Multiview Registration of 3D Scenes by Minimizing Error between Coordinate Frames

Gregory C. Sharp[1], Sang W. Lee[2], and David K. Wehe[3]

[1] Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48109
[2] Department of Media Technology
Sogang University, Seoul, Korea
[3] Department of Nuclear Engineering and Radiological Sciences
University of Michigan, Ann Arbor, MI 48109

Abstract. This paper addresses the problem of large scale multiview registration of range images captured from unknown viewing directions. To reduce the computational burden, we decouple the local problem of pairwise registration on neighboring views from the global problem of distribution of accumulated errors. We define the global problem over the graph of neighboring views, and we show that this graph can be decomposed into a set of cycles such that the optimal transformation parameters for each cycle can be solved in closed form. We then describe an iterative procedure that can be used to integrate the solutions for the set of cycles across the graph. This method for error distribution does not require point correspondences between views, and therefore can be used together with robot odometry or any method of pairwise registration. Experimental results demonstrate the effectiveness of this technique on range images of an indoor facility.

1 Introduction

To build three dimensional models using a range scanner, multiple scans are usually required due to occluded surfaces and limitations in the field of view of the scanner. These multiple scans must then be registered within a common coordinate frame before a coherent parametric description of the object can be formed. The registration of two views is most often performed through a variant of the iterative closest point registration algorithm [4,7,30,17,8,6,24,14], although methods based on matching features [10] or maximum likelihood search [28,18] are also often used. However, registration of more than two views is a somewhat more difficult problem, and there is not yet consensus on the best approach. Chen and Medioni, and Masuda both incrementally register views against a growing global union of view points [7,17]. Pulli also performs incremental registration against a growing set, but includes a backtracking step when global error becomes unacceptable [21]. Pennec describes a method that alternates between computing an average shape for the set of images, and registration of the scans against the mean shape [19]. Bergevin et al. place all views into a global frame of

A. Heyden et al. (Eds.): ECCV 2002, LNCS 2351, pp. 587–597, 2002.

reference, and then repeatedly select a view and register it against all others [3]. Blais and Levine use simulated annealing to simultaneously minimize a cost metric based on the total distance between all matches in all views [5]. Stoddart et al. find pairwise correspondences between points in all views, and then iteratively minimizing correspondence errors over all views using a descent algorithm [27]. Eggert et al. improve on this technique, adding special processing for boundary points and a multiresolution framework [9]. Williams and Bennamoun suggested a further refinement by including individual covariance weights for each point [13]. Sawhney et al. and Shum and Szeliski perform the global alignment of two dimensional mosaics by non-linear minimization of distance between point corrispondences [22,26]. Benjemaa and Schmitt propose a nonlinear solution based on the quaternion representation [1]. Their formulation is a multiview extension of the pairwise solution proposed by Horn [11], and using distance between pairwise correspondences as the optimization criteria.

When there are large numbers of views, or when information such as odometry is used in conjunction with point correspondences, the global registration parameters can be solved as the parameters that minimize error with respect to the estimates of the relative motion between view pairs. Lu and Milios solve this problem by linearizing the rotational component [16]. This formulation is useful when the total rotational error is small. In this paper, we propose an analytic method for solving the global registration parameters using the relative motion between view pairs as the error criteria. This criteria does not require linearization, and therefore can be used even when the accumulated rotational error is large. Furthermore, since this criteria does not require point correspondences, our multiview registration method can be used together with robot odometry and any pairwise registration. Our overall approach is to build up a graph that describes the neighboring of views, and then decompose the graph into basis cycles. We solve the nonlinear system over each basis cycle in closed form, and the solutions for the constituent basis cycles are merged using an averaging technique.

2 Error Distribution in a Cycle

We shall take as given that the translation can be decoupled from the rotation, and solved using a linear method [16,1,23]. This effectively decouples the rotation and translation problem, allowing us to analyze rotation and translation separately. For the remainder of the paper, we consider only the rotations.

