A Fast Automatic Method for Registration of Partially-Overlapping Range Images

Chu-Song Chen, Yi-Ping Hung

Institute of Information Science
Academia Sinica, Taipei, Taiwan
email: hung@iis.sinica.edu.tw

Jen-Bo Cheng

Department of Comp. Sci. and Info. Eng.
National Taiwan University
Taipei, Taiwan

Abstract

A popular approach for 3D registration of partially-overlapping range images is the ICP (iterative closest point) method and many of its variations. The major drawback of this type of iterative approaches is that they require a good initial estimate to guarantee that the correct solution can always be found. In this paper, we propose a new method, the RANSAC-based DARCES (data-aligned rigidity-constrained exhaustive search) method, which can solve the partially-overlapping 3D registration problem efficiently and reliably without any initial estimation. Another important characteristic of our method is that it requires no local features in the 3D data set. An extra characteristic is that, for the noiseless case, the basic algorithm of our DARCES method can guarantee that the solution it finds is the true one, due to its exhaustive-search nature. Even with the nature of exhaustive search, its time complexity can be shown to be relatively low. Experiments have demonstrated that our method is efficient and reliable for registering partially-overlapping range images.

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. In the past, a popular type of approaches to solve the 3D registration problem is the *iterative approach* [1][3]. Iterative approaches have the advantages that they are fast and easy-to-implement. However, the drawbacks are that (i) they require a good initial estimate 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* [4][7]. 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 (see Fig. 11), and (ii) a large percentage of computation time is usually spent on preprocessing, which include the extraction of invariant features [4][7] and the organization of the extracted feature-primitives (e.g., sorting [4], hashing[6], etc.). In this paper, a new approach is introduced for 3D registration of range images, which requires neither complex preprocessing nor initial transformation.

This paper is organized as follows. Section 2 presents our new approach, the DARCES (Data-Aligned Rigidity-Constrained Exhaustive Search) method, for solving the fully-contained 3D registration problem. Section 3 then extends the basic procedure to the RANSAC-based DARCES method for solving the partially-overlapping 3D registration problem. Section 4 shows some experimental results, and Section 5 gives some conclusions and discussions.

2 DARCES Method for fully-contained 3D Registration

Given two data sets, namely, the *scene data set* and the *model data set*, the problem of 3D registration is to find the 3D rigid transformation that can make the overlapping region of the two data sets as large as possible. In this section, we first 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. First, we select some *reference points* from the scene surface, as shown in Fig. 1(a). For example, we can select the ref-

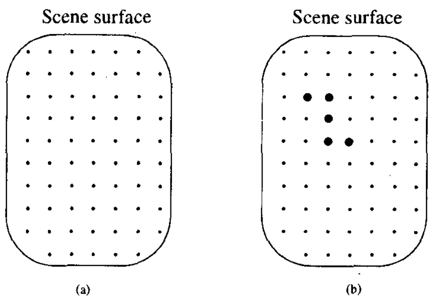

Figure 1: (a) Selection of the *reference points* in the scene surface. (b) Selection of a set of *control points* from these reference points.

erence points by uniform sampling from the indexing grids of the range images, or we can use all the data points contained in the scene data set as the reference points (which will be less efficient). In the subsequent processing, a set of *control points* (at least three) are selected from these reference points, as shown in Fig. 1(b).

2.1 Use of Three Control Points

In this section, we consider only the case where three control points are used. The three selected control points will be called the *primary point* S_p, the *secondary point* S_s, and the *auxiliary point* S_a, respectively.

First, in the model data set, consider the possible corresponding points of the primary point S_p. Without the knowledge of feature attributes, every 3D point contained in the model data set can be a 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. Notice that our method can be easily extended so that available feature attributes associated with each 3D data point, (e.g., 3D curvature or image luminance) can be used to reduce the number of possible correspondence.

Now, suppose S_p is hypothesized to correspond to a model point M_p. Then, in the model data set, we try to find some candidate points corresponding to the secondary point S_s using the rigidity constraint. Let the distance between S_p and S_s be 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 the corresponding model point of the primary point S_p is hypothesized to be M_p, 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.

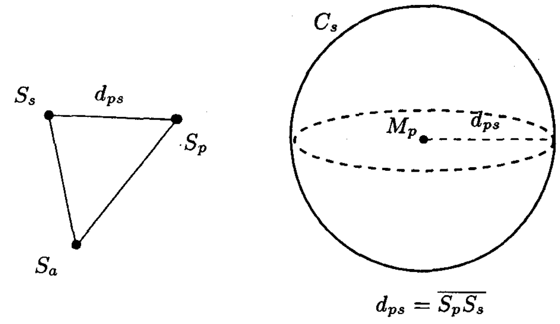

Figure 2: (a) The triangle formed by the three control points selected from the scene data set. (b) The search region (in the model data set) for finding the correspondence of the secondary control point is restricted to the surface of a sphere.

