RANSAC-Based DARCES: A New Approach to Fast Automatic Registration of Partially Overlapping Range Images

Chu-Song Chen, *Member*, *IEEE Computer Society*,
Yi-Ping Hung, *Member*,
IEEE Computer Society, and
Jen-Bo Cheng

Abstract—In this paper, we propose a new method, the RANSAC-based DARCES method, which can solve the partially overlapping 3D registration problem without any initial estimation. For the noiseless case, the basic algorithm of our method can guarantee that the solution it finds is the true one, and its time complexity can be shown to be relatively low. An extra characteristic is that our method can be used even for the case that there are no local features in the 3D data sets.

Index Terms—Computer vision, range data, range image, registration, 3D imaging.

1 Introduction

Registration of two partially overlapping range images taken from different views is an important task in 3D computer vision. In the past, a popular type of approach to solving the 3D registration problem has been the *iterative approach* [1], [6]. However, the drawbacks are that 1) they require a good initial estimate to prevent the iterative process from being trapped in a local minimum and 2) there is no guarantee of getting the correct solution even for the noiseless case. Many approaches have modified the two approaches proposed in [1] and [6] to obtain more reliable correspondence in each iteration [8], [11].

Another popular type of method is the *feature-based approach* [10], [7], [9]. Feature-based approaches have the advantage that they do not require initial estimates of the rigid-motion parameters. Their drawbacks are mainly that 1) they can not solve the problem in which the 3D data sets contain no prominent/salient local features and 2) a large percentage of the computation time is usually spent on preprocessing, which includes extraction of invariant features [7], [10] and organization of the extracted feature-primitives (e.g., sorting [7]). In addition, Blais and Levine expressed the 3D registration task as an optimization problem [2]. The very-fast-simulated-reannealing technique was used to find the global minimum of the error function.

Our goal in this paper is to solve the 3D registration problem in a fast and reliable manner. We propose a new method—the *data-aligned rigidity-constrained exhaustive search* (DARCES), which can check all possible data-alignments of two given 3D data sets in an efficient way while requiring no preprocessing and no initial estimates of the 3D rigid-motion parameters. Furthermore, to solve the partially overlapping 3D registration problem, the *random sample consensus* (RANSAC) scheme is integrated into the DARCES procedure.

• *The authors are with the Institute of Information Science, Academia Sinica, Nankang, Taipei, Taiwan.*
E-mail: {song, hung}@iis.sinica.edu.tw.

Manuscript received 1 June 1998; revised 9 Mar. 1999.
Recommended for acceptance by G.G. Medioni.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 107684.

0162-8828/99/$10.00 © 1999 IEEE

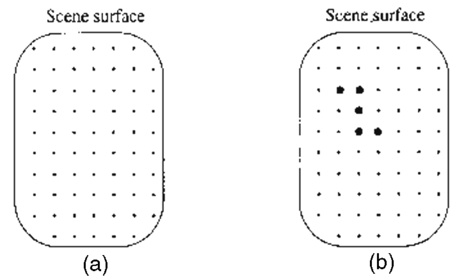


Fig. 1. (a) Selection of the *reference points* in the scene surface. (b) Selection of at least three *control points* from these reference points.

2 Fully-Contained Case—DARCES

Let two data sets, namely, the *scene data set* and the *model data set*, be given. In this section, we first consider a simpler 3D registration problem where the shape of the scene data set is *fully contained* in the shape of the model data set. In the beginning, we have to select some *reference points* from the scene surface, as shown in Fig. 1a. 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 this will be less efficient). In the subsequent processing, a set of (at least three) *control points* is selected from these reference points, as shown in Fig. 1b.

2.1 Using Three Control Points

In this section, we will only consider the case where *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.

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 as corresponding to each of the n_m points in the model data set, where n_m is the number of model points.

Suppose S_p is hypothesized as corresponding to a model point M_p. Then, in the model data set, we will try to find some candidate points corresponding to the secondary point S_s. 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 is 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, now correspond to the model points M_p and M_s. The

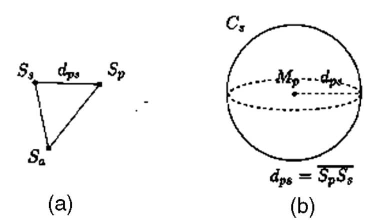


Fig. 2. (a) The triangle formed by the three control points selected from the scene data set. (b) The search region of the secondary control point in the model data set is restricted to the surface of a sphere.

