Computer Vision and Image Understanding **87,** 104–115 (2002)
doi:10.1006/cviu.2002.0986

Pair-Wise Range Image Registration: A Study in Outlier Classification

Gerald Dalley[1]

Department of Electrical Engineering, The Ohio State University, 205 Dreese Lab, 2015 Neil Ave., Columbus, Ohio 43210
E-mail: dalleyg@ieee.org

and

Patrick Flynn

Department of Computer Science and Engineering, University of Notre Dame, 384 Fitzpatrick Hall of Engineering, Notre Dame, Indiana 46556
E-mail: flynn@nd.edu

Received August 31, 2001; accepted September 6, 2002

In this paper, we present a robustness study on several popular techniques for performing fine registration of partially overlapping 2.5D range image pairs, with a focus on model building. In our first set of tests, we qualitatively evaluate the output of several iterative closest point (ICP) variants on real-world data. Our second set of tests expands to include additional ICP variants and an implementation of Chen and Medioni's point-to-plane minimizing algorithm. These tests evaluate quantitatively how well these algorithm variants are able to correct initial simulated rigid rotation and translation errors. The aim of these variants in both sets of tests is to classify as outliers particular point pairs containing vertices outside of the region of overlap of the two range images. In addition to testing these variants with different parameter settings, we also study how performing topologically uniform subsampling of the meshes affects the registration quality.

Key Words: registration; range image; model building; ICP; iterative closest point.

[1] Supported in part by the National Science Foundation under Grant EIA-9818212, the Electrical Engineering Department of The Ohio State University, the Dayton Area Graduate Studies Institute, and the Air Force Research Laboratory Sensors Directorate.

1077-3142/02 $35.00

1. INTRODUCTION

In recent years, there has been growing interest in techniques for building 3D computer models of real-world objects and scenes without requiring humans to manually produce these models using laborious and error-prone CAD-based approaches. Using range sensors, users are able to capture 3D images of objects from different viewpoints that may be combined to form the final model of the object or scene [2]. These models then may be used for a variety of purposes such as building 3D maps for robot navigation, providing training data for computer vision experiments, and digitizing historical buildings for restoration planning [16, 19].

After acquiring a set of range images, a coarse alignment is generally known—either from some type of positional sensors or via a feature-matching registration step. After refining the registration, the data are combined to produce a single surface description [4, 10, 12, 17]. In this paper, we will study the robustness of algorithms for refining the initial coarse registration when partially overlapping range image pairs are being registered. A preliminary study of this problem appears in [5].

The first class of registration algorithms we will consider is based on the iterative closest point (ICP) algorithm popularized by Besl and McKay [1]. Consider a set of source points, $\mathtt{P}$, being registered to a set of destination points $\mathtt{X}$ by using a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{T}$. The ICP registration process minimizes the objective function

$$\mathtt{f}_n(\mathbf{R}_n;\mathbf{T}_n) = \sum_{i_n} \mathtt{w}_{i_n} \left\| \mathbf{x}_{i_n} - \left(\mathbf{R}_n \mathbf{p}_{i_n} + \mathbf{T}_n\right)\right\|^2 = \sum_{i_n} \mathtt{w}_{i_n}, \tag{1}$$

where $\mathbf{x}_{i_n}$ is the closest point in data set $\mathtt{X}$ to the point $\mathbf{p}_{i_n}$, and $\mathtt{n}$ is the current iteration number. Besl uses a unit quaternion-based approach to find the values for $\mathbf{R}_n$ and $\mathbf{T}_n$ that minimize this function (for each iteration), assuming $\mathtt{w}_{i_n} = 1$ for all $\mathtt{j}_n$ at iteration $\mathtt{n}$. These calculations are performed iteratively by transforming the source points by the calculated $\mathbf{R}_n$ and $\mathbf{T}_n$ until the registration converges. This base ICP process is guaranteed to converge to some (local) minimum for any starting registration when $\mathtt{P}$ contains a subset of the points in $\mathtt{X}$.

