Autonomous Robots 4, 333–349 (1997)

Globally Consistent Range Scan Alignment for Environment Mapping

F. LU* AND E. MILIOS
Department of Computer Science, York University, North York, Ontario, Canada
eem@cs.yorku.ca

Abstract. A robot exploring an unknown environment may need to build a world model from sensor measurements. In order to integrate all the frames of sensor data, it is essential to align the data properly. An incremental approach has been typically used in the past, in which each local frame of data is aligned to a cumulative global model, and then merged to the model. Because different parts of the model are updated independently while there are errors in the registration, such an approach may result in an inconsistent model.

In this paper, we study the problem of consistent registration of multiple frames of measurements (range scans), together with the related issues of representation and manipulation of spatial uncertainties. Our approach is to maintain all the local frames of data as well as the relative spatial relationships between local frames. These spatial relationships are modeled as random variables and are derived from matching pairwise scans or from odometry. Then we formulate a procedure based on the maximum likelihood criterion to optimally combine all the spatial relations. Consistency is achieved by using all the spatial relations as constraints to solve for the data frame poses simultaneously. Experiments with both simulated and real data will be presented.

Keywords: sensor-based mobile robotics, laser range scanning, mapping, range scan registration, range scan alignment

1. Introduction

1.1. Problem Definition

The general problem we want to solve is to let a mobile robot explore an unknown environment using range sensing and build a map of the environment from sensor data. In this paper, we address the issue of consistent alignment of data frames so that they can be integrated to form a world model. However, the issue of building a high-level model from registered sensor data is beyond the scope of this paper.

A horizontal range scan is a collection of range measurements taken from a single robot position. In previous robot navigation systems, range scans have often been used for robot self-localization in known environments (Cox, 1991). Using range measurements (sonar or laser) for modeling an unknown environment has also been studied in the past (Leonard et al., 1990; Crowley, 1989; Gonzalez et al., 1994). A range scan represents a partial view of the world. By merging many such scans taken at different locations, a more complete description of the world can be obtained. Figure 1 gives an example of a single range scan and a world model consisting of many scans.

The essential issue here is to align the scans properly so that they can be merged. But the difficulty is that odometry information alone is usually inadequate for determining the relative scan poses (because of odometry errors that accumulate). On the other hand, we are unable to use pre-mapped external landmarks to correct pose errors because the environment is unknown. A generally employed approach of building a world model is to incrementally integrate new data to the model. When each frame of sensor data is obtained, it is aligned to a previous frame or to a cumulative global model. Then the new frame of data is integrated into the global model by averaging the data or using a Kalman

*Currently with PCI Enterprises Inc., Richmond Hill, Ontario, Canada. A major part of this work was carried out when the author was at University of Toronto.

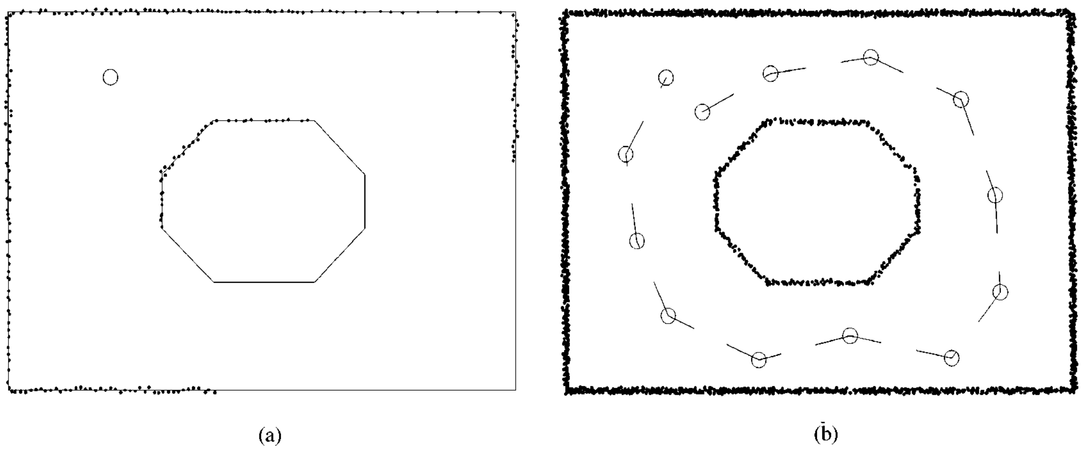


Figure 1. Building world model from range scans: (a) one range scan in a simulated world; (b) model consisting of many scans. The small circles show the poses at which the scans are taken.

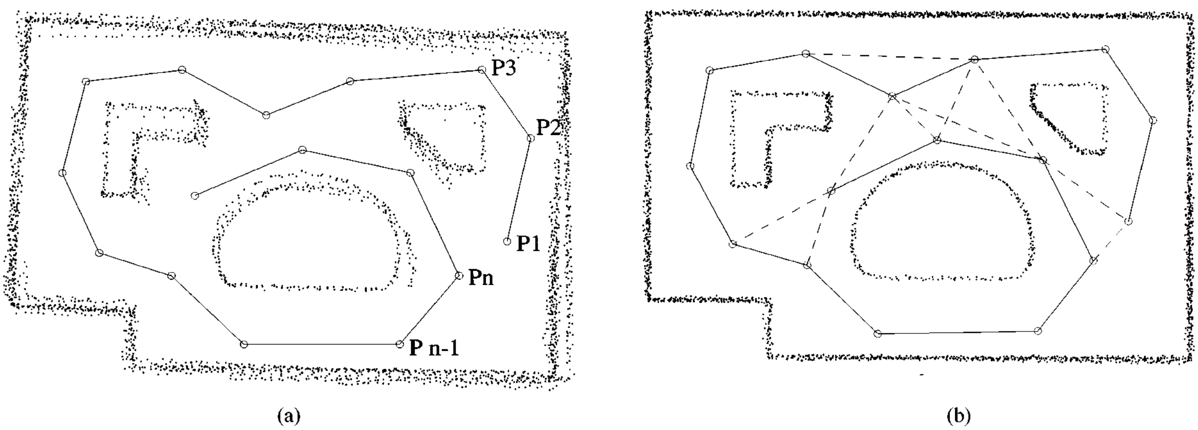


Figure 2. An example of consistently aligning a set of simulated scans: (a) original scans badly misaligned due to accumulated pose errors; (b) the result of aligning these scans based on a network of relative pose constraints. The constraints are indicated by line segments connecting pairs of poses. Two types of constraints are used: those derived from aligning a pair of scans (marked by both solid and dotted lines), and those from odometry measurements (marked by solid lines).

filter (Ayache and Faugeras, 1989; Kriegman et al., 1989; Leonard et al., 1990; Crowley, 1989; Gonzalez et al., 1994). A major problem with this approach is that the resulting world model may eventually become inconsistent as different parts of the model are updated independently. Moreover, it may be difficult to resolve such inconsistency if the data frames have already been permanently integrated.

To be able to resolve inconsistency once it is detected at a later stage, we need to maintain the local frames of data together with their estimated poses. In addition, we need a systematic method to propagate pose corrections to all related frames.

Consider an example as shown in Fig. 2(a). The robot starts at P_1 and returns to a nearby location P_n after visiting $P_2, \ldots, P_{n-1}$ along the way. By registering the scan taken at P_n against scan P_{n-1}, the pose of P_n can be estimated. However since P_n is close to P_1, it is also possible to derive pose P_n based on P_1 by matching these two scans. Because of errors, the two estimates of P_n could be conflicting. If a weighted average of the two is used as the estimate of P_n, the pose of P_{n-1} should also be updated as otherwise the relation $P_{n-1}P_n$ will be inconsistent with its previous estimate. This inconsistency could be significant if the looped path is long. Similarly, other poses along the path should also be updated. In general, the result of matching pairwise scans is a complex, and possibly conflicting, network of pose relations. We need a uniform framework to integrate all these relations and resolve the conflicts.

In this paper, we present such a framework for consistently registering multiple range scans. The idea of our approach is to maintain all the local frames of data as well as a network of spatial relations among the local frames. Here each local frame is defined as the collection of sensor data measured from a single robot pose. The robot pose, in some global reference frame, is also used to define the local coordinate system of the data frame. Spatial relations between local frames are derived from matching pairs of scans or from odometry measurements. We treat the history of robot poses in a global coordinate system (which define all the local frame positions) as variables. Our goal is to estimate all these pose variables using the network of constraints, and register the scans based on the solved poses. Consistency among the local frames is ensured as all the spatial relations are taken into account simultaneously.

