

Predetermination of ICP Registration Errors And Its Application to View Planning

Kok-Lim Low
National University of Singapore
lowkl@comp.nus.edu.sg

Anselmo Lastra
University of North Carolina at Chapel Hill
lastra@cs.unc.edu

Abstract

We present an analytical method to estimate the absolute registration error bounds if two surfaces were to be aligned using the ICP (Iterative Closest Point) algorithm. The estimation takes into account (1) the amount of overlap between the surfaces, (2) the noise in the surface points' positions, and (3) the geometric constraint on the 3D rigid-body transformation between the two surfaces. Given a required confidence level, the method of estimation enables us to predetermine the registration accuracy of two overlapping surfaces. This is very useful for automated range acquisition planning where it is important to ensure that the next scan to be acquired can be registered to the previous scans within the desired accuracy. We demonstrate a view-planning system that incorporates our estimation method in the selection of good candidate views for the range acquisition of indoor environments.

1. Introduction

Since its introduction by Besl and McKay [Besl92], and by Chen and Medioni [Chen92], the *Iterative Closest Point* (ICP) algorithm and its many variants have become the most widely-used approaches for aligning three-dimensional surfaces, especially for surfaces created from range images. In the ICP algorithm described in [Besl92], each point in one data set is paired with the closest point in the other data set to form correspondence pairs. Then, using a *point-to-point error metric*, the sum of the squared distance between points in each correspondence pair is minimized by rigidly transforming one of the data sets. The process is iterated until the error becomes smaller than a threshold or it stops changing.

In autonomous and semi-automated range acquisition, each new scanning pose is computed by a view planning subsystem based on a partial model of the environment or object, and taking into account the required reconstruction quality and the acquisition constraints [Low06b]. When the scanner is being positioned at the planned pose, errors in the positioning and pose measurement often cause the actual pose to be different from the planned pose. As a result of this

unknown pose error, the range scan acquired from the new scanner pose will be misaligned with the current partial scene model. In order to correctly merge the new scan into the partial scene model, the former must first be registered or aligned with the latter. For autonomous range acquisition, where the scanner is mounted on a mobile robot, the registration is able to provide a much more accurate localization of the mobile robot, and greatly reduces the effect of accumulated drift in the robot's location.

However, registration of two surfaces using the ICP algorithm is not guaranteed to be successful. Registration failures can occur for several reasons, for example,

- (1) when there is insufficient overlap between the two surfaces,
- (2) when the range measurement errors are too large,
- (3) when there is insufficient geometric constraint on the 3D rigid-body transformation between the two surfaces (for example, when a plane is being registered to another plane, the former can "slide" and "spin" on the latter without being constrained in those motions), and
- (4) when the initial relative pose between the two surfaces is too large.

A reliable view planner must consider the above factors to ensure that the new range scan acquired from the planned view can be successfully registered with the current scene model to within a certain error tolerance.

In this paper, we present a novel *registration accuracy metric* that allows the view planner to analyze each candidate view for Factors (1), (2) and (3), so that the next scan will have a *high probability* of being registered successfully to within a specified error bound.

Factor (4) is not considered here. It should not pose a problem for us since in autonomous acquisition we can obtain a rough estimate of the relative pose from the positioning system. In other applications using ICP, the rough relative pose estimate may be obtained from user's inputs.

Our registration accuracy metric consists of two *registration accuracy conditions*, which are built upon the registration constraint analysis in [Simon96]. Simon presented a means to measure the relative amount of constraint on the 3D rigid-body transformation exerted

by the shape of a surface. More specifically, his method is able to compute, for a set of surface points, a value that represents the amount of constraint on the 3D rigid-body transformation when the surface is being aligned with itself. Being a relative measurement, this value is only useful for comparing with the values of other point sets, so as to determine which point set has the best constraint on the transformation. It does not indicate the absolute accuracy that can be achieved with each point set during registration. Therefore it is not useful for view planning where one wants to determine whether a view can produce scan that can be accurately registered, and not to find the view that can allow the best registration accuracy. Our registration accuracy conditions extend Simon’s constraint analysis to estimate absolute registration accuracies.

The paper is organized as follows. Section 2 reviews some related work. Section 3 reviews Simon’s constraint analysis in detail and presents the derivation of our registration accuracy conditions. Section 4 shows how the registration accuracy conditions are used in a view planning system for range acquisition of indoor environments. Section 5 concludes the paper and discusses some potential extensions of this work.

2. Related Work

The registration constraint analysis method of Simon has been used for the selection of correspondence points on range data to improve geometric constraint during ICP registration [Gelfand03].

Of the existing view planning algorithms for range acquisition, only a few consider the registration constraint when computing the next view. However, the only criterion they consider is the amount of overlap between surfaces, and the registration constraint is satisfied when the amount of overlap is beyond a certain arbitrary threshold [Pito96, Sanchiz99, Scott01, González-Baños99].

3. Registration Accuracy Conditions

We present Simon’s registration constraint analysis first and then show partial derivation of our two registration accuracy conditions—one for translational error, and the second for rotational error. The complete derivation of the conditions can be found in [Low06a].

3.1. Simon’s Constraint Analysis

The basic idea of the constraint analysis is as follows. Let P be a set of points on a surface. Suppose a small 3D rigid-body transformation is applied to P to produce the point set Q , in which some of the points may no longer be on the original surface. Let E_Q be the sum of the squared distance between each point in Q and the original surface. The constraint analysis tries to quantify the sensitivity of E_Q to each component of the 3D rigid-body transformation. The higher the sensitivity is with respect to a certain component, the stronger the constraint on the motion corresponding to

the component.