Unfortunately, when building new models from range images, $\mathtt{P}$ partially overlaps $\mathtt{X}$ instead of being a subset of it. Schütz *et al.* propose a simple heuristic method of determining which corresponding point pairs belong to overlapping regions [15]. They theorize that point pairs whose distance is much greater than the separation of the centers of mass of two partially registered range images must be outliers. They calculate a binary weighting factor for each point pair as follows

$$\mathtt{w}_{i_n} = \begin{cases} 1 & \text{if } \left\| \mathbf{x}_{i_n} - \left(\mathbf{R}_n \mathbf{p}_{i_n} + \mathbf{T}_n\right)\right\|^2 < (\mathtt{c}\cdot\mathtt{s})^2 \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where $\mathtt{s}$ is the range scanner sampling resolution and $\mathtt{c}$ is an empirically determined threshold based on the separation of the centers of mass of the data sets. Those pairs whose two points are separated by more than a specified value are considered outliers and given a weight of zero.

Zhang has created a statistical model for classifying outlier point pairs [19]. He theorizes that the distances between corresponding point pairs are distributed as a Gaussian when the sample mean of these distances is small. Given a coarse registration of two range images

with significant overlap, the closer a pair of points is to each other, the more likely that they belong to the overlapping region. Zhang uses a heuristic method to set a threshold based on the estimated shape of the Gaussian distribution of the point pair distances. Like Schütz, Zhang sets the pair weights $\mathtt{w}_{i_n}$ to 0 or 1 based on whether a pair's points are closer than the threshold.

In addition to ICP-based registration approaches, many researchers use other iterative whole-surface registration techniques. The most popular variants are based on Chen and Medioni's work published the same year as Besl and McKay's ICP paper [2, 6, 12], which minimizes the distance along point normals of one surface to tangent planes of another surface using the following minimization function

$$\mathcal{g}_n(\mathbf{R}_n;\mathbf{T}_n) = \sum_{i_n} \mathtt{w}_{i_n} \left\| \mathbf{n}_{i_n}^T \left(\mathbf{x}_{i_n} - \left(\mathbf{R}_n \mathbf{p}_{i_n} + \mathbf{T}_n \right) \right) \right\|^2 = \sum_{i_n} \mathtt{w}_{i_n}, \tag{3}$$

where $\mathbf{n}_{i_n}$ is the surface normal at point $\mathbf{x}_{i_n}$ [2]. Under small angle assumptions, the minimization of $\mathcal{g}_n(\mathbf{R};\mathbf{T})$ may be linearized [3, pp. 125–127].[2]

In addition to developing novel registration algorithms, several existing comparative analyses of registration algorithms have been made. Some of the most notable to date include the following. Lorusso *et al.* have evaluated four closed-form solutions to Eq. (1) [11]. Pulli suggests when ICP versus point-to-plane minimization (both to be defined in the next section) should be used when performing multiview registration and introduces a few additional outlier classification heuristics [12]. Rusinkiewicz and Levoy have evaluated registration algorithms [13], focusing on computation efficiency in search of real-time performance [9]. Their experiments are performed on three synthetic range images being registered to exact copies of themselves. The primary evaluation metric used was the root mean squared (RMS) distance between point pairs using the known true correspondences.

2. IMPLEMENTATION

To facilitate making comparisons between different registration algorithms and variants, we have developed a registration test-bed environment. Our range image registration test-bed software uses the Visualization Toolkit, an open-source library for the manipulation and visualization of 2D, 3D, and higher-dimensional data [14]. The library contains an object hierarchy built to support componentized visualization pipelines. Our software builds upon the base library by supplying a pairwise ICP registration algorithm with pluggable variants to the base algorithm. The key ICP variants currently implemented and tested include:

- ICP iteration control: This feature uses Besl's criterion requiring the change in the mean squared surface between two ICP iterations to drop below a prespecified level [1]. Throughout the rest of this paper, we will refer to this level as the "exit criterion."
- Subsampling: As we load the source and destination meshes, we uniformly subsample the vertices topologically. For example, a subsampling factor of 4 means we select every fourth vertex in each mesh direction, or we evenly select $1/16$ of the points.

[2] Unfortunately, the rotation matrix produced by a naïve implementation of the method described by Chen is not orthogonal and shrinks the object if applied directly. To correct for errors introduced by making the small angle assumptions, we use an intermediate unit quaternion [18] to extract the rotation components of the $\mathbf{R}$ and produce a corrected version which does not scale the object.

• Outlier point classification: Individual point pairs may be either rejected as outliers or have their weights (w_{i_n} in Eq. (1)) otherwise adjusted based on some confidence criterion. To date, we have only explored outlier rejection and not pair weighting.