Let the captured views lie within coordinate frames $\mathbf{V}_1, \cdots, \mathbf{V}_n$, where two frames $\mathbf{V}_i$ and $\mathbf{V}_j$ are related by the rotation and translation $(\mathbf{R}_{i,j}, \mathbf{t}_{i,j})$. A point $\mathbf{p}_j$ in frame $\mathbf{V}_j$ can be described in frame $\mathbf{V}_i$ according to the change of coordinates

$$\mathbf{p}_i = \mathbf{R}_{i,j}\mathbf{p}_j + \mathbf{t}_{i,j}. \tag{1}$$

This change of coordinates between neighboring views can be found using a pairwise registration procedure; we shall call these *measurements.* When each view is connected only to its immediate neighbors in a cycle of views (see fig. 1, the measurements $\{(\mathbf{R}_{1,2}, \mathbf{t}_{1,2}), \cdots, (\mathbf{R}_{n,1}, \mathbf{t}_{n,1})\}$ are an overparameterization of the space with six extra degrees of freedom, and therefore the composition of these changes of coordinates about the cycle may not be identity. We define a cycle *consistent* if its associated rotations or rigid transformations compose to identity. Our goal is to find a new set of rigid transformations, which we call *estimates* and denote $\{(\hat{\mathbf{R}}_{1,2}, \hat{\mathbf{t}}_{1,2}), \cdots, (\hat{\mathbf{R}}_{n,1}, \hat{\mathbf{t}}_{n,1})\}$, that is consistent and that minimizes the mean squared error in frame space between the estimate and measurement.

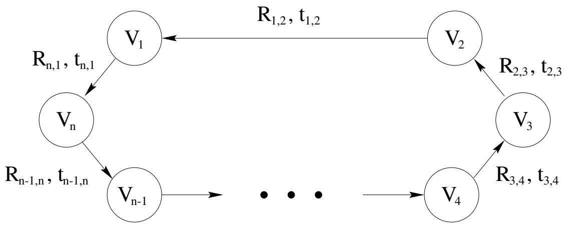

Fig. 1. A cycle of view-pairs where view V_1 has been aligned with V_2, V_2 aligned with V_3, and so on.

Considering only rotation, the estimates that minimizes the mean squared angular error between the estimate and measurement can be solved in closed form [23]. Define the matrix $\mathbf{E}_{k,k+1}$ to be the rotation such that

$$\mathbf{R}_{1,2} \cdots \mathbf{R}_{k,k+1}\mathbf{E}_{k,k+1}\mathbf{R}_{k+1,k+2} \cdots \mathbf{R}_{n,1} = \mathbf{I}.$$

That is, $\mathbf{E}_{k,k+1}$ removes the error that occurs when composing the measurement rotations about the cycle from frame $k+1$ all the way back to frame $k+1$ (see fig. 2 left). We represent fractional portions of this error matrix using the notation $\mathbf{E}_{k,k+1}^{<\alpha>}$, where $\mathbf{E}_{k,k+1}^{<\alpha>}$ shares the same axis of rotation as $\mathbf{E}_{k,k+1}$, but the angle of rotation has been scaled by α. The best estimate for the rotation $\mathbf{R}_{k,k+1}$ is given by

$$\hat{\mathbf{R}}_{k,k+1} = \mathbf{E}_{k-1,k}^{<1/n>}\mathbf{R}_{k,k+1} = \mathbf{R}_{k,k+1}\mathbf{E}_{k,k+1}^{<1/n>}. \qquad (2)$$

This idea is shown in the right half of fig. 2. This estimate is both correct, in the sense that it commutes to identity about the cycle, and optimal in the sense of minimizing mean squared angular error with respect to the measurements [23].

3 Error Distribution in a Graph

Given that we know how to find optimal relative rotations which satisfy the consistency constraint about a single cycle, we now consider a finite number of views connected pairwise by edges in a graph structure. We seek to satisfy

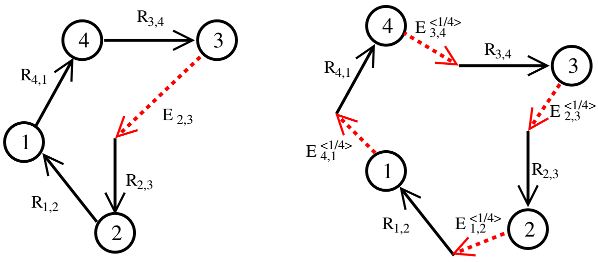

Fig. 2. (Left) $\mathbf{E}_{k,k+1}$ applied to $\mathbf{R}_{k,k+1}$ create a consistent cycle. (Right) $\mathbf{E}_{k,k+1}^{<1/n>}$ applied to each $\mathbf{R}_{k,k+1}$ creates an optimal set of estimates.

