Proceedings of 2004 IEEE/RSJ International Conference on
Intelligent Robots and Systems
September 28 - October 2, 2004, Sendai, Japan

Simultaneous Localization and Mapping with Active Stereo Vision

J. Diebel, K. Reuterswärd, and S. Thrun
Stanford University
Stanford, California, USA
[diebel, reutersward, thrun]@stanford.edu

J. Davis, R. Gupta
Honda Research Institute
Mountain View, California, USA
[jedavis, rgupta]@honda-ri.com

***Abstract*— We present an algorithm for creating globally consistent three-dimensional maps from depth fields produced by camera-based range measurement systems. Our approach is specifically suited to dealing with the high noise levels and the large number of outliers often produced by such systems. Range data is filtered to reject outliers within each scan. The point-to-plane variant of ICP is used for local alignment, including weightings that favor nearby points and a novel outlier rejection strategy that increases the robustness for this class of data while eliminating the burden of user-specified thresholds. Global consistency is imposed on cycles by optimally distributing the cyclic discrepancy according to the local fit correlation matrices. The algorithm is demonstrated on a dataset collected by an active unstructured-light space-time stereo vision system.**

Keywords-active; stereo; SLAM; ICP

I. INTRODUCTION

Developing robust algorithms for generating accurate maps from range measurements is widely seen as one of the most important challenges impeding the development of truly autonomous robots. Such maps are critical for localization and path planning and could be used as an important information source in broader frameworks of artificial intelligence, providing concrete sets of measurements to associate with the notions of objects or locales. The two problems of mapping and localization are generally seen as two facets of the same fundamental problem and a great deal of work has been done in this field over the course of the past two decades.

Past work has predominately focused on using time-of-flight laser scanners to map interior environments. Laser scanners have the advantage of producing reliable data with well understood noise characteristics, and robotic mapping schemes based on this technology have been successfully applied by several researchers [1-4].

These systems are, however, fundamentally limited. The primary problem with time-of-flight systems is spatial accuracy. The uncertainty in measurements from such systems is multiple centimeters, regardless of the distance to the target. This large absolute error makes the mapping of small-scale structure impossible and strains algorithms for finding planes and other global structure in the data. A second major problem is speed. Laser scanners operate by scanning a single line with every pass so scanning over a large surface area requires a great deal of time, thereby placing limitations on how fast a robot equipped with such a scanner can move through its environment. A third problem is that laser scanners cost much more than some recently developed alternatives.

Stereo vision systems are a natural choice for such an application but conventional stereo systems provide very noisy data with large gaps in low-texture regions. Recent independent research in the computer graphics and computer vision communities has led to the development of camera-based range measurement systems that are capable of producing dense depth fields at up to 60 frames per second. There are presently several variants of active stereo and active triangulation systems under development but they all work on the same principle of achieving or improving correspondence by projecting non-uniform light into the scene. These systems are particularly accurate in the near-field and have been used to produce impressive models of hand-held objects [5-8]. This sort of scanner is precisely what is needed to overcome the limitations of the laser-scanner mapping systems currently in use, but little work has been done to apply this technology to the mapping and localization problem [9].

The present work combines active stereo vision with simultaneous mapping and localization. When operated over long ranges and in challenging environments, data from camera-based systems has fundamentally different noise and outlier characteristics than both laser scanner and near-field stereo systems, and the direct application of existing methods fails to yield a robust algorithm. False measurements in the data are common and a large number of seemingly meaningful but incongruous measurements can be made when viewing reflective or partially transparent surfaces like windows. Furthermore, camera-based range finding systems depend on deriving accurate spatial information from camera images and therefore come with all the attendant problems of camera calibration, made even more difficult when considering a large working volume. Overcoming all of these challenges is necessary for the application of camera-based range finding techniques to the mapping and localization problem.

In this paper we extend the iterative closest point (ICP) family of local registration algorithms to deal with the challenges associated with aligning long-range camera-based scans of difficult environments. In particular, we propose three changes to the standard algorithm: (1)

0-7803-8463-6/04/$20.00 ©2004 IEEE

individual scans are processed to eliminate outliers by dynamically thresholding local similarity metrics, (2) point pairings are aggressively filtered using a related dynamic threshold, and (3) a weighting function is applied that favors near-field measurements. The new thresholding technique, which is based on a statistical model of the error metrics, has the notable advantages of improving robustness while reducing the number of user-specified input parameters required to process the data.

We integrate this approach into a global SLAM algorithm capable of creating globally consistent maps of cyclic environments. This problem has received a lot of attention in two dimensional mapping [10-12], and in this paper we apply the same techniques to three dimensions. In particular we applied a constrained optimization technique to optimally redistribute cyclic discrepancies over the entire cycle. Here we use the fit-confidence information encoded in the covariance matrix produced by each point-to-plane ICP alignment. This framework provides an effective method for correcting the error accumulated by chaining together uncertain local motion estimates.