The following outlier point classification schemes have been implemented and tested:

• Schütz's distance thresholder: This classifier identifies outliers as those point pairs that are separated by "too much" space in an attempt to solve the problem of partially overlapping data sets [15]. The value used for s in Eq. (2) will henceforth be referred to as the classifier's *parameter*. For our experiments, we always set c in Eq. (2) to be the Euclidean distance between the centers of mass of the source and destination range images.

• Zhang's statistical outlier classifier: This classifier examines the statistical distribution of unsigned point pair distances to estimate which pairs are outliers [19]. s is used as this classifier's parameter.

In addition to these variant classes, our test-bed contains infrastructure to support additional ICP variants and other, non-ICP registration techniques. We also ported parts of our test-bed to Matlab for rapid prototyping of some of the variants. Our current implementation of the point-to-plane minimization algorithm uses this port.

3. QUALITATIVE EXPERIMENTS

In [5], we evaluated the aforementioned ICP variants to determine the effects of outlier classifier, uniform subsampling, and exit criterion on registration quality. We have expanded on that initial set of experiments by evaluating the point-to-plane minimization algorithm in addition to the ICP variants.

We found that for range image pairs that approximate Besl and McKay's requirement of full overlap, using no outlier classifier generally yielded the best results. For those pairs that had significant nonoverlapping regions, both of the classifiers generally yielded good results, with the classifier based on Zhang's work performing slightly better in most cases than the one based on the work of Schütz. The point-to-plane minimization sometimes performed even worse than not using an outlier classifier with the ICP minimization and sometimes was competitive with Zhang's and Schütz's classifiers.

We also found that although decimated data could be registered, those registrations tend to only be "good" in the context of their decimation. Once the range image pair is viewed undecimated, the registration imperfections readily manifest themselves. The greatest speed benefits relative to quality degradation occurred when we only decimated the source mesh. Additionally, we found that modifying the exit criterion had predictable results. As that criterion is lowered, the sequence simply goes further along the path it is following unless it first encounters numerical or algorithmic instabilities.

4. QUANTITATIVE EXPERIMENTS

4.1. Experimental Setup

We also wanted to perform some experiments with numerical results because our qualitative results are imprecise. Since the ground-truth registration is unobtainable for physically scanned range images using our range sensor, we developed a synthetic range image generator. We first selected two models built using commercially available software and range

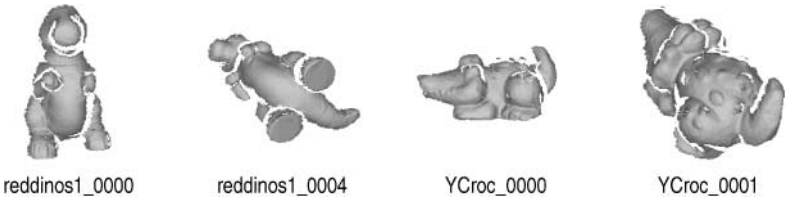

FIG. 1. Renderings of some of the range images used for the quantitative tests. For these tests, the *reddinos1_0000* and *YCroc_0000* range images are the destination images.

scans taken with our sensor. We then generated a set of synthetic range images from those complete models, recording the positions of the virtual camera (see Fig. 1). Table 1 gives pertinent data about the objects from which the range images were generated.

Next, we randomly introduced perturbations in rotation and translation independently. These transformations were made by rotating about the centroid of the source range image followed by translation (see Table 2). We performed tests on all combinations of 0°, 2°, and 8° of rotation error with 0, 2, and 8 mm of translation error. Five pairs of random vectors were generated for each combination. The first vector defines the direction of the normal about which the rotation was introduced, and the second random vector gives the direction of the introduced translation error. For the cases where no rotation or translation error was introduced, only one test was performed, not 10. The total number of times a subexperiment was performed with different error values was $3 \cdot 3 \cdot 5 - 4 = 41$. The nonclassifier parameter settings used for these quantitative tests are tabulated in Table 3.

Based on our qualitative test results and initial quantitative results, we made several modifications to our experimental setup.

First, our qualitative tests indicated that there is a significantly greater penalty in decimating the destination mesh as opposed to decimating the source mesh. As a result, we always used only nonsubsampled destination range images.

Second, for the ICP tests we experimented with throwing out any point pair matches where either point lay on its mesh edge, as suggested by Turk and Levoy [17]. We will refer to this process as *edge pruning* in this paper. Surprisingly, we found that this produced only extremely small differences in our results, as will be discussed in Section 4.3.

