A Frequency Domain Approach to Registration Estimation in Three-Dimensional Space

Phillip Curtis, *Student Member, IEEE*, and Pierre Payeur, *Member, IEEE*

***Abstract*—Autonomous robotic systems require automatic registration of data that are collected by on-board sensors. Techniques requiring user intervention are unsuitable for autonomous robotic applications, whereas iterative-based techniques do not scale well as the data set size increases and, additionally, tend toward locally minimal solutions. To avoid the latter problem, an accurate initial estimation of the transformation is required for iterative algorithms to properly perform. However, in some situations, an initial estimate of the transformation may not be readily available; hence, a method that does not require such an initial estimate nor descends into local minima is desirable. The method presented in this paper takes advantage of the multidimensional Fourier transform, which inherently decouples the estimation of the rotational parameters from the estimation of the translational parameters, to compute 3-D registration between range images without requiring an initial estimation of the transformation and avoiding problems of the classical iterative techniques. Using the magnitude of the Fourier transform, an axis of rotation is estimated by determining the line that contains the minimal energy differential between two rotated 3-D images. A coarse-to-fine approach is used to determine the angle of rotation from the minimal sum of the squared difference between the two rotated images. Due to the Hermitian symmetry introduced by the Fourier transform, two possible solutions for the angle of rotation exist. The proper solution is identified through the use of a phase-correlation technique, and the estimate of translation is simultaneously obtained. Experimental results and an extended performance evaluation illustrate the accuracy that can be achieved by the proposed registration technique on simulated and on real range images. Last, a comparison of computational stability with that of the classical iterative closest point method is presented.**

***Index Terms*—Autonomous robotics, data fusion, frequency domain, pose estimation, range imaging, registration, 3-D modeling.**

I. Introduction

THE proliferation of low-cost high-quality range sensing systems has led to their use for many different purposes such as creating virtual objects for virtual museums, space exploration, and games. Registration estimation involves determining the rotations and translations that are required to align one image with another. The prevalent methods for registering range images involve relying on the positional sensors of the data acquisition apparatus, relying on complex feature detection, and matching algorithms or iterative algorithms that may converge to an incorrect solution. These methods offer limited scalability, and their execution time dramatically increases with an increase in the number of points in the data sets to be registered.

This paper, which is an extended version of [1], presents a registration estimation algorithm for 3-D measurements that is automatic, does not need any initial estimate of the transformation parameters, and requires no *a priori* knowledge of the object that is being registered, with the only assumption being that there is sufficient overlap between data sets for the algorithm to produce an accurate estimate. The proposed algorithm takes advantage of the fact that the Fourier transform decouples the estimation of the rotational parameters from the estimation of the translational parameters. This is accomplished by separating the frequency information into a magnitude and a phase component. The axis of rotation is determined, followed by the angle of rotation. Due to the Hermitian property of the Fourier transform, there are two possible solutions for the rotational parameters, separated by 180°. To determine the correct rotational parameters, as well as estimate the translational parameters, a phase correlation utilizing the frequency domain is applied to each possible solution. The solution that produces the most impulsive phase correlation is selected as the correct solution, with the location of the impulse corresponding to the translational parameters.

The proposed algorithm has been tested on 3-D data sets and also provides the framework for a multiple dimension extension. The theoretical description in the following section is generalized to the multiple dimensional case. Experimental results are presented for the 3-D case using pure range data, and an analysis of the quality of the registration estimates is presented, illustrating the differences between data clouds that are obtained through a real-life data acquisition setup and data clouds that are obtained through a simulated environment. Furthermore, the results are evaluated according to precision in the registration as well as execution time.

II. Review of Techniques

Traditionally, registration estimation has been performed in the space domain. The most widely explored approach to solve the registration estimation problem is the iterative approach, and the most widely adopted approach is the iterative closest point (ICP) algorithm and its various reincarnations and modifications. The ICP algorithm, initially proposed by Besl and McKay [2], describes a method for registering a set of 3-D data (P) with a reference or model data set (X). The method

Manuscript received July 15, 2006; revised September 21, 2007. This work was supported by the Canadian Foundation for Innovation, the Ontario Innovation Trust, and the Natural Science and Engineering Research Council of Canada.

The authors are with Vision, Imaging, Video, and Autonomous Systems Research Laboratory, School of Information Technology and Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: pcurtis@site.uottawa.ca; ppayeur@site.uottawa.ca).

Digital Object Identifier 10.1109/TIM.2007.909499

operates by calculating the closest points in set P with those in set X. From this matching of closest points, an estimate for the registration parameters is made. P is transformed by this estimate, and the mean squared error (MSE) between the transformed P and X is computed. If the MSE is not beneath a predefined threshold, the estimation process is repeated using the newly transformed data set P; otherwise, the current estimate of the registration is the solution.

The ICP algorithm, as with most iterative convergence algorithms, tends toward the closest local minima when using discrete data sets. The solution obtained may or may not correspond to the global minimizing solution, and hence, the ICP algorithm requires an initial estimate that is close to the actual registration parameters or a particularly transformed data set to obtain a precise solution. The estimation of the registration parameters is performed in two stages—by first calculating the rotational parameters, and then using the rotational parameters to calculate the translational parameters; hence, if the rotation is incorrect, the translation will be incorrect. The advantages of the ICP algorithm are its precision, flexibility, and ease of use, and its main disadvantage is its tendency to converge to local minima solutions instead of the proper global minimum solution without a close initial estimate.

There are several papers available in the literature [3]–[5] that discuss modifications to the ICP algorithm to provide better convergence toward the global minimum. This is achieved either by transforming the data, altering the matching criterion, and providing a close initial estimate of the registration parameters, or by acceleration of the algorithm at various steps.

Another alternative to using the ICP class of algorithms for registration estimation are those that rely on matched features or points. A closed-form solution for the estimation of registration parameters exists for any set of three matched noncollinear points [6]. The process involves finding first an estimate of the translation based on the difference in position of the centroids, and then estimating the rotation about the centroid using standard linear-system-solving techniques. The solution resulting from such a situation tends to contain errors due to the inexactness of features and matches between features, and, as a result, typically many more points are used utilizing a least squares approach. Other techniques using this approach involve refining the extraction and matching of these features [7], [8] to account for, and reduce, sources of error. For this class of algorithms, the majority of processing occurs in the feature extraction and feature matching phases. Numerous applications of feature-based registration estimation are found in artifact modeling [9] and mobile robot localization [10].

Alternatives to registration in the space domain have been proposed to avoid the need to match features, and to deal with unorganized data clouds, by taking advantage of some characteristics of the Fourier transform. By avoiding the feature detection, extraction, and matching steps that occur in classical registration techniques, frequency domain algorithms avoid possible imprecision and poor matches that are inherent to this sort of process.

Lucchese *et al.* [11] extended their previous work with frequency domain registration estimation in the 2-D case [12] to the 3-D case. As previously stated, the Fourier transform decouples the rotational parameter estimation from the translation estimation. To prevent the impulsive nature of the effect of singularities on the frequency spectrum, the data set is convolved with a spherical Gaussian kernel with standard deviation of 0.05 and diameter of seven voxels. This creates a spherical solid region about each point, with density decaying with distance from the point.

