = O(n_r)$, the total time required for the DARCES method using three control points is $T^{\{S_p,S_s,S_a\}} = O(n_m({n_d}^2 + n_d(n_d + n_r))) = O(n_m \cdot {n_d}^2 + n_m \cdot n_d \cdot n_r)$.

According to the above analysis, the computation time of the DARCES method using three control points is dominated by three variables, n_m, d, and n_r. The smaller are the three variables, the faster the DARCES method is executed. Here, n_m (the number of the model data points) can not be controlled in our implementation. On the other hand, d (the

edge lenth of the triangle) and n_r (number of the reference points) can be appropriately selected to increase the efficiency of the DARCES method. In the following, to choose an appropriate d, we will give an analysis that how to select an *acceptable minimal triangle* formed by the three control points. We will also indicate that the time complexity becomes less sensitive to n_r if more than three control points are used (see Section 3.3).

3.2.4 Selection of the Acceptable Minimal Triangle

According to the complexity analysis, smaller d will make the DARCES method more efficient. However, if d is too small, the induced rigid-transformation will be very sensitive to noise. In this section, we try to find an acceptable minimal triangle to be used in the DARCES approach if three control points are used. In fact, the acceptable minimal size of the triangle can be theoretically derived if some of the parameters (e.g., the average data-acquisition error, and the error tolerance) are given in advance. Here, we also assume that the selected three control points form a regular triangle as shown in Fig. 6. Let the average position error of the data points (includes both the data acquisition error of the range-finder and the quantization error) be e, and let c be the center of the triangle. Notice that the rigid-transformation T_c computed in our approach is based on the three point correspondences (s_p, m_p), (s_s, m_s) and (s_a, m_a). However, the true positions of each model point m_p, m_s, and m_a are within spheres whose radii are e and centers are m_p, m_s, and m_a, respectively. Hence, for a scene point P whose distance to c is t, the alignment error caused by e will be enlarged to a x. Then, it can be easily derived that x is $\sqrt{3}te/d$. Here, we define the *enlarge ratio* to be $h = x/t$. If we want that the enlarged ratio to be smaller than a threshold H, d should be larger than $d_{min} = \sqrt{3}e/H$, where d_{min} is referred to as the edge length of the acceptable minimal triangle. For example, assume that $e = 1.0$ millimeter (mm) and we hope to control the enlarge ratio to be smaller than $h = 0.1$. Then, $d_{min} = 17.32$ mm. In our work, the size of the triangle is fixed to a small constant, which is determined by the above-mentioned theoretical analysis. Thus, the time required for search can be significantly reduced.

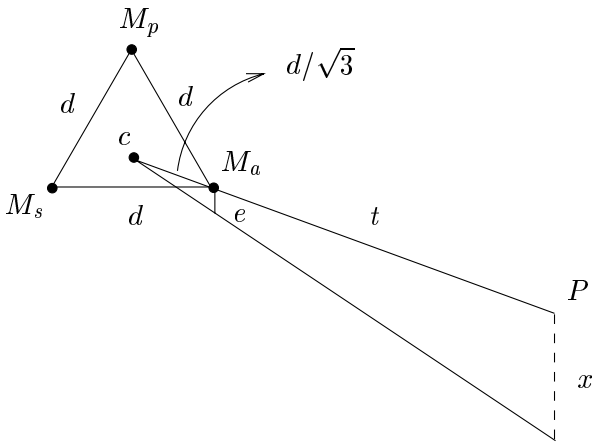


Figure 6: Derivation of the smallest triangle.

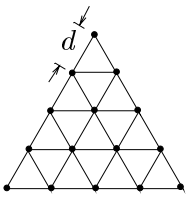


Figure 7: An example of selecting fifteen control points in the index plane.

3.3 Using More Than Three Control Points to Achieve More Efficient Search

In this subsection, we consider the case that if *more than three* control points are used. Assume that there are n_c ($n_c > 3$) control points selected from the scene data set, namely, the first three control points, S_p, S_s, S_a, and the other control points, S_4, S_5, ..., S_{n_c}, respectively. It should be noticed that the first three control points, S_p, S_s, S_a, should also be selected according to the principle that they form an acceptable minimal triangle. In our work, the other control points are usually selected according to that their distances to their neighbors are approximate to that of the edge length of the acceptable minimal triangle. For example, Fig. 7 shows the case that if 15 control points are selected.

3.3.1 Search Range Reduction By Further Using the Rigidity Constraints

First we use the same search procedure introduced in the previous section for the first three control points. Once the first three control points are successfully aligned to some model points M_p, M_s and M_a, respectively, a rigid-transformation T_c can then be computed. Remember that the first constraint of the DARCES approach is to successfully align *all* of the control points on the model data set. To achieve this goal, we use T_c to sequentially transform each of the other control points S_i to new positions, $T_cS_4, T_cS_5, T_cS_6, ..., T_cS_{n_c}$. During the sequence of transformations, once there is a $T_cS_i, (i = 1, 2, ..., n_c)$ violating the alignment constraint (i.e., its distance to the model data set is larger than a given threshold), then it is not required to try other transformations $T_cS_i, T_cS_{i+1}, ..., T_cS_{n_c}$ because we can directly assert that the current transformation T_c can not successfully align all the control points on the model data set. We call this speedup effect the *early jump-out* effect in this paper. Furthermore, in this case, it also require not to further verify T_c by the transformations of all the reference points, $S_{r_1}, S_{r_2}, ..., S_{r_i}$, to new positions. Therefore, the time required for verification can be completely saved.