Finally, we illustrate the capabilities of our algorithm with an example based on an active space-time stereo vision system [8]. The outer perimeter of a room was surveyed by a mobile robot following a rectangular path with rounded corners. Two hundred and thirty depth fields were recorded, each containing about 25,000 points. Local alignment was performed only between consecutive frames and global registration was performed once, upon completion of the cycle. The resulting combined dataset was down-sampled with a point-clustering technique, yielding the rendered image shown in the Experimental Results section.

II. Iterative Closest Point Algorithm

Our work is based on the extensive literature on the iterative closest point algorithm. This algorithm is used to find the full six-degree-of-freedom transformation between overlapping but unaligned scans. There are several variants of this algorithm and we focus on the point-to-plane formulation [13], which tends to have better convergence characteristics and yields more accurate results than the various point-to-point alternatives [14]. Broadly speaking the algorithm works by finding corresponding points in the two meshes and applying a transformation to minimize an error metric, a function of the point pairs, that quantifies the relative misalignment. After the transformation, new correspondences are found and the process is repeated, either for a fixed number of iterations or until the cumulative transformation stops changing with each iteration. Modified to include a prior estimate based on independent odometry, the error metric reads

$$E = \sum_i w_i \left[\left(Rp_i + T - q_i \right) \cdot n_i \right]^2 + \left[w_{odo} \cdot \left(y - \hat{y} \right) \right]^2 .$$

Here, p_i and q_i are points in the two meshes under consideration; n_i is the normal corresponding to p_i; R is a linearized rotation matrix, parameterized by a set of three Euler angles; T is a translation vector; w_i is the weight associated with a given pair; y is the six-dimensional state vector, of which the first three elements are the Euler angles and the second three elements are the translation parameters; $\hat{y}$ is the estimated state vector implied by the odometry; and w_{odo} is the vector of odometry weights. Differentiating this expression with respect to each of the state variables leads to a system of six linear equations in six unknowns. Solving this linear system leads to an optimal set of transformation parameters for a given set of point pairs.

The weights w_i are selected as some function of the quality of the match, either penalizing large point-to-point distances or large disparities between paired normals [14]. The index i in the error metric summation is the sum over all sampled point pairs that have not been rejected as unacceptable. There are several common adaptive rejection criteria: (1) reject the worst *n%* of pairs, usually about 10%; (2) reject pairs with point-to-point distances that exceed some multiple of the standard deviation of the distances, typically 2.5 times [15]; (3) reject pairs with incompatible normals, and (4) reject pairs that lie on the mesh boundaries. The last of these is considered critical to good alignment and is used in the vast majority of ICP implementations, including ours. The rejection based on incompatible normals was also included in all the ICP variations we considered.

In the present work, a challenging dataset was collected that includes incidental views of glass surfaces and range measurements of up to seventy times the baseline of the stereo system. Full details of the experimental configuration are given in the Experimental Results section, including sample scans. Here, it suffices to say that the data pushes the limits of scan alignment. In particular, the glass surfaces yielded large numbers of outliers that are not distinguishable from the good data based on traditional filtration techniques. The ICP algorithm, modified to include prior estimates based on odometry, was run on this dataset with the standard outlier-rejection techniques and it was found to be very brittle.

For example, experiments were run using the common filtration technique of rejecting the worst *n%* of point pairs. Here, we considered rejection rates of 10%, 20%, 25%, and 30%. Even with a relatively strong prior estimate of the robot's motion, the algorithm failed consistently on the most difficult pairings. The results of these trials are summarize in Table 1 and three illustrative failures are shown in Fig. 1. A point-cloud rendering of a sample misalignment is shown in Fig. 5.

In order to understand why these failures occur we look at the distribution of paired-point distances in our dataset. Fig. 4 shows the initial distribution of paired-point distances for a sampling of consecutive-frame pairings. Here we can see the root cause of the problem. Taking too small a rejection rate leads to not filtering enough point pairs in cases for which there are a large number of outliers, and taking too large a rejection rate throws away

too many points in fairly clean cases. Thus, for this rejection scheme to work, we must tune the rejection rate parameter for a given dataset to find if there is some suitable rejection rate that exhibits neither mode of failure. Even when this is possible it is clearly less than ideal.

Another common rejection strategy was also considered. Rejecting point pairs with distances that exceed some multiple of the standard deviation of the distances was found to be even less effective. This is due to the fact that the standard deviation is strongly effected by outliers and our distributions are highly non-Gaussian with heavy tails. The distribution resulting from a cut of 2.5σ is shown in Fig. 4. The threshold for cases with large numbers of outliers is far too high, leading to systematic biasing of the final alignment estimate.

In summary, traditional ICP is failing on this sequence because the noise characteristics are more challenging than the algorithm was designed to accommodate. The most notable characteristic is that the level of noise varies significantly between frames, so it is unlikely that a single rejection rate will successfully filter all the data. It is certainly possible to find sets of traditional ICP parameters that will successfully match any given pair and it is even conceivable that a particular set will succeed on the entire sequence. But any change in one of the parameters will lead to a failure, making any such solution brittle.