the consistency constraint about all circuits in the graph while minimizing error from the initial measurements. We limit our attention to finite, connected, and simple undirected graphs (i.e. having neither loops nor parallel edges), for which each edge lies within at least one cycle. Graphs that are not connected may be processed component by component, and edges that do not lie within any cycle may be given estimates equal to their initial measurements.

Our method for distributing the rotation error is described in algorithm 1. First, choose a set of cycles that form a cycle basis, which will be described in proposition 1. Next, distribute the error about each cycle independently, recording the best rotation for each edge. Finally, since some edges may be contained in more than one cycle, average all rotation estimates for a given edge [20]. Because of this averaging, consistency is not guaranteed. However, it will be shown in proposition 2 that a repetition of this process will reduce the inconsistency to zero.

We define a *circuit* to be a closed path that may contain repeated edges or nodes, and a *cycle* to be a circuit that contains neither repeated edges nor repeated nodes. From graph theory we know that any circuit in a graph can be represented as a linear combination of a set of *basis cycles* in the edge space of the graph $\mathcal{T}$ [2]. This result is used to prove the validity of loop analysis in circuit theory, where the superposition principle allows currents to be added together in linear combinations. For non-commutative groups such as rotations, we establish a slightly weaker result to guarantee consistency throughout all circuits in the graph.

Proposition 1 *Let $\mathcal{T}$ be a spanning tree for a graph $\mathcal{G}$, and let $\mathcal{B}$ be the set of cycles formed by adding a single edge not in $\mathcal{T}$ to $\mathcal{T}$. If the cycles in $\mathcal{B}$ are consistent then all circuits in $\mathcal{G}$ are consistent.*

Proof: Let us represent a circuit as a string, each token in the string representing a node traversed from start to finish. The initial and final token in the string must be the same. For example, the circuit going from a to b to c back to a would be represented *abca*. The circuit *abca* is consistent if $\mathbf{R}_{a,b}\mathbf{R}_{b,c}\mathbf{R}_{c,a} = \mathbf{I}$.

We define four operations that operate on circuits: (a) node insertion, (b) node deletion, (c) circuit insertion, and (d) circuit deletion. If X and Y are substrings, a, b, and c are nodes, and C is a circuit that starts and finishes on c, these operations are described grammatically as follows:

$$\begin{aligned} XabaY &\leftarrow XaY \qquad &(a)\\ XaY &\leftarrow XabaY \qquad &(b)\\ XCY &\leftarrow XcY \qquad &(c)\\ XcY &\leftarrow XCY \qquad &(d). \end{aligned}$$

Node insertion and node deletion preserve the consistency of a circuit, since going a to b back to a inserts or deletes $\mathbf{R}_{ab}\mathbf{R}_{ba} = \mathbf{I}$ into the composition of rotations. Similarly, if the circuit C is consistent, then circuit insertion and circuit deletion preserve the consistency of a circuit because the operation inserts or deletes $\mathbf{R}_C = \mathbf{I}$.

Now suppose that we are given an initial graph which is a spanning tree. All circuits in the graph can be formed by a sequence of node insertion operations from of a single node and therefore the graph is consistent. Proceeding inductively, we assume that a partial graph containing the initial spanning tree is consistent, and we add an edge ab. The edge ab, when added to the initial spanning tree, forms a single cycle $C = ab...a$ in $\mathcal{B}$. By hypothesis, both C and it's reverse circuit $ba...b$ are consistent.