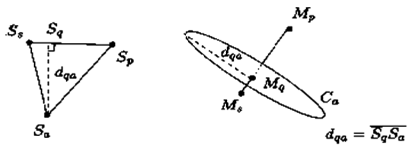


Fig. 3. The search region of the auxiliary control point in the model data set is restricted to the contour of a 3D circle.

candidates of M_a, the corresponding model points 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 onto the line segment $\overline{S_pS_s}$, as shown in Fig. 3. 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, $C_a = \{p = (x, y, z) \mid \|p - M_q\| = d_{qa}$ and $\overline{pM_q}$ is perpendicular to $\overline{M_pM_s}\}$.

After all three control points are successfully aligned on the model surface, a unique rigid-transformation, namely T_c, can be determined by using the three point correspondence pairs: (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, $S_{r_1}, S_{r_2}, \ldots, S_{r_{n_r}}$, can be brought to new positions $S'_{r_1}, S'_{r_2}, \ldots, S'_{r_{n_r}}$. We count the number of occurrences, namely, N_o, when 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, $i = 1, 2, \ldots, n_r$. 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. Finally, the rigid-transformation with the largest overlapping number is selected as the solution of our registration task.

In principle, the three control points form a triangle. 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. To simplify the determination process, we assume that the three selected control points form a regular triangle as shown in Fig. 4. Let the average position error of the data points (including both the data acquisition error of the range-finder and the error caused by limit image resolution) be e and let c be the center of the triangle. For a scene point P whose distance to c is t, the alignment error caused by e will be enlarged to an x. It can be found that x is $\sqrt{3}te/d$. Here, we define the *enlarge ratio* as $h = x/t$. If we want the enlarged ratio to be smaller than a threshold H, d should be larger than $d_{min} = \sqrt{3}e/H$, where d_{min} refers to the edge length of the acceptable minimal triangle. For example, assume that $e = 1.0$ mm and that we hope to keep the enlarge ratio 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 means of the above-mentioned theoretical 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.[1]

1. Details of the complexity analysis can be found in [5].

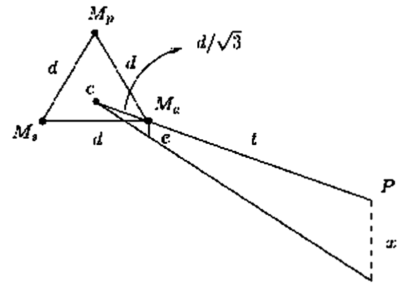


Fig. 4. Derivation of the smallest triangle.

In general, the DARCES method can already provide the registration result accurate to a considerable extent. An ICP-based method [11] is used in our work for further refinement. For implementation purposes, direct searches in the 3D space on the surface of a sphere or on the boundary of a circle may not be trivial. Hence, we can exploit the fact that a range image can be treated as the projection of the 3D points onto an *index plane*. 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. 5a. To simplify the implementation, a square search region is used instead of a 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 approximately 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 $\mathbf{P}_M$, thus forming a 2D ellipse on $\mathbf{P}_M$, as shown in Fig. 5b. The search can then be restricted to the 3D points corresponding to the boundary of the 2D ellipse on the indexing plane. However, indexing the boundary of a 2D ellipse is also not an easy task. In our work, to make the implementation easier, we do not use the projection of the 3D circle C_a; instead, the projections of two other 3D spheres (C_{pa} and C_{sa}) are used, as shown in Fig. 5c, where C_{pa} is the sphere whose center is M_p and whose radius is $\overline{S_pS_a}$, and C_{sa} is the sphere whose center is M_s and whose 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. 5c.