Lucchese *et al.* estimated the axis of rotation by determining the radial projections of the difference image through the (0, 0, 0) frequency point (DC). By performing this step, the determination of the axis of rotation is reduced to determining the minimum of a 2-D function. The estimate is further refined by resampling the function in a thin cylinder around the estimated axis of rotation, followed by determining the angular histograms of the projection of the cylinder onto the three orthogonal elementary planes, and then, finally, determining the angles corresponding to the maxima for each histogram. These maxima are used to determine the optimal axis of rotation. With the axis of rotation now determined, the coordinate system is transformed such that the problem becomes a 2-D rotation problem about the w-axis, as illustrated in Fig. 1.

After the problem is reduced to a 2-D problem of estimating the rotation about w, a 2-D image is created from the 3-D image by integrating along the w-axis. The complexity of finding the angle of rotation is further reduced to the 1-D case by changing to polar coordinates and integrating along the distance parameter. Last, the angle of rotation is determined by the peak of the correlation between the corresponding 1-D functions of each range image.

Due to the Hermitian symmetry of the Fourier transform, there are two complementary solutions, which are separated by 180°. The ambiguity between solutions is resolved in the estimation of translation. To estimate the translation, the original data are rotated by each solution and transformed into the frequency domain. A phase correlation between images is performed. The phase correlation function corresponding to the correct solution will be impulsive in nature, and the location of the impulse corresponds to the translation. The phase correlation function corresponding to the incorrect solution will not be impulsive in nature.

To avoid the computational penalty of performing a 3-D phase correlation, the authors perform three 1-D phase correlation functions based on the projections onto the primary axes. Additionally, to minimize the numerical errors that are involved in the computation of differences between small numbers, a logarithmic difference function is used for the estimation of the rotation axis. Last, to reduce the effects of discrimination in estimating the angle of rotation, a minimal search-based windowing function is used along each plane, and these minima are added together to form the 2-D image.

The method is used to produce an initial estimate to be refined by the ICP algorithm. The disadvantages of this algorithm are the computational cost of applying the fast Fourier transform (FFT) several times (one time for each image for the estimation of the axis of rotation, one more time for one image in the estimation of the angle of rotation, and, finally, two more times on one image for the estimation of translation), as well as the need for computing several histograms. On the other hand,

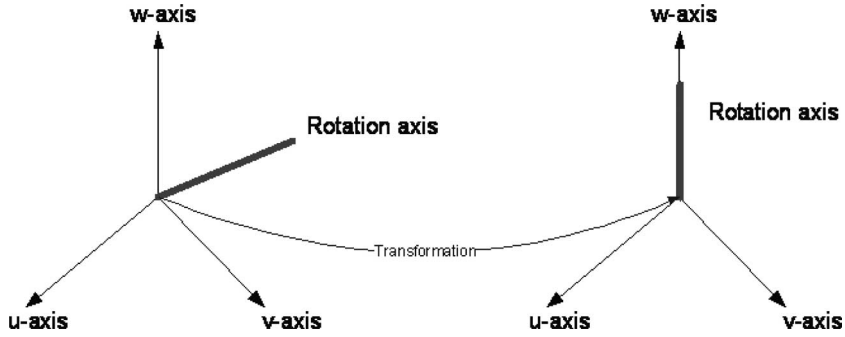

Fig. 1. Axis transformation.

this frequency domain algorithm eliminates the need to extract and match features, and avoids local minima solutions that may occur with traditional iterative algorithms.

III. Registration in the Frequency Domain

This section describes the theory that is required for frequency domain registration techniques, expanding on the mathematical arguments put forth by Lucchese *et al.* [11].

Let N be the number of dimensions of the signal and M be the diagonal matrix containing the reciprocal of the size of each dimension. Let $\mathrm{Im}_1[\vec{n}]$ be the space domain samples from one viewpoint, and let $\mathrm{Im}_2[\vec{n}]$ be the same space domain samples from a different viewpoint, with the vector $\vec{n}$ indicating the sample location in discrete Cartesian coordinates, with respect to the origin of the image. Let the rigid transformation between Im_1 and Im_2 be represented by

$$\mathrm{Im}_1[\vec{n}] = \mathrm{Im}_2[R\vec{n} + T] \tag{1}$$

where R is the $N \times N$ rotation matrix, and T is the $N \times 1$ translation vector.

The discrete Fourier transform (DFT) representations of Im_1 and Im_2 are

$$F_{\mathrm{Im}_1}[\vec{k}] = \sum_{n_N=0}^{M_N-1} \sum_{n_{N-1}=0}^{M_{N-1}-1} \cdots \sum_{n_1=0}^{M_1-1} \mathrm{Im}_1[\vec{n}] e^{-j2\pi \vec{k}^T M \vec{n}} \tag{2}$$

$$F_{\mathrm{Im}_2}[\vec{k}] = \sum_{n_N=0}^{M_N-1} \sum_{n_{N-1}=0}^{M_{N-1}-1} \cdots \sum_{n_1=0}^{M_1-1} \mathrm{Im}_2[\vec{n}] e^{-j2\pi \vec{k}^T M \vec{n}} \tag{3}$$

with

$$\vec{n} = \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_N \end{bmatrix}, \quad \vec{k} = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_N \end{bmatrix}, \quad M = \begin{bmatrix} \frac{1}{M_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{M_2} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{M_N} \end{bmatrix} \tag{4}$$

where $\vec{k}$ is the vector representing the N-dimensional discrete frequency index; $n_1, n_2, \ldots, n_N$ are the components of the N-dimensional vector $\vec{n}$; and $M_1, M_2, \ldots, M_N$ are the discrete size of each of the respective N dimensions.

If the rigid transformation depicted in (1) is applied to (3), the relationship between F_{Im_1} and F_{Im_2} can be found. Note that the F_{Im_2} notation is slightly modified here to introduce the rigid linear transformation, i.e.,

$$\begin{aligned} F_{\mathrm{Im}_2}[\vec{k}] &= \sum_{n_N=0}^{M_N-1} \sum_{n_{N-1}=0}^{M_{N-1}-1} \cdots \sum_{n_1=0}^{M_1-1} \mathrm{Im}_2[R\vec{n} + T] \\ &\quad \times e^{-j2\pi \vec{k}^T M(R\vec{n}+T)} \\ F_{\mathrm{Im}_2}[\vec{k}] &= \sum_{n_N=0}^{M_N-1} \sum_{n_{N-1}=0}^{M_{N-1}-1} \cdots \sum_{n_1=0}^{M_1-1} \mathrm{Im}_1[\vec{n}] \\ &\quad \times e^{-j2\pi \vec{k}^T M(R\vec{n})} e^{-j2\pi \vec{k}^T MT} \\ F_{\mathrm{Im}_2}[\vec{k}] &= \left(\sum_{n_N=0}^{M_N-1} \sum_{n_{N-1}=0}^{M_{N-1}-1} \cdots \sum_{n_1=0}^{M_1-1} \mathrm{Im}_1[\vec{n}] e^{-j2\pi \vec{k}^T MR\vec{n}} \right) \\ &\quad \times e^{-j2\pi \vec{k}^T MT}. \end{aligned} \tag{5}$$