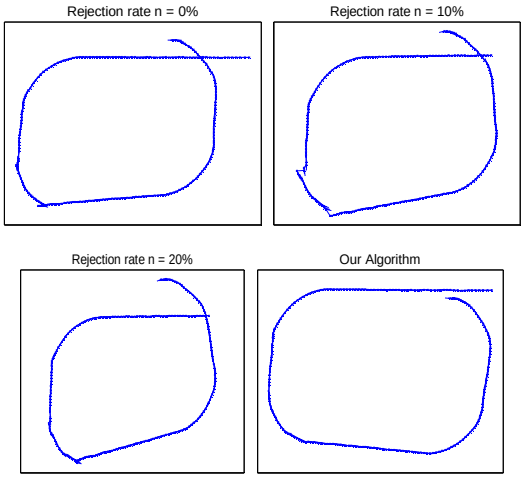

Figure 1. Robot paths predicted by standard ICP algorithm with typical *n%* rejection strategy, compared to result from our algorithm. Notice that all of the existing ICP variations fail on at least one frame.

TABLE I. FAILURE OF TRADITIONAL ICP WITH N% REJECTION

Rejection rate	Failures
10%	4 pairs due to rejecting too few matches
20%	1 pair due to rejecting too few matches
25%	1 pair due to rejecting too few matches and 1 pair due to rejecting too many matches
30%	1 pair due to rejecting too many matches

III. APPROACH

The application of camera-based range measurement systems to the mapping and localization problem promises to overcome many of the limitations of the laser range scanners used in existing mapping systems. It is clear, however, that the noise characteristics of these new systems, when operated at long ranges and in difficult environments, are fundamentally different from both camera-based systems at near range and laser-scanner systems in general. In order to overcome these problems we propose three changes to the standard ICP algorithm. In this section we discuss each of these changes, as well as the global registration technique used to impose cyclic consistency.

A. Outlier Rejection

Stereo algorithms are particularly prone to noise, so strong outlier rejection is critical to local registration. Considering only a single mesh, our approach is to characterize each vertex by two measures. The first is the maximum edge length attached to it and the second is the maximum angle between the normal of the vertex in question and the normals of its neighbors. A vertex is rejected as an outlier if it is part of the intersection of the outliers within each category.

In order to define an outlier we consider what kind of distributions we expect for each of these parameters. The maximum edge length parameter will be highly non-Gaussian with a positive skew and a heavy tail. Fig. 2 shows a likely distribution for this parameter. In such a distribution the median, as computed by random sampling, is a computationally inexpensive approximation of the mode and twice the mode is, in turn, an effective threshold for outlier rejection. The maximum angular difference between neighboring normals has a similar distribution to that of the maximum edge length associated with each vertex and we apply the same model to calculate the threshold.

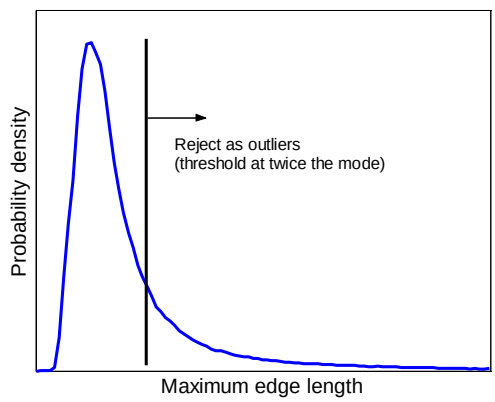

Figure 2. Expected distribution of maximum edge lengths and other error metrics.

We validated the accuracy of our model by plotting the actual distributions observed in the present dataset. Fig. 3 shows the edge length distributions for a sampling of the frames, before and after normalization with respect to the dynamic threshold. Also shown is the distribution of maximum angle between neighboring normals before and after normalization. In both cases, approximating the mode with the median was an effective way of making computationally efficient dynamic cuts.

As discussed in the Iterative Closest Point Algorithm section, traditional ICP pair rejection strategies fail in datasets for which there is a great deal of variability in the noise characteristics between frames. This is exactly what we expect in data collected by camera-based range systems in difficult environments (e.g., containing reflective surfaces). We solve this problem in a manner similar to our solution to the single-scan filtration techniques discussed above. We expect a similar heavy-tailed, positively-skewed distribution of initial paired-point distances, but we are more concerned with avoiding the possibility of initially rejecting too many pairs, so we threshold at a slightly larger cut of three times the median. We are willing to accept a higher cut here because as ICP converges, the threshold will shrink as the median shrinks and more outliers will be rejected. The resulting normalized distribution is shown in Fig.4.