The following describes how E_Q is computed. Since, given an arbitrary surface, there is no closed-form analytical expression for the distance between a point \mathbf{x} , and the surface, a first-order approximation of the true point-to-surface distance is used:

$$D(\mathbf{x}) = \frac{F(\mathbf{x})}{\|\nabla F(\mathbf{x})\|} \quad (1)$$

where $F(\mathbf{x}) = 0$ is the implicit equation of the surface, $\|\nabla F(\mathbf{x})\|$ is the magnitude of the gradient to the surface, \mathbf{x} is a point which may or may not lie on the surface and $D(\mathbf{x})$ is the approximate distance.

Let \mathbf{x}_s be a point that lies on the surface, i.e. $D(\mathbf{x}_s) = 0$. This point can be perturbed by applying a differential transformation \mathbf{T} to it. \mathbf{T} can be represented by a homogeneous transformation which is a function of the six parameters $(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$, where $(\omega_x, \omega_y, \omega_z)$ are rotations about the x, y and z axes, respectively, and (t_x, t_y, t_z) are the translations along the newly rotated x, y and z axes. The rate of change of D with respect to an arbitrary transformation \mathbf{T} of the point \mathbf{x}_s is given by

$$\mathbf{V}(\mathbf{x}_s) = \frac{\partial}{\partial \mathbf{t}} D(\mathbf{T}(\mathbf{x}_s)) = \begin{bmatrix} \mathbf{n}_{\mathbf{x}_s} \\ \mathbf{x}_s \times \mathbf{n}_{\mathbf{x}_s} \end{bmatrix} \quad (2)$$

where $\mathbf{t} = [t_x, t_y, t_z, \omega_x, \omega_y, \omega_z]^T$ and $\mathbf{n}_{\mathbf{x}_s}$ is the unit normal to the surface evaluated at the point \mathbf{x}_s . Equation (2) can be written as

$$D(\mathbf{T}(\mathbf{x}_s)) = \mathbf{V}^T(\mathbf{x}_s) \mathbf{t} = \begin{bmatrix} \mathbf{n}_{\mathbf{x}_s} \\ \mathbf{x}_s \times \mathbf{n}_{\mathbf{x}_s} \end{bmatrix}^T \mathbf{t}. \quad (3)$$

By squaring Equation (3), we get

$$D^2(\mathbf{T}(\mathbf{x}_s)) = d\mathbf{t}^T \mathbf{V}(\mathbf{x}_s) \mathbf{V}^T(\mathbf{x}_s) d\mathbf{t} = d\mathbf{t}^T \mathbf{M}(\mathbf{x}_s) d\mathbf{t} \quad (4)$$

where $\mathbf{M}(\mathbf{x}_s) = \mathbf{V}(\mathbf{x}_s) \mathbf{V}^T(\mathbf{x}_s)$ is a symmetric, positive semi-definite 6×6 matrix. By summing the quantity in Equation (4) over a set P of discrete surface points, we get

$$\begin{aligned} E_p(\mathbf{T}(\mathbf{x}_s)) &= \sum_{\mathbf{x}_s \in P} D^2(\mathbf{T}(\mathbf{x}_s)) \\ &= d\mathbf{t}^T \left(\sum_{\mathbf{x}_s \in P} \mathbf{M}(\mathbf{x}_s) \right) d\mathbf{t} = d\mathbf{t}^T \mathbf{\Psi} d\mathbf{t}. \end{aligned} \quad (5)$$

The E_Q that was mentioned earlier is actually $E_Q = E_p(\mathbf{T}(\mathbf{x}_s))$, where $Q = \{\mathbf{T}(\mathbf{x}_s) : \mathbf{x}_s \in P\}$. The matrix $\mathbf{\Psi}$ is a scatter matrix that contains information about the distribution of the original $\mathbf{V}(\mathbf{x}_s)$ over all points in P . Now, by performing principal component analysis to $\mathbf{\Psi}$, we can factorize it into the form

$$\Psi = \mathbf{Q} \Lambda \mathbf{Q}^T = \mathbf{Q} \cdot \begin{bmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & & \\ & & \lambda_3 & & & \\ & & & \lambda_4 & & \\ & & & & \lambda_5 & \\ & & & & & \lambda_6 \end{bmatrix} \cdot \mathbf{Q}^T \quad (6)$$

where $\mathbf{Q} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{q}_3 \ \mathbf{q}_4 \ \mathbf{q}_5 \ \mathbf{q}_6]$. $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 \geq \lambda_5 \geq \lambda_6$ are the eigenvalues of Ψ , and \mathbf{q}_i are the corresponding unit eigenvectors. Each eigenvector \mathbf{q}_i represents a differential transformation where the first three elements are the translation components and the last three elements are the rotation components.

Each eigenvalue λ_i is proportional to the rate of change of the error E_p induced by a transformation in the direction specified by \mathbf{q}_i . This implies that \mathbf{q}_1 , which corresponds to the largest eigenvalue, represents the transformation of maximum constraint. When an eigenvalue λ_i is close to or equal to zero, transforming the points in the direction specified by \mathbf{q}_i will not change E_p . This means that the set of points has no constraint on the transformation in the direction specified by \mathbf{q}_i .