Third, we controlled the outlier classifier parameter settings more adaptively for the Schütz and Zhang implementations. For the base value of `s`, we used one half the "average sampling distance" of the destination mesh shown in Table 1. These values were found

TABLE 1
Mesh Statistics for the Destination Range Images Used in the Quantitative Experiments

Range image object	Number of points	Total surface area	Sampling density[a]	Average sampling distance[b]
YCroc	20,653	5,900 mm^2	3.50 points/mm^2	0.641 mm (0.641 mm)
reddinos1	14,239	15,178 mm^2	0.938 points/mm^2	1.18 mm (1.175 mm)

[a] The sampling density is calculated as the number of points divided by the total surface area. Note that this density gives a sense of how smooth the original object was before the synthetic range images were generated.

[b] The average sampling distance is calculated by finding the longest edge in each mesh polygon and averaging their lengths. The values in parentheses are the average sampling distance of the destination range image generated from the object.

TABLE 2
Exit Criterion for All Experiments

Criterion name	Qualitative		Quantitative	
	ICP	Point-to-plane	ICP	Point-to-plane
Change in RMS pair distance	0.3, 0.03, 0.003 mm	0.3, 0.03, 0.003 mm	N/A	N/A
Maximum number of iterations	N/A	N/A	25	25
Maximum time	30 min	30 min	15 min	30 min

by taking the longest edge of each mesh triangle and averaging these lengths [19]. Mesh triangles whose surface normals face more than 45° away from the camera are not used to calculate these values. In our experiments that used Zhang's classifier, we used this value, one fourth of this value, and four times this value as the parameter settings. For Schütz's classifier, we used this value, one half of this value, and two times this value. Additionally, we ran a batch of tests for the Zhang classifier with edge pruning enabled and the parameter set to zero to force it into a fall-back histogram peak-finding mode (see [19, pp. 126–127]).

Fourth, we chose a different set of criteria for terminating the registration iterations for the qualitative tests, as detailed in Table 2. In our initial tests, we found that often a sequence of ICP iterations would continue to make appreciable progress toward the correct solution, even when its progress was slow in terms of change in the RMS point pair distance. By changing the iteration stopping criteria, we allowed the registration to continue to progress. In all cases, we terminate the iterations when any of the criteria are met. For the quantitative ICP tests, we lowered the time limit due to increased implementation efficiencies. Our average running time was 20 s and our maximum running time was 2 min. Implementing an accelerated ICP algorithm, as described in [1], would be another alternative.

Fifth, for a limited set of additional tests, we experimented with adding isotropic Gaussian noise to each point in the destination range image to better simulate real-world data. The noise had a standard deviation proportional to the average sampling distance. Note that in addition to this Gaussian noise, all tests (even the "noiseless" ones) have quantization noise when the source and destination range images are different because the range images do not necessarily sample exactly the same surface locations.

4.2. Analysis Methodology

To evaluate our results, we used an error measure similar to the one employed by Rusinkiewicz and Levoy [13]. Specifically, we used the RMS distance between the source

TABLE 3
Nonclassifier Settings Chosen for the Quantitative Experiments

Object	Number of view pairs	Source image subsampling factors	Errors introduced		Number of tests performed
			Rotation	Translation	
reddinos1	6	1, 2, 4, 8	0°, 2°, 8°	0, 2, 8 mm	14,703 (222)[a]
YCroc	3	1, 2, 4, 8	0°, 2°, 8°	0, 2, 8 mm	7,360 (136)[a]
				Total	22,063 (358)[a]

[a] Values in parentheses represent those which are different for the point-to-plane tests (vs. the point-to-point minimization tests).

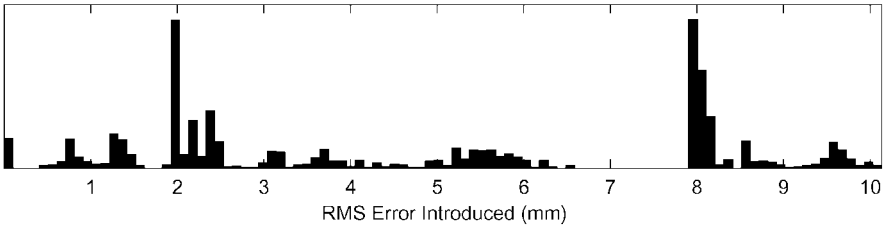

FIG. 2. Distribution of RMS errors introduced for all of the quantitative tests.

range image points in their final location to the same points in their correct location. Although we could have chosen to find the rotation and translation errors by reversing the process used to introduce the initial errors, this unified measure allowed us to use a single number for evaluation. The distribution of RMS errors that were introduced is shown in Fig. 2.