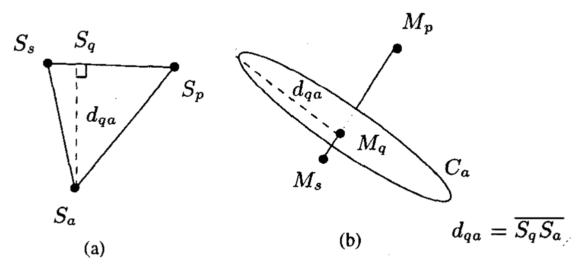

Figure 3: (a) S_q is the orthogonal projection of S_a to the line segment $\overline{S_pS_s}$. (b) The search region in the model data set for finding the correspondence of the auxiliary control point is restricted to the contour of a 3D circle.

After the corresponding model points of both the primary point S_p and the secondary point S_s are hypothesized to be M_p and M_s, respectively, we can then consider the constrained search range of the auxiliary point S_a. The candidates of M_a, the corresponding model pont of the auxiliary scene data point S_a, should be found (if M_a exists) 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 must 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 $\overline{\mathbf{p}M_q}$ is perpendicular to $\overline{M_pM_s}\}$.

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

For each possible three-point correspondence, an overlapping number can be computed as described above. Then, the rigid transformation with the largest overlapping number is selected as the solution of our registration task. In general, the registration result provided by the DARCES method is accurate to some extent. This result can be further refined by some well-known iterative procedure [1][3]. In our method, the ICP method [1] is used for refinement. When implementing the DARCES algorithm, direct search in the 3D space on the surface of a sphere or on the boundary of a circle may not be a trivial task. Therefore, we exploit the fact that a range image can be indexed by projecting the 3D data points onto an *index plane.* With some deliberation, it is not hard to find that the search of the matching candidates of the secondary and the auxiliary points can be restricted within some squared search regions in the index plane, and hence, the implementation can be considerably simplified.

Next, consider the triangle formed by the three control points. 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 to be too small, the computed rigid transformation will be very sensitive to noise. Hence, how to determine an *acceptable minimal triangle* for the DARCES procedure is an important issue. Here, we assume that 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 (see Fig. 4). For a scene point P whose distance to c is t, the alignment error caused by e will be enlarged to x. Here, we define the *enlargement ratio* to be $h = x/t$. It can be shown that if we want the enlargement 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 enlargement

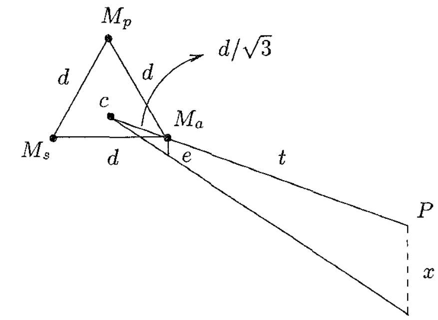

Figure 4: Illustration for determining the acceptable minimal triangle.

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 analysis. Thus, the time required for search can be significantly reduced.

Let n_d be the equivalent number of pixels (in the index plane) for an edge segment of length d_{min} in the 3D space. The time complexity of the DARCES method using three control points can be shown to be $O(n_m(n_d{}^2 + n_d(n_d + n_r))) = O(n_m \cdot n_d{}^2 + n_m \cdot n_r \cdot n_d)$, where n_m is the number of the model data points and n_r is the number of the reference points chosen from the scene data points. In practice, n_r is usually chosen to be a small fraction of n_m, and n_d is typically even smaller. For example, for the case shown in Fig. 9, $n_m = 6000$, $n_r = 360$, and $n_d = 8$.