Consider any circuit in this new partial graph. By a sequence of node deletions, it is equivalent to a circuit that contains substrings of a and b of length at most two. For example, two node deletions applied to the substring $ababab$ yields ab, which has length of two. Next, each substring of length two can be modified, using circuit expansion with $C = ab...a$ (or $ba...b$) into a substring that contains two substrings of a and b, one of length three and one of length one. For example, $xbay$ is equivalent to $xbCy = xbab...ay$. Finally, we apply node deletion to the length three substrings of a and b to find an equivalent circuit that contains substrings of a and b of length at most one. This final circuit does not contain the substring ab or ba, and therefore was a circuit in the partial graph before adding the new edge. Hence all circuits in the new partial graph are consistent, and by induction $\mathcal{G}$ is consistent. ∎

This proposition is weaker than the results found using superposition. Although $\mathcal{B}$ is indeed a set of basis cycles [2], there are some sets of basis cycles that cannot be constructed using the spanning tree method. However, this property is not restricted to graphs of rotations. This property holds for graphs of any non-commutative group, including affine and projective transforms.

Proposition 2 *The rotation error distribution algorithm (algorithm 1) converges to a consistent graph.*

Proof: Let the graph have n edges and k basis cycles. Consider some edge i, with measurement $\mathbf{R}_i$, and estimates from $m > 1$ different cycles: $\hat{\mathbf{R}}_i^1, \cdots, \hat{\mathbf{R}}_i^m$.

The measurement error of edge i in cycle j before distribution is $\angle\hat{\mathbf{R}}_i^j(\mathbf{R}_i)^{-1}$. When we average the estimates to find $\hat{\mathbf{R}}_i$, because the averaging minimizes deviation from the estimates, its total contribution to deviation over all cycles is decreased

$$\sum_{j=1}^{m} \angle\hat{\mathbf{R}}_i^j(\hat{\mathbf{R}}_i)^{-1} < \sum_{j=1}^{m} \angle\hat{\mathbf{R}}_i^j(\mathbf{R}_i)^{-1}.$$

Furthermore, the edges that are not in the spanning tree now contribute zero deviation. Hence, the deviation is reduced by at least a fixed percentage per iteration, and the sum deviation from identity about all basis cycles monotonically decreases. From proposition 1, we conclude that this procedure converges to a consistent graph. ■

Algorithm 1 Error Distribution

Given graph, basis cycles, and measurements
$\mathcal{R}^0 \leftarrow$ measurements
repeat
  for all basis cycles in graph **do**
    Use $\mathcal{R}^{i-1}$ to compute rotation estimates for each edge in cycle
  end for
  for all edges in graph **do**
    Average all rotation estimates for this edge and add to $\mathcal{R}^i$
  end for
until convergence

4 Multiview Registration

The global error distribution method described in section 3 is used together with any pairwise registration algorithm to perform completely automatic globally consistent multiview range image registration. As described in algorithm 2, the pairwise registration method is run for a while to find locally good fits for each pair. Then the error distribution method is run to convert the local fits into a global fit. This global fit then provides feedback into the local algorithm, giving it new initial conditions for the next set of local fits.

5 Experimental Results

Figure 3 shows 9 views from an sequence of range images taken from on board of a mobile robot in an indoor environment. The range sensor is a structured light range camera that uses a DLP data projector and CCD camera to compute range points as described by Trobina [29]. The range of the sensor is between

Algorithm 2 Multiview Registration

Let $\mathcal{T}$ be the initial set of pairwise transformations over a graph of views
repeat
  Use pairwise registration to update $\mathcal{T}$
  Use error distribution to update $\mathcal{T}$
until convergence

about 3 to 8 meters, with a noise standard deviation of about 1 cm at 5 meters. The total volume captured by the 9 views is about $5 \times 5 \times 5$ meters.

A graph of neighboring views, shown in figure 4, was constructed manually for these images. In future work we wish to consider automatic construction of the graph, building on the work of Sawhney et al. [22], Kang et al. [15], and Huber [12]. After the graph is found, a spanning tree for the graph was found, as shown in bold on the graph. These neighboring views were registered pairwise using a maximum likelihood matching procedure that uses both valid and missing data points to align the pair of views [25].