Given two different sets of discrete points on a surface, Simon compares the quantity $\lambda_6/\sqrt{\lambda_1}$ (called the *noise amplification index*) of one point set to that of the other point set. The point set with the higher noise amplification index is considered to have a better "overall" constraint on the transformation.

One problem with the result in Equation (6) is that the rotation components are dependent on the scale of the surface being analyzed. This is the consequence of having \mathbf{x}_s in the cross product in Equation (3). Simon addressed this problem by shifting the centroid of the point set to the origin and scaling all points so that the average distance from the points to the origin is 1.

Equation (6) and the noise amplification index cannot be used to estimate absolute registration errors, especially when the points in the point set have been perturbed by noise. In the next subsection, we present our extension to Simon's constraint analysis to estimate absolute bounds on the registration errors.

3.2. Estimating Absolute Registration Errors

In the following derivations, we consider translation and rotation separately. This makes sense because translation and rotation are parameterized by different entity types, i.e. translation is parameterized by distance and rotation by angle. By considering them separately, as we will see, the arbitrary scale of the object is no longer a problem for the rotation analysis.

3.2.1. Translational Alignment Error

Following the derivation in Section 3.1, the translation component of $V(\mathbf{x}_s)$ is

$$V_\tau(\mathbf{x}_s) = \frac{\partial}{\partial \tau} D(\mathbf{T}_\tau(\mathbf{x}_s)) = \mathbf{n}_{\mathbf{x}_s} \quad (7)$$

where $\tau = [t_x, t_y, t_z]^T$. Then

$$D(\mathbf{T}_\tau(\mathbf{x}_s)) = V_\tau^T(\mathbf{x}_s) d\tau. \quad (8)$$

By squaring Equation (8), we get

$$\begin{aligned} D^2(\mathbf{T}_\tau(\mathbf{x}_s)) &= d\tau^T V_\tau(\mathbf{x}_s) V_\tau^T(\mathbf{x}_s) d\tau \\ &= d\tau^T \mathbf{M}_\tau(\mathbf{x}_s) d\tau \end{aligned} \quad (9)$$

where $\mathbf{M}_\tau(\mathbf{x}_s) = V_\tau(\mathbf{x}_s) V_\tau^T(\mathbf{x}_s)$ is a symmetric, positive semi-definite 3×3 matrix. By summing the quantity in Equation (9) over the set P of discrete surface points, we get

$$\begin{aligned} E_p(\mathbf{T}_\tau(\mathbf{x}_s)) &= \sum_{\mathbf{x}_s \in P} D^2(\mathbf{T}_\tau(\mathbf{x}_s)) \\ &= d\tau^T \left(\sum_{\mathbf{x}_s \in P} \mathbf{M}_\tau(\mathbf{x}_s) \right) d\tau = d\tau^T \Psi_\tau d\tau. \end{aligned} \quad (10)$$

Using principal component analysis, Ψ_τ can be factorized into

$$\Psi_\tau = \mathbf{Q}_\tau \Lambda_\tau \mathbf{Q}_\tau^T = \mathbf{Q}_\tau \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \mathbf{Q}_\tau^T \quad (11)$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of Ψ_τ , and the columns of \mathbf{Q}_τ are the corresponding unit eigenvectors. Since the sum of the eigenvalues is equal to the trace of the original matrix,

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= \text{tr}(\Psi_\tau) \\ &= \text{tr} \left(\sum_{\mathbf{x}_s \in P} \mathbf{M}_\tau(\mathbf{x}_s) \right) = \text{tr} \left(\sum_{\mathbf{x}_s \in P} V_\tau(\mathbf{x}_s) V_\tau^T(\mathbf{x}_s) \right) \\ &= \sum_{\mathbf{x}_s \in P} \|V_\tau(\mathbf{x}_s)\|^2 = \sum_{\mathbf{x}_s \in P} \|\mathbf{n}_{\mathbf{x}_s}\|^2 = \sum_{\mathbf{x}_s \in P} 1 = |P| = M \end{aligned} \quad (12)$$

where M is the number of points in P .

For the alignment of the surface to be successful, λ_i must be greater than 0 for $i = 1, 2$, and 3. Ideally, we wish to select the set of points for P such that $\lambda_1 = \lambda_2 = \lambda_3 = M/3$, which means that the surface alignment is constrained equally in all three orthogonal translation directions. The minimum requirement for alignment is that $M = 3$, and that the three surface normals must span the 3D space. In practice, because of errors in the range measurements, it is necessary to have $M \geq M_{\min} > 3$, such that the translational alignment can be performed to a certain desired accuracy, assuming the surface is already correctly oriented.

The translation resulting from the alignment can be decomposed into three orthogonal directions. In the following, we investigate the relationship between M_{\min} and the translational alignment accuracy

by looking at the alignment errors in the three orthogonal directions. Without loss of generality, we choose the x , y , and z directions.

Let $[x_i, y_i, z_i]^T$ be the 3D coordinates of the i th true surface point, where $i = 1, 2, \dots, M$. Suppose there are two sets of measurements of the surface points, producing the point set A with coordinates $[x_i + u_{Ai}, y_i + v_{Ai}, z_i + w_{Ai}]^T$ and the point set B with coordinates $[x_i + u_{Bi}, y_i + v_{Bi}, z_i + w_{Bi}]^T$, where (u_{Ai}, v_{Ai}, w_{Ai}) , (u_{Bi}, v_{Bi}, w_{Bi}) are measurement errors in the x , y , and z directions, respectively.