Figure 2 shows an example of consistently aligning a set of simulated scans. Part (a) shows the original scans badly misaligned due to accumulated pose errors. Part (b) shows the result of aligning these scans based on a network of relative pose constraints (with edges indicated by line segments).

1.2. Related Work

The first project that systematically studied the consistency issue in dynamic world modeling is the HILARE project (Chatila and Laumond, 1985). In this system, range signals are segmented into objects which are associated with local object frames. Each local frame is referenced in an absolute global frame along with the uncertainty on the robot's pose at which the object frame is constructed. New sensor data are matched to the current model of individual object frames. If some object which has been discovered earlier is observed again, its object frame pose is updated (by averaging). In circumstances that the uncertainty of some object frame is less than the uncertainty of the current robot pose, as it happens when the object frame is created earlier, and later the robot sees the object again, the robot's pose may be corrected with respect to that object frame. After correcting the current robot pose, the correction is propagated backwards with a "fading" effect to correct the previous poses. Although the HILARE system addressed the issue of resolving model inconsistency, its solution has the following potential problems. First of all, the system associates local frames with "objects". But if the results of segmenting sensor data or matching the data with model are imperfect, the "objects" and therefore the local frames may not be defined or maintained consistently. When a previously recorded object is detected again, the system only attempts to update the poses (and the associated frames) along the path between the two instances of detecting this object, while the global consistency among all frames in the model may not be maintained. HILARE uses a scalar random variable to represent the uncertainty of a three-degree-of-freedom pose, therefore it can not model the confidences in the individual pose components.

Moutarlier and Chatila presented a theoretical framework for fusing uncertain measurements for environment modeling (Moutarlier and Chatila, 1989). They first discussed two types of representations: relation-based and location-based. In relation-based representation, an object is related to another by the uncertain transform between their reference frames. A network of relationships is maintained as the world model. When new observations are made, all the relationships need to be updated to preserve consistency. In location-based representation, the global references of individual object frames are maintained together with their uncertainties. When objects are re-observed, these object frames and other related frames are updated with respect to the global reference frame. After comparing these two approaches, Moutarlier and Chatila choose to use the location-based approach. They treat the object and robot locations as state variables and maintain all the object variance/covariance matrices as state information. A stochastic-based formulation for fusing new measurements and updating the state variables is introduced. In addition to a global updating approach, they also introduced a relocation-fusion approach which first updates the robot position based on the new observations and then updates the object frames. The relocation-fusion approach reduces the influence of sensor bias in the estimation, although the algorithm is suboptimal.

In a series of work by Durrant-Whyte (1987, 1988a, 1988b), the problem of maintaining consistency in a network of spatial relations was studied thoroughly. In their formulation, the environment model is represented by a set of spatial relations between objects. A probabilistic fusion algorithm similar to the Kalman filter is employed to integrate new measurements to the *a priori* model. When some relations are updated as a result of new observations, the consistency among all relations are enforced by using explicit constraints on the loops of the network. The updating procedure is formulated as constrained optimization and it allows new observations to be propagated through the network while consistency between prior constraints and observed information is maintained. In another similar approach, Tang and Lee (1992) formulated a geometric feature relation graph for consistent sensor data fusion. They proposed a two-step procedure for resolving inconsistency in a network of measurements of relations. In the first step, a compromise between the conflicting measurements of relations is achieved by the fusion of these measurements. Then in the second step, corrections are propagated to other relations in the network.

The difficulty in maintaining model consistency in a relation-based representation is that the relations are not independent variables. Therefore additional constraints are needed in formulating an updating procedure. The constrained optimization approach seems very complicated and difficult to apply in practice.

In view of the previous methods, we present a new approach which has the following distinctive characteristics:

1. We use an unambiguous definition of an object frame as the collection of sensor measurements observed from a single robot position. Thus we avoid the difficult task of segmenting and recognizing objects (which the previous methods rely on in order to create and update object frames). It is also important to note that we use a robot pose to define the reference for an object frame. In a local frame, the relative object positions with respect to the robot pose are fixed (whose uncertainty is no more than bounded sensing errors). During the estimation process, when the robot position in the global reference frame is updated, effectively the global coordinates of all objects in the current frame are updated accordingly. Therefore by maintaining the history of robot poses, we also maintain the spatial relationships among the object frames.
2. Our approach uses a combination of relation-based and location-based representations. We treat relations as primitives, but treat locations as free variables. This is different from the pure relation-based approach in that we do not directly update the existing relations in the network when new observations are made. We simply add new relations to the network. All the relations are used as constraints to solve for the location variables which, in turn, define a set of updated and consistent relations. On the other hand, our approach is different from the location-base approach by Moutarlier and Chatila (1989) in that we do not assume the covariance matrices between the object frames as known. Our state information is the entire set of raw relations. We derive the covariance matrices at the same time as we solve for the position variables.
3. We obtain direct spatial relations between object frames. Because our object frames are tied to robot poses, odometry measurements directly provide spatial relations between the frames. More importantly, we may align two overlapping frames of data (in our case range scans) to derive more accurate relations between frames. In previous approaches, the robot typically relies on odometry to first determine its new pose. Then the detection of objects allows the robot pose as well as the object locations to be updated. Since the relations between object frames are updated rather indirectly through the robot pose, biases in odometry measurements may lead to divergence in the estimation of object positions, as reported in Moutarlier and Chatila (1989). Moutarlier and Chatila propose an algorithm that is supposed to address the divergence problem at the expense of a sub-optimal solution. Our formulation does not have this problem, as we obtain direct spatial relations between object frames by aligning the data, and therefore we are less sensitive to odometry biases.

2. Overview of Approach

We formulate our approach to multiple scan registration as one of estimating the global poses of the scans, by using all the pose relations as constraints. Here the scan poses are considered as variables. A pose relation is an estimated spatial relation between the two poses which can be derived from matching two range scans. We also obtain pose relations from odometry measurements. Finally, we estimate all the poses by solving an optimization problem. The issues involved in this approach are discussed in the following subsections.

2.1. Deriving Pose Relations

Since we use a robot pose to define the local coordinate system of a scan, pose relations between scans can be directly obtained from odometry which measures the relative movement of the robot. In Section 4.2, we will discuss the representation of odometry pose constraint and its uncertainty.

More accurate relations between scan poses are derived from aligning pairwise scans of points. Here any pairwise scan matching algorithm can be used. One possible choice is the extension to Cox's algorithm (Cox, 1991) where line segments are first fit to one scan and then points in another scan are matched to the derived line segments. In our previous studies, we proposed another scan matching algorithm which is based on direct point to point matching (Lu, 1995; Lu and Milios, 1997). In either case, the scan matching algorithm takes two scans and a rough initial estimate of their relative pose (for example from odometry information) as input. The output is a much improved estimate of the relative pose.

After aligning two scans, we can record a set of corresponding points on the two scans. This correspondence set will form a constraint between the two poses. In Section 4.3, we will formulate this type of constraint and its uncertainty as used in the estimation algorithm.

When we match two scans, we first project one scan to the local coordinate of the other scan, and discard the points which are likely not visible from the second pose. The amount of overlap between two scans is estimated empirically from the spatial extent of the matching parts between the two scans. A pose relation is only derived when the overlap is significant enough (larger than a given threshold).

2.2. *Constructing a Network of Pose Relations*

Given the pairwise pose relations, we can form a network. Formally, the network of constraints is defined as a set of nodes and a set of links between pairs of nodes. A node of the network is a pose of the robot on its trajectory at which a range scan is taken. Here a pose is defined as a three dimensional vector $(x, y, \theta)^t$ consisting of a 2D robot position and the home orientation of the rotating range sensor. We then define two types of links between a pair of nodes. First, if two poses are adjacent along the robot path, we say that there is a *weak link* between the two nodes which is the odometry measurement of the relative pose. Second, if the range scans taken at two poses have a sufficient overlap, we say that there is a *strong link* between the two nodes. To decide whether there is sufficient overlap between scans, we use an empirical measure. The spatial extent in the overlapping part of two scans should be larger than a fixed percentage of the spatial extent covered by both scans.

For each strong link, a constraint on the relative pose is determined by the set of corresponding points on the two scans given by the matching algorithm. It is possible to have multiple links between two nodes. Figure 3 shows an environment and the constructed network of pose relations.

2.3. *Combining Pose Relations in a Network*

The pose relations in a network can be potentially inconsistent because they are not independent variables (the number of relations may be more than the degrees of freedom in the network), while the individually estimated relations are prone to errors. Our task is to combine all the pose relations and resolve any inconsistency. This problem is formulated as one of optimally estimating the global poses of nodes in the network. We do not deal with the relations directly. Rather, we first solve for the nodes which constitute a set of free variables. Then a consistent set of relations which represents a compromise of all *a priori* relations is defined by the poses on the nodes.