The transpose of a rotation matrix being its inverse, one can write:

$$\begin{aligned} F_{\mathrm{Im}_2}[\vec{k}] &= \left(\sum_{n_N=0}^{M_N-1} \sum_{n_{N-1}=0}^{M_{N-1}-1} \cdots \sum_{n_1=0}^{M_1-1} \mathrm{Im}_1[\vec{n}] e^{-j2\pi M \vec{k}^T (R^{-1})^T \vec{n}} \right) \\ &\quad \times e^{-j2\pi \vec{k}^T MT} \\ F_{\mathrm{Im}_2}[\vec{k}] &= \left(\sum_{n_N=0}^{M_N-1} \sum_{n_{N-1}=0}^{M_{N-1}-1} \cdots \sum_{n_1=0}^{M_1-1} \mathrm{Im}_1[\vec{n}] e^{-j2\pi M (R^{-1}\vec{k})^T \vec{n}} \right) \\ &\quad \times e^{-j2\pi \vec{k}^T MT}. \end{aligned} \tag{6}$$

If a change in variables to the above equation is made ($\vec{k} \rightarrow R\vec{k}$), it then becomes

$$\begin{aligned} F_{\mathrm{Im}_2}[R\vec{k}] &= \left(\sum_{n_N=0}^{M_N-1} \sum_{n_{N-1}=0}^{M_{N-1}-1} \cdots \sum_{n_1=0}^{M_1-1} \mathrm{Im}_1[\vec{n}] e^{-j2\pi \vec{k}^T M \vec{n}} \right) \\ &\quad \times e^{-j2\pi (R\vec{k})^T MT}. \end{aligned} \tag{7}$$

It is observed that in (7), the translation component is independent of $\vec{n}$, and that the part of the equation in brackets

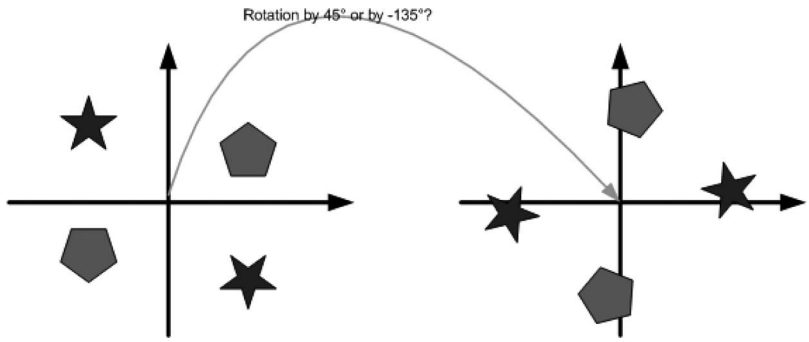

Fig. 2. Example of Hermitian symmetry of the magnitude of the Fourier transform.

is equal to F_{Im_1} [see (2)]. This leads to the deduction of the relationship between F_{Im_1} and F_{Im_2}, i.e.,

$$F_{\mathrm{Im}_2}[R\vec{k}] = F_{\mathrm{Im}_1}[\vec{k}]e^{-j2\pi(R\vec{k})^T MT}. \tag{8}$$

From (8), two separate equations can be developed—one for solving for the rotation matrix R, and one for solving for the translation vector T when the rotation matrix is given. This is accomplished by separating the equation into amplitude and phase components, i.e.,

$$\left|F_{\mathrm{Im}_2}[R\vec{k}]\right| = \left|F_{\mathrm{Im}_1}[\vec{k}]\right| \tag{9}$$

$$\angle F_{\mathrm{Im}_2}[R\vec{k}] = \angle F_{\mathrm{Im}_1}[\vec{k}] - 2\pi(R\vec{k})^T MT. \tag{10}$$

From (9), it is possible to deduce the rotation matrix R using the amplitude spectra of Im_1 and Im_2. Once the rotation matrix is known, the translation vector T can be solved for through the use of a phase correlation method using the phase spectra of Im_1 and a derotated Im_2 [see (10)].

A. Determining the Rotation Matrix

It is well known that all rotations in the 3-D space can be represented as a rotation about an axis of rotation, a fact that the frequency domain technique takes advantage of. By rotating an object in space about an axis, the positions of all points change—except those belonging to the axis of rotation. This, when mathematically described, corresponds to multiplying a scaled version of an eigenvector of a matrix with the matrix itself (note that the axis of rotation corresponds to the eigenvector with an eigenvalue equal to one for the rotation matrix).

This holds true in the frequency domain since the rotation is not affected by a Fourier transform; hence, by determining the zero line in the difference function of the magnitude of the Fourier transforms of the images to be registered, the axis of rotation can be determined as the locations where

$$\left|F_{\mathrm{Im}_2}[R\vec{k}]\right| = \left|F_{\mathrm{Im}_2}[\vec{k}]\right| = \left|F_{\mathrm{Im}_1}[\vec{k}]\right|. \tag{11}$$

To determine the angle of rotation, one can then rotate F_{Im_1} about the axis of rotation until the rotated F_{Im_1} is equal to F_{Im_2}. This can be accomplished by minimizing the MSE between transformed magnitude images, i.e., $|F_{\mathrm{Im}_1}|$ and $|F_{\mathrm{Im}_2}|$, for each value of the angle of rotation.

To determine the rotation matrix, there are a few issues that must be addressed. First and foremost, the Fourier transform of real data is Hermitian symmetric. In other words, the frequency domain spectrum for the frequency values between zero and π is also represented by the complex conjugate of the values contained between 2π and π, i.e.,

$$F(w) \rightarrow F^*(2\pi - w). \tag{12}$$

In the discrete mapping of the DFT, where $\vec{K}$ is the column vector containing the discrete size along each dimension, and $\vec{k}$ is a discrete frequency location, the Hermitian symmetry effect is represented by

$$F[\vec{k}] = F^*[\vec{K} - \vec{k}], \qquad \vec{K} = \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_N \end{bmatrix}. \tag{13}$$

This mapping is beneficial and detrimental to determining the rotation matrix. The benefit is that only half of the frequency data are unique; therefore, only half of the DFT needs to be computed. The drawback to the Hermitian symmetric mapping is that two solutions for rotation are obtained with a separation of 180°. This is exemplified in Fig. 2.

B. Determining the Translation Vector

Once the rotational parameters have been estimated, the translation vector can be determined. Due to the presence of two complementary solutions for the angle of rotation, a method of determining the proper solution is needed. Fortunately, (10) provides solutions to the problem of solution selection as well as estimation of the translational parameters. The application of (10) is equivalent to a phase correlation of Im_1 with Im_2. The secondary solution for R (which is denoted by R') will provide a reflection, in addition to the rotation, about the axis of rotation.

The solution corresponding to the proper solution will provide an impulse-like response, whereas the complementary solution will provide a nonimpulsive response. The solution in each case will be rarely strictly impulsive; however, one solution will be more impulse-like than the other solution. This

enables us to select the solution based on whether the cross correlation between Im_1 and the derotated versions of Im_2 produces an impulse-like function, with the location of the impulse signifying the estimation of the translational parameters.