The number of control points n_c can be chosen to be any number between 3 and n_r. If we use more control points (i.e., a larger n_c), then the probability of "early jump-out" will be higher. Accordingly, in the noiseless case, the fastest way for solving the fully-contained 3D registration (FC3DR) problem is treating all the reference points as the control points in the DARCES procedure (i.e., choosing $n_c = n_r$). In this situation, once some control/reference points (except for the first three control points) can not be successfully aligned using the current transformation, the transformations of all the other control/reference points can then be omitted.

3.3.2 Complexity Analysis: Case of Using More Than Three Control Points

The algorithm of the DARCES method using more than three control points is similar to that of using three ones introduced in Section 3.2.3, except for the following modifications:
1. Select n_c control points, including the first three ones, S_p, S_s, S_a and the others S_4,

S_5,..., S_{n_c}.

7'. [a new step adding between steps 7 and 8]

for(j=4; j$\leq n_c$, j++) {

if($M_j = T_c S_j$ is not successfully aligned on the model data set) {

break; /* that is, break the above loop */

}

}

8. for($S_i \leftarrow$ each reference point contained in the scene data set, except S_p, S_s, S_a, and S_4, S_5, ..., S_{n_c})

The exact complexity analysis of the case that $n_c < n_r$ is difficult because different data sets and different distribution of the control points may result in different early jump-out effects. In general, $\varsigma_1 \subset \varsigma_2$ implies $T^{\varsigma 1_1} < T^{\varsigma_2}$, where ς_1 and ς_2 are two selected sets of control points and T^{ς_1} and T^{ς_2} are the time complexity of using the DARCES approach to solve the FC3DR problem with control point sets ς_1 and ς_2, respectively. Hence, a lower bound of the complexity can be derived if we consider the case that $n_c = n_r$ (i.e., all reference points are selected to be the control points). Basically, the verification steps (i.e., steps 8-10) will not be executed in this case. In an ideal situation, if the early jump-outs always occur while dealing with the fourth control point, then the time complexity is $T^{all} = O(n_m \cdot {n_d}^2 + n_r)$ because only the correct solution can allow the completion of the loop of step 7'. Hence, T^{all} is an lower bound of the complexity of the DARCES approach. In principle, if C is a set of control points in which the first three ones are $\{S_p, S_s, S_a\}$, then $T^{\{S_p,S_s,S_a\}} \leq T^{\varsigma} \leq T^{all}$ (where $T^{\{S_p,S_s,S_a\}}$ has been derived in Section 3.2.3). It is worth noting that that n_r (the number of reference points) is no longer a critical term in T^{all}. This implies an interesting fact that the more control points used, the less influence of n_r on the time complexity. Therefore, it is not required to reduce n_r (as mentioned in Section 3.2.3), and the only critical term is d (the length of the triangle formed by the first three control points). [3]

[3]However, the exact form of T^{ς} for $4 \leq n_c \leq n_r$ is difficult to be written as an analytical form, because

Unfortunately, while the strategy of using as many control points as possible is better for solving the FC3DR problem, it is not always better for the partially-overlapping 3D registration problem. In principle, to solve the partially-overlapping 3D registration problem, it is required that all the control points lie on the overlapping region of the two data sets. However, the more control points used, the more likely that some of the control points will fall outside the overlapping region. Hence, it is an important issue to choose a good number of control points having good distribution. In general, determining the optimal number of control points is a difficult problem. Also, the optimal configuration of the control points depends on the size and the shape of the overlapping region of the two data sets, and thus, is quite data dependent. In our approach, we use a random-selection strategy to select of the control points, which will be introduced in Section 4.

3.4 Discussion

In a theoretical point of view, the DARCES algorithm is an exhaustive-search method for solving the FC3DR problem. That is, all possible cases of the successful alignments of two data sets are verified. Due to the exhaustive nature, the DARCES method has the following two advantages: (i) it can promise to obtain the solution for the noiseless case if there is a unique rigid-transformation which allows an overlapping area larger than the given threshold, and (ii) it can also be applied even the data sets have no prominent/salient local features.

On the other hand, the DARCES method is a highly-efficient exhaustive-search method. It is because that many useful constraints generated by the rigidity-relation of the data points are sufficiently used including (i) the constrained search range of the possible corresponding positions of the second and the third control points (Section 3.2.1), (ii) the use of an acceptable minimal triangle (Section 3.2.4) for the first three control points, and (iii) the use of more than three control points for the early jump-out of the verification process (Section 3.3). Compared to the optimization strategy proposed in [6], the DARCES

it depends on the shape and the distributions of the data sets.

method does not search in the parameter space without any constraints. Instead, it only search the parameters which make the two data sets aligned.

The use of the DARCES method can be very flexible as discussed in the following.
(1) If some initial estimations of the rigid-transformation are available, it can be easily utilized in the DARCES approach. In principle, we can use it to restrict the search range of the primary control point. First, the scene data can be transformed in advance using the initial estimation of the rigid-transformation. Then, the search range of the primary control point can be restricted within a neighbor region.
(2) The DARCES method can be applied for either the original 3D data or the extracted feature data (e.g., the 3D points in the invariant curves [7][21]), contained in range images. It is obvious that the DARCES method can also be easily extended to use some feature attributes associated with each 3D data point, e.g., 3D curvature or image luminance.