2.2 Using 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 $S_4, S_5, \ldots, S_{n_c}$ (all the other control points), respectively. 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, are respectively hypothesized as corresponding to the model points, 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, $S_i, i = 4, 5, \ldots, n_c$, to a new position, T_cS_i, and to check if T_cS_i satisfies the alignment constraint (i.e., if its distance to the

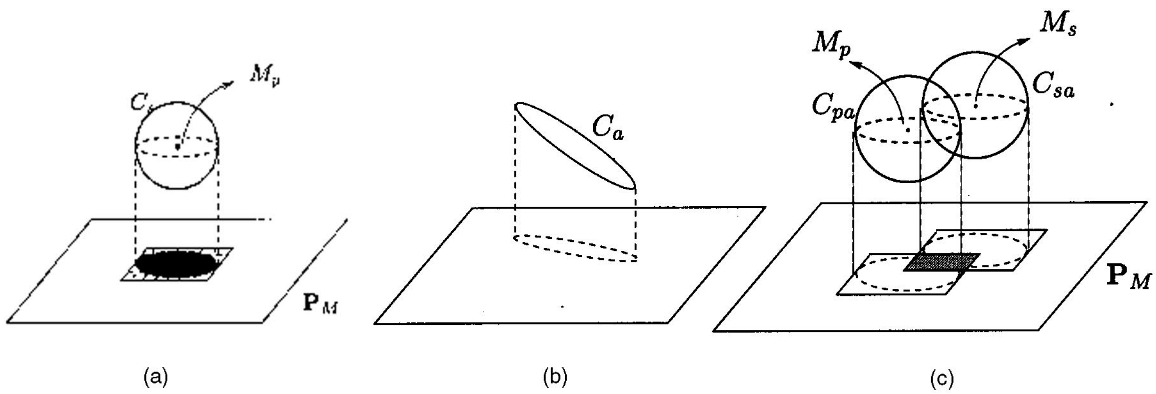


Fig. 5. (a) The 2D circular region of the projection of the 3D sphere onto the index plane and the search of M_s is performed in a squared search region containing this circle. (b) The 2D ellipse of the projection of the 3D circle onto the index plane. (c) The search is performed in the intersection of the two squared regions.

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. Hence, in the noiseless case, *treating all the reference points as the control points* (i.e., choosing $n_c = n_r$) will be the fastest way for solving the fully contained problem. In general, if ς_1 and ς_2 are two selected sets of control points and T^{ς_1} and T^{ς_2} are their time complexities, respectively, then $\varsigma_1 \subseteq \varsigma_2$ implies that $T^{\varsigma 1_1} \leq T^{\varsigma_2}$. Therefore, a lower bound of the complexity can be derived if we consider the case where $n_c = n_r$. Ideally, if early jump-outs always occur (except for the correct solution) when dealing with the fourth control point, then the time complexity can be shown to be $T^{all} = O(n_m \cdot {n_d}^2 + n_r)$ (see [5]), which is a lower bound of the complexity of the DARCES approach.

3 Partially Overlapping Case—RANSAC-Based DARCES

While the strategy of using as many control points as possible is better for solving the fully contained problem, unfortunately, 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 the first (i.e., the primary) control point, which will be introduced in the following.

The RANSAC-based DARCES approach starts by randomly selecting a primary control point from the scene data set.[2] In our approach, the secondary and the auxiliary control points are selected such that they approximately form an acceptable minimal triangle. The other control points are selected around the acceptable minimal triangle such that they gradually form a larger triangle. For instance, Fig. 6 shows an example in which 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 the overlapping number larger than a threshold, then that transformation is regarded 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.

A statistical analysis of the required number of random trials can be found [3], which shows that our method can solve the partially overlapping 3D registration problem with only a few random trials. The RANSAC-based DARCES procedure described above is referred to as the *basic algorithm* of our approach. It can also be further speeded up by incorporating with a coarse-to-fine search structure (however, not all the possible alignments are searched in this situation). The coarse-to-fine search strategy we have adopted is the *three-step algorithm*, which is popular in the field of image/video coding. A detailed description of the three-step algorithm used in our method can be found in [3].

4 Experimental Results

Fig. 7a shows four range data sets of a model head obtained from different view points. The range images were grabbed using a stereo range finder similar to that described in [4]. Each of them contains roughly 4,200 data points. The RANSAC-based DARCES method was used to register the data sets contained in the two range images, which process was referred to as *coarse registration*. In this experiment, 15 control points are used and the three-step algorithm was also used to further speed up the RANSAC-based DARCES approach. After that, a modified ICP approach [11] was used to refine the obtained 3D rigid-transformation, which was referred to as *fine registration* in our experiment. The registered and integrated 3D data set is shown in Fig. 7b and Fig. 7c.[3] Fig. 7d and Fig. 7e show the shaded and the texture-mapped images of Fig. 7c, respectively. The CPU time of total registration of the three pairs formed by the four data sets was 61.98 seconds (using an SGI O^2 workstation), where coarse registration took 58.73 seconds and fine registration took 3.25 seconds. Notice that the computation time was measured for the entire 3D registration task, instead of treating some procedures as off-line processes (such as the feature-extraction procedure and the feature-organization procedure in a