4.3. Results

We have broken the analysis of our quantitative results into the following sections:

1. Effects of outlier classifier type and parameter settings
2. Effects of subsampling
3. Timing and iteration control
4. Effects of noise

4.3.1. Effects of outlier classifier type and parameter settings. One of the most interesting observations we made was that many of the experiments we performed resulted in the registration process making the RMS error (see Section 4.2) much worse than it originally was.

Figure 3 contains a set of histograms showing the distribution of improvement of the RMS error for four collections of tests using range images from the noiseless *YCroc* object. In this graph, we have included one histogram from each ICP classifier type and one for the

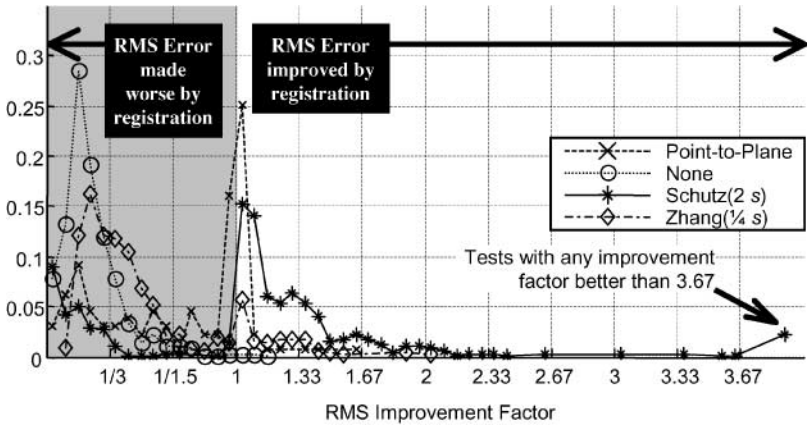

FIG. 3. Distribution of the RMS effects of select sets of tests. The vertical axis is the proportion of tests falling within the given histogram bin.

point-to-plane minimizer. For Schütz's and Zhang's classifiers, we chose parameter settings of 2s and 0.25s, respectively, because these represented the best results out of all of the parameter settings chosen for them (refer back to Section 4.1 for a detailed description of how s is calculated). Histogram bins on the left, shaded portion of the graph represent cases where the RMS error was made worse. Registration tests contributing to bins to the far left would have been classified as catastrophic failures in our qualitative tests. Bins on the right, unshaded side represent cases where the RMS error after registration was better than it was beforehand.

Using no outlier classifier almost always resulted in a decrease in registration quality, as we had expected. Also as we expected, using the Schütz classifier generally resulted in an improved registration. The two most significant features of its distribution are that it was most likely to have a small improvement, and that there are a significant number of tests for which the RMS error after registration was much more than four times better than the initial RMS error before registration.

Somewhat surprising was that using a point-to-plane minimizer often made the registration worse, but most often resulted in only minor net changes in terms of RMS error.

Even more surprising to us was that Zhang's classifier performed nearly as poorly as not using any outlier classifier at all in terms of RMS error for these noiseless tests. This is in contrast with our findings in our qualitative tests (see Section 3). We believe that this reversal in performance is due to a number of factors:

1. *Noise*: Our qualitative tests used only range images taken from an actual, imperfect physical range sensor, but our quantitative tests used only noiseless synthetic range images. Section 4.3.4 contains preliminary results for added Gaussian noise.
2. *Data warping*: Included in our qualitative tests were range images of two human heads, and humans are not capable of remaining perfectly rigid when repositioning for different scans.
3. *Different error criteria*: For the qualitative tests, our evaluations were based on seeing if there was a high degree of interpenetration for regions of low curvature and if good visual correspondence of high-curvature feature regions existed. Improvements in these subjective measures do not always correspond to improvements in an RMS measure.

A second surprising general observation is that performing edge pruning had an almost imperceptible effect on our results. Sometimes the results improved with edge pruning enabled, sometimes they became worse. In examining individual test cases, we found that although the RMS errors were different with and without this feature, the differences were often in the third or fourth significant digit of the RMS error.