As a search strategy, the DARCES method can also be explained from the *interpretation tree* (IT) point of view. The IT was used in many model-based recognition applications [2][15][22]. In particular, the DARCES method is one of the *prediction and verification of hypotheses* (PVH) strategies [15] which can be used to prune an IT. A PVH strategy can reduce the search complexity of an IT by pruning both the breadth and the height of the tree (see Faugeras [15]). For every path explored from the root in the interpretation tree corresponding to a partial recognition $((M_1, S_{i_1}), (M_2, S_{i_2}), ..., (M_k, S_{i_k}))$ (where M and S denote model and scene data sets, respectively), we compute the best rigid transformation T_k from model to scene. Then only those unmatched scene primitives that are sufficiently close to $T_k(M_{k+1})$ should be considered; and hence, the breadth of the tree can be reduced. On the other hand, when we compute the best rigid transformation, we also obtain a residual error e_k, we do not want to generate recognition sequences R_n with too large an error e_n, so we want to be able to stop the search when e_k grows larger than some threshold; and hence, the heights of the tree can be reduced.

Although the DARCES method is one of the PVH strategies, it has the following distinctive characteristics:

1. First, the DARCES method is designed for solving the 3D registration problem, in particular, without using local features. The exact constrained search ranges of the secondary and the auxiliary control points can both be precisely derived from a geometrical point of view. We can then approximate these search ranges with square grids in the index plane (because range images are used) to simplify the implementation. We will also introduce in Section 5 that, since the data distribution is regular if local features are not used, the DARCES approach can be easily incorporated in a coarse-to-fine search structure which can further considerably increase the search speed.

2. Second, we have indicated that even the PVH strategy is used to prune a search tree, efficient search can not be achieved if an inappropriate set of control points are selected (and hence an inappropriate search tree may be generated), especially for the case that no local attributes are used. Because the constrained search regions can be precisely obtained, it allows us to derive that the critical terms of the search complexity are d^2 and n_r, where d is the edge length of the triangle formed by the first three control points and n_r is the number of the reference points, as introduced in Section 3.2.3. Furthermore, we have also indicated that n_r is not critical anymore if more control points are used, as introduced in Section 3.3.2. Therefore, how to select the acceptable minimal triangle to be used in the DARCES approach (i.e., to minimize d) is the most important issue for selecting an appropriate set of control points, and has been introduced in Section 3.2.4.

3. Third, to solve the partially-overlapping 3D registration problem, the DARCES method is incorporated in a RANSAC scheme. In this scheme, the first-level vertices in the IT are selected randomly, instead of selecting them sequentially. Such a strategy makes the search more fast while the correct solution can still be reliable solved. We will introduce in Section 4 that, using this strategy, only a few random trials are required to find a successful partial search.

4 RANSAC-Based DARCES Approach for Partially-Overlapping Case

In the last section, we have introduced the DARCES method for solving the FC3DR problem. In this section, to solve the general partially-overlapping 3D registration problem, we integrate the DARCES method with a robust estimation method, the RANSAC scheme [16]. RANSAC and LeMedS were proposed as a general robust estimation method for surface or model fitting. Robust estimation means that model fitting is not influenced by *outliers*. Accordingly, they can also be applied to solve the partial-matching problem. In the LeMedS strategy, a sorting procedure is required to sort the alignment errors of all the reference points. On the other hand, in the RANSAC scheme, the sorting procedure is replaced with the counting of the number of *inliers*. Hence, the computation complexity of the RANSAC scheme is lower. However, in practice, our method can be modified to use different types of robust estimators. The RANSAC-based DARCES approach starts by randomly selecting a primary control point from the scene data set. Notice that only the primary control point S_p is selected randomly, while the others are determined based on S_p and the length of the minimal triangle, d_{min}, as introduced in Section 3.3. Once the control points are selected, the DARCES procedure is performed to find possible alignments of these two data sets. If the rigid transformation found by the DARCES procedure has overlapping number larger than a threshold, then that transformation is regard as the solution of our 3D registration task; otherwise, we select another primary point randomly from the scene data set, and perform the above procedure again, until it successfully finds a rigid transformation having a sufficiently large overlapping number.

The RANSAC-based DARCES algorithm is a modification of the DARCES one using more than three control points. Those modifications are:

1. Randomly select the primary control point, S_p. The other control points are then selected according to the principle that their distances to their neighbors are approximate to d_{min}.

11. If the largest counts of N_{T_c} is larger than a pregiven threshold Ω, then output T_c and stop;

Else, goto 1.

Step 11 gives a *stopping rule* of the RANSAC-based DARCES algorithm. The threshold Ω is proportional to the allowed minimal overlapping ratio of the two partially-overlapping data sets which has to be pre-given by users.