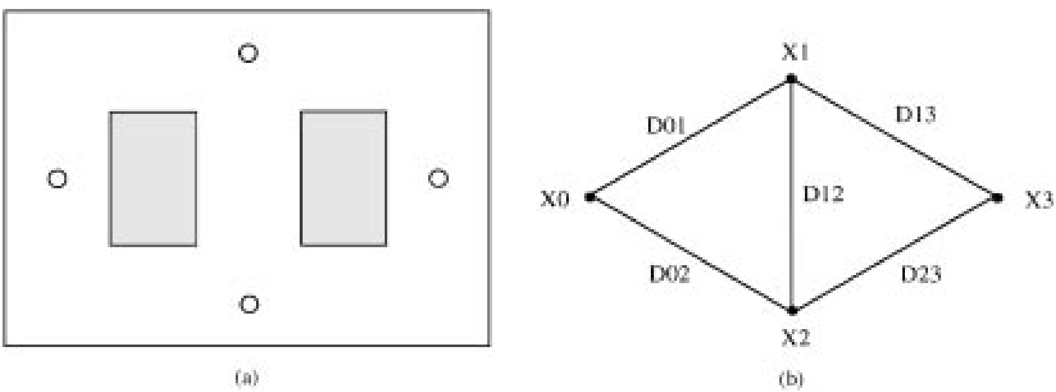


Figure 3. Example of constructing a network of pose relations from matching pairwise scans: (a) a simulated environment where the scan poses are labeled by circles; (b) the network of pose relations constructed from matching overlapping scans.

An optimization problem is defined as follows. We construct an objective function from the network with all the pose coordinates as variables (except one pose which defines our reference coordinate system). Every link in the network is translated into a term in the objective function which can be conceived as a spring connecting two nodes. The spring achieves minimum energy when the relative pose between the two nodes equals the measured value (either from matching two scans or from odometry). Then the objective function represents the total energy in the network. We finally solve for all the pose variables at once by minimizing this total energy function.

2.4. The Three-Node Example

Using the 3-node example, we illustrate the difference of our formulation from previous approaches.

Assume that the network consists of three nodes: P_1, P_2, P_3, and three relations $T_1 = P_1 P_2$, $T_2 = P_2 P_3$, $T_3 = P_3 P_1$. When there is new measurement $\bar{T}_1$ for relation T_1, the algorithm by Durrant-Whyte (1988a) updates the three relations to T_1', T_2', T_3' based on an optimization criterion which is subject to the constraint $T_1' T_2' T_3' = I$.

In our approach, we pool together all the relations T_1, T_2, T_3, as well $\bar{T}_1$ to form an optimization problem and solve for a new estimate for the nodes: P_1', P_2', P_3'. These node positions define a consistent set of relations: $T_1' = P_1' P_2'$, $T_2' = P_2' P_3'$, $T_3' = P_3' P_1'$. Note that the node positions are free variables so we do not need to solve a complex constrained system.

Moutarlier and Chatila (1989) also treat the node positions as variables when updating the network with new measurements. But they assume the knowledge of covariance matrices among the *a priori* estimates of P_1, P_2, P_3. However, we only require the variances of individual measurement errors on the relations T_1, T_2, T_3, $\bar{T}_1$, which are directly available from sensor models.

The rest of the paper is organized as follows. In Section 3, we present the optimization criterion by considering a generic optimal estimation problem. We derive a closed-form solution in a linear special case. In Section 4, we formulate the pose relations as well as the objective function in the context of range scan registration. The closed-form solution derived in Section 3 is applied to solve for the scan poses. In Section 5, we present experimental results.

3. Optimal Estimation from a Network of Relations

In this section, we formulate a generic optimal estimation algorithm which combines a set of relations in a network. This algorithm will later be applied in Section 4 in the context of robot pose estimation and scan data registration.

3.1. Definition of the Estimation Problem

We consider the following generic optimal estimation problem. Assume that we are given a network of uncertain measurements about $n+1$ nodes $X_0, X_1, \ldots, X_n$. Here each node X_i represents a d-dimensional position vector. A link D_{ij} between two nodes X_i and X_j represents a measurable difference of the two positions. Generally, D_{ij} is a (possibly nonlinear) function of X_i and X_j and we refer to this function as the measurement equation. Especially interesting to us is the simple linear case where the measurement equation is $D_{ij} = X_i - X_j$.

We model an observation of D_{ij} as $\bar{D}_{ij} = D_{ij} + \Delta D_{ij}$ where ΔD_{ij} is a random Gaussian error with zero mean and known covariance matrix C_{ij}. Given a set of measurements $\bar{D}_{ij}$ between pairs of nodes and the covariance C_{ij}, our goal is to derive the optimal estimate of the position X_i's by combining all the measurements. Moreover, we want to derive the covariance matrices of the estimated X_i's based on the covariance matrices of the measurements.

Our criterion of optimal estimation is based on the *maximum likelihood* or *minimum variance* concept. The node position X_i's (and hence the position difference D_{ij}'s) are determined in such a way that the conditional joint probability of the derived D_{ij}'s, given their observations $\bar{D}_{ij}$'s, is maximized. If we assume that all the observation errors are Gaussian and mutually independent, the criterion is equivalent to minimizing the following Mahalanobis distance (where the summation is over all the given measurements):

$$W = \sum_{(i,j)} (D_{ij} - \bar{D}_{ij})^t C_{ij}^{-1} (D_{ij} - \bar{D}_{ij}). \qquad (1)$$

Even if the observation errors are not independent, a similar distance function can still be formed. However, it will involve the correlation matrices of the measurements. The assumption of independence is actually not necessary in our formulation. The assumption makes practical sense as the covariances of errors are difficult to estimate.

A typical application of the optimal estimation problem is in mobile robot navigation, where we want to estimate the robot pose and its uncertainty in three degrees of freedom (x, y, θ). The observations are the relative robot poses from odometry, and also possible from matching sensor measurements. We want to utilize all the available measurements to derive the optimal estimate of the robot poses. Note that in this application, the measurement equation is non-linear because of the θ component in the robot pose.

Our approach above differs from the one typically used within a Kalman filter formulation, in which only the current pose is estimated, while the history of previous poses and associated measurements is collapsed into the current state of the Kalman filter. Our objective, however, is not simply getting from A to B safely and accurately, but also building a map of the environment. It is, therefore, meaningful to use all the measurements obtained so far in the mapping process. The old poses themselves are not particularly useful. But we are using the poses to define local object frames. Thus maintaining the history of robot poses is equivalent to maintaining the structure of the environment model. The advantage of using a pose to define a data frame is that it is unambiguous and it avoids the difficult segmentation and object identification problem present in other work.

Next, we study the case when the measurement equation is linear and we derive closed-form solutions for the optimal estimates of the nodes and their covariances. Later, we will solve the non-linear robot pose estimation problem by approximately forming linear measurement equations.

3.2. *Solution of Optimal Linear Estimation*

We consider the special case where the measurement equation has the simple linear form: $D_{ij} = X_i - X_j$. Here $X_i, i = 0, 1, \ldots, n$ are the nodes in the network which are d-dimensional vectors and the D_{ij}'s are the links of the network. Without loss of generality, we assume that there is a link D_{ij} between every pair of nodes X_i, X_j. For each D_{ij}, we have an observation $\bar{D}_{ij}$ which is assumed to have Gaussian distribution with mean value D_{ij} and known covariance C_{ij}. In case the actual link D_{ij} is missing, we can simply let the corresponding C_{ij}^{-1} be 0. Then the criterion for the optimal estimation is to minimize the following Mahalanobis distance:

$$W = \sum_{0 \le i < j \le n} (X_i - X_j - \bar{D}_{ij})^t C_{ij}^{-1} (X_i - X_j - \bar{D}_{ij}). \tag{2}$$

Note that W is a function of all the position X_i's. Since we can only solve for relative positions given the relative measurements, we choose one node X_0 as a reference and consider its coordinate as constant. Without loss of generality, we let $X_0 = 0$ and then $X_1, X_2, \ldots, X_n$ will represent the relative positions from X_0.

We can express the measurement equations in a matrix form as

$$\mathbf{D} = \mathbf{HX} \tag{3}$$

where $\mathbf{X}$ is the nd-dimensional vector which is the concatenation of $X_1, X_2, \ldots, X_n$; $\mathbf{D}$ is the concatenation of all the position differences of the form $D_{ij} = X_i - X_j$; and $\mathbf{H}$ is the incidence matrix with all entries being $1, -1$, or 0. Then the function W can be represented in matrix form as:

$$W = (\bar{\mathbf{D}} - \mathbf{HX})^t \mathbf{C}^{-1} (\bar{\mathbf{D}} - \mathbf{HX}) \tag{4}$$

where $\bar{\mathbf{D}}$ is the concatenation of all the observations $\bar{D}_{ij}$ for the corresponding D_{ij} and $\mathbf{C}$ is the covariance of $\bar{\mathbf{D}}$ which is a square matrix consists of C_{ij}'s as submatrices.