Suppose point set B is to be translated so that it is aligned with point set A . For the alignment, the correspondences between points in point sets A and B are known. The alignment uses the least-squares (least-sum-of-squares) error metric, where we want to find the translation vector $\boldsymbol{\tau} = [t_x, t_y, t_z]^T$ to translate point set B to minimize

$$SSE = \sum_{i=1}^M \left\| \begin{bmatrix} x_i + u_{Ai} \\ y_i + v_{Ai} \\ z_i + w_{Ai} \end{bmatrix} - \begin{bmatrix} x_i + u_{Bi} \\ y_i + v_{Bi} \\ z_i + w_{Bi} \end{bmatrix} - \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \right\|^2. \quad (13)$$

Given that the measurement errors u_{Ai} , u_{Bi} , v_{Ai} , v_{Bi} , w_{Ai} , and w_{Bi} are independent and normally distributed with mean 0 and standard deviations not greater than e_{RMS} , we can reach the following condition:

Translational Alignment Error Condition. We can be $(1 - \alpha)^3$ 100% confident that the translational alignment error in *any* direction will not exceed $\sqrt{\varepsilon_x^2 + \varepsilon_y^2 + \varepsilon_z^2} = \sqrt{3} \varepsilon_\tau$ when $\lambda_i \geq M_{\min}/3$, $\forall i = 1, 2, 3$, where $M_{\min} = 6(z_{\alpha/2} e_{\text{RMS}} / \varepsilon_\tau)^2$.

The value $z_{\alpha/2}$ is a value of a random variable Z that has the standard normal distribution such that

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha.$$

Given that a confidence interval has been selected, the translational alignment error condition can be used to estimate the translational error bound. For example, suppose we are given point sets A and B , which are two different sets of measurements of the same set of true surface points, and the RMS measurement error is e_{RMS} . We first use one of the two point sets to compute the values of λ_1 , λ_2 and λ_3 as in Equations (7)–(11). Then, let $M_{\min} = 3\lambda_3$, and solve for the value of ε_τ in the equation $M_{\min} = 6(z_{\alpha/2} e_{\text{RMS}} / \varepsilon_\tau)^2$. According to the error condition, we can be $(1 - \alpha)^3$ 100% confident that, if we register point sets A and B to each

other, the translational alignment error in *any* direction will not exceed $\sqrt{3} \varepsilon_\tau$.

For analyzing candidate views for registration accuracies, e_{RMS} and ε_τ are already specified, and these allow M_{\min} to be calculated. If a view produces $\lambda_3 < M_{\min}/3$, then it is rejected. Later, we describe how to compute, for our view planner, a value for e_{RMS} , and show, given a candidate view, how λ_1 , λ_2 and λ_3 are computed.

3.2.2. Rotational Alignment Error

The rotation component of $V(\mathbf{x}_s)$ can be written as

$$\mathbf{V}_0(\mathbf{x}_s) = \frac{\partial}{\partial \boldsymbol{\theta}} D(\mathbf{T}_0(\mathbf{x}_s)) = \mathbf{x}_s \times \mathbf{n}_{\mathbf{x}_s} \quad (14)$$

where $\boldsymbol{\theta} = [\omega_x, \omega_y, \omega_z]^T$. Then

$$D(\mathbf{T}_0(\mathbf{x}_s)) = \mathbf{V}_0^T(\mathbf{x}_s) d\boldsymbol{\theta}. \quad (15)$$

By squaring Equation (15), we get

$$\begin{aligned} D^2(\mathbf{T}_0(\mathbf{x}_s)) &= d\boldsymbol{\theta}^T \mathbf{V}_0(\mathbf{x}_s) \mathbf{V}_0^T(\mathbf{x}_s) d\boldsymbol{\theta} \\ &= d\boldsymbol{\theta}^T \mathbf{M}_0(\mathbf{x}_s) d\boldsymbol{\theta}. \end{aligned} \quad (16)$$

By summing the quantity in Equation (16) over the set P , we get

$$\begin{aligned} E_P(\mathbf{T}_0(\mathbf{x}_s)) &= \sum_{\mathbf{x}_s \in P} D^2(\mathbf{T}_0(\mathbf{x}_s)) \\ &= d\boldsymbol{\theta}^T \left(\sum_{\mathbf{x}_s \in P} \mathbf{M}_0(\mathbf{x}_s) \right) d\boldsymbol{\theta} = d\boldsymbol{\theta}^T \boldsymbol{\Psi}_0 d\boldsymbol{\theta}. \end{aligned} \quad (17)$$