The final output of three iterations of algorithm 2 is shown in figure 5, where each of the nine views is rendered in a different color. Errors due to registration

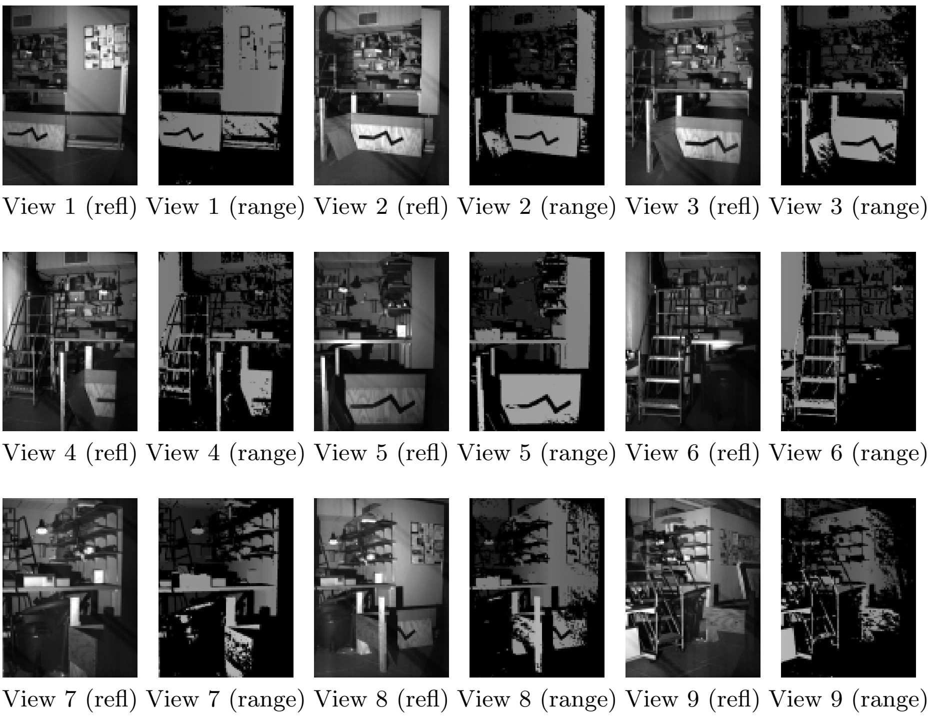

Fig. 3. Reflectance and range images of an indoor scene.

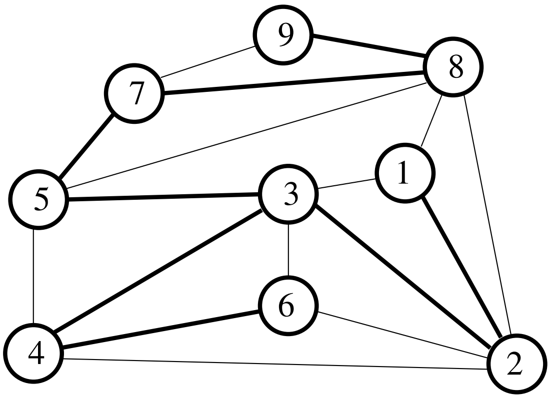

Fig. 4. Graph of neighboring views for range images in figure 3. The spanning tree for the graph is shown in bold.

errors are consistently lower than 0.5 cm, half of the noise standard deviation of the sensor.

6 Discussion

The error distribution method described in equation 2 is easily modified to perform error distribution on a simply weighted graph, where the weights are real valued coefficients that reflect our relative confidence in the rotation estimates. However, it is not clear how this framework can be modified to include covariance, or unequal weights for different directions of rotation. Using real covariance matrices is possible, but this implies linearization of the rotations which limits utility to small rotational error. Using point correspondences to determine the covariance is also problematic. Because the correspondences are only valid in a small region, any covariance information obtained through point correspondences is suspect.

The chief benefit of the graph representation is that it allows updates to proceed locally on each cycle. We believe that this approach may offer insight into related high dimensional sensing problems, including multiview two dimensional image registration, structure from motion, and multisensor fusion. Related sensor readings can be put into a graph structure, and cycles within the graph can be analyzed independently. The nonlinear analysis will be facilitated by the simpler structure of a single cycle compared to the entire graph.