The insight of this paper is that while the distribution of the various error metrics vary substantially between frames, they are all of sufficiently similar character that we can collapse them into a single curve by normalizing with respect to the proper threshold. This can be seen clearly in Fig. 3, in which raw and normalized distributions are shown together. In Fig. 4 we show the improved performance of our algorithm relative to the standard ICP outlier rejection algorithms. Here we focus on the outlying curves. With all three existing methods there are clearly cases where the existing thresholds fail to reject a large number of outliers; it is even possible to see second modes that have been allowed in under the threshold. In the traditional ICP framework, the only way to get rid of these outliers is to turn up the rejection rate, which, as in the case of a rejection rate of 30%, leads to cutting out too much good data in fairly clean distributions.

B. Weighting

The noise of a single triangulation in camera-based range-finding algorithms can be shown by geometric reasoning (under reasonable assumptions) to vary linearly with depth. We therefore reformulate the weights associated with each point pair to linearly favor nearby points over distant ones. Specifically we define a weighting function that varies linearly from one at zero depth to zero at some specified maximum depth of field. A reasonable value for this would be some integer multiple of the baseline of the system in question. In the present work a maximum depth of field of 10 meters was assumed.

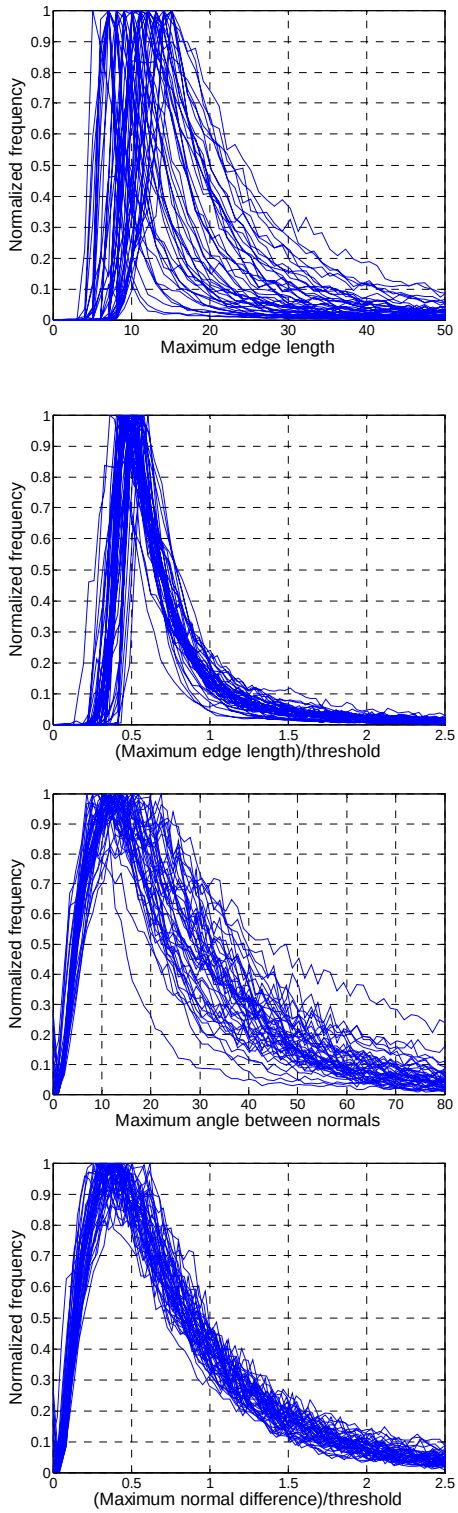

Figure 3. The distribution of the maximum edge length (top), the same data normalized with respect to the threshold selected as twice the approximate mode (middle top), the distribution of maximum angle between neighboring normals (middle bottom), and the same data normalized with respect to the threshold selected as twice the approximate mode (bottom).