Using principal component analysis, $\boldsymbol{\Psi}_0$ can be factorized into

$$\boldsymbol{\Psi}_0 = \mathbf{Q}_0 \boldsymbol{\Lambda}_0 \mathbf{Q}_0^T = \mathbf{Q}_0 \begin{bmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma_3 \end{bmatrix} \mathbf{Q}_0^T \quad (18)$$

where $\gamma_1 \geq \gamma_2 \geq \gamma_3$ are the eigenvalues of $\boldsymbol{\Psi}_0$, and the columns of \mathbf{Q}_0 are the corresponding unit eigenvectors. Since the sum of the eigenvalues is equal to the trace of the original matrix,

$$\begin{aligned} S &= \gamma_1 + \gamma_2 + \gamma_3 = \text{tr}(\boldsymbol{\Psi}_0) = \text{tr} \left(\sum_{\mathbf{x}_s \in P} \mathbf{M}_0(\mathbf{x}_s) \right) \\ &= \text{tr} \left(\sum_{\mathbf{x}_s \in P} \mathbf{V}_0(\mathbf{x}_s) \mathbf{V}_0^T(\mathbf{x}_s) \right) = \sum_{\mathbf{x}_s \in P} \|\mathbf{V}_0(\mathbf{x}_s)\|^2 \\ &= \sum_{\mathbf{x}_s \in P} \|\mathbf{x}_s \times \mathbf{n}_{\mathbf{x}_s}\|^2 = \sum_{\mathbf{x}_s \in P} \|\mathbf{x}_s\|^2 \sin^2 \phi_{\mathbf{x}_s} \end{aligned} \quad (19)$$

where $\phi_{\mathbf{x}_s}$ is the angle between the vector \mathbf{x}_s and $\mathbf{n}_{\mathbf{x}_s}$.

For the alignment of the surface to be successful, γ_i must be greater than 0 for $i = 1, 2$, and 3. Ideally, we wish to select the set of points for P such that

$\gamma_1 = \gamma_2 = \gamma_3 = S/3$, which means that the surface alignment is constrained equally in the rotations about all three orthogonal axes. The minimum requirement for alignment is that $M = 3$, where M is the number of points in P , and that the three vector in the set $\{\mathbf{x}_s \times \mathbf{n}_{x_s} \mid \mathbf{x}_s \in P\}$ must span the 3D space. In practice, with errors in the range measurements, it is necessary to have $M \geq M_{\min} > 3$ and $S \geq S_{\min} > 0$, such that the rotational alignment can be performed to a certain desired accuracy, assuming the surface is already correctly translated. In the following, we investigate the conditions necessary to attain a specified angular accuracy in the rotational alignment.

Given the two point sets A and B in Section 3.2.1, and suppose point set B is to be rotated about the origin so that it is aligned with point set A . The alignment uses the least-squares error metric, where we want to find the rotation vector $\boldsymbol{\theta} = [\omega_x, \omega_y, \omega_z]^T$ to rotate point set B to minimize

$$SSE = \sum_{i=1}^M \left\| \begin{bmatrix} x_i + u_{Ai} \\ y_i + v_{Ai} \\ z_i + w_{Ai} \end{bmatrix} - \mathbf{R}(\omega_x, \omega_y, \omega_z) \cdot \begin{bmatrix} x_i + u_{Bi} \\ y_i + v_{Bi} \\ z_i + w_{Bi} \end{bmatrix} \right\|^2 \quad (20)$$

where $\mathbf{R}(\omega_x, \omega_y, \omega_z)$ is the composite rotation matrix for rotations of ω_x , ω_y and ω_z radians about the x -axis, y -axis and z -axis, respectively. Since the rotations will be small, we can approximate the matrix $\mathbf{R}(\omega_x, \omega_y, \omega_z)$ by using the approximations $\sin\theta \approx \theta$ and $\cos\theta \approx 1$ when $\theta \approx 0$. With this approximation, and given that the measurement errors u_{Ai} , u_{Bi} , v_{Ai} , v_{Bi} , w_{Ai} , and w_{Bi} are independent and normally distributed with mean 0 and standard deviations not greater than e_{RMS} , we can reach the following condition:

Rotational Alignment Error Condition. We can be $(1 - \alpha)^3 100\%$ confident that the rotational alignment errors about the x -axis, y -axis and z -axis, respectively, will not exceed $\varepsilon_0 > 0$ radians when $\gamma_i \geq S_{\min}/3$, $\forall i = 1, 2, 3$, where $S_{\min} = 6(z_{\alpha/2} e_{\text{RMS}} / \varepsilon_0)^2$.

Similar to the application of the translational alignment error condition for analyzing candidate views, if a view produces $\gamma_3 < S_{\min}/3$, then it is rejected.

3.2.3. Remarks on the Two Conditions

The three registration accuracy factors listed Section 1 are taken into account in the two alignment error conditions. For Factor (1), if there is a lot of overlap between the two surfaces, and thus a large number of point pairs that can be used in the registration, the values of the eigenvalues λ_i and γ_i will likely be

large. Larger values for λ_i and γ_i allow M_{\min} and S_{\min} to be higher in value, which in turn result in smaller error bounds ε_τ and ε_θ . For Factor (2), the range measurement errors are incorporated into e_{RMS} . For Factor (3), the values of λ_i and γ_i reflect the amount of constraint the points have on the different components of the 3D rigid-body transformation, and each eigenvalue is proportional to the registration accuracy that can be achieved in the corresponding motion.

The derivations of the alignment error conditions assume that the point correspondences between the two point sets are known. In practice, these point correspondences may not be available, such as when these point sets come from range images of the same surface scanned from different scanner positions. Fortunately, in many cases, the ICP algorithm is able to provide a good approximation of the point correspondences. As the ICP algorithm iteratively refines the alignment of the two surfaces, the matching of point pairs becomes more accurate. When the two surfaces are almost correctly aligned, the accuracy of the closest-point matching is limited by the sampling spacing between points in each point set. The errors caused by the sampling spacing are considered quantization errors, and can be incorporated into e_{RMS} when using the two alignment error conditions.