7 Summary

We have described a general multiview registration algorithm that distributes pairwise error accumulated over multiple views back to its constituent views in

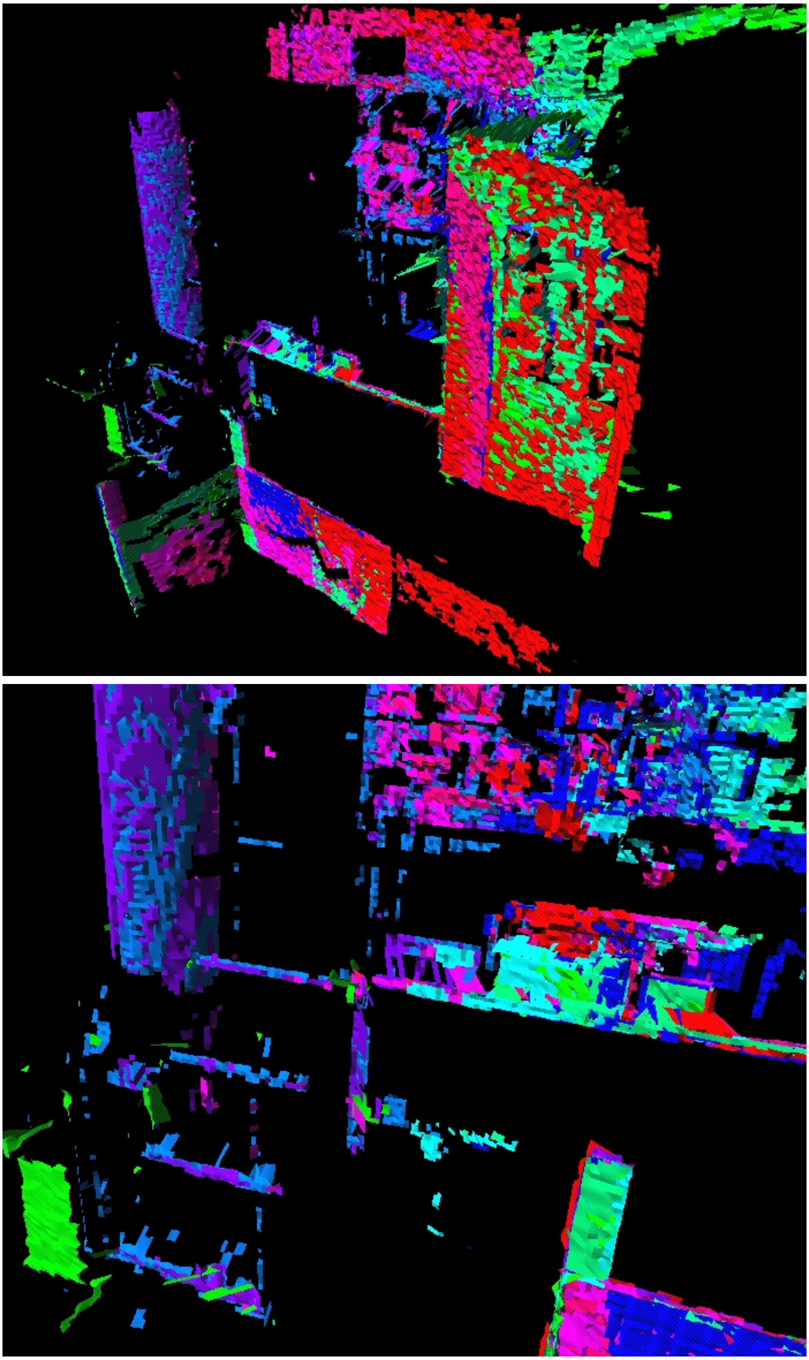

Fig. 5. (Top) Overview of registration results. Each of the nine views is rendered in a different color. (Bottom) Detail of registration results near the step ladder.

a fair manner. Any pairwise registration algorithm may be used to generate the estimates of relative motion between each pair of views, and the accumulated error is distributed to all views. By posing the problem as the minimization of error between coordinate frames, this method is useful for large scale registration problems that may involve hundreds of views.

Acknowledgments. This material is based upon work supported by DOE under Award No. DE-FG04-86NE37969.