For noiseless data we found that Schütz's classifier generally performed the best, followed by using point-to-plane minimization, then by Zhang's classifier, and using no outlier classifier produced the worst results.

4.3.2. Effects of subsampling. After observing general trends based on the outlier classification method, we examined the effects of subsampling on noiseless data. Tables 4, 5, and 6 summarize these results, broken down by classifier type and parameter settings. These tables indicate how often a particular group of tests improved the registration in terms of the RMS error. Table 4 summarizes the results for all of our tests, Table 5 contains data only for range images from the *YCroc* object, and Table 6 is for the *reddinos1* object. Note that the results for the *YCroc* object are significantly better than those for the *reddinos1* object.

TABLE 4
Percentage of Experiments Where the Registration Algorithm Improved the Actual Registration for Range Images from All Objects

Classifier	Parameter	Edges pruned	Source subsampling factor 1	2	4	8	Total
None	N/A	Y	*0.00%*	*0.00%*	*0.00%*	1.36%	0.34%
		N	*0.00%*	*0.00%*	*0.00%*	*1.08%*	*0.27%*
Zhang	0	Y	5.16%	5.69%	7.32%	16.80%	8.75%
	s/4	Y	6.78%	9.76%	10.03%	15.99%	10.64%
		N	6.78%	9.49%	10.30%	15.18%	10.43%
	s	Y	7.32%	9.76%	8.40%	15.72%	10.30%
		N	7.05%	9.49%	8.67%	14.91%	10.03%
	4s	Y	4.07%	4.61%	4.35%	6.23%	4.81%
		N	4.07%	4.61%	4.07%	5.42%	4.54%
Point-to-plane minimization	N/A	N/A	N/A	33.33%	23.81%	11.68%	13.97%
Schütz	s/2	Y	57.77%	53.28%	**52.35%**	43.22%	51.73%
		N	**58.20%**	**53.55%**	50.97%	44.07%	**51.76%**
	s	Y	50.95%	52.32%	47.68%	43.60%	48.64%
		N	51.23%	52.04%	48.50%	**44.69%**	49.11%
	2s	Y	44.44%	43.63%	40.76%	40.65%	42.37%
		N	45.26%	44.72%	40.49%	40.92%	42.85%

Note. The largest value in each column has been printed in boldface and the smallest has been italicized.

TABLE 5
Percentage of Experiments Where the Registration Algorithm Improved the Actual Registration for Range Images from the *YCroc* Object

Classifier	Parameter	Edges pruned	Source subsampling factor 1	2	4	8	Total
None	N/A	Y	*0.00%*	*0.00%*	*0.00%*	3.25%	0.81%
		N	*0.00%*	*0.00%*	*0.00%*	*2.44%*	*0.61%*
Zhang	0	Y	7.38%	8.94%	8.13%	16.26%	10.18%
	s/4	Y	13.82%	14.63%	15.45%	17.89%	15.45%
		N	13.82%	14.63%	15.45%	16.26%	15.04%
	s	Y	17.07%	17.07%	13.82%	20.33%	17.07%
		N	16.26%	17.07%	14.63%	18.70%	16.67%
	4s	Y	6.50%	6.50%	5.69%	8.13%	6.71%
		N	6.50%	6.50%	5.69%	6.50%	6.30%
Point-to-plane minimization	N/A	N/A	N/A	N/A	39.39%	28.16%	30.88%
Schütz	s/2	Y	76.42%	77.24%	76.86%	61.21%	73.08%
		N	77.87%	77.24%	76.03%	61.21%	73.24%
	s	Y	**79.67%**	**78.86%**	76.42%	69.92%	76.22%
		N	**79.67%**	78.05%	**78.86%**	**73.17%**	**77.44%**
	2s	Y	76.42%	73.17%	73.98%	63.41%	71.75%
		N	76.42%	72.36%	73.98%	62.60%	71.34%

Note. The largest value in each column has been printed in boldface and the smallest has been italicized.