A limitation of the alignment error conditions is that the errors in the surface normals are not taken into consideration. The true surface normals at the measured points are needed to compute the eigenvalues in the alignment error conditions, but they are usually not available, and estimated normals are used instead. For a range image, we can estimate these normals using the sample points by fitting a plane to the points in the neighborhood of each candidate point. The estimation is more accurate if the points are from a relatively smooth region on the surface.

4. Application to View Planning

This section describes how the alignment error conditions are used to test candidate views during view planning. It is important to note that the validity of using the conditions to test candidate views is based on the assumption that the eigenvalues λ_i and γ_i at the planned pose will not be significantly different from those at the unknown actual pose. This is generally not an issue, as the scanner pose errors are very small in comparison to the size of the scene to be acquired.

To apply the two conditions, we have to first choose and determine the values of the several parameters, i.e. (1) the confidence interval $(1 - \alpha)^3 100\%$ and the confidence limit $z_{\alpha/2}$, (2) the points' RMS position error e_{RMS} , (3) the translation error tolerance ε_τ , and (4) the rotation angle error tolerance ε_θ . These allow

M_{\min} and S_{\min} to be calculated. Then, if a candidate view has $\lambda_3 < M_{\min}/3$ or $\gamma_3 < S_{\min}/3$, it is rejected. These form our registration accuracy metric for view planning. It is not straightforward to select values for some of the parameters, such as e_{RMS} and ε_0 . The following subsections describe how we compute the eigenvalues λ_i and γ_i , and select the appropriate values for the parameters.

The registration accuracy metric has been incorporated in our view planning system [Low06a, Low06b]. To compute the next view, the view planner evaluates a set of acquisition constraints (not including the registration accuracy metric) and quality requirements for every possible view. The views are ranked by their scores, and starting from the highest-score view, the registration accuracy metric is evaluated until the first view that satisfies it. That view is chosen as the view for the next scan. Our view planning system uses a volumetric octree to represent the partial scene model, and each octree leaf node that intersects any actual surface of the scene is called a *surface voxel* and it contains a surface normal.

4.1. Computing the Eigenvalues

Given a candidate view, in order to compute the eigenvalues λ_i and γ_i , we need to find out which surface points will be used in the registration should the view be used to make the next range scan. In the view planner, this new range image will be registered with the center points of surface voxels in the octree partial scene model. Each surface voxel's center will be paired with the nearest 3D point on the new range image. Therefore, the maximum number of point pairs that can be used in the registration is the total number of surface voxels in the scene model. However, those surface voxels that will not be overlapped by the new range image will not be able to form point pairs to be used in the registration.

Even though the new scan has not been acquired, we can still estimate which surface voxels would be used in the registration. To do that, for each surface voxel in the current partial scene model, we test whether it satisfies all the scanning constraints (e.g. visibility) with respect to the candidate view. If satisfied, the surface voxel is selected for the computation of λ_i and γ_i . The center points and the surface normals of the selected surface voxels are then used to compute λ_i as stated in Equations (7)–(11), and γ_i as in Equations (14)–(18).

4.2. Selecting the RMS Error e_{RMS}

The main sources of errors that contribute to e_{RMS} are (1) the range measurement errors, (2) the quantization errors caused by the voxelization of the surface, and (3) the error caused by pairing each voxel's center to a new range image point that does not correspond to the same true surface point as the voxel's center.

The range measurement errors are related to the precision or measurement uncertainty of the range sensor. To be conservative about estimating the registration errors, we should use the lowest precision within the working range of the sensor or the precision at the furthest surface to be measured. The precision is usually specified as the RMS (or standard deviation) of the range errors and it is denoted by σ_{range} here.

The voxelization of the surface introduces a maximum error of $\pm 1/2$ voxel width at each voxel's center in each of the x , y and z directions. This quantization error is uniformly distributed between $\pm 1/2$ voxel width, has a mean of zero, and a standard deviation of

$$\sigma_{\text{voxel-quant}} = (\text{voxel width})/\sqrt{12}.$$

As pointed out in Section 3.2.3, when two surfaces are almost correctly aligned, the accuracy of the closest-point matching is limited by the spacing between adjacent points in the range image. This error can be reduced to be approximately equal to the range errors, and therefore their standard deviation,

$$\sigma_{\text{matching}} = \sigma_{\text{range}}.$$

Finally, the RMS error, e_{RMS} , for the alignment error conditions is $e_{\text{RMS}} = \sqrt{\sigma_{\text{range}}^2 + \sigma_{\text{voxel-quant}}^2 + \sigma_{\text{matching}}^2}$.

4.3. Choosing Rotation Error Tolerance ε_0

Sometimes, it is more convenient to specify the rotational alignment error tolerance as a distance instead of as an angle ε_0 . Since in the translational alignment error condition, the translation error tolerance is $\sqrt{3} \varepsilon_\tau$ in any direction, it is natural to want to also limit ε_0 so that the maximum rotation-induced translation error is $\sqrt{3} \varepsilon_\tau$. However, this value of ε_0 can be too small and results in a very large S_{\min} that makes the rotational alignment error condition too stringent. To derive a practical value for ε_0 , we should consider the rotation-induced translations of all the points that would be used in the registration.