IV. Proposed Algorithm

A. Determining the Axis of Rotation

To determine the axis of rotation, the difference between the two magnitudes of the FFT must be calculated. The straight difference is not reliable in practice since the magnitudes of the FFT may be small, as well as the effects caused by noise in the images may introduce incorrect minima in the difference function. In [11], a normalized logarithmic difference is used to rectify this problem. Although this does effectively work, calculating logarithmic differences is quite processor intensive compared to calculating straight differences. The difference function that is developed for the proposed algorithm is the normalized percentage difference, which ensures that the values with a large relative difference, even when the magnitudes are small, produce a large difference, and that the values with a small relative difference, even when the magnitudes are large, produce a small difference, while having a lower processing cost compared to the normalized logarithmic difference. The normalized percentage difference is defined as

$$SE[\vec{k}] = \left(\frac{\left| \frac{F_{\mathrm{Im}_1}[\vec{k}]}{F_{\mathrm{Im}_1}[\vec{0}]} \right| - \left| \frac{F_{\mathrm{Im}_2}[\vec{k}]}{F_{\mathrm{Im}_2}[\vec{0}]} \right|}{\max\left\{ \left| \frac{F_{\mathrm{Im}_1}[\vec{k}]}{F_{\mathrm{Im}_1}[\vec{0}]} \right|, \left| \frac{F_{\mathrm{Im}_2}[\vec{k}]}{F_{\mathrm{Im}_2}[\vec{0}]} \right| \right\}} \right)^2. \tag{14}$$

In this difference, the frequency domain images are normalized with respect to the zero location, as the zero location provides a good indication of the energy present in the images. The divisor is then chosen to be the maximum of the two points in the difference to ensure that the values fall between zero and one.

The minimal weight zero crossing line, which corresponds to the axis of rotation, is determined through the use of a moving-window search algorithm. This algorithm determines the minimal weight path crossing the frequency domain origin within a small window. The window is successively moved away from the origin along the minimal weight path, resulting in the axis of rotation being determined with higher precision as the window moves further away.

B. Determining the Angle of Rotation

The angle of rotation is determined by first subsampling the frequency domain. The frequency domain is further reduced such that the only remaining frequency locations lie between $-\pi/2$ and $\pi/2$ along each dimension. This step is performed to minimize the number of computations to be performed.

The angle of rotation is then coarsely iterated between $-\pi$ and π. Due to possible numerical errors made at various stages of the calculation, the full range of $-\pi$ to π is used to determine the optimal angle, as opposed to the minimal required range of 0 to π. The frequency locations selected in F_{Im_2} are rotated, followed by calculating the squared error normalized magnitude percentage difference, which is defined as

$$SE'[\mathrm{Angle}] = \sum_{\forall \vec{k}} \left(\frac{\left| \frac{F_{\mathrm{Im}_1}[\vec{k}]}{F_{\mathrm{Im}_1}[\vec{0}]} \right| - \left| \frac{F_{\mathrm{Im}_2}\left[R(\vec{R}_{\mathrm{Axis}}, \mathrm{Angle})\vec{k}\right]}{F_{\mathrm{Im}_2}[\vec{0}]} \right|}{\max\left\{ \left| \frac{F_{\mathrm{Im}_1}[\vec{k}]}{F_{\mathrm{Im}_1}[\vec{0}]} \right|, \left| \frac{F_{\mathrm{Im}_2}\left[R(\vec{R}_{\mathrm{Axis}}, \mathrm{Angle})\vec{k}\right]}{F_{\mathrm{Im}_2}[\vec{0}]} \right| \right\}} \right)^2. \tag{15}$$

The angle corresponding to the minimal error is selected for the new midpoint in the coarse-to-fine search. The selected points in F_{Im_2} are now rotated more finely between $(\mathrm{angle} - \pi/2)$ and $(\mathrm{angle} + \pi/2)$, and again, the error is calculated, and the minimal error angle is selected. This task continues until the desired angular precision is reached. The angular search range is cut in half and centered on the minimal error angle from the previous coarser iteration, and the range is divided up according to how many frequency divisions are desired.

Once the angle has been determined, it should be noted that due to the Hermitian symmetrical nature of the Fourier transform, there exists a second solution to the angle of rotation that differs by 180° (π radians) from the determined angle, i.e.,

$$\mathrm{Angle}' = \mathrm{Angle} \pm \pi. \tag{16}$$

The selection of the correct solution is performed in the subsequent section.

In the study of Lucchese *et al.* [11], the angle of rotation is determined involving a 1-D correlation technique after integrating the images along the axis of rotation and along the radius. These steps are complex and require an additional forward and inverse Fourier transform, as well as determining a maximum of a noisy 1-D function. The coarse-to-fine approach eliminates the need for the complex correlation technique and zooms in on the least squared error solution for the frequency locations selected. This ensures that accuracy is maintained while keeping the algorithm simple and easy to understand. As the number of frequency points increases, so does the accuracy of the algorithm. Also, as the number of frequency divisions increases, the precision of the algorithm increases. The computation time of the angle of rotation increases as the previously mentioned parameters increase.

C. Solution Selection

Due to the previously mentioned Hermitian symmetry of the frequency domain, there exist two possible solutions for the rotation. To properly disambiguate the solutions, a phase correlation of each solution must be performed, as suggested by Lucchese *et al.* [11]. For this to occur, Im_2 must be derotated by each of the complementary rotational solutions, producing $\mathrm{Im}_2^a[\vec{n}]$ and $\mathrm{Im}_2^b[\vec{n}]$, i.e.,

$$\mathrm{Im}_2^a[\vec{n}] = \mathrm{Im}_2\left[R^{-1}(\vec{R}_{\mathrm{Axis}}, \mathrm{Angle}) \cdot \vec{n}\right] \tag{17}$$

$$\mathrm{Im}_2^b[\vec{n}] = \mathrm{Im}_2\left[R^{-1}(\vec{R}_{\mathrm{Axis}}, \mathrm{Angle} - \pi) \cdot \vec{n}\right]. \tag{18}$$

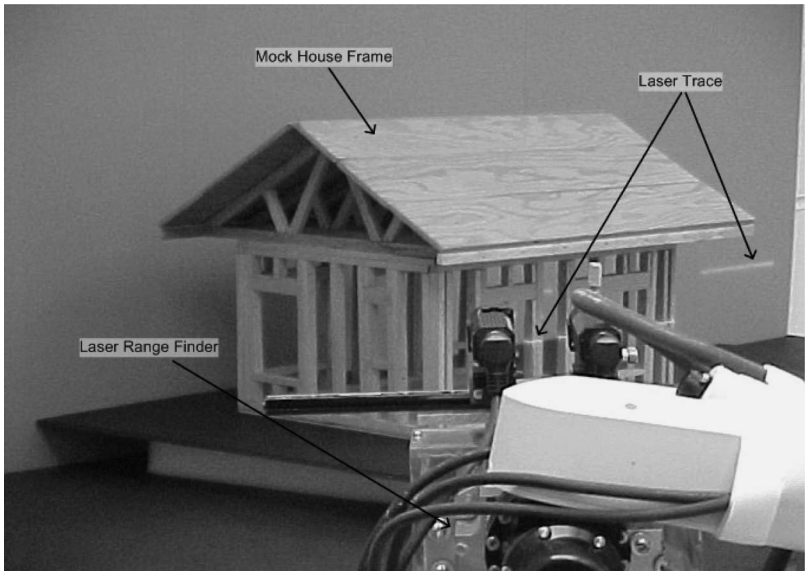

Fig. 3. Robotic laser range data acquisition system scanning the mock house frame.