References

1. R. Benjemaa and F. Schmitt. A solution for the registration of multiple 3-d point sets using unit quaternions. In *European Conference on Computer Vision*, 1998.
2. C. Berge. *The Theory of Graphs and its applications*. Barnes & Noble, Inc., 1962.
3. R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau. Towards a general multi-view registration technique. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):540–547, May 1996.
4. P.J. Besl and N.D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.
5. G. Blais and M.D. Levine. Registering multiview range data to create 3d computer objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):820–824, August 1995.
6. C.S. Chen, Y.P. Hung, and J.B. Cheng. Ransac-based darces: A new approach to fast automatic registration of partially overlapping range images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(11):1229–1234, November 1999.
7. Y. Chen and G.G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.
8. C. Dorai, J. Weng, and A.K. Jain. Optimal registration of object views using range data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(10):1131–1138, October 1997.
9. D.W. Eggert, A.W. Fitzgibbon, and R.B. Fisher. Simultaneous registration of multiple range views for use in reverse engineering of cad models. *Computer Vision and Image Understanding*, 69(3):253–272, March 1998.
10. O.D. Faugeras and M. Hebert. The representation, recognition, and locating of 3-d objects. *International Journal of Robotics Research*, 5(3):27–52, 1986.
11. B.K.P. Horn. Closed form solutions of absolute orientation using unit quaternions. *Journal of the Optical Society of America-A*, 4(4):629–642, April 1987.
12. D.F. Huber. Automatic 3d modeling using range images obtained from unknown viewpoints. In *International Conference on 3D Digital Imaging and Modeling*, pages 153–160, May 2001.
13. Williams J. and Bennamoun M. A multiple view 3d registration algorithm with statistical error modeling. *IEICE Transactions on Information and Systems*, 83-D(8):1662–1670, August 2000.
14. A.E. Johnson. Surface landmark selection and matching in natural terrain. In *Computer Vision and Pattern Recognition*, volume 2, pages 413–420, 2000.

15. E.Y. Kang, I. Cohen, and G.G. Medioni. A graph-based global registration for 2d mosaics. In *International Conference on Pattern Recognition*, pages Vol I: 257–260, 2000.
16. F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, 4(4):333–349, 1997.
17. T. Masuda and N. Yokoya. A robust method for registration and segmentation of multiple range images. *Computer Vision and Image Understanding*, 61(3):295–307, May 1995.
18. C.F. Olson. Probabilistic self-localization for mobile robots. *IEEE Transactions on Robotics and Automation*, 16(1):55–66, February 2000.
19. X. Pennec. Multiple Registration and Mean Rigid Shapes - Application to the 3D case. In *Image Fusion and Shape Variability Techniques (16th Leeds Annual Statistical (LASR) Workshop)*, pages 178–185, july 1996.
20. X. Pennec. Computing the mean of geometric features: Application to the mean rotation. Technical Report RR-3371, INRIA, 1998.
21. K. Pulli. Multiview registration for large data sets. In *International Conference on 3D Digital Imaging and Modeling*, pages 160–168, 1999.
22. H.S. Sawhney, S. Hsu, and R. Kumar. Robust video mosaicing through topology inference and local to global alignment. In *European Conference on Computer Vision*, pages 103–119, 1998.
23. G.C. Sharp, S.W. Lee, and D.K. Wehe. Toward multiview registration in frame space. In *IEEE International Conference on Robotics and Automation*, 2001.
24. G.C. Sharp, S.W. Lee, and D.K. Wehe. Invariant features and the registration of rigid bodies. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):90–102, January 2002.
25. G.C. Sharp, S.W. Lee, and D.K. Wehe. Registration of range images in the presence of occlusions and missing data. Technical report, University of Michigan, 2002.
26. H.Y. Shum and R. Szeliski. Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision*, 36(2):101–130, February 2000.
27. A.J. Stoddart and A. Hilton. Registration of multiple point sets. In *International Conference on Pattern Recognition*, page B6A.5, 1996.
28. S. Thrun, D. Fox, and W. Burgard. A probabilistic approach to concurrent mapping and localization for mobile robots. *Machine Learning*, 31(1-3):29–53, 1998.
29. M. Trobina. Error model of a coded-light range sensor. Technical Report BIWI-TR-164, Communication Technology Laboratory, Image Science Group, ETH Zurich, 1995.
30. Z.Y. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, October 1994.