A statistical analysis of the required number of random trials is given below to show that our method can solve the partially-overlapping 3D registration problem with only a few random trials. First, consider the case that three control points are used. To simplify our analysis, we assume that the *overlapping region* (OR) in the index plane of the scene data set is a square region whose edge length is l as shown in Fig. 8. Suppose the shape and size of the triangle used in our approach is fixed, all of the three control points will lie on the overlapping region if the primary control point falls into the *eroded overlapping region* (EOR), as shown in Fig. 8. Assume that the number of data points contained in the OR of the scene data set is n_o. Then, $r = n_o/n_s$ is referred to as the *overlapping ratio* of OR, where n_s is the number of the scene data points. From Fig. 8, the ratio of the area of the EOR to that of the OR can be shown to be $(l-d)^2/l^2$. Therefore, in a single random selection, the probability that the primary control point lies on the EOR is $p = r \cdot (l-d)^2/l^2$. Hence, the expected times of random trials is $E = 1 \cdot p + 2 \cdot (1-p)p + 3(1-p)^2 p + ... = 1/p$. Similar derivations can also be used for the case of using more than three control points. For instance, consider the case of using 15 control points. If the edge length of the triangle formed by the first three control points is 17.32 mm (which is the same as the ones given in Section 3.2), then the edge length formed by the 15 control points is 69.28 mm (see Fig. refFigFifteen). Assume that the edge length of the overlapping square is 120 mm, and the overlapping ratio is 0.75. Then, the expected times of random trials is 7.46.

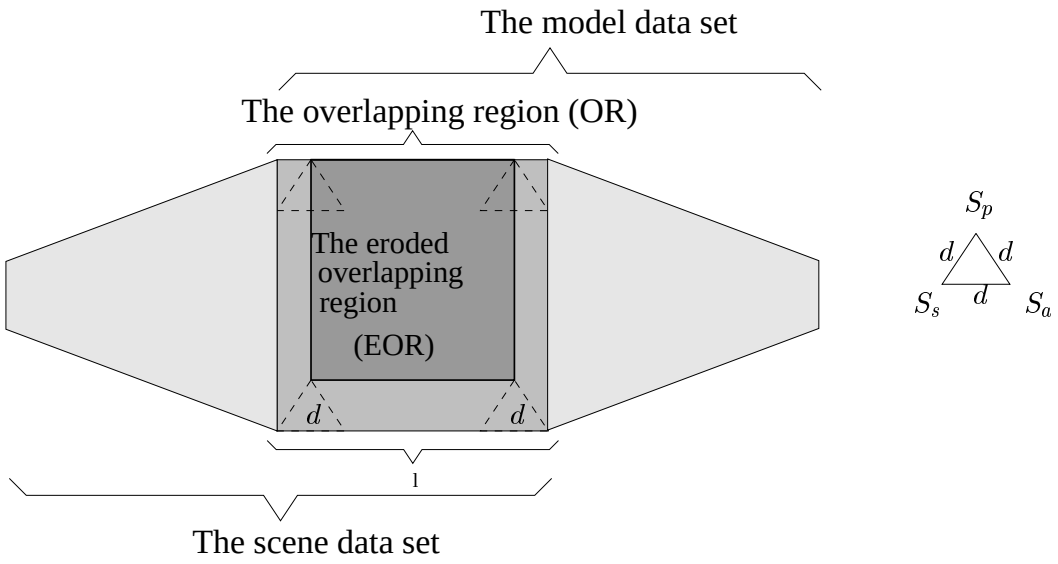


Figure 8: If the primary control point is selected to lie on the eroded overlapping region, then all of the three control points will lie on the overlapping region.

5 Coarse To Fine Scheme: Three-Step Algorithm

The DARCES procedure can reduce the time required by the exhaustive search for the FC3DR problem by using the rigidity constraint. However, due to its exhaustive-search nature, the computation time is difficult to be further reduced without using other constraint. Consequently, if we want to further speed up the DARCES method without using other constraint, the restriction of *exhaustive* search may have to be appropriately loosened. That is, by allowing not to search all the possible alignments, the speed can be considerably increased (hopefully, without affecting the outcome of the search in most cases). If no local features are used, the regularity of the data distribution makes the DARCES approach be easily incorporated in a coarse-to-fine search structure. The speedup strategy we have adopted is the *three-step algorithm*, which is popular in the field of image/video coding. The three-step algorithm is an n level coarse-to-fine method, where n is typically (but not restricted to be) three. In our approach, three-step algorithm is used to further constrain the search ranges of the primary control point, S_p. First, the correspondences of S_p are searched on the grids of the coarsest level (i.e., level 1) in the index plane, as shown in Fig. 9. The best correspondence obtained from level 1 is then

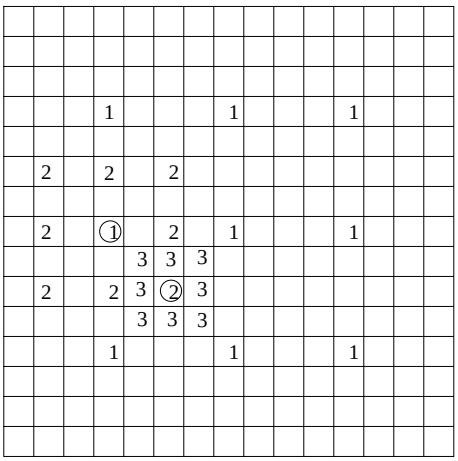


Figure 9: The three-step algorithm is used in the DARCES approach to find the corresponding positions of the primary control point in a coarse to fine manner.

used as an initial estimate for the next level. In level 2, the search range for possible correspondences of S_p can be restricted to a local region around the initial estimate obtained from level 1. Then, the best correspondence of S_p obtained in level 2 can be used as an initial estimate for searching the best correspondence of M_p in level 3. Notice that in the three-step algorithm, only S_p is searched in a coarse-to-fine manner. Once S_p is hypothesized to correspond to a model point, the correspondence candidates of all the other control points are searched in the finest level.