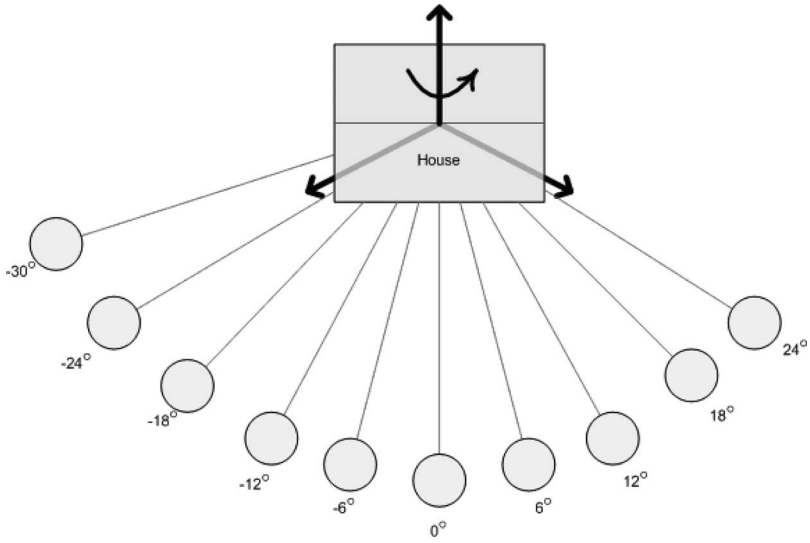

Fig. 4. Illustrating scanning positions that are used for experimentation and evaluation of the registration approach.

This step requires going back to the space domain due to the phase discontinuities that are present in the frequency domain from the sparse data sets.

In the proposed approach, the phase correlation between Im_1 and Im_2^a, and Im_1 and Im_2^b is to be performed in the frequency domain, appropriately zero padded to ensure that the correlation performed is not a cyclic correlation. The phase of the Fourier transform, as previously stated, contains the translation component. Taking advantage of this fact, the phases between image 1 and the candidates for the correct solution of image 2 are subtracted, leaving only the phase difference and, after performing the inverse Fourier transform, the translational shift between the two images. This is formally described as follows:

$$P_1^{2a}[d,n] = \mathrm{IFFT}\left\{\angle\mathrm{FFT}\left\{\mathrm{Im}_1[d,n]\right\} - \angle\mathrm{FFT}\left\{\mathrm{Im}_2^a[d,n]\right\}\right\} \tag{19}$$

$$P_1^{2b}[d,n] = \mathrm{IFFT}\left\{\angle\mathrm{FFT}\left\{\mathrm{Im}_1[d,n]\right\} - \angle\mathrm{FFT}\left\{\mathrm{Im}_2^b[d,n]\right\}\right\}. \tag{20}$$

With phase correlation functions now computed, the solution selection process may be started. The solution corresponding to the correct rotation will be more impulsive in nature compared to the other solution due to the nature of the correlation. This is performed by determining the summation over each dimension of the ratios of the gain corresponding to the highest peak encountered in the collapsed phase correlation function versus the variance of the phase correlation function, i.e.,

$$\mathrm{SPGRa} = \sum_{\forall d} \frac{\underset{\text{Over } n}{\max}\left\{\left(P_1^{2a}[d,n]\right)^2\right\}}{\underset{\text{Over } n}{\mathrm{var}}\left\{\left(P_1^{2a}[d,n]\right)^2\right\}} \tag{21}$$

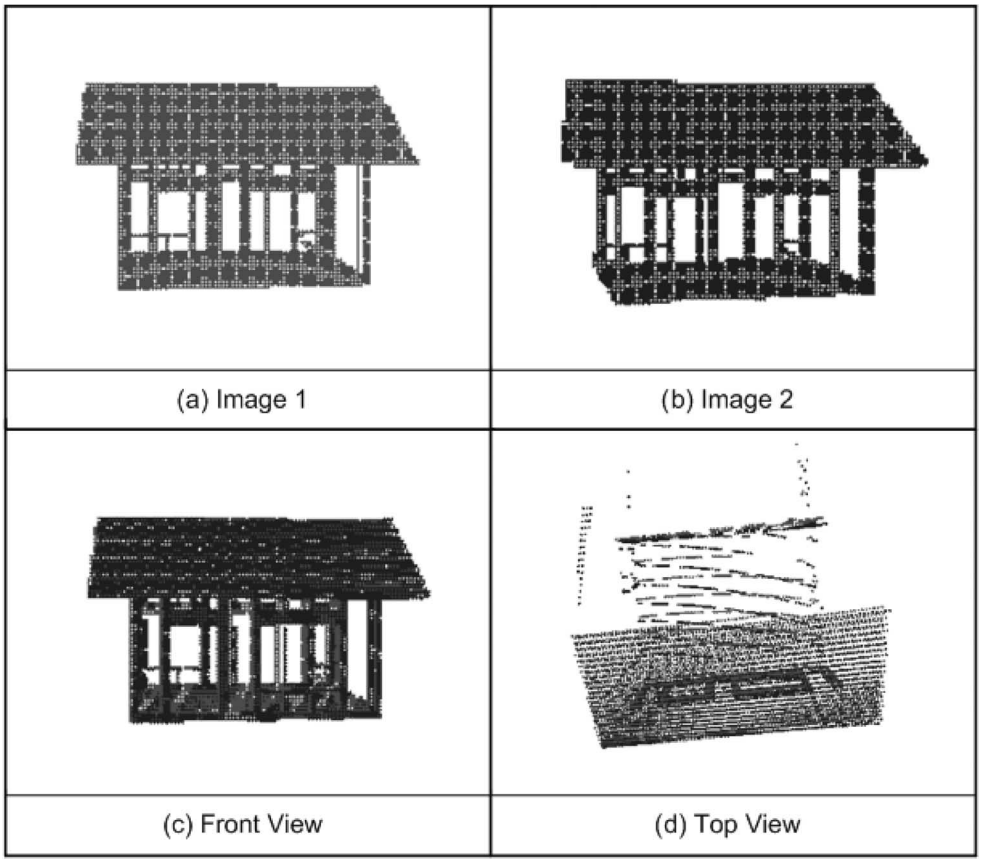

Fig. 5. Point cloud representation of (a) the first image to be registered, (b) the second image to be registered, (c) the front view of the registration results, and (d) the top view of the registration results for a simulated house frame.

$$\text{SPGRb} = \sum_{\forall d} \frac{\max\limits_{\text{Over } n} \left\{ \left(P_1^{2b}[d,n] \right)^2 \right\}}{\text{var}\limits_{\text{Over } n} \left\{ \left(P_1^{2b}[d,n] \right)^2 \right\}}. \tag{22}$$

The above function was chosen as a good measure of the impulsiveness of a data set since it provides a direct measure of the peak energy compared to the averaged energy of the data. The proper rotational solution corresponds to which of SPGRa or SPGRb is higher. If SPGRa is higher, the rotational solution is $R(\vec{R}_{\text{Axis}}, \text{Angle})$; otherwise, $R(\vec{R}_{\text{Axis}}, \text{Angle} - \pi)$ is the solution.

D. Determining the Translation

The translational parameters correspond to the location of the peaks in each dimension used in the previous section to choose the correct solution. If SPGRa was the correct solution, then the translational parameters correspond to the location in normal space of the peaks found in $P_1^{2a}[d,n]$, and, conversely, if SPGRb was the correct solution, then the translational parameters correspond to the location in normal space of the peaks found in $P_1^{2b}[d,n]$.