Then the solution for $\mathbf{X}$ which minimizes W is given by

$$\mathbf{X} = (\mathbf{H}^t \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^t \mathbf{C}^{-1} \bar{\mathbf{D}}. \tag{5}$$

The covariance of $\mathbf{X}$ is

$$\mathbf{C_X} = (\mathbf{H}^t \mathbf{C}^{-1} \mathbf{H})^{-1}. \tag{6}$$

If the measurement errors are independent, C will be block-diagonal and the solution can be simplified. Denote the $nd \times nd$ matrix $\mathbf{H}^t \mathbf{C}^{-1} \mathbf{H}$ by $\mathbf{G}$ and expand the matrix multiplications.

We can obtain the $d \times d$ sub-matrices of $\mathbf{G}$ as

$$\begin{aligned} G_{ii} &= \sum_{j=0}^{n} C_{ij}^{-1} \\ G_{ij} &= -C_{ij}^{-1} \quad (i \neq j) \end{aligned} \tag{7}$$

Denote the nd-dimensional vector $\mathbf{H}^t\mathbf{C}^{-1}\bar{\mathbf{D}}$ by $\mathbf{B}$. Its d-dimensional sub-vectors are the following (let $\bar{D}_{ij} = -\bar{D}_{ji}$):

$$B_i = \sum_{j=0;\ j\neq i}^{n} C_{ij}^{-1}\bar{D}_{ij}. \tag{8}$$

Then the position estimates and covariance can be written as

$$\mathbf{X} = \mathbf{G}^{-1}\mathbf{B}; \quad \mathbf{C}_\mathbf{X} = \mathbf{G}^{-1}. \tag{9}$$

The above algorithm requires $\mathbf{G} = \mathbf{H}^t\mathbf{C}^{-1}\mathbf{H}$ to be invertible. If the network is fully connected and the individual error covariances are normally behaved, we believe it is possible to prove that $\mathbf{G}$ is invertible. Note the dimension of $\mathbf{G}$ (number of free nodes) is less than or equal to the dimension of $\mathbf{C}$ (number of edges) in a fully connected graph.

3.3. Special Networks

We will apply the formula in Eq. (9) to two interesting special cases as in Fig. 4. First, if the network consists of two serially connected links, D_{01} and D_{12}, the derived estimate of X_2 and its covariance matrix are

$$X_2 = D_{01} + D_{12} \tag{10}$$

$$C_2 = C_{01} + C_{12} \tag{11}$$

Another case to consider is the network which consists of two parallel links D' and D'' between two nodes X_0 and X_1. If the covariance of the two links is C' and C'', the estimate of X_1 and its covariance are given by

$$X_1 = (C'^{-1} + C''^{-1})^{-1}(C'^{-1}D' + C''^{-1}D'') \tag{12}$$

$$C = (C'^{-1} + C''^{-1})^{-1} \tag{13}$$

The solution is equivalent to the Kalman filter formulation. The above two cases correspond to the compounding and merging operations given by Smith and Cheeseman (1986), which are used to reduce a complex network to a single relation. Smith and Cheeseman's algorithm has a limitation that it only applies to networks formed by serial and parallel connections.

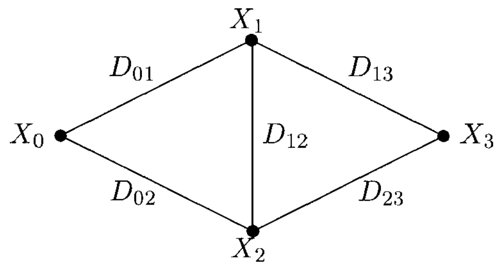


Figure 5. A wheatstone bridge network.

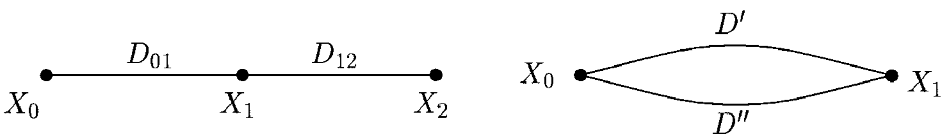
Consider the network in the form of a Wheatstone bridge (Fig. 5). The estimate of X_3 can not be obtained through compounding and merging operations. Therefore, the method by Smith and Cheeseman can not be directly applied to simplify this network,[1] while in our method, the variables X_1, X_2, X_3 can be solved from the linear system $\mathbf{GX} = \mathbf{B}$ where

$$\mathrm{G} = \begin{pmatrix} C_{01}^{-1} + C_{12}^{-1} + C_{13}^{-1} & -C_{12}^{-1} & -C_{13}^{-1} \\ -C_{12}^{-1} & C_{02}^{-1} + C_{12}^{-1} + C_{23}^{-1} & -C_{23}^{-1} \\ -C_{13}^{-1} & -C_{23}^{-1} & C_{13}^{-1} + C_{23}^{-1} \end{pmatrix} \tag{14}$$

$$\mathrm{B} = \begin{pmatrix} C_{01}^{-1}\bar{D}_{01} + C_{12}^{-1}\bar{D}_{12} + C_{13}^{-1}\bar{D}_{13} \\ C_{02}^{-1}\bar{D}_{02} - C_{12}^{-1}\bar{D}_{12} + C_{23}^{-1}\bar{D}_{23} \\ -C_{13}^{-1}\bar{D}_{13} - C_{23}^{-1}\bar{D}_{23} \end{pmatrix}. \tag{15}$$

The covariance matrix for the estimated position X_3 has a nice symmetric form (derived by expanding $\mathbf{G}^{-1}$):

$$\begin{aligned} \mathrm{C}_3^{-1} = &\begin{pmatrix} C_{01}^{-1} & C_{02}^{-1} \end{pmatrix} \\ &\times \begin{pmatrix} C_{01}^{-1} + C_{12}^{-1} + C_{13}^{-1} & -C_{12}^{-1} \\ -C_{12}^{-1} & C_{02}^{-1} + C_{12}^{-1} + C_{23}^{-1} \end{pmatrix}^{-1} \\ &\times \begin{pmatrix} C_{13}^{-1} \\ C_{23}^{-1} \end{pmatrix} \end{aligned} \tag{16}$$



(a) (b)

Figure 4. (a) Serial connection; (b) parallel connection.

4. Derivation of Pose Relations

In this section, we will apply the optimal estimation algorithm, as derived in Section 3, to the robot pose estimation and scan data registration problem. To do this, we need to derive linearized measurement equations for the pose relations. In the following subsections, we study a constraint on pose difference given by matched scans or odometry measurements. For each constraint, we formulate a term in the form of Mahalanobis distance. For convenience in discussions of pose measurements, we will first define a pose compounding operation.

4.1. Pose Compounding Operation

Assume that the robot starts at a pose $V_b = (x_b, y_b, \theta_b)^t$ and it then changes its pose by $D = (x, y, \theta)^t$ relative to V_b, ending up at a new pose $V_a = (x_a, y_a, \theta_a)^t$. Then we say that pose V_a is the compounding of V_b and D. We denote it as:

$$V_a = V_b \oplus D. \tag{17}$$

The coordinates of the poses are related by:

$$x_a = x_b + x \cos\theta_b - y \sin\theta_b \tag{18}$$

$$y_a = y_b + x \sin\theta_b + y \cos\theta_b \tag{19}$$

$$\theta_a = \theta_b + \theta. \tag{20}$$

This is the same compounding operation as defined by Smith and Cheeseman (1986). If we consider that an absolute pose defines a coordinate system (the xy coordinates of the origin and the direction of one axis), and a relative pose defines a change of coordinate system (a translation followed by a rotation), then the compounding operation gives the pose which defines the new coordinate system after the transformation. The compounding operation is not commutative, but it is associative. We can thus define the compounding of a series of poses.

It is also useful to define the inverse of compounding which takes two poses and gives the relative pose:

$$D = V_a \ominus V_b. \tag{21}$$

The coordinates are related by the following equations:

$$x = (x_a - x_b) \cos\theta_b + (y_a - y_b) \sin\theta_b \tag{22}$$

$$y = -(x_a - x_b) \sin\theta_b + (y_a - y_b) \cos\theta_b \tag{23}$$

$$\theta = \theta_a - \theta_b. \tag{24}$$

If D_{ab} is the relative pose $V_a \ominus V_b$, the reversed relative pose $D_{ba} = V_b \ominus V_a$ can be obtained from D_{ab} by a unary operation:

$$D_{ba} = \ominus D_{ab} = (0, 0, 0)^t \ominus D_{ab}. \tag{25}$$

We can verify that $(\ominus D) \oplus V = V \ominus D$.