2.2 Use of More Than Three Control Points

In this section, we consider the case where *more than three* control points are used. Let the n_c ($n_c > 3$) control points selected from the scene data set be denoted by S_p, S_s, S_a (the first three control points), and S_4, S_5, ..., S_{n_c} (all the other control points). Here, the search procedure is similar to the one described in Section 2.1 for the case of using three control points, except that all the n_c control points (instead of only three control points) have to find their possible candidates before computing the overlapping number – which is a relatively time-consuming process having complexity of $O(n_r)$. That is, once the first three control points, S_p, S_s, and S_a, find their possible candidates, M_p, M_s, M_a, during the search process, the rigid transformation T_c computed with those three possible matches will be used to sequentially transform each of the remaining control point, $S_i, i = 4, 5, ..., n_c$, to a new position, $T_c S_i$, and check if

T_cS_i satisfies the alignment constraint (i.e., if its distance to the model data set is smaller than a given threshold). As long as any one of the remaining control points does not satisfy the alignment constraint, we jump out immediately and search for another new set of candidate matches for the control points. By using this early jump-out strategy, the time for verifying a T_c with all the n_r reference point can be largely 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. In the noiseless case, treating all the reference points as the control points in the DARCES procedure (i.e., choosing $n_c = n_r$) will be the fastest way for solving the fully-contained 3D registration (FC3DR) problem.

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 the following.

3 RANSAC-based DARCES Method for Partially-Overlapping 3D Registration

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 [5]. The RANSAC-based DARCES approach starts by randomly selecting a primary control point from the scene data set. In our approach, the secondary and the auxiliary control points are selected such that they approximately form an acceptable minimal triangle described at the end of Section 2.1. The other control points are selected around

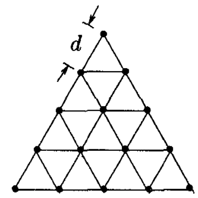

Figure 5: An example of selecting fifteen control points in the index plane.

the acceptable minimal triangle such that they gradually form a larger triangle. For example, Fig. 5 shows the case where 15 control points are selected. 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.

3.1 Statistical Analysis of the Required Random Trials

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. 6. 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. 6. 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. 6, 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 $E = 1\cdot p+2\cdot(1-p)p+3(1-p)^2p+... = 1/p$. 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

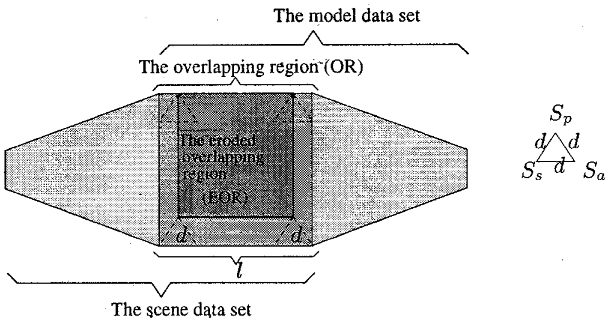

Figure 6: If the primary control point selected falls within the eroded overlapping region (EOR), then all of the three control points will fall within the overlapping region (OR).

is 17.32 mm (which is the same as the ones given in Section 2.1), then the edge length formed by the 15 control points is 69.28 mm. 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.

3.2 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). 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. 7. The best correspondence obtained from level 1 is then 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 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. This strategy can make the combination of the RANSAC scheme and the three-step algorithm more efficient.

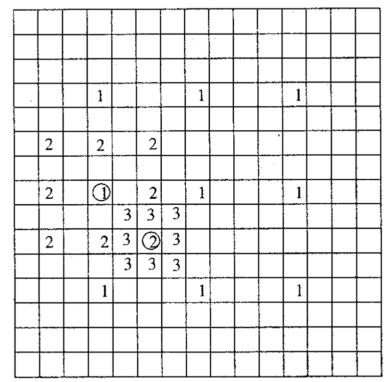

Figure 7: The three-step algorithm can be used in the DARCES approach to speed up the searching for the correspondence of the primary control point in a coarse-to-fine manner.

4 Experimental Results

Figs. 8(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-

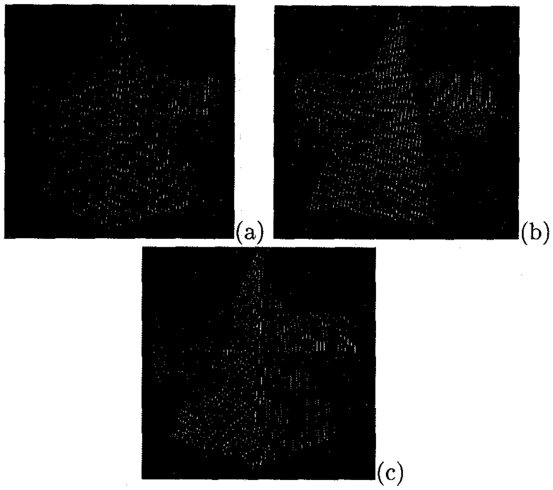

Figure 8: (a) and (b) show the range data of an object taken from two different view points. (c) shows the 3D data set obtained by registering and integrating the 3D data sets of (a) and (b).