V. Experimental Results

Data acquisition was accomplished through the use of a Jupiter laser range finder [13] mounted on a seven-degree-of-freedom robotic arm. This ensures that the registration of the range images is known for the purpose of evaluation, and also guarantees precision and reliability. A more detailed description of the range acquisition system that is used for this experimentation can be found in [14].

The calculation of error is separately performed for rotation and translation, as these are estimated in the algorithm in different steps, and, as a result, errors in the estimation of rotation affect the estimation in translation. The rotation error $\varepsilon_{\text{Rotation}}$ is calculated as the Euclidean distance to the unitary matrix from the actual rotation matrix R_{Actual} derotated by the calculated rotation matrix $R_{\text{Calculated}}$, i.e.,

$$\varepsilon_{\text{Rotation}} = \left\| I - R_{\text{Actual}} R_{\text{Calculated}}^T \right\|. \tag{23}$$

The translation error $\varepsilon_{\text{Translation}}$ is the Euclidean distance to the actual translation vector T_{Actual} from the calculated translation vector $T_{\text{Calculated}}$, i.e.,

$$\varepsilon_{\text{Translation}} = \| T_{\text{Actual}} - T_{\text{Calculated}} \|. \tag{24}$$

To evaluate the performance of the frequency domain algorithm, the comparison of the registration estimates is made upon ten different views of a mock house frame, as shown in Fig. 3, acquired with the laser range scanning system, as well as through a simulation of a similar environment. The simulated views and the real-world acquired views were chosen

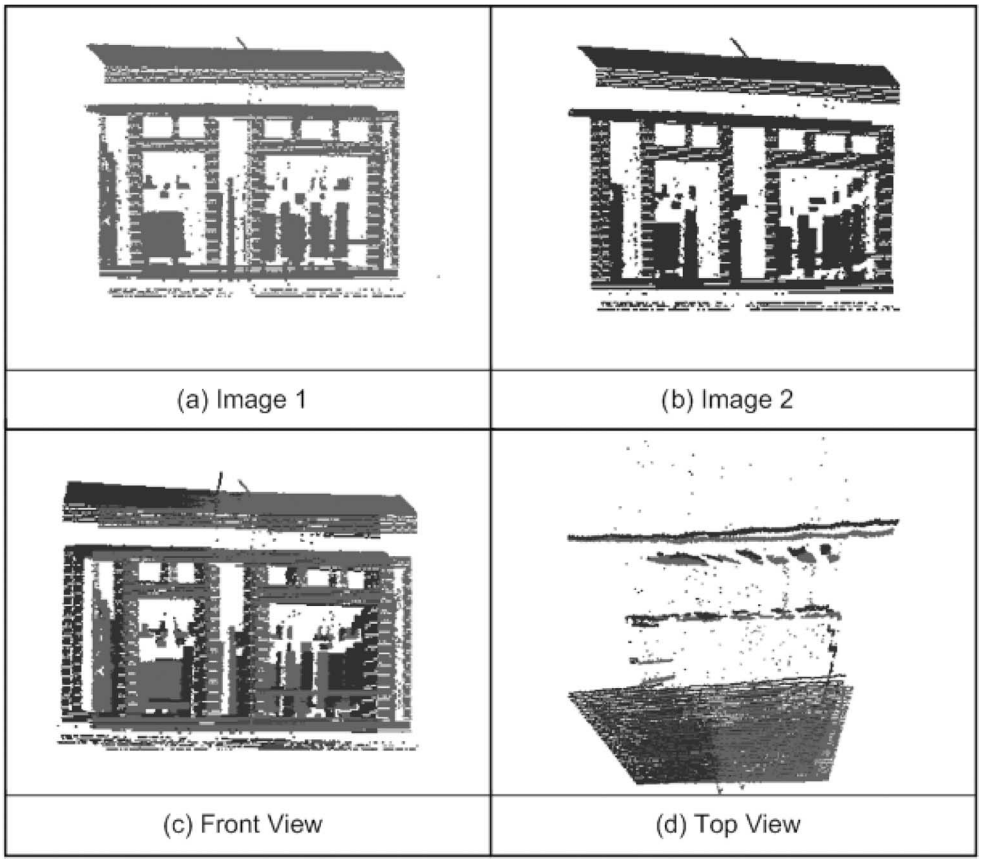

Fig. 6. Point cloud representation of (a) the first image to be registered, (b) the second image to be registered, (c) the front view of the registration results, and (d) the top view of the registration results for a real model of a house frame.

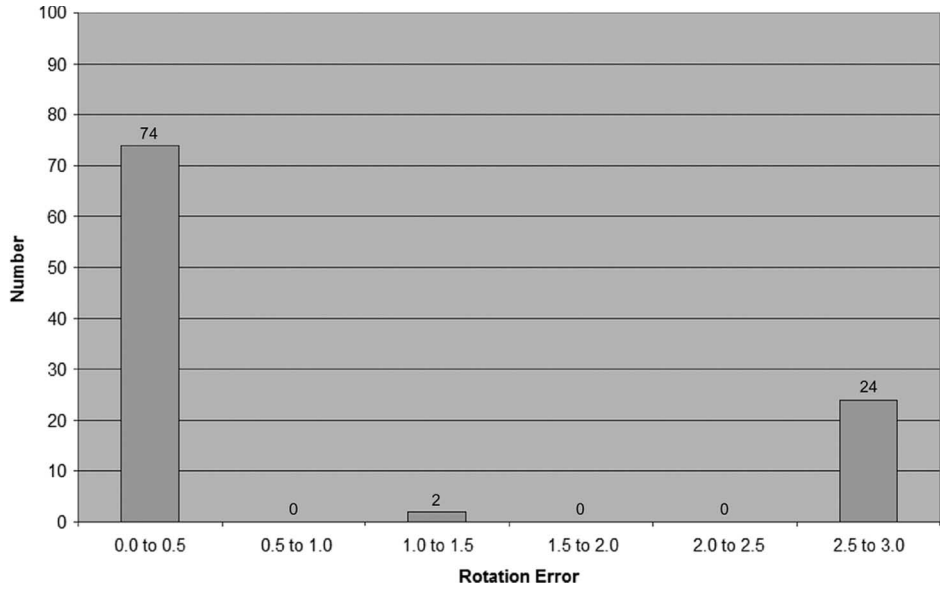

Fig. 7. Registration error on the rotation estimation for 100 pairwise registration attempts on simulated range images of the house frame.

to be approximately the same. The rotations were equivalent and spaced 6° apart, whereas the translations varied due to various constraints, as shown in Fig. 4. The acquired range images were then registered pairwise, ending up with 100 registration attempts to be analyzed. Fig. 5(c) and (d) illustrates an experimental result that is obtained by applying the described frequency domain registration algorithm upon 3-D data point clouds shown in Fig. 5(a) and (b) that represent a model of a house frame that is generated using a laser range finder simulator operating on a virtual representation of the object. As can be observed with this superposition of the point clouds (respectively in dark and light shades) after registration, under low noise conditions, the accuracy of registration estimates is high, with very little visible error in the merge.