We also want to define a compounding operation between a full 3D pose $V_b = (x_b, y_b, \theta_b)$ and a 2D position vector $u = (x, y)^t$. The result is another 2D vector $u' = (x', y')^t$. We still denote the operation as

$$u' = V_b \oplus u. \tag{26}$$

The coordinates for u' are given by the first two equations of the full 3D pose compounding (Eqs. (18, 19)). This 2D compounding operation is useful for transforming an non-oriented point (typically from a range sensor) from its local sensor coordinate system to the global coordinate system.

4.2. Pose Relations from Matched Scans

Let V_a and V_b be two nodes in the network and assume there is a strong link connecting the two poses. From the pairwise scan matching algorithm, we get a set of pairs of corresponding points: $u_k^a, u_k^b, k = 1, \ldots, m$, where the 2D non-oriented points u_k^a, u_k^b are from scan S_a and S_b, respectively. Each pair (u_k^a, u_k^b) corresponds to the same physical point in the robot's environment while they are represented in different local coordinate systems. If we ignore any sensing or matching errors, two corresponding points are related by:

$$\Delta Z_k = V_a \oplus u_k^a - V_b \oplus u_k^b = 0. \tag{27}$$

If we take the random observation errors into account, we can regard ΔZ_k as a random variable with zero mean and some unknown covariance C_k^Z. From the correspondence pairs, we can form a constraint on the pose difference by minimizing the following distance function:

$$F_{ab}(V_a, V_b) = \sum_{k=1}^{m} \left\| \left(V_a \oplus u_k^a\right) - \left(V_b \oplus u_k^b\right) \right\|^2. \tag{28}$$

If we notice that a pose change is a rigid transformation under which the squared Euclidean distance $\|\cdot\|^2$ is invariant, we can rewrite the function in an equivalent form:

$$F_{ab}(V_a, V_b) = \sum_{k=1}^{m} \left\| \left((V_a \ominus V_b) \oplus u_k^a\right) - u_k^b \right\|^2. \tag{29}$$

Thus F_{ab} is a function of $D' = V_a \ominus V_b$. The solution of D' which minimizes F_{ab} can be derived in closed-form (see Lu, 1995). The relation $D' = V_a \ominus V_b$ is the measurement equation.

In order to reduce F_{ab} into the Mahalanobis distance form, we linearize each term ΔZ_k. Let $\bar{V}_a = (\bar{x}_a, \bar{y}_a, \bar{\theta}_a)^t$, $\bar{V}_b = (\bar{x}_b, \bar{y}_b, \bar{\theta}_b)^t$ be some close estimates of V_a and V_b. Denote $\Delta V_a = \bar{V}_a - V_a$ and $\Delta V_b = \bar{V}_b - V_b$. Let $u_k = (x_k, y_k)^t = V_a \oplus u_k^a \approx V_b \oplus u_k^b$ (the global coordinates of a pair of matching points). Then for small ΔV_a and ΔV_b, we can derive from Taylor expansion:

$$\begin{aligned}
\Delta Z_k &= V_a \oplus u_k^a - V_b \oplus u_k^b \\
&= (\bar{V}_a - \Delta V_a) \oplus u_k^a - (\bar{V}_b - \Delta V_b) \oplus u_k^b \\
&\approx \left(\bar{V}_a \oplus u_k^a - \bar{V}_b \oplus u_k^b\right) \\
&\quad - \left(\begin{pmatrix} 1 & 0 & \bar{y}_a - y_k \\ 0 & 1 & -\bar{x}_a + x_k \end{pmatrix} \Delta V_a \right. \\
&\quad \left. - \begin{pmatrix} 1 & 0 & \bar{y}_b - y_k \\ 0 & 1 & -\bar{x}_b + x_k \end{pmatrix} \Delta V_b \right) \\
&= \left(\bar{V}_a \oplus u_k^a - \bar{V}_b \oplus u_k^b\right) \\
&\quad - \begin{pmatrix} 1 & 0 & -y_k \\ 0 & 1 & x_k \end{pmatrix} (\bar{H}_a \Delta V_a - \bar{H}_b \Delta V_b)
\end{aligned} \tag{30}$$

where

$$\bar{H}_a = \begin{pmatrix} 1 & 0 & \bar{y}_a \\ 0 & 1 & -\bar{x}_a \\ 0 & 0 & 1 \end{pmatrix}, \quad \bar{H}_b = \begin{pmatrix} 1 & 0 & \bar{y}_b \\ 0 & 1 & -\bar{x}_b \\ 0 & 0 & 1 \end{pmatrix}. \tag{31}$$

We can rewrite Eq. (30) as

$$\Delta Z_k \approx \bar{Z}_k - M_k D \tag{32}$$

where

$$\bar{Z}_k = \bar{V}_a \oplus u_k^a - \bar{V}_b \oplus u_k^b \tag{33}$$

$$M_k = \begin{pmatrix} 1 & 0 & -y_k \\ 0 & 1 & x_k \end{pmatrix} \tag{34}$$

$$D = (\bar{H}_a \Delta V_a - \bar{H}_b \Delta V_b). \tag{35}$$

Thus we can now regard D in Eq. (35) as the pose difference measurement equation to replace $D' = V_a \ominus V_b$. For the m correspondence pairs, we can form m equations as in Eq. (32). If we concatenate the $\bar{Z}_k$'s to form a $2m \times 1$ vector $\mathbf{Z}$, and stack the M_k's to form a $2m \times 3$ matrix $\mathbf{M}$, then F_{ab} can be rewritten as a quadratic function of D:

$$F_{ab}(D) = \sum_{k=1}^{m} (\Delta Z_k)^t (\Delta Z_k) \tag{36}$$

$$\approx (\mathbf{Z} - \mathbf{M}D)^t (\mathbf{Z} - \mathbf{M}D). \tag{37}$$

We can then solve for the $D = \bar{D}$ which minimizes F_{ab} as

$$\bar{D} = (\mathbf{M}^t \mathbf{M})^{-1} \mathbf{M}^t \mathbf{Z}. \tag{38}$$

The criterion of minimizing $F_{ab}(D)$ constitutes a least-squares linear regression. In Eq. (32), M_k is known and $\bar{Z}_k$ is observed with an error ΔZ_k having zero mean and unknown covariance C_k^Z. If we assume that all the errors are independent variables having the same Gaussian distribution, and further assume that the error covariance matrices have the form:

$$C_k^Z = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}, \tag{39}$$

then the least squares solution $\bar{D}$ has the Gaussian distribution whose mean value is the true underlying value and whose estimated covariance matrix is given by $C_D = s^2(\mathbf{M}^t\mathbf{M})^{-1}$, where s^2 is the unbiased estimate of σ^2:

$$s^2 = (\mathbf{Z} - \mathbf{M}\bar{D})^t (\mathbf{Z} - \mathbf{M}\bar{D})/(2m-3) = \frac{F_{ab}(\bar{D})}{2m-3}. \tag{40}$$

Moreover, we notice that Eq. (37) can be rewritten as

$$F_{ab}(D) \approx (\bar{D} - D)^t (\mathbf{M}^t\mathbf{M})(\bar{D} - D) + F_{ab}(\bar{D}). \tag{41}$$

We can define the energy term W_{ab} corresponding to the pose relation which is equivalent to a Mahalanobis distance:

$$W_{ab} = (F_{ab}(D) - F_{ab}(\bar{D}))/s^2 \tag{42}$$

$$\approx (\bar{D} - D)^t C_D^{-1} (\bar{D} - D) \tag{43}$$

where

$$C_D = s^2 (\mathbf{M}^t \mathbf{M})^{-1} \tag{44}$$

is the estimated covariance of $\bar{D}$. Note that D (as given in Eq. (35)) is the linearized pose difference measurement equation.

In deriving the covariance matrix C_D, we made assumptions that the matrix is diagonal and the individual components of errors are zero mean Gaussian. It is probably difficult to justify these assumption. However, we believe that they are reasonable ones in practice. If any other estimates of the covariance matrices are available, they can certainly also be incorporated in our global estimation formulation.