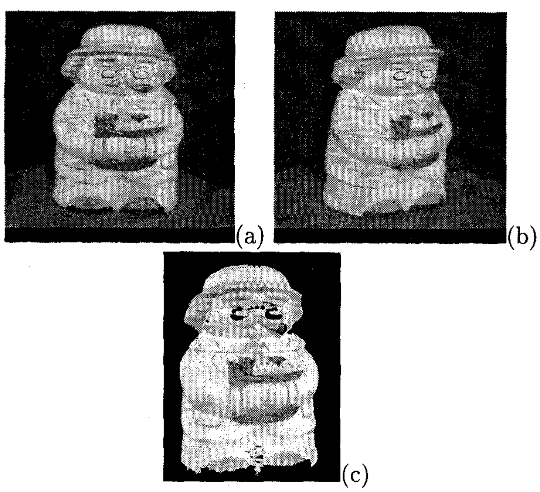

Figure 9: (a) and (b) are the intensity images of a toy taken from two different views where their range data are taken. The two range data sets are then registered and integrated into a single data set. (c) shows the texture-mapped images by mapping and blending the intensity images onto the integrated 3D data set.

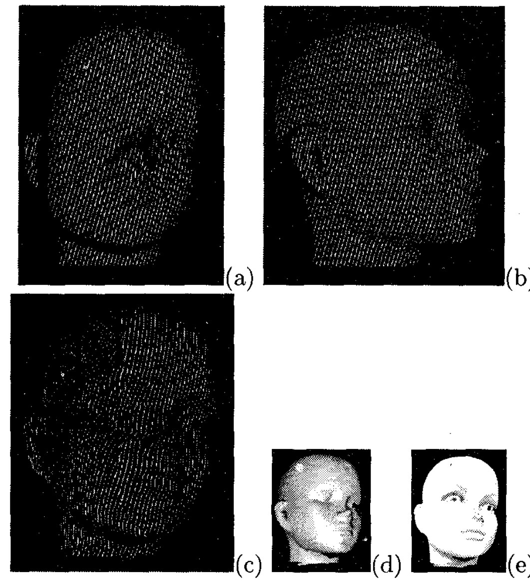

Figure 10: (a) and (b) show the range data of a model head taken from different view points. (c) shows the 3D data set obtained by registering and integrating the 3D data sets of (a) and (b). (d) shows the shaded image of (c), and (e) shows the texture-mapped image of (c).

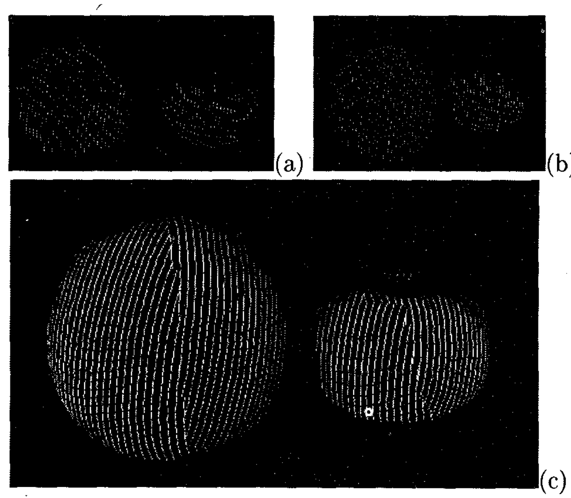

Figure 11: (a) and (b) show the range data of a pair of fruits taken from different view points. (c) shows the 3D data set obtained by registering and integrating the 3D data sets of (a) and (b).

line processes (such as feature-extraction and feature-organization in a feature-based approach). Fig. 8 (c) shows the integration result of the two overlapped data sets.

Figs. 9(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. 9(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. 9(c).

Figs. 10(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. 10(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. 10(d) and (e) show the shaded and the texture-mapped images of Fig. 10(c), respectively.

In Fig. 11 (a) and (b), range images of a pair of fruits are taken from two different views. Fig. 11 (a) is the right view and Fig. 11 (b) is the left view. Their viewing angles differ 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. 11 (c) shows the registered data set which takes only 3.95 seconds with two random trials.

5 Conclusion and Discussions

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. [1][3].

2. It can only be used if the data sets contain sufficient local features [4][7].

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. 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. We have also shown how to determine the acceptable minimal triangle formed by the control points if three control points are used, 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.

Acknowledgements

This work was supported in part by National Science Council, Republic of China under Grant NSC 87-2213-E-011-018.

References

[1] 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.

[2] 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.

[3] 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.

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

[5] 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.

[6] 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.

[7] 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.