2. Notice that only the primary control point, S_p, is selected randomly.

3. We implemented the zippering method [11] to integrate the two overlapping range data sets.

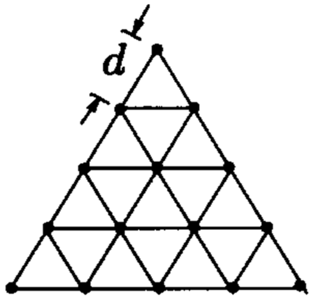


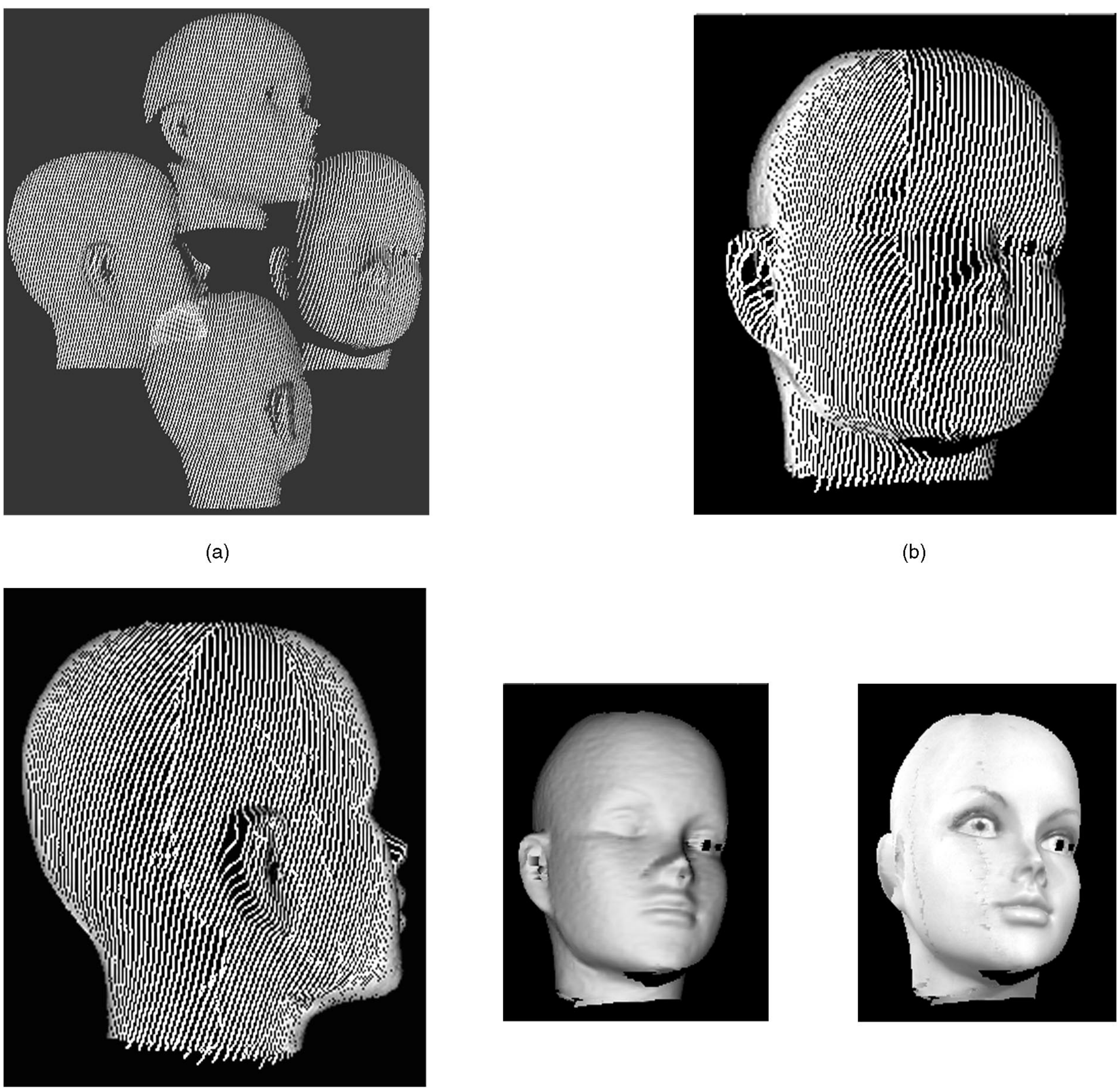
Fig. 6. An example of selecting 15 control points in the index plane.