4.3. *Pose Relations from Odometry*

We also form an energy term in the objective function for each weak link. Suppose odometry gives a measurement $\bar{D}'$ of the relative pose D' as the robot travels from pose V_b to pose V_a. The measurement equation is:

$$D' = V_a \ominus V_b. \tag{45}$$

We define the energy term in the objective function as follows:

$$W_{ab} = (\bar{D}' - D')^t C'^{-1} (\bar{D}' - D') \tag{46}$$

where C' is the covariance of the odometry error in the measurement $\bar{D}'$.

The covariance of measurement error is estimated as follows. Consider that a cycle of pose change consists of: (1) the robot platform rotation by an angle α to face towards the new target position; (2) the robot translation by a distance L to arrive at the new position; (3) the sensor rotation by a total cumulative angle β (usually 360°) to take a scan of measurements while the platform is stationary. We model the deviations σ_α, σ_L, σ_β, of the errors in the variables α, L, and β as proportional to their corresponding values, while the constant ratios are determined empirically. The 3D pose change $D' = (x, y, \theta)^t$ is derived as:

$$x = L\cos\alpha; \quad y = L\sin\alpha; \quad \theta = \alpha + \beta \tag{47}$$

Then the covariance C' of the pose change D' can be approximated as:

$$C' = J \begin{pmatrix} \sigma_\alpha^2 & 0 & 0 \\ 0 & \sigma_L^2 & 0 \\ 0 & 0 & \sigma_\beta^2 \end{pmatrix} J^t \tag{48}$$

where J is the Jacobian matrix consisting of the partial derivatives of $(x, y, \theta)^t$ with respect to $(\alpha, L, \beta)^t$:

$$J = \begin{pmatrix} -L\sin\alpha & \cos\alpha & 0 \\ L\cos\alpha & \sin\alpha & 0 \\ 1 & 0 & 1 \end{pmatrix}. \tag{49}$$

We would also like to linearize and transform the measurement equation of D' to make the pose difference representation for odometry measurements consistent with that for matched sensing data. Consider the observation error $\Delta D' = \bar{D}' - D'$ of odometry. Let $\bar{V}_a = (\bar{x}_a, \bar{y}_a, \bar{\theta}_a)^t$, $\bar{V}_b = (\bar{x}_b, \bar{y}_b, \bar{\theta}_b)^t$ be some close estimates of V_a and V_b. Denote $\Delta V_a = \bar{V}_a - V_a$ and $\Delta V_b = \bar{V}_b - V_b$. Then through Taylor expansion, the observation error $\Delta D'$ becomes:

$$\begin{aligned} \Delta D' &= \bar{D}' - D' = \bar{D}' - (V_a \ominus V_b) && (50)\\ &= \bar{D}' - ((\bar{V}_a - \Delta V_a) \ominus (\bar{V}_b - \Delta V_b)) && (51)\\ &\approx \bar{D}' - (\bar{V}_a \ominus \bar{V}_b) + \bar{K}_b^{-1}(\Delta V_a - \bar{H}_{ab}\Delta V_b) && (52) \end{aligned}$$

where

$$\begin{aligned} \bar{K}_b^{-1} &= \begin{pmatrix} \cos\bar{\theta}_b & \sin\bar{\theta}_b & 0 \\ -\sin\bar{\theta}_b & \cos\bar{\theta}_b & 0 \\ 0 & 0 & 1 \end{pmatrix}; \\ \bar{H}_{ab} &= \begin{pmatrix} 1 & 0 & -\bar{y}_a + \bar{y}_b \\ 0 & 1 & \bar{x}_a - \bar{x}_b \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{53}$$

Notice that $\bar{H}_{ab} = \bar{H}_a^{-1}\bar{H}_b$ where $\bar{H}_a$ and $\bar{H}_b$ are defined in Eq. (31). If we define a new observation error $\Delta D = -\bar{H}_a \bar{K}_b \Delta D'$, then we can rewrite Eq. (52) as

$$\Delta D = \bar{D} - (\bar{H}_a \Delta V_a - \bar{H}_b \Delta V_b) = \bar{D} - D \tag{54}$$

where we denote

$$\bar{D} = \bar{H}_a \bar{K}_b((\bar{V}_a \ominus \bar{V}_b) - \bar{D}') \tag{55}$$

$$D = \bar{H}_a \Delta V_a - \bar{H}_b \Delta V_b. \tag{56}$$

Notice that now we are dealing with the measurement equation for D which is consistent with that for matched sensing data. $\bar{D}$ can be considered as an observation of D. The covariance C of $\bar{D}$ can be computed from the covariance C' of $\bar{D}'$ as:

$$C = \bar{H}_a \bar{K}_b C' \bar{K}_b^t \bar{H}_a^t. \tag{57}$$

The energy term in the objective function now becomes:

$$W_{ab} \approx (\bar{D} - D)^t C^{-1} (\bar{D} - D). \tag{58}$$

4.4. Optimal Pose Estimation

Once we have uniformly formulated the two types of measurements, we can apply the estimate algorithm in Section 3 to solve for the pose variables. Denote the robot poses as $V_i, i = 0, 1, \ldots, n$. The total energy function from all the measurements is:

$$W = \sum_{(i,j)} (\bar{D}_{ij} - D_{ij})^t C_{ij}^{-1} (\bar{D}_{ij} - D_{ij}) \tag{59}$$

where D_{ij} is the linearized pose difference between V_j and V_i:

$$D_{ij} = \bar{H}_i \Delta V_i - \bar{H}_j \Delta V_j \tag{60}$$

and $\bar{D}_{ij}$ is an observation of D_{ij} ($\bar{D}_{ij}$ is derived from the true observations, either range data or odometry measurements). The covariance C_{ij} is also known.

By regarding $X_i = \bar{H}_i \Delta V_i$ as the state vector corresponding to a node of the network as in Section 3.2, we can directly apply the closed-form linear solution to solve for the X_i's as well as their covariance C_i^X. The formulas are in Eqs. (5) to (9). Then the pose V_i and its covariance C_i can be updated as:

$$V_i = \bar{V}_i - \bar{H}_i^{-1} X_i, \quad C_i = \left(\bar{H}_i^{-1}\right) C_i^X \left(\bar{H}_i^{-1}\right)^t. \tag{61}$$

Note that the pose estimate V_i and the covariance C_i is given based on the assumption that the reference pose $V_0 = 0$. If, in fact, $V_0 = (x_0, y_0, \theta_0)^t$ is non-zero, the solution should be transformed to

$$V_i' = V_0 \oplus V_i; \quad C_i' = K_0 C_i K_0^t \tag{62}$$

where

$$K_0 = \begin{pmatrix} \cos\theta_0 & -\sin\theta_0 & 0 \\ \sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{63}$$

4.5. Sequential Estimation

The estimation algorithm we previously discussed is a one-step procedure which solves for all the pose variables at the same time. The algorithm is to be applied only after collecting all the measurements. Yet it will be more practically useful if we have a sequential algorithm which continuously provides estimates about the current or past pose variables after getting each new measurement. Here we will describe such a sequential procedure.

Our sequential algorithm maintains the current best estimate about the poses of previously visited places. At each step, a new location is visited and measurements about the new location as well as the previous locations are gathered. By using these new measurements, the current pose can be estimated while the previous poses can be updated.

Let $X_1, \ldots, X_{n-1}$ be the pose vectors which we have previously estimated and let X_n be the current new pose which we are about to measure. Let $\mathbf{X}$ represent the concatenation of $X_1, \ldots, X_{n-1}, X_n$. Assume that we currently have an estimate $\mathbf{X}_0$ of $\mathbf{X}$ whose inverse covariance matrix is $\mathbf{C}_{X_0}^{-1}$. Because we have no knowledge about X_n yet, the X_n component in $\mathbf{X}_0$ contains an arbitrary value and the matrix $\mathbf{C}_{X_0}^{-1}$ has all zeros in the last d rows and d columns, where $d = 3$. Now consider the addition of a set of new measurements relating X_n to some of the past pose variables. Let the measurement equation, in matrix form, be $\mathbf{D} = \mathbf{HX}$ ($\mathbf{H}$ is a constant matrix). Assume that the set of measurements is $\bar{\mathbf{D}}$ which is an unbiased observation of $\mathbf{D}$ whose error has Gaussian distribution with covariance matrix $\mathbf{C_D}$. The updated estimate of $\mathbf{X}$ after using the new measurements is the one which minimizes the following function, using the maximum likelihood criterion, and assuming independent errors:

$$\begin{aligned} W = {} & (\mathbf{X} - \mathbf{X}_0)^t \mathbf{C}_{X_0}^{-1} (\mathbf{X} - \mathbf{X}_0) \\ & + (\bar{\mathbf{D}} - \mathbf{HX})^t \mathbf{C_D}^{-1} (\bar{\mathbf{D}} - \mathbf{HX}). \end{aligned} \tag{64}$$