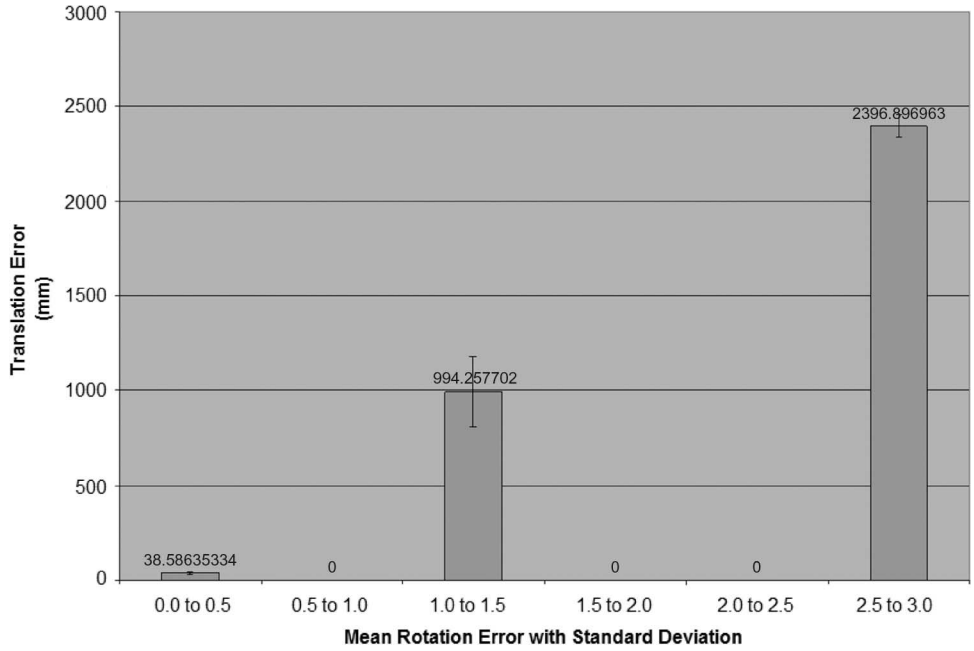

Fig. 8. Translation error versus rotation error for 100 pairwise registration attempts on simulated range images of the house frame.

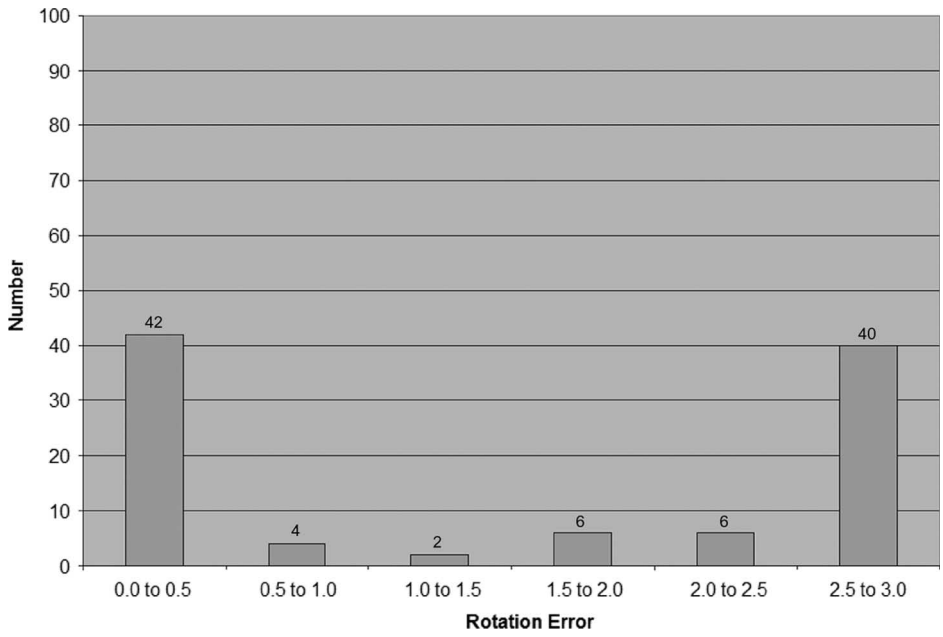

Fig. 9. Registration error on the rotation estimation for 100 pairwise registration attempts on real range images of the actual house frame.

To evaluate the robustness of the approach to noise and incomplete data sets, the experimentation was reproduced using a real range image acquisition system on an actual mock house frame. Fig. 6(c) and (d) illustrates a similar superposition of two point clouds, shown in Fig. 6(a) and (b), collected on this real setup. The viewing angles were chosen to coincide with those shown in Fig. 5. This evaluation demonstrates that the registration estimation is also of high quality when applied on data that are collected under realistic operational conditions. The main perceptible errors reside in a small rotation error that is visible in the top and front views, and a small translation error that is visible in the front view. These are mainly due to the background plane that appears in the real scene but was absent in the simulated case, introducing some extra symmetry in the real data set. Also, the real range sensor tends to produce outliers in the data sets that influence the estimation of registration parameters. As expected, the results from the registration of the simulated data set are better than those of the real-world data set due to the reasons mentioned above. However, the proposed framework demonstrates enough stability to operate in typical 3-D sensing environments.

Fig. 7 shows the division of results according to the rotation error for the simulated range data set. The results originating from 100 pairwise registration attempts have been classified into six bins associated with different error levels, computed with (23). There are three primary sources of error when

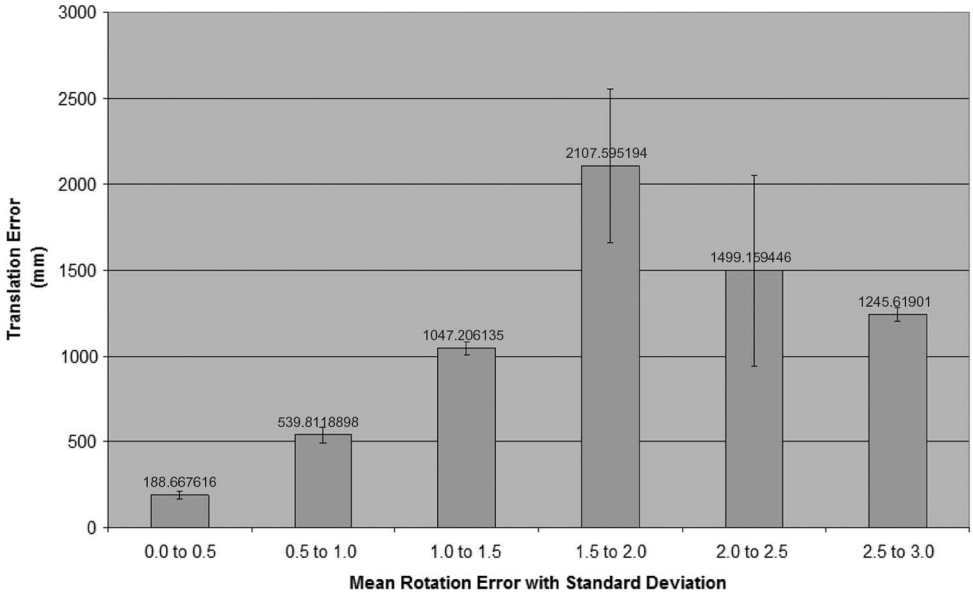

Fig. 10. Translation error versus rotation error for 100 pairwise registration attempts on real range images of the actual house frame.