Given a candidate view, let V be the set of surface voxels that satisfy all the scanning constraints with respect to the view, and let $v_i \in V$ for $i=1, 2, \dots, |V|$. We wish to limit the rotation-induced translation errors of these voxels to within a user-specified RMS value τ_{RMS} , i.e.

$$\sqrt{\frac{1}{N} \sum_{i=1}^N \tau_i^2} \leq \tau_{\text{RMS}} \Rightarrow \sum_{i=1}^N \tau_i^2 \leq N \tau_{\text{RMS}}^2$$

where $N = |V|$, and τ_i is the rotation-induced translation error of the center of voxel v_i . With some manipulations, we arrive at the angle tolerance

$$\varepsilon_0 = 2 \sin^{-1} \sqrt{\frac{N \tau_{\text{RMS}}^2}{4 \sum_{i=1}^N d_i^2}}$$

where d_i is the distance between the origin and the center of voxel v_i .

4.4. Experimental Results

To demonstrate the usefulness of our registration accuracy metric for automated view planning, we present example results from a simulated acquisition of a synthetic scene and from the acquisition of a real scene.

4.4.1. Simulated Acquisition of Synthetic Scene

Figure 1(a) shows a synthetic indoor scene whose size is about $35 \times 20 \times 10$ cubic feet. A doorway connects the two rooms. Both rooms have flat ceilings that are not shown in the figure. The first scan is made from Room 1, and Figure 1(b) shows the triangle mesh constructed from the range image. Again, the ceiling is not shown. The scan is made with range precision $\sigma_{\text{range}} = 0.5$ inch. Next, for the purpose of view planning computation, the triangle mesh is voxelized to create a volumetric partial scene model, where each surface voxel is 2 inches wide.

To compute the view for the next scan, the initial view evaluation produces a set of candidate views that have the highest scores. The views are then tested for the registration constraint by evaluating our registration accuracy metric. The highest-score view fails the constraint, and 21 other views are tested until one near the doorway is found to satisfy the constraint. These are shown in Figure 1(c).

In fact, the scan acquired from the highest-score view does not have enough shape constraint in the overlapping region with the first scan, and registering them using our ICP algorithm actually fails. The range scan made from the highest-score view is shown in Figure 2(a). The final planned view produces the range scan shown in Figure 2(b), which can be registered to the first scan well within the specified error tolerance of $\varepsilon_{\tau} = 0.5$ inch.

In the above example, the following values are provided for computing the alignment error conditions:

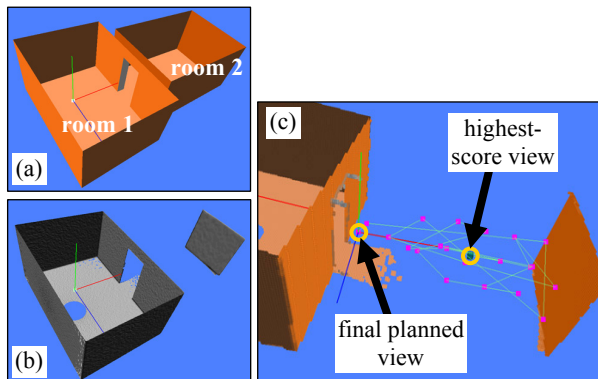


Figure 1: (a) A synthetic scene. (b) The scan made from the first room. (c) The highest-score view and the final planned view.

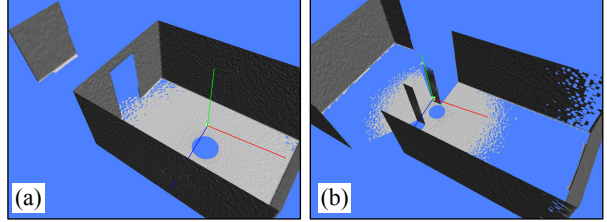


Figure 2: (a) Range scan made from the highest-score view. (b) Range scan made from the final planned view.

- $z_{\alpha/2} \approx 2.712$, corresponding to a 99% confidence interval;
- surface voxel width is 2 inches;
- range precision, $\sigma_{\text{range}} = 0.5$ inch;
- $\varepsilon_{\tau} = 0.5$ inch;
- maximum RMS value of the rotation-induced translation errors, $\tau_{\text{RMS}} = 0.5$ inch.

Since $\sigma_{\text{voxel-quant}} = (\text{voxel width})/\sqrt{12}$ and $\sigma_{\text{matching}} = \sigma_{\text{range}}$, the value of e_{RMS} is computed as

$$e_{\text{RMS}} = \sqrt{\sigma_{\text{range}}^2 + \sigma_{\text{voxel-quant}}^2 + \sigma_{\text{matching}}^2} \approx 0.9129.$$

In Table 1, we can see the actual values of ε_0 , λ_i , γ_i , $M_{\min}/3$ and $S_{\min}/3$ computed at the highest-score view and the final planned view. The eigenvectors are not shown here. All the values have been rounded to 4 significant figures. The highest-score view is rejected because $\lambda_3 < M_{\min}/3$ and $\gamma_3 < S_{\min}/3$, whereas, at the final planned view, $\lambda_3 \geq M_{\min}/3$ and $\gamma_3 \geq S_{\min}/3$, and it is therefore accepted.