Notice that when combining the RANSAC scheme and the three-step algorithm, a sequence of increasing thresholds, Ω_1, Ω_2,...Ω_k were respectively given in advance for the coarse to fine levels. If the overlapping number computed at a coarser level is smaller than the threshold of this level, then we give up the search in the finer levels and immediately start another random trial. Compared to the strategy that only a single threshold is used for the finest level, this strategy can make the combination of the RANSAC scheme and the three-step algorithm more efficient.

6 Experimental Results

Figs. 10(a) and (b) show two range data sets of an object taken from two different views. Their viewing angles differ by about 30^o. Each of them contains roughly 3650 data points. The RANSAC-based DARCES method is used to register the data sets contained in the two range images. The three-step algorithm is also used to further speed up the RANSAC-based DARCES approach. The average registration error in this experiment is 0.21 millimeters (mm). In our experiment, 15 control points are used. Only two random trials are required for finding the correct registration, and the CPU time is 5.85 seconds including both the coarse registration and the fine registration (using a SGI O^2 workstation). Notice that the computation time is measured for the entire 3D registration task, rather than treating some procedures as off-line processes (such as feature-extraction and feature-organization in a feature-based approach). Fig. 10 (c) shows the integration result of the two overlapped data sets.

Figs. 11(a) and (b) show two intensity images of a toy taken from different views. Their viewing angles also differ by about 30^o. Each of them contains roughly 6000 data points. The range images are taken from the same views as those for taking the intensity images of Figs. 11(a) and (b). Then, the two range images are registered and integrated into a single data set. The CPU time taken for registration is 7.54 seconds, and the registration error is 0.78 mm. The texture-mapped image obtained by mapping and blending the intensity images onto the integrated data set is shown in Fig. 11(c).

Figs. 12(a) and (b) show two range data sets of a model head taken from different view points. Their viewing angles differ by about 45^o. Each of them contains roughly 4200 data points. The registered and integrated 3D data set is shown in Fig. 12(c). The CPU time for registration is 58.61 seconds because it takes 20 random trials before obtaining the good transformation, and the average registration error is 1.47 mm. Figs. 12(d) and (e) show the shaded and the texture-mapped images of Fig. 12(c), respectively.

In Fig. 13 (a) and (b), range images of a pair of fruits are taken from two different views. Fig. 13 (a) is the right view and Fig. 13 (b) is the left view. Their viewing angles differ

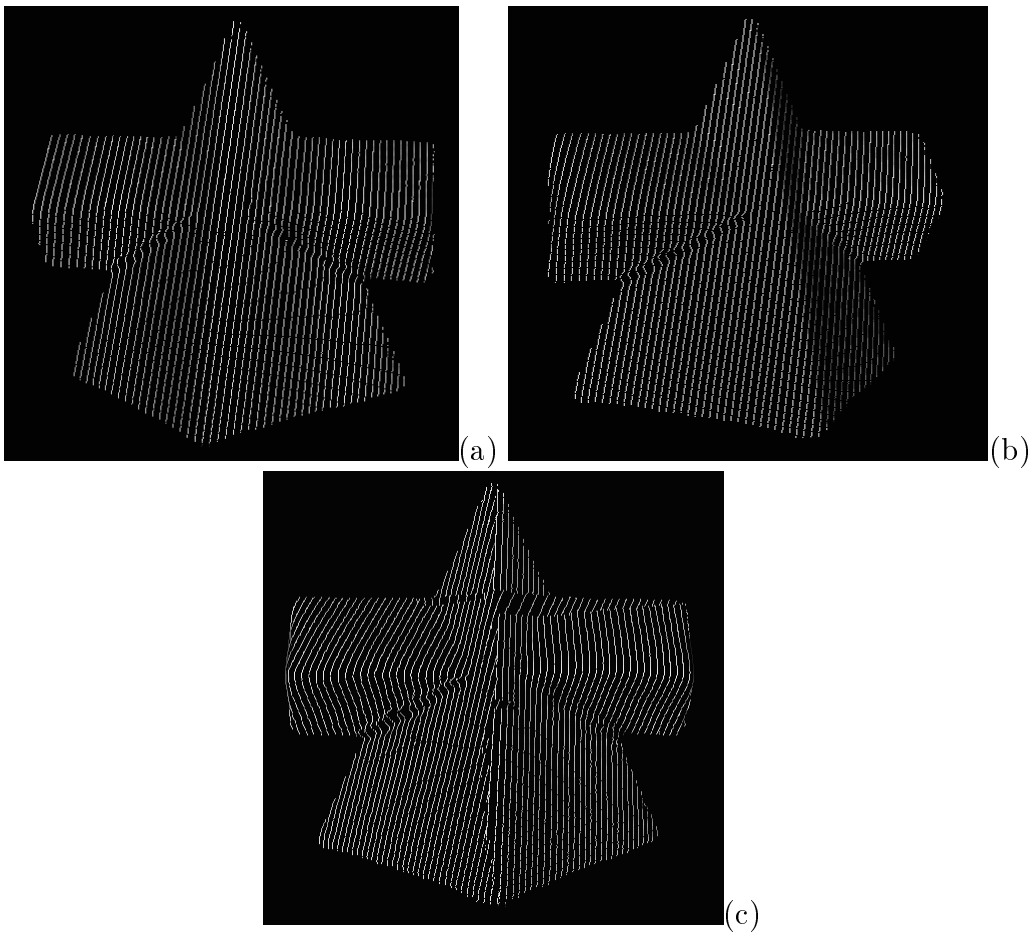


Figure 10: (a) and (b) show the range data of an object taken from two different view points. (c) shows the 3D data set obtained by registering and integrating the 3D data sets of (a) and (b).

by about 30^o. Each of them contains roughly 2400 data points. Notice that in this case the two range data sets contain no good local features. Hence, in general, it is difficult to solve this 3D registration problem if we use a feature-based method. Nevertheless, by using the RANSAC-based DARCES approach, the two data sets can be successfully registered. Fig. 13 (c) shows the registered data set which takes only 3.95 seconds with two random trials.