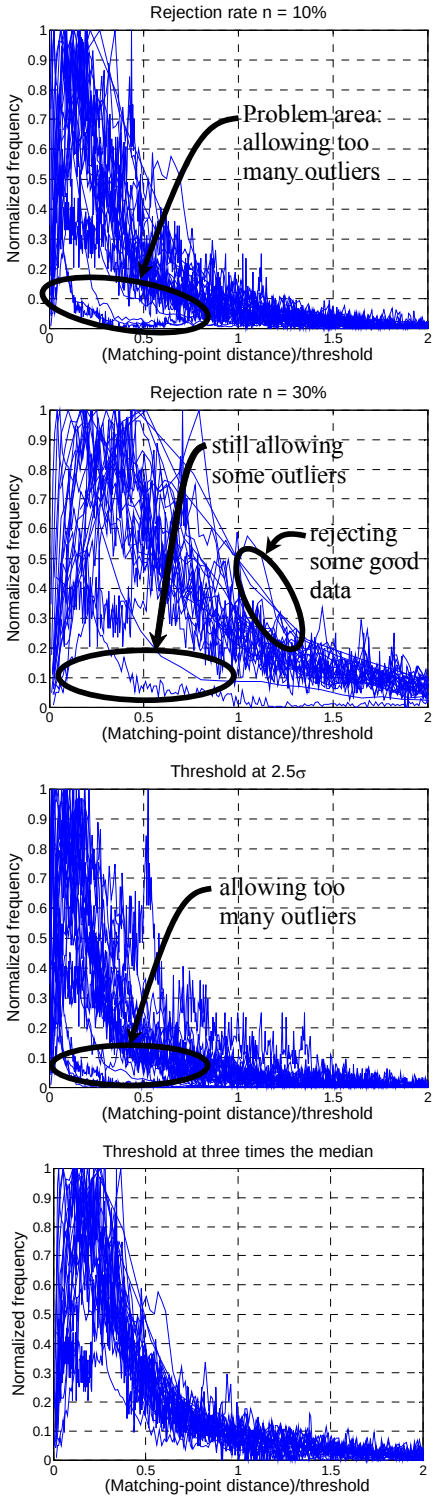

Figure 4. Distributions of paired-point distances normalized with respect to a threshold defined by rejecting the worst 10% of matched pairs (top), rejecting the worst 30% of matched pairs (middle top), rejecting pairs with distances greater than 2.5 times the standard deviation (middle bottom), and cutting at three times the median value (bottom).

C. *Modified ICP Results*

The combination of advanced outlier rejection and appropriate weighting solved the misalignment problems ICP was experiencing. An interesting way to visualize the effect of the new filtering technique is to consider the distribution of the equivalent rejection rate. That is, for each pairing we compute the *n%* rejection rate that would yield the same threshold as the new algorithm; the histogram of this equivalent threshold is shown in Fig. 5. Notice that for most scan pairs a threshold of 5-15% is computed, but that for a few particularly noisy scans the threshold is equivalent to rejecting 30% of the points. Also shown is the final point-pair distance distribution in our algorithm. Notice that the curves all collapse nicely.

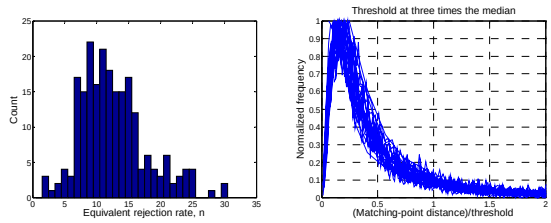

Figure 5. The distribution of equivalent rejection rates (left) and the final (converged) distribution of paired-point distances normalized with respect to the threshold of our algorithm (right).

In order to illustrate the types of alignments that were effected by the changes to the algorithm, a before-and-after comparison is made with one of the critical local registrations. Fig. 6 compares a match based on a traditional ICP scheme with a 25% rejection rate to the same match based on the new algorithm. This particular combination is a relatively noise-free pair that is being over-cut by the large rejection rate. To reiterate, turning down the rejection rate does solve this pair's problems, but only at the cost of creating a misalignment between other pairs with higher outlier content.

Note that in this example the source of the misalignment is noise due to a glass-wall that is out of view to the right. The misaligned planes do not attract each other because they have been flagged as outliers due to the high rejection rate, so spurious correlations in the noise draw the alignment apart.

D. *Global Registration*

As the robot moves through a scene it will inevitably encounter regions that it has already crossed, such as when completing a loop of hallways or scanning the outer perimeter of a large room. In this case the robot is faced with a contradiction. The chaining together of local motion estimates based on local scan alignment produces an estimate of the current position that is not the same as the direct estimates provided by comparing the current scan directly with the previous scans of the same region. Resolving this discrepancy is the global registration problem, which is solved by optimally distributing the discrepancy over all the local motion estimates according to our measures of local alignment quality from each point-to-plane alignment.

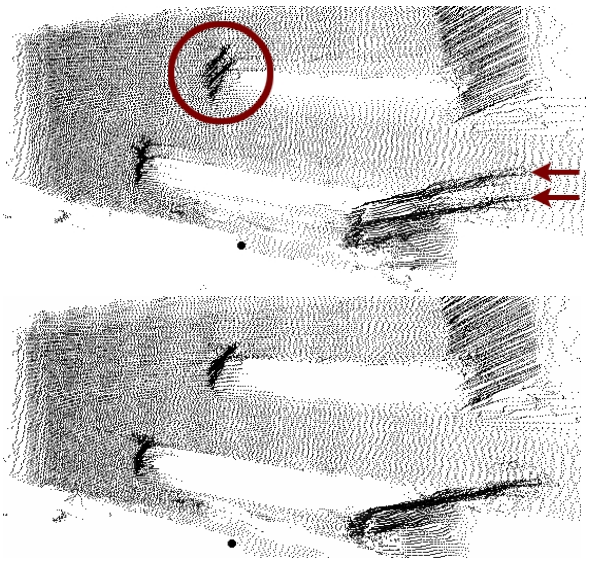

Figure 6. A sample misalignment due to treating too many accurate point pairs as outliers (top). Notice the poorly aligned features in the annotated circle, and also near the two arrows. The same match as computed by the new algorithm has features which are much better aligned (bottom).