TABLE I
PERFORMANCE COMPARISON WITH THE ICP APPROACH

	Average Number of Points	Average Time for ICP (s)	Average Time for FFT (s)
Data set 1	7526.82	247.20	10.30
Dataset 2	3668.00	52.01	8.87
Factor	2.05	4.75	1.16

estimating the rotation using the described frequency domain algorithm. The first source is an error on the estimation of the axis of rotation, the second source is an error on the estimation of the angle of rotation, and the third error occurs if the wrong solution is chosen. When there is a little error in the estimation of the rotation axis and angle, and when the correct solution is chosen, the registration error falls into the first error bin, identified 0.0–0.5. When there is a large error in the rotation axis and angle, the registration error falls into the second through fifth error bins, with values from 0.5 to 2.5. When there is a small error in the rotation axis and angle, but the wrong solution is selected, the registration error falls into the sixth bin, marked 2.5–3.0, as a solution on the opposite side of the unit sphere yields the maximum error. The errors on rotation estimation are unitless [see (23)]. These results demonstrate that the estimation of the axis and angle of rotation is correct in most of the circumstances for the simulated data sets (98%), with errors mainly occurring during the solution selection phase of the algorithm. The translation error, as would be expected, increases with an increase in the rotation error, as shown in Fig. 8, over the same set of 100 pairwise registration attempts on the simulated house frame.

Fig. 9 shows the division of results according to the rotation error for the real range data set. The results have been classified into six bins, as previously described. Due to the presence of a background, outliers, and other sources of noise in the actual laser range acquisition, there are more results in the middle four bins compared to the simulated results; however, the majority of the results (82%) lie in the first and sixth bins, demonstrating the robustness of the proposed registration scheme under realistic operational conditions. The translation error also increases with respect to the rotation error, as shown in Fig. 10, although the sixth bin is smaller when compared to the fourth and fifth bins, which is apparently due to the high level of symmetry present in the scans leading to a lower than expected translation estimate.

During experimentation, the proposed frequency domain registration algorithm was also compared against the ICP algorithm. Due to the relatively poor performance of the basic "vanilla" ICP algorithm on these data sets (as most cases converged to local minima due to the symmetries), the detailed registration results are omitted. As there exist improvements to the ICP algorithm to deal with these local minima issues, this analysis rather focuses on execution time, as this is a common factor of comparison among the ICP-based algorithms. A dual-processor Pentium III 600-MHz Windows 2000 PC was used to provide the results. As reported in Table I, it is observed that when the number of points in the range image is doubled, the average completion time for ICP more than doubles, whereas the completion time for the proposed algorithm remains

relatively constant for a particular parameter set. The proposed approach also appears as being less sensitive to the data set size while performing much faster than the classical ICP.

VI. Conclusion

This paper demonstrates that the frequency domain registration is practical and reliable, even when applied on large data sets due to its scalability. The proposed registration technique extends previous work and provides strategies to achieve efficiency gains, in particular, those pertaining to the determination of the axis and angle of rotation, without deteriorating the registration parameters' accuracy. In addition, the described algorithm is able to start the registration estimation process from a raw cloud of 3-D points without requiring any initial estimate and does not tend to converge toward local minima solutions as is observed when using the ICP class of techniques. A set of experimental results demonstrates the robustness, the scalability, and the computational stability of the new frequency domain registration algorithm when applied to simulated and real data sets.

References

[1] P. Curtis and P. Payeur, "A frequency-domain approach to registration estimation in 3-D space," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, Sorrento, Italy, Apr. 24–27, 2006, pp. 293–298.

[2] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.

[3] G. C. Sharp, S. W. Lee, and D. K. Wehe, "ICP registration using invariant features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 1, pp. 90–102, Jan. 2002.

[4] M. A. Rodrigues and Y. Liu, "Registering two overlapping range images using a relative registration error histogram," in *Proc. IEEE Int. Conf. Image Process.*, Rochester, NY, Jun. 24–28, 2002, vol. 3, pp. 841–844.

[5] R. Benjemaa and F. Schmitt, "Fast global registration of 3D sampled surfaces using a multi-Z-buffer technique," in *Proc. IEEE Int. Conf. Recent Adv. 3-D Digital Imag. Model.*, Ottawa, ON, Canada, May 12–15, 1997, pp. 113–120.

[6] E. Trucco and A. Verri, *Introductory Techniques for 2-D Computer Vision*. Toronto, ON, Canada: Prentice-Hall, 1998.

[7] C.-S. Chen, Y.-P. Hung, and J.-B. Cheng, "A fast automatic method for registration of partially-overlapping range images," in *Proc. IEEE 6th Int. Conf. Comput. Vis.*, Bombay, India, Jan. 4–7, 1998, pp. 242–248.

[8] E. Gagnon, J.-F. Rivest, M. Greenspan, and N. Burtnyk, "A computer-assisted range image registration system for nuclear waste cleanup," *IEEE Trans. Instrum. Meas.*, vol. 48, no. 3, pp. 758–762, Jun. 1999.

[9] G. Godin, J.-A. Beraldin, J. Taylor, L. Cournoyer, M. Rioux, S. El-Hakim, R. Baribeau, F. Blais, P. Boulanger, J. Domey, and M. Picard, "Active optical 3D imaging for heritage applications," *IEEE Comput. Graph. Appl.*, vol. 22, no. 5, pp. 24–35, Sep./Oct. 2002.

[10] D. Laurent, M. El Mustapha, P. Claude, and V. Pascal, "A mobile robot localization based on a multisensor cooperation approach," in *Proc. IEEE IECON 22nd Int. Conf.*, Taipei, Taiwan, R.O.C., Aug. 5–10, 1996, vol. 1, pp. 155–160.

[11] L. Lucchese, G. Doretto, and G. M. Cortelazzo, "A frequency domain technique for range data registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 11, pp. 1468–1484, Nov. 2002.

[12] L. Lucchese and G. M. Cortelazzo, "A noise-robust frequency domain technique for estimating planar roto-translations," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1769–1786, Jun. 2000.

[13] Servo-Robot Inc., *Jupiter 3-D Laser Vision Camera Installation and Operation Manual*, 1996, St-Bruno, QC, Canada.

[14] P. Curtis, C. S. Yang, and P. Payeur, "An integrated robotic multi-modal range sensing system," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, Ottawa, ON, Canada, May 17–19, 2005, pp. 1991–1996.

Phillip Curtis (S'03) received the M.A.Sc. degree in electrical engineering in 2005 from the University of Ottawa, Ottawa, ON, Canada, where he is currently working toward the Ph.D. degree.

His research interests include registration algorithms, modeling, and machine vision.

Pierre Payeur (S'90–M'98) received the Ph.D. degree in electrical engineering from the Université Laval, Quebec City, QC, Canada, in 1999.

In 1998, he joined the University of Ottawa, Ottawa, ON, Canada, where he is currently an Associate Professor with the School of Information Technology and Engineering. He is a Founding Member of the Vision, Imaging, Video, and Autonomous Systems Research Laboratory and a member of the Sensing and Modeling Research Laboratory. His current research interests include 3-D imaging and modeling, range data processing, robot guidance, teleoperation, motion capture, and integration of computer vision in autonomous systems.

Dr. Payeur is a member of the IEEE Robotics and Automation Society, the IEEE Instrumentation and Measurement Society, and the Ordre des Ingénieurs du Québec.