feature-based approach). The average registration error of coarse registration is 1.66 mm, and that of fine registration is 1.45 mm.

In Fig. 8c and Fig. 8d, range images of a pair of fruits are shown from two different views. Fig. 8c is the right view and Fig. 8d is the left view. Each image contains roughly 2,400 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, using the RANSAC-based DARCES approach, the two data sets can be successfully registered. Fig. 8e shows the registered data set, which needed only 3.95 seconds with two random trials.

5 Conclusions and Discussion

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



(a) (b) (c) (d) (e)

Fig. 7. (a) shows four range images grabbed from different view points. (b) and (c) are the results of registration and integration, observed from two different viewing directions. (d) is the shaded image of the integrated 3D data set. (e) is the textured mapped image of the integrated 3D data set.

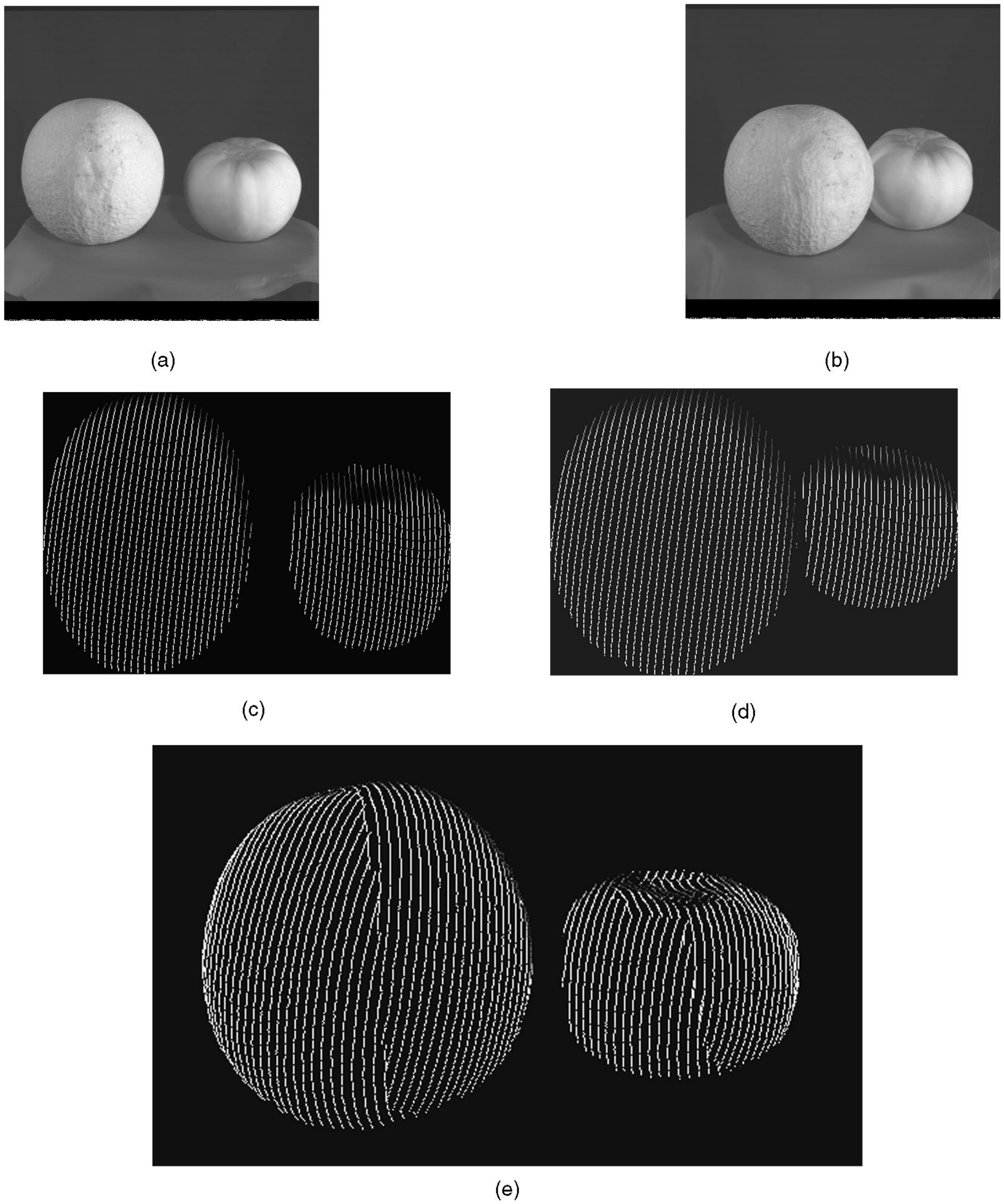


Fig. 8. (a) and (b) are a pair of fruits observed from different view points. (c) and (d) are their range data. (e) shows the 3D data set registered using the RANSAC-based DARCES approach. Notice that, in this case, the two range data sets contain no good local features. Hence, in general, it is difficult to solve this 3D registration problem if we use a feature-based method. Nevertheless, using the RANSAC-based DARCES approach, the two data sets can be successfully registered.

1. They cannot ensure a correct solution even for the noiseless case [1], [6], [8], [11].
2. They require a good initial estimate of the rigid transformation between the two data sets [1], [6].
3. They can only be used if the data sets contain sufficient local features [7], [10].

In this paper, we have proposed the RANSAC-based DARCES approach, which has none of the above three limitations. The basic algorithm of our approach can guarantee that the solution it finds is the true one and it can be used for the featureless case while requiring no initial estimates. Also, our method is faster than most of the existing methods without using initial estimations. Our approach simply treats the 3D registration problem as a partial-matching problem and uses the rigidity constraint among some preselected control points to restrict the search range used for matching. Although some approaches have also used rigidity constraints to facilitate the matching processes [7], [8], our approach is the first one to show that the 3D registration problem can be solved in a relatively low-order computation time by carefully using all of the constraints provided by the rigidity.