The problem is illustrated schematically in Fig. 7. Here the pose vectors are denoted P_i and the relative transformations between the ith pose vector and the $(i+1)^{\text{th}}$ pose vector is denoted y_i.

We seek a solution that reconfigures the pose vectors to minimized the total error associated with the discrepancy between the chosen transformations and those predicted by ICP. That is, the error due to a perturbation in any given relative transformation is given by

$$\varepsilon_i = \left(dy_i\right)^T C_i \left(dy_i\right),$$

where C_i is the ICP correlation matrix associated with alignment b_i. We write the perturbation away from the ICP alignment as a non-linear function of the associated poses

$$dy_i = \tilde{y}_i\left(P_i, P_{i+1}\right) - y_i,$$

which can be linearized using a Taylor series expansion

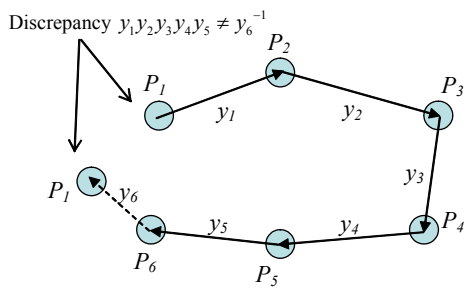

Figure 7. A schematic illustration of the global registration problem.

$$dy_i = \tilde{y}_i\big|_0 + \begin{bmatrix} \left.\dfrac{\partial y_i}{\partial P_i}\right|_0 & \left.\dfrac{\partial y_i}{\partial P_{i+1}}\right|_0 \end{bmatrix} \begin{bmatrix} P_i - P_i\big|_0 \\ P_{i+1} - P_{i+1}\big|_0 \end{bmatrix} - y_i .$$

Here, evaluation at zero indicates values computed by local alignment alone. Thus we have a linear equation for the perturbation away from the ICP alignment that is caused by a given change in the associated pose vectors. This equation can be inserted into our expression for the related error contribution, yielding terms of the form

$$\varepsilon_i = \left[\begin{bmatrix} J_i^1 & J_i^2 \end{bmatrix} \begin{bmatrix} P_i \\ P_{i+1} \end{bmatrix} - d_i\right]^T C_i \left[\begin{bmatrix} J_i^1 & J_i^2 \end{bmatrix} \begin{bmatrix} P_i \\ P_{i+1} \end{bmatrix} - d_i\right]$$

and the whole quadratic system can be optimized by solving the linear system

$$\left(J^T C\, J\right)\vec{P} = \left(J^T C\right)\vec{d},$$

where

$$C = \begin{bmatrix} C_1 & 0 & \cdots & \cdots & 0 \\ 0 & C_2 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & C_n \end{bmatrix},$$

$$J = \begin{bmatrix} J_1^1 & J_1^2 & 0 & \cdots & \cdots & 0 \\ 0 & J_2^1 & J_2^2 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \vdots & & \ddots & J_{n-1}^1 & J_{n-1}^2 \\ J_n^2 & 0 & \cdots & \cdots & 0 & J_n^1 \end{bmatrix},$$

$$\vec{d} = \begin{bmatrix} d_1 & d_2 & \cdots & d_n \end{bmatrix}^T, \quad \text{and}$$

$$\vec{P} = \begin{bmatrix} P_1 & P_2 & \cdots & P_n \end{bmatrix}^T.$$

The resulting system is of size $n{\cdot}dof$, where n is the number of poses in the cycle and dof is the number of degrees of freedom allowed to each transformation.

IV. Experimental Results

We illustrate our algorithm on a set of data collected by a mobile robot equipped with an active space-time stereo system, as shown in Fig. 8. We used a 2 *dof* (rotation, translation) Nomad robot base with SRI stereo cameras in a room of size of 8 meters by 8 meters. The robot was moved in steps of one inch in a closed loop by using encoders to

correct for motion error. The ranging sub-system has a 9 cm baseline and a standard data projector to display patterns. The cameras are synchronized and calibrated using a standard technique [16,17]. At each waypoint along the robot path, the ranging system captures a depth map.

Although the space-time stereo ranging method is described in greater detail elsewhere [7,8], we summarize our particular implementation here. The data projector is used to display a sequence of 25 patterns which are imaged on the SRI stereo camera. The patterns consist of random vertical black and white stripes. Although it is not necessary, the projector and cameras are synchronized. Depth is recovered using a 1x1x25 matching vector, i.e. the stereo epipolar search has a spatial extent of a single pixel, and a temporal extent of all 25 frames. Normalized cross correlation is used to compare matching vectors. Since the active illumination allows for good correspondence to be determined, no global regularization method such as graph cuts or dynamic programming is used. Two sample range fields are shown in Fig. 9.