The solution can be derived as

$$\mathbf{X} = \left(\mathbf{C}_{X_0}^{-1} + \mathbf{H}^t \mathbf{C_D}^{-1} \mathbf{H}\right)^{-1} \left(\mathbf{C}_{X_0}^{-1} \mathbf{X}_0 + \mathbf{H}^t \mathbf{C_D}^{-1} \bar{\mathbf{D}}\right) \tag{65}$$

and the new covariance of $\mathbf{X}$ is

$$\mathbf{C_X} = \left(\mathbf{C}_{X_0}^{-1} + \mathbf{H}^t \mathbf{C_D}^{-1} \mathbf{H}\right)^{-1}. \tag{66}$$

A convenient way of updating $\mathbf{X}$ and $\mathbf{C_X}$ is to maintain a matrix $\mathbf{G} = \sum \mathbf{H}^t \mathbf{C_D}^{-1} \mathbf{H}$ and a vector $\mathbf{B} = \sum \mathbf{H}^t \mathbf{C_D}^{-1} \bar{\mathbf{D}}$ (the summation is over different sets of measurements). Then at each step, the updating algorithm is the following: First increase the dimensions of $\mathbf{G}$ and $\mathbf{B}$ to include the new pose X_n. Update $\mathbf{G}$ and $\mathbf{B}$ as

$$\mathbf{G} \leftarrow \mathbf{G} + \mathbf{H}^t \mathbf{C_D}^{-1} \mathbf{H} \tag{67}$$

$$\mathbf{B} \leftarrow \mathbf{B} + \mathbf{H}^t \mathbf{C_D}^{-1} \bar{\mathbf{D}}. \tag{68}$$

Then the new $\mathbf{X}$ and $\mathbf{C_X}$ are given by

$$\mathbf{X} = \mathbf{G}^{-1} \mathbf{B}; \quad \mathbf{C_X} = \mathbf{G}^{-1}. \tag{69}$$

One potential problem with the above sequential updating procedure is that the state variable $\mathbf{X}$ keeps expanding as it is augmented by a new state at each step. In case the robot path is very long, the variable size may become too large, causing storage or performance problems. A possible solution is to delete some of the old variables while adding the new ones.

We propose a strategy of reducing the number of state variables as follows. In order to choose a pose to be deleted, we check all pairs of poses and find a pair (X_i, X_j) where the correlation between the two poses is the strongest. We then force the relative pose between X_i and X_j to be fixed as a constant. Then X_i can be deleted from the state variables as it can be obtained from X_j. When deleting the node X_i from the network, we transform any link (X_i, X_k) into a link from X_j to X_k. Note that the covariance matrix $\mathbf{C_X}$ contains all the pairwise covariance between any two poses. A correlation ratio between two poses can be computed from the covariance and variance.

By only fixing some relative poses, the basic structure in the network is still maintained. Thus we are still able to consistently update all the pose variables once given new measurements. This strategy is more flexible than the simple strategy of fixing selected absolute poses as constants.

Another approach to reducing the size of the system is to decompose the large network into smaller components. The estimation algorithm is to be applied to each sub-network. The relative pose between two nodes in different sub-networks can be obtained through pose compounding. If there is a single link connecting two parts of a network, the poses in two parts can be estimated separately and then combined with compounding, without loss of information. If, however, the network is strongly connected that there are two or more links between any two nodes, then a decomposition could give a sub-optimal estimation.

5. Implementation and Experiments

5.1. *Implementation of Estimation Procedure*

The implementation of the estimation algorithm is as follows. After building the network, we obtain the initial pose estimates $\bar{V}_1, \ldots, \bar{V}_n$ by compounding the odometry measurements. Then for each link, we compute a measurement vector $\bar{D}_{ij}$ and a covariance matrix C_{ij} according to Eqs. (38), (44) or Eqs. (55), (57). Finally, we form a large linear system $\mathbf{GX} = \mathbf{B}$ as explained in Section 3.2 and solve for the pose variables $\mathbf{X}$.

The components needed to build $\mathbf{G}$ and $\mathbf{B}$ are C_{ij}^{-1} and $C_{ij}^{-1}\bar{D}_{ij}$. In the case of a strong link (from matching a pair of scans), these components can be readily computed as $C_{ij}^{-1} = (\mathbf{M}^t\mathbf{M})/s^2$; $C_{ij}^{-1}\bar{D}_{ij} = (\mathbf{M}^t\mathbf{Z})/s^2$ which can be expanded into simple summations by noting the regularity in the matrix $\mathbf{M}$. In the case of a weak link (from odometry), these components can be computed by multiplications of small matrices (3×3). The most expensive operation in the estimation process is to compute the inverse of a $3n \times 3n$ matrix $\mathbf{G}$ which gives the covariance of $\mathbf{X}$.

The network is stored as a list of links and a list of nodes. Each link contains the following information: type of link, labels of the two nodes, the computed measurement (relative pose), and the covariance matrix of the measurement. Each node contains a range scan.

Note that we made linear approximations in the measurement equations in formulating the optimization criterion. The first order approximation error is proportional to the error in the initial pose estimate. Clearly, if we employ the newly derived pose estimate to formulate the linear algorithm again, a even more accurate pose estimate can be obtained.

The iterative strategy based on this observation converges very fast. Typically, the first iteration corrects over 90% of the total pose error correctable by iterating the process. It usually takes four or five iterations to converge to the limit of machine accuracy.

5.2. *Experiments with Simulated and Real Scan Data*

We now present experiments of applying our algorithm to register simulated and real range scan data. We first show an example using a simulated environment and measurements. This is useful because ground truth is known. Then an example using real data is presented.

In the first example, we simulate a rectangular environment with a width of 10 units. The robot travels around a central object and forms a loop in the path. There are 13 poses along the path at which simulated range scans are generated (with random measurement errors). We also simulate a random odometry error (which is the difference between a pose change the robot thinks it made and the actual pose change) at each leg of the trajectory. The magnitude of the accumulated odometry error is typically in the range of 0.5 units.

We apply our iterative global pose estimation algorithm to correct the pose errors. In Fig. 6(a), we show all the scans recorded in the initial coordinate system where the pose of each scan is obtained by compounding odometry measurements. Due to the accumulation of odometry error the scan data are aligned poorly. In Fig. 6(b), we show the result of correcting the pose

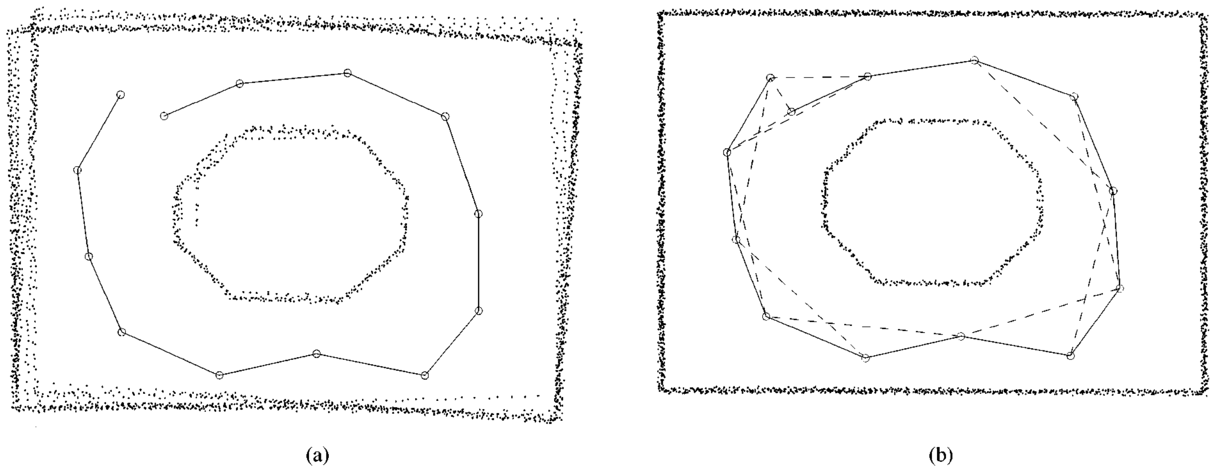


Figure 6. Global registration of multiple scans using simulated scan data: (a) scans recorded in a global coordinate system where the pose of each scan is obtained from compounding odometry measurements. The scans align poorly because of accumulation of odometry error; (b) the result of correcting pose errors. Both the dashed lines and solid lines show the constraints from matching scan pairs. The solid lines also give the robot path and odometry constraints.