In addition, we have shown how the acceptable minimal triangle formed by the first three control points can be determined to greatly reduce the computation time. We have also shown that how additional control points can be used to speed up the search process. 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. Although the RANSAC-based DARCES approach is designed to solve the 3D registration problem, it also has great potential for application in general 3D object recognition (perhaps through a combination of some feature-primitive extractions).

ACKNOWLEDGMENTS

This work was supported in part by the 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 Trans. 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 Trans. Pattern Analysis and Machine Intelligence,* vol. 17, no. 8, pp. 820-824, Aug. 1995.
[3] C.S. Chen et al., "A Fast Automatic Method for Registration of Partially-Overlapping Range Images" *Proc. Int'l Conf. Computer Vision,* Bombay, India, pp. 242-248, 1998.
[4] C.S. Chen et al., "Range Data Acquisition Using Color Structured Lighting and Stereo Vision," *Image and Vision Computing,* vol. 15, pp. 445-456, 1997.
[5] C.S. Chen et al., ftp.iis.sinica.edu.tw/pub/paper/DARCES.ps (anonymous ftp).
[6] 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.
[7] C.S. Chua, "3D Free-Form Surface Registration and Object Recognition," *Int'l J. Computer Vision,* vol. 17, pp. 77-99, 1996.
[8] C. Dorai et al., "Registration and Integration of Multiple Object Views for 3D Model Construction," *IEEE. Trans Pattern Analysis and Machine Intelligence,* vol. 20, no. 1, pp. 83-89, Jan. 1998.
[9] K. Higuchi et al., "Building 3-D Models from Unregistered Range Images," *Graphical Models and Image Processing,* vol. 57, no. 4, pp. 315-333, 1995.
[10] F. Stein and G. Medioni, "Structural Indexing: Efficient 3-D Object Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence.* vol. 14, no. 2, pp. 125-145, 1992.
[11] G. Turk and M. Levoy, "Zippered Polygon Meshes from Range Images," *Proc. SIGGRAPH,* pp. 311-318, 1994.