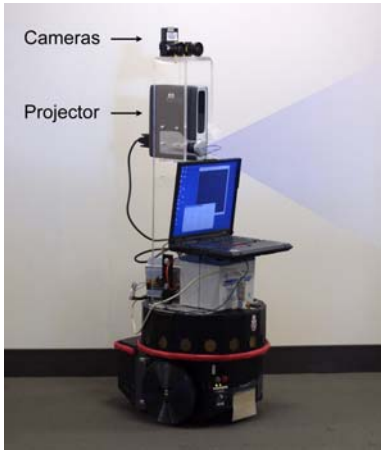

Figure 8. Experimental setup.

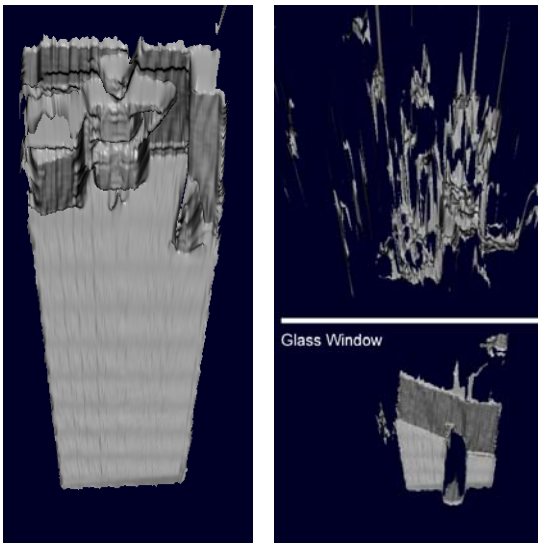

Figure 9. A clean range field (left) and one with a lot of noise behind an intervening glass window due to a reflections (right).

The robot was commanded to follow a rectangular path with rounded corners about the outer perimeter of the room. Two hundred and thirty frames were collected during the course of the robot's motion and the data was processed according to our algorithm. After the fact, each ICP alignment was examined by eye to confirm that the modified ICP algorithm did indeed functioned without mistakes for the entire run.

The full circuit, before and after global registration is shown in Fig. 10. For reference, the actual final position of the robot was also just to the right of the first segment of the robots path, so the path shown here is very accurate. The resulting three-dimensional model of the room was down-sampled with the point clustering techniques of Pauly [18] and is visualized in Fig. 11.

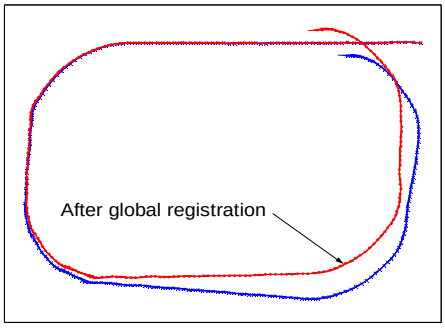

Figure 10. Robot paths predicted by our algorithm before and after global registration is performed.

V. Summary And Conclusions

In this paper, we proposed a new method for the problem of large-scale 3-D mapping with stereoscopic sensors. We modify the ICP algorithm in two major ways: (1) by weighting nearby points more heavily and (2) by using a new dynamic threshold that better filters the data without requiring user-specified cuts. Global consistency is imposed by optimally redistributing cyclic discrepancy according to the local fit covariance matrices. We successfully acquired a large 3-D map of a cyclic environment, using a data set for which previous ICP algorithms fail.

Alternative solutions to this problem certainly exist. For example, applying a carefully selected fixed paired-point distance threshold of 0.15 meters on top of a 5% cut was found to be an effective strategy for yielding consistently accurate alignments. Such a strategy depends, however, on a user-specified fixed threshold and is therefore less desirable than a strategy that selects the appropriate threshold automatically. It is also conceivable to improve the fit we make to the statistical model. Applying a correction after initial thresholding to take into account the standard deviation of the distances between the good pairings could lead to an even more robust algorithm. One should note here that these thresholds are being applied many times per second on large volumes of data, so we must be careful to design a technique that does not damage the overall computational performance of the algorithm. We believe that the algorithm presented here