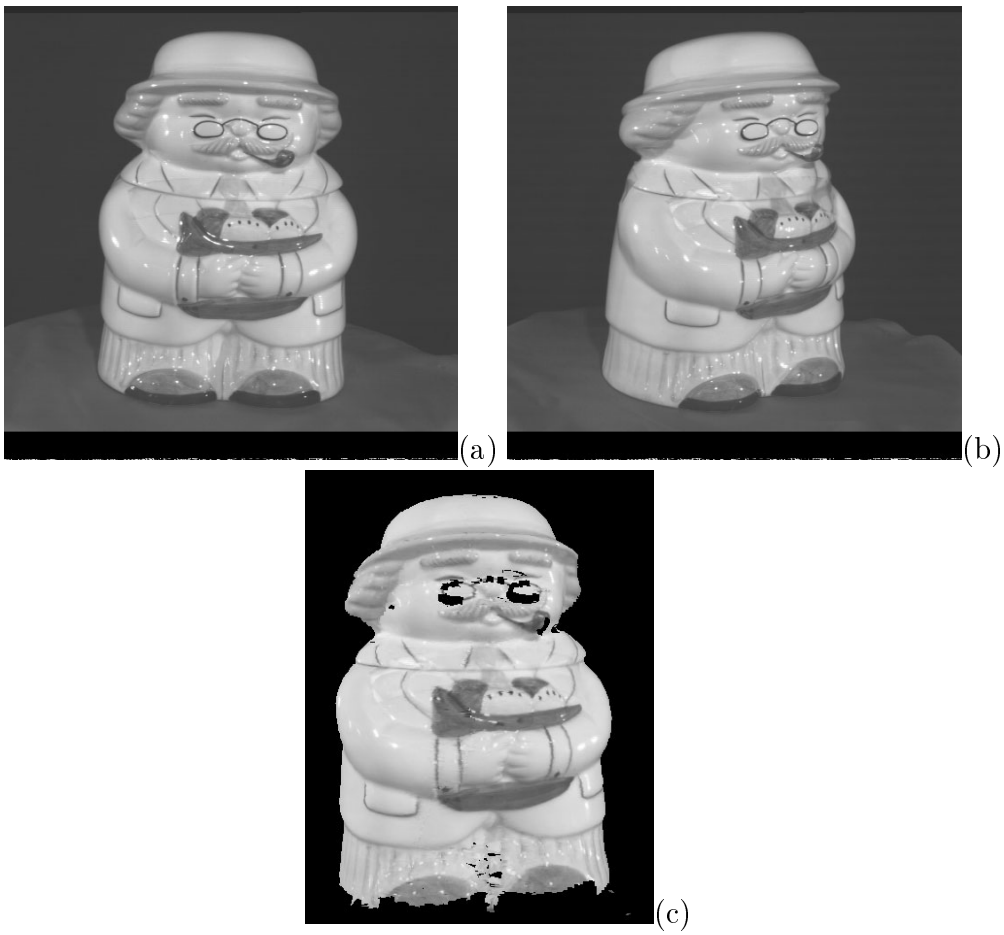


Figure 11: (a) and (b) are the intensity images of a toy taken from two different views where their range data are taken. The two range data sets are then registered and integrated into a single data set. (c) shows the texture-mapped images by mapping and blending the intensity images onto the integrated 3D data set.