TABLE 6
Percentage of Experiments Where the Registration Algorithm Improved the Actual Registration for Range Images from the *reddinos1* Object

Classifier	Parameter	Edges pruned	Source subsampling factor 1	2	4	8	Total
None	N/A	Y	*0.00%*	*0.00%*	*0.00%*	*0.41%*	*0.10%*
		N	*0.00%*	*0.00%*	*0.00%*	*0.41%*	*0.10%*
Zhang	0	Y	4.07%	4.07%	6.91%	17.07%	8.03%
	s/4	Y	3.25%	7.32%	7.32%	15.04%	8.23%
		N	3.25%	6.91%	7.72%	14.63%	8.13%
	s	Y	2.44%	6.10%	5.69%	13.41%	6.91%
		N	2.44%	5.69%	5.69%	13.01%	6.71%
	4s	Y	2.85%	3.66%	3.67%	5.28%	3.87%
		N	2.85%	3.66%	3.25%	4.88%	3.66%
Point-to-plane minimization	N/A	N/A	N/A	33.33%	6.67%	2.66%	3.60%
Schütz	s/2	Y	**48.36%**	41.15%	**40.00%**	34.45%	**41.04%**
		N	**48.36%**	**41.56%**	38.33%	**35.71%**	**41.04%**
	s	Y	36.48%	38.93%	33.20%	30.33%	34.73%
		N	36.89%	38.93%	33.20%	30.33%	34.84%
	2s	Y	28.46%	28.86%	24.08%	29.27%	27.67%
		N	29.67%	30.89%	23.67%	30.08%	28.59%

Note. The largest value in each column has been printed in boldface and the smallest has been italicized.

When no classifier is used, the only times that the registration improved were when one out of every 64 original vertices were used. We believe that these cases where the registration improved were "lucky" tests where most of the points outside of the overlapping region were subsampled out. Interestingly, the same trend applies to our tests that use Zhang's outlier classifier: increasing the subsampling factor improves the results. Additionally, we noticed that for the *YCroc* tests using $\mathtt{s}$ as the parameter yielded the best results. For *reddinos1*, using $\mathtt{s} = 4$ was best with low subsampling factors, and at high subsampling factors forcing the classifier into its fall-back mode using 0 as the parameter yielded the best results.

In contrast, the point-to-plane minimization technique tended to be the most sensitive to subsampling of the source range image. For the *YCroc* object, too few tests to be statistically reliable were successfully completed when a subsampling factor less than four was used. For a factor of four, the percentage of improved registrations lies between the best results of Zhang's classifier and the worst results of Schütz's. For a factor of eight, the results are still better than those from Zhang's classifier, but are significantly worse than those from the denser mesh. The *reddinos* object generated similar results, but with more dramatic penalties for excessively subsampling the source range image. As noted by Rusinkiewicz and Levoy [13], nonuniform decimation procedures are more appropriate such as normal-space sampling and feature-based decimation. We believe that using one of these techniques would greatly improve these results.

Once again, Schütz's classifier generated results closer to what we expected. For the range images from the *YCroc* object, the results remain relatively stable until a subsampling factor of eight is used. For the *reddinos1* object, the results are a little more mixed, but the general trend of high subsampling resulting in fewer tests with improvements remains.

4.3.3. Timing and iteration control. Although the focus of our experiments was not on improving execution speed and numerous performance improvements could readily be made to our implementations, we will briefly discuss the speed characteristics of our ICP implementation. Our point-to-plane minimization implementation was extremely inefficient computationally, so we will omit it from our discussion.

For our tests, each ICP iteration took approximately one second to perform on a 450 MHz Pentium II processor running Linux. Approximately half of that time was spent building a kd-tree [7] of the destination mesh. A more efficient implementation would only build the tree once or would use a different method for performing the nearest-neighbor searches. A few examples of more efficient search techniques are presented by Chen and Medioni [2], Rusinkiewicz and Levoy [13], and Greenspan and Godin [8]. Because the bulk of the computational time spent in an ICP iteration is on the nearest neighbor search, the classifier type had little effect on performance.

In terms of the number of iterations used, all of the classifiers used all 25 allowed iterations nearly always, except for the Schütz classifier. It virtually always terminated the ICP loop early due to degeneracies that destabilized eigenvector extraction algorithm used to minimize Eq. 1 [1, 11].

4.3.4. Effects of noise. After observing that Schütz's classifier performed better than Zhang's for noiseless data, we created a set of experiments on the undecimated *YCroc* object. We added Gaussian noise with a standard deviation equal to $\frac{1}{4}$ the average sampling distance to each point in the destination range image. We then tested the ICP variants and determined what percentage of the tests resulted in an improved registration. We found that, consistent with the qualitative experiments, Zhang's classifier performed significantly better than Schütz's, and it was also much better than no outlier classifier.