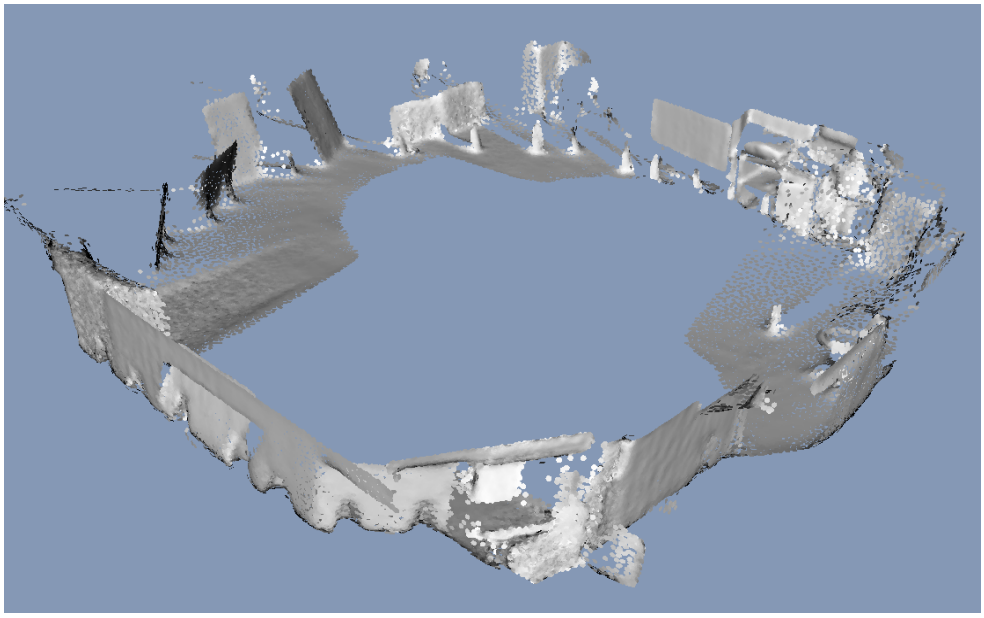

Figure 11. Rendered view of three-dimensional scene.

offers a suitable balance between computational efficiency and statistical rigor.

Despite these positive results there are a number of questions that warrant future research. An effective algorithm for detecting when a cycle has been completed is essential to the autonomous application of SLAM. The same algorithm should be applied to detect all overlapping frames that can be compared to improve global alignment. In addition, we have found that good camera calibration is critical to successful scan alignment. A robust algorithm for calibrating short-baseline stereo at long ranges would also improve alignment.

[1] J.-S. Gutmann and K. Konolige. Incremental Mapping of Large Cyclic Environments. Proceedings of the IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA). 2000.

[2] Burgard, W., Fox, D., Hennig, D., & Schmidt, T. (1996). Estimating the absolute position of a mobile robot using position probability grids. Proceedings of the Thirteenth National Conference on Artificial Intelligence (pp. 896–901). Menlo Park, CA: AAAI Press/MIT Press.

[3] Andreas Nüchter, Hartmut Surmann, Kai Lingemann, Joachim Hertzberg, and Sebastian Thrun. 6D SLAM with Application in Autonomous Mine Mapping, in Proceedings IEEE 2004 International Conference Robotics and Automation (ICRA 2004), New Orleans, USA, April 2004

[4] Hähnel, D. and Burgard, W. and Thrun, S., Learning Compact 3D Models of Indoor and Outdoor Environments with a Mobile Robot, Robotics and Autonomous Systems, Volume 44, pages 15-27, 2003.

[5] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy. Real-time 3D model acquisition. ACM Trans. on Graphics (SIGGRAPH 2002 proceedings), 21(3):438–446, 2002.

[6] L. Zhang, B. Curless, and S. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In IEEE 3D Data Processing Visualization and Transmis-sion, 2002.

[7] L. Zhang, B. Curless, and S. Seitz. Spacetime stereo: Shape recovery for dynamic scenes. In Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR), 2003.

[8] J. Davis, R. Ramamoorthi, and S. Rusinkiewicz. Spacetime stereo: A unifying framework for depth from triangulation. In Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR),, 2003.

[9] Se, S. and Lowe, D. and Little, J. Vision-based Mobile robot localization and mapping using scale-invariant features, Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Seoul, Korea, pages 2051--2058, May 2001.

[10] Lu, F. and Milios, E., Globally Consistent Range Scan Alignment for Environment Mapping. Autonomous Robots. v. 4, pp. 333-349. 1997

[11] G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte, and M. Csorba. An experimental and theoretical investigation into simultaneous localisation and map building (SLAM). Lecture Notes in Control and Information Sciences: Experimental Robotics VI, Springer, 2000.

[12] Hähnel, D. and Burgard, W. and Wegbreit, B. and Thrun, S., Towards Lazy Data Association in SLAM. Proceedings of the 11th International Symposium of Robotics Research (ISRR'03). Springer. 2003.

[13] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. Proc. IEEE Conf. on Robotics and Automation, 1991.

[14] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. Proc. International Conference on 3D Digital Imaging and Modeling (3DIM), 2001.

[15] Masuda, T., Sakaue, K., and Yokoya, N. Registration and Integration of Multiple Range Images for 3-D Model Construction, Proc. CVPR, 1996.

[16] Jean-Yves Bouguet, Visual methods for three-dimensional modeling, PhD Dissertation, California Institute of Technology, 1999.

[17] Z. Zhang, Flexible Camera Calibration by Viewing a Plane from Unknown Orientations, International Conference on Computer Vision, ICCV, 1999.

[18] Mark Pauly, Markus Gross, and Leif P. Kobbelt. Efficient simplification of point-sampled surfaces. In Proceedings IEEE Visualization 2002, pages 163-170. Computer Society Press, 2002..