	Highest-Score View	Final Planned View
ε_0	0.001647 radian	0.002264 radian
$M_{\min}/3$	49.03	49.03
λ_1	5,521	9,683
λ_2	38.54	2,219
λ_3	13.34	364.3
$S_{\min}/3$	4,520,000	2,392,000
γ_1	6,935,000	37,540,000
γ_2	6,483,000	32,280,000
γ_3	48,920	4,403,000

Table 1

The problem with the highest-score view is that it is too far from the doorway and cannot acquire enough surfaces of the other room to overlap the surfaces acquired in the first scan. The only overlapping surfaces cannot constrain the translational motion along the y and z directions and the rotational motion about the x -axis, which is evident in the small values of λ_2 , λ_3 and γ_3 .

However, at the final planned view, the amount of

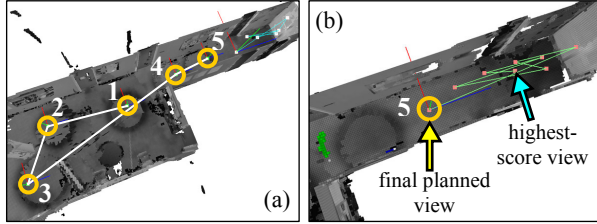


Figure 3: (a) A real scene acquired with five planned views. (b) The fifth view was selected after the highest-score view failed the registration constraint.

overlapping surfaces with the first scan is significantly higher, especially the surfaces that can constrain the translational motion along the y and z directions, and the rotational motion about the x -axis.

4.4.2. Acquisition of Real Scene

The registration accuracy metric has also been tested in the acquisition of real scenes. In this example, the real scene is part of an office floor. Figure 3(a) shows the model reconstructed from scans made from five automatically-planned views. During the planning of the fifth view in the corridor, the initial view evaluation produced a set of candidate views that have high scores. The views are then tested for the registration constraint by evaluating our registration accuracy metric. The eight highest-score views fail the constraint before the final one is selected (see Figure 3(b)). Those rejected views are actually too far away from the previous planned view, and because the corridor is quite narrow, their overlapping surfaces with the scan produced at the previous planned view are small and have very little constraint on the translational direction along the corridor.

5. Conclusion and Future Work

The ability to predetermine registration accuracy between the existing scans and the next scan is essential to a reliable automated view planning system. We have presented the derivation of a registration accuracy metric that can be used for such purpose. The metric consists of two alignment error conditions—one for the translational error, and one for the rotational error. We have also provided details about the selection of the parameter values for the metric and how to evaluate the metric. Finally, we use two examples to demonstrate the effectiveness of the metric in practical automated view planning.

Currently, for each candidate view, the metric is evaluated using all the points in the partial scene model. This can be quite inefficient when the scene model is very high resolution and a great number of candidate views have to be evaluated. A possible im-

provement to this is to use a hierarchical approach to exploit the coherence in the view space and the scene surfaces [Low06b].

Acknowledgements

We thank Young Truong and Ping Bai for their help with the statistical derivations. This work is supported by NSF grant number ACI-0205425.

References

- [Besl92] P. J. Besl and N. D. McKay. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 14(2), pp. 239–256, 1992.
- [Chen92] Yang Chen, and Gerard Medioni. Object Modeling by Registration of Multiple Range Images. *International Journal of Image and Vision Computing*, 10(3), pp. 145–155, 1992.
- [Gelfand03] Natasha Gelfand, Leslie Ikemoto, Szymon Rusinkiewicz, and Marc Levoy. Geometrically Stable Sampling for the ICP Algorithm. *Proceedings of the International Conference on 3-D Digital Imaging and Modeling (3DIM)*, pp. 260–267, 2003.
- [González-Baños99] H. González-Baños, E. Mao, J.-C. Latombe, T. M. Murali and A. Efrat. Planning Robot Motion Strategies for Efficient Model Construction. *Proceedings of International Symposium on Robotics Research*, pp. 345–352, 1999.
- [Low06a] Kok-Lim Low. View Planning for Range Acquisition of Indoor Environments. Ph.D. Dissertation, Department of Computer Science, University of North Carolina at Chapel Hill, North Carolina, United States, 2006.
- [Low06b] Kok-Lim Low and Anselmo Lastra. An Adaptive Hierarchical Next-Best-View Algorithm for 3D Reconstruction of Indoor Scenes. *Proceedings of 14th Pacific Conference on Computer Graphics and Applications (Pacific Graphics 2006)*, October 2006.
- [Low06c] Kok-Lim Low and Anselmo Lastra. Efficient Constraint Evaluation Algorithms for Hierarchical Next-Best-View Planning. *Proceedings of the 3rd Symposium on 3D Data Processing, Visualization and Transmission (3DPVT 2006)*, June 2006.
- [Pito96] R. Pito. A Sensor Based Solution to the Next Best View Problem. *Proceedings of IEEE International Conference on Pattern Recognition*, pp. 941–945, 1996.
- [Sanchiz99] J. M. Sanchiz and R. B. Fisher. A Next-Best-View Algorithm for 3D Scene Recovery with 5 Degrees of Freedom. *Proceedings of British Machine Vision Conference*, 1999.
- [Scott01] W. R. Scott, G. Roth, and J.-F. Rivest. View Planning with a Registration Component. *Proceedings of IEEE International Conference on 3D Digital Imaging and Modeling (3DIM)*, 2001.
- [Simon96] David A. Simon. Fast and Accurate Shape-Based Registration. Ph. D. Dissertation, Carnegie Mellon University, CMU-RI-TR-96-45, 1996.