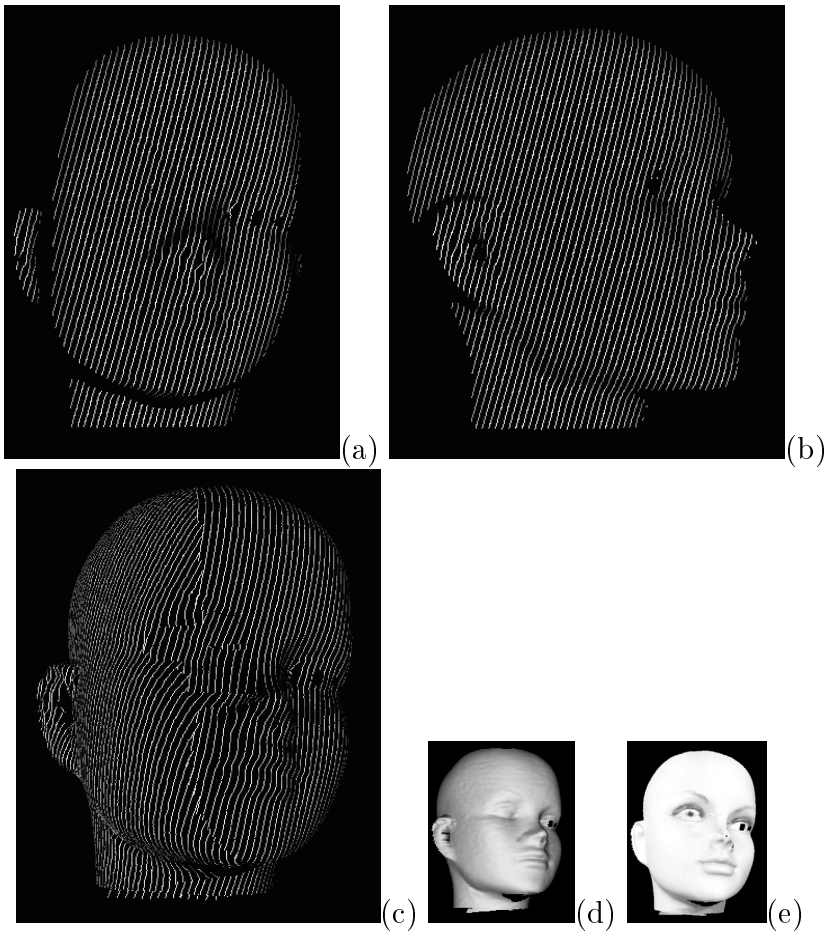
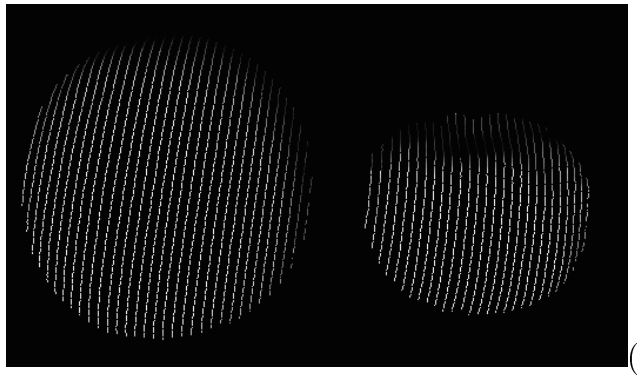
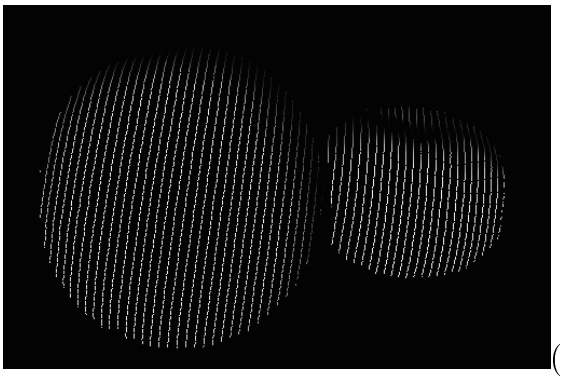


Figure 12: (a) and (b) show the range data of a model head taken from different view points. (c) shows the 3D data set obtained by registering and integrating the 3D data sets of (a) and (b). (d) shows the shaded image of (c), and (e) shows the texture-mapped image of (c).


(a)


(b)

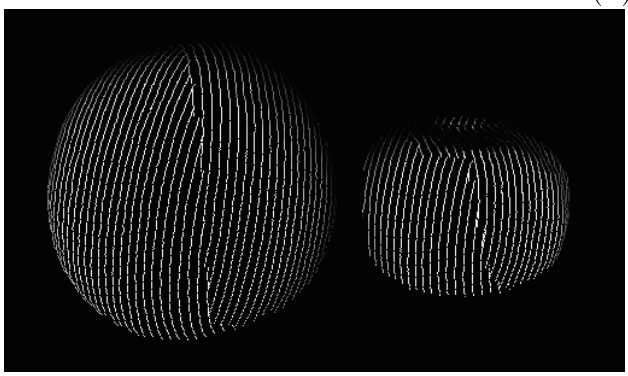

(c)

Figure 13: (a) and (b) show the range data of a pair of fruits taken from different view points. (c) shows the 3D data set obtained by registering and integrating the 3D data sets of (a) and (b).

7 Conclusions and Discussion

The existing techniques developed for solving the 3D registration problem usually have one of the following limitations:

1. It requires a good initial estimation of rigid-transformation parameters between the two data sets [5][10].
2. It can only be used if the data sets contain sufficient local features [11][17][24].

Most of the existing techniques for solving the partially-overlapping 3D registration problem have either one of the following limitations:

1. It requires a good initial estimate of the rigid transformation between the two data sets. [5][10].
2. It can only be used if the data sets contain sufficient local features [11][24].

In this paper, we propose the RANSAC-based DARCES approach which has neither of the above two limitations. Also, our method is faster than most of the existing methods which do not require initial estimations. Our approach simply treats the 3D registration problem as a partial-matching problem, and uses the rigidity constraint among some pre-selected control points to restrict the search range for matching. Although some approaches have used rigidity constraints to facilitate the matching processes [11][12][25] too, our approach is the first one to use it in a prediction-and-verification of hypotheses way among some randomly-selected control points, and shows that the 3D registration problem can be solved in a relative low-order computation time by carefully using all of the constraints provided by the rigidity. In addition, we have indicated that by appropriately selecting the number and the distribution of the control points, the computation time can be greatly reduced. Therefore, we have shown how to determine the acceptable minimal triangle formed by the first three control points, and how to use the additional control points to speed up the search process. Finally, we integrate our method with the three-step algorithm, and show that the computation time can be further reduced while the registration can be still reliable with the help of the RANSAC scheme. Although the

principle used in our approach is simple and easy-to-implement, to the best knowledge of the authors, no one have adopted similar ideas to solve the 3D registration problem before. Experiments have demonstrated that our method is efficient and reliable for registering partially-overlapping range images.