5. CONCLUSIONS

Out of the algorithms we tested, we make the following conditional recommendations:

1. If registering partially overlapping range images from potentially noisy data and initializing the registration manually, we recommend using Zhang's outlier classifier with small parameter values. In our qualitative tests, we found that it was best able to avoid catastrophic failures and was able to provide a higher degree of interpenetration and matching of important feature areas when compared to using Schütz's classifier or no classifier at all.

2. If registering partially overlapping range images from noiseless data with the initial registration containing small random errors in rigid rotation and translation, we recommend using Schütz's classifier with a parameter value near the average sampling distance of the destination mesh.

3. Irrespective of the technique chosen, the setting of parameters specific to the algorithm must be done with care, as these settings can affect performance significantly. Moreover, the noise level of the data must be assessed critically through experimentation prior to commitment to a specific technique. While truly noise-free data are impossible to obtain from real sensors, some types of sensors can produce much lower noise levels than others.

REFERENCES

1. P. J. Besl and N. D. McKay, A method for registration of 3-d shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 1992, 239–256.
2. Y. Chen and G. G. Medioni, Object modeling by registration of multiple range images, *Image Vision Comput.* **10**, 1992, 145–155.
3. Y. Chen, *Description of Complex Objects Using Multiple Range Images*, Ph.D. dissertation, University of Southern California, August 1994.
4. B. Curless and M. Levoy, A volumetric method for building complex models from range images, in *Proc. SIGGRAPH, August 1996*, pp. 303–312.
5. G. Dalley and P. Flynn, Range image registration: A software platform and empirical evaluation, in *Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling, 28 May–1 June, 2001*, Vol. 1, pp. 246–253.
6. C. Dorai, G. Wang, A. K. Jain, and C. Mercer, Registration and integration of multiple object views for 3d model construction, *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, January 1998, 83–89.
7. J. H. Friedman, J. L. Bentley, and R. A. Finkel, An algorithm for finding best matches in logarithmic expected time, *ACM Trans. Math. Software* **3**, 1977, 209–226.
8. M. Greenspan and G. Godin, A nearest neighbor method for efficient icp, in *Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling, 28 May–1 June 2001*, Vol. 1, pp. 161–168.
9. O. Hall-Holt and S. Rusinkiewicz, Stripe boundary codes for real-time structured-light range scanning of moving objects, in *Proceedings of the Eighth IEEE International Conference on Computer Vision, July 2001*, Vol. 2, pp. 359–366.
10. H. Hoppe, T. DeRose, T. Duchamp, M. Halstead, H. Jin, J. McDonald, J. Schweitzer, and W. Stuetzle, Piecewise smooth surface reconstruction, in *Proceedings of SIGGRAPH, July 1994*, pp. 295–302.
11. A. Lorusso, D. W. Eggert, and R. B. Fisher, A comparison of four algorithms for estimating 3-d rigid transmformation, in *BMVC95, 1995*.
12. K. Pulli, Mulitiview registration for large data sets, in *Second International Conference on 3-D Digital Imaging and Modeling, October 1999*.
13. S. Rusinkiewicz and M. Levoy, Efficient variants of the ICP algorithm, in *Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling, 28 May–1 June, 2001*, pp. 145–152.
14. W. Schroeder, K. Martin, and B. Lorensen, *The Visualization Toolkit*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1998.
15. C. Schütz, T. Jost, and H. Hügli, Semi-automatic 3D object digitizing system using range images, in *Proceedings of Asian Conference on Computer Vision, Jan. 1998*.
16. I. Stamos and P. Allen, 3d model construction using range and image data, in *Proceedings of Computer Vision and Pattern Recogntion, June 2000*, I:531–536.
17. G. Turk and M. Levoy, Zippered polygon meshes from range images, in *Proceedings on SIGGRAPH, July 1994*, pp. 311–318.
18. M. W. Walker, L. Shao, and R. A.Volz, Estimating 3-d location parameters using dual number quaternions, *Comput. Vision Graphics Image Process. Image Understanding* **54**, 1991, 358–367.
19. Z. Y. Zhang, Iterative point matching for registration of free-form curves and surfaces, *Internat. J. Comput. Vision* **13**, 1994, 119–152.