References

[1] *MPEG standard draft ISO-IEC/JTC1 SC29 on 22*, November, 1991.

[2] F. Arman and J. K. Aggarwal, "Model-Based Object Recognition in Dense-Range Images – A Review,", ACM Computing Survey, Vol. 25, No. 1, pp. 125-145, 1993.

[3] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-Square Fitting of Two 3-D Point Set," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, pp. 698-700, 1987.

[4] R. Bergevin, D. Laurendeau, and D. Poussart, "Registering Range Views of Multipart Objects," *Computer Vision and Image Understanding*, Vol. 61, No. 1, pp. 1-16, 1995.

[5] P. J. Besl and N. D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, pp. 239-256, 1992.

[6] G. Blais and M. D. Levine, "Registering Multiview Range Data to Create 3D Computer Objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 8, pp. 820-824, 1995.

[7] C. S. Chen, Y. P. Hung, and J. L. Wu, "Extraction of Corner-Edge-Surface Structure from Range Images Using Mathematical Morphology," *IEICE Transactions on Information and Systems*, Vol. E78-D, Special Issue on Machine Vision Applications, pp. 1636-1641, 1995.

[8] C. S. Chen, Y. P. Hung, and J. L. Wu, "A 3D Range Data Acquisition System Combining Laser Lighting and Stereo Vision," *Journal of the Chinese Institute of Electrical Engineering*, pp. 157-168, 1996.

[9] C. S. Chen, Y. P. Hung, and J. L. Wu, "Model-Based Object Recognition with Range Images by Combining Morphological Feature Extraction and Geometric Hashing," *Proceedings of the 13th International Conference on Pattern Recognition*, Vienna, Austria, pp. 565-569, 1996.

[10] Y. Chen and G. Medioni, "Object Modeling by Registration of Multiple Range Images," *Image and Vision Computing*, Vol. 10, No. 3, pp. 145-155, 1992.

[11] C. S. Chua, "3D Free-Form Surface Registration and Object Recognition", *International Journal of Computer Vision*, 17, pp. 77-99, 1996.

[12] C. Dorai, J. Weng, and A. K. Jain, "Optimal Registration of Multiple Range Views," *Proceedings of the International Conference on Pattern Recognition*, pp. 569-571, 1994.

[13] C. Dorai, G. Wang, A. K. Jain, and C. Mercer, "From Images to Models: Automatic 3D Object Model Construction from Multiple Views," *Proceedings of the International Conference on Pattern Recognition Vol. 1*, pp. 770-774, 1996.

[14] D. Eggert, A. W. Fitzgibbon, and R. B. Fisher, "Simultaneous Registration of Multiple Range Views for Use in Reverse Engineering," *Proceedings of International Conference on Pattern Recognition*, Vienna, Austria, pp. 243-247, 1996.

[15] O. Faugeras, *Three-Dimensional Computer Vision*, The MIT Press, London, England, 1993.

[16] M. A. Fischler and R. C. Bolles, "Random Sample Consensus, A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Communications of the ACM*, Vol. 24, No. 6, pp. 381-395, 1981.

[17] A. Gueziec and N. Ayache, "Smoothing and Matching of 3D Space Curves," *International Journal of Computer Vision*, 12:1, pp. 79-104, 1994.

[18] K. Higuchi, M. Hebert, and K. Ikeuchi, "Building 3-D Models from Unregistered Range Images," *Graphical Models and Image Processing*, Vol. 57, No. 4, pp. 315-333, 1995.

[19] Y. Lamdan, J. T. Schwartz, and H. J. Wolfson, "Affine Invariant Model-Based Object Recognition," *IEEE Transactions on Robotics and Automation*, Vol. 6, pp. 578-589, 1990.

[20] T. Masuda and N. Yokoya, "A robust Method for Registration and Segmentation of Multiple Range Images," *Computer Vision and Image Understanding*, Vol. 61, pp. 295-307, 1995.

[21] O. Monga, S. Benayoun, and O. D. Faugeras, "Using Partial Derivatives of 3D Images to Extract Typical Surface Features," *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, Urbana Champaign, IL, 1992.

[22] M. Suk and S. M. Bhandarkar, *Three-Dimensional Object Recognition from Range Images*, Springer Verlag, 1992.

[23] M. Soucy and D. Laurendeau, "A General Surface Approach to the Integration of a Set of Range Views," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, pp. 344-358, 1995.

[24] F. Stein and G. Medioni, "Structural Indexing: Efficient 3-D Object Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligences*, Vol. 14, No. 2, pp. 125-145, 1992.

[25] J. P. Thirion, "New Feature Points Based on Geometric Invariants for 3D Image Registration," *International Journal of Computer Vision*, 18(2), pp. 121-137, 1996.

[26] G. Turk and M. Levoy, "Zippered polygon Meshes from Range Images," *Computer Graphics Proceedings, Annual Conference Series, SIGGRAPH*, pp. 311-318, 1994.