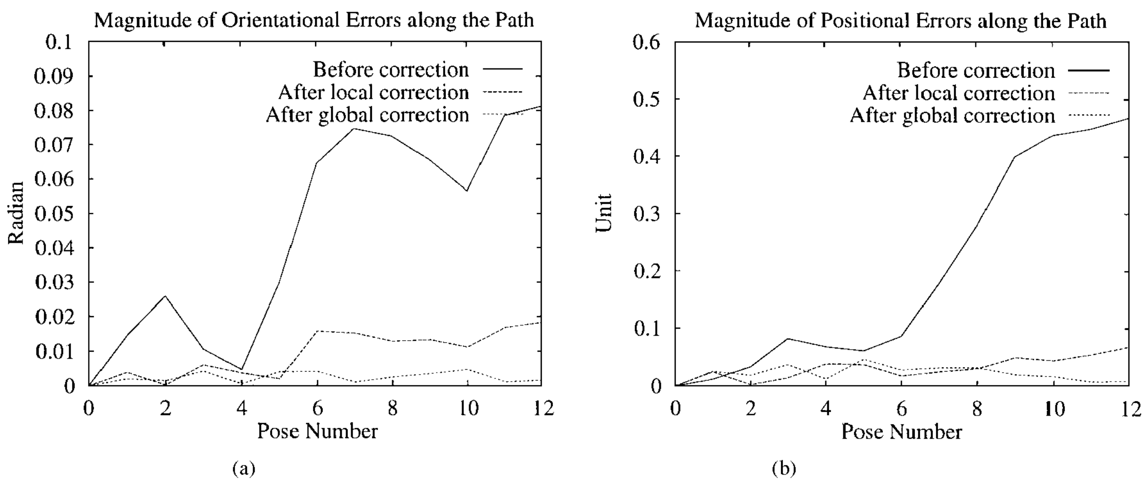
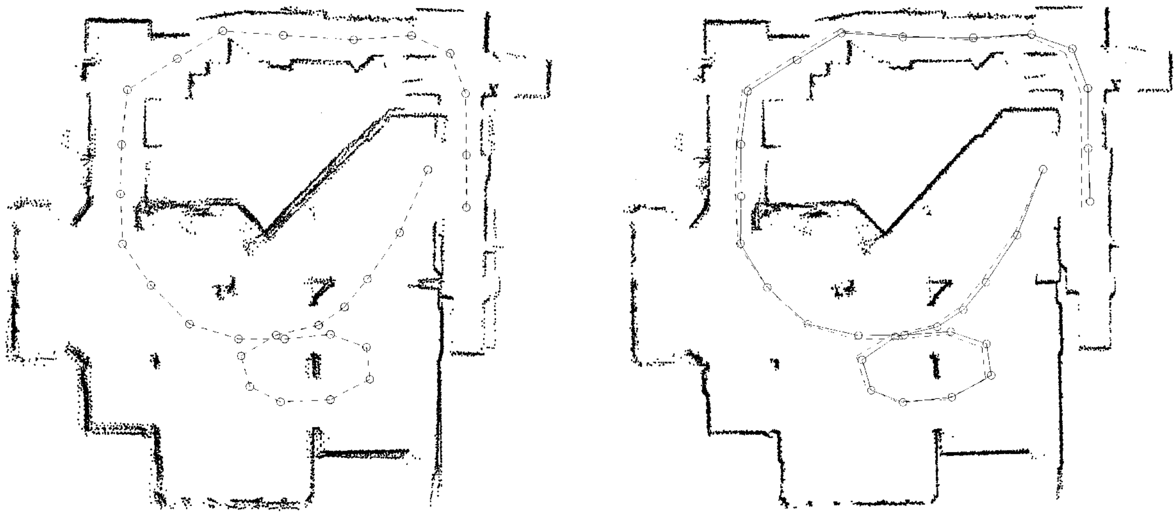


Figure 7. Pose errors along the path, before correction, after local correction, and after global correction: (a) orientational errors; (b) positional errors.

errors and realigning the scan data. Each line segment (either dashed or solid) in the figure represents a strong link obtained from matching two scans. In addition, the solid lines show the robot path which corresponds to the weak links. A plot of orientational and positional errors of the poses along the path, both before and after the correction, is given in Fig. 7. Pose errors are accumulated along the path while the corrected pose errors are bounded. For comparison, we also apply a local registration procedure which matches one scan only to the previous scan. The pose errors along the path after this local correction are also shown in Fig. 7. Although pose errors are also significantly reduced after local corrections, they can still potentially grow without bound. In this example, global registration produces more accurate results than local correction.

Then we present the experiment using real range scans and odometry data. The testing environment is the cafeteria and nearby corridor in FAW, Ulm, Germany. The robot travels through the environment following a given path. A sequence of 30 scans which were taken by the robot with an interval of about 2 meters between scan poses were obtained. The laser sensor is a Ladar 2D IBEO Lasertechnik which is mounted on the AMOS robot. This laser sensor has a maximum viewing angle of 220 degrees. Thus having only the

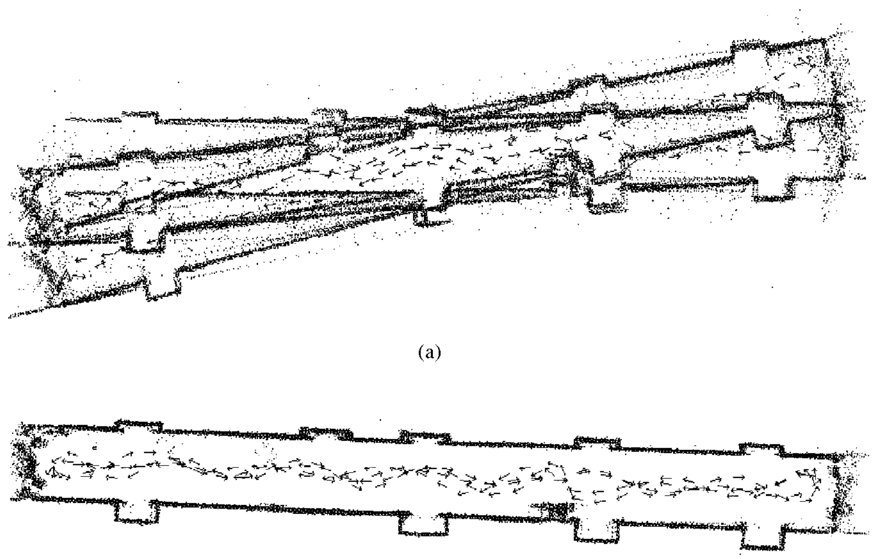


(a) (b)

Figure 8. Consistent global registration of 30 real range scans which are collected by a robot at FAW, Ulm, Germany: (a) unregistered scans whose poses are subject to large odometry errors; (b) registered scans after correcting the pose errors. The robot path estimated from odometry is shown in dashed lines. The corrected path is shown in solid lines.

2D positions of two poses close together does not necessarily ensure a sufficient overlap between the scans taken at the two poses; we also need the sensor heading directions to be similar. Among the 30 scans, 84 links from matching overlapping scan pairs are constructed. Some of these pairwise scan matching results have been shown in (Lu, 1995). In Fig. 8, we show (a) the unregistered scans and (b) the globally registered scans in part (b).

Further experimental results with a variant of our algorithm are reported in (Gutmann and Schlegel, 1996). Figure 9 contains experimental results which



(a)

(b)

Figure 9. Mapping of a Hallway using the RWI Pioneer platform and a SICK laser range scanner: (a) raw laser range scans; (b) aligned laser range scans.

are obtained using our global registration procedure together with a modified version of Cox's pairwise scan matching algorithm[2]. The laser data are collected on the RWI Pioneer platform using the SICK laser ranging device (http://www.sick.de). The Pioneer is a low-cost platform with odometry error significantly higher than the much more expensive platforms used in our other experiments. The hallway environment shown in Fig. 9 is poor in features that allow localization of the robot along the hallway. The data was collected by a robot that went up and down the hallway several times. A large rotation error was introduced by the large turns at the ends of the hallway.

6. Discussion

In this paper, we formulated the problem of consistent range data registration as one of optimal pose estimation from a network of relations. The main ideas are as follows. We associate a robot pose to a range scan to define an unambiguous object frame. By consistently maintaining the history of robot poses, we effectively allow all object frames to be consistently registered in the global reference frame. We use a combination of relation-based and location-based approach to represent the world model. It can be viewed as a two-step procedure. First, spatial relations between object frames are directly derived from odometry measurements and matching pairwise frames. These relations, along with their uncertainties, constitute all the information in the model. In the second step, the relations are converted to object frame locations based on an optimization criterion. This formulation avoids the use of complex constrained optimization. Furthermore, it does not require the assumption of known *a priori* covariance between object frames.

We also derived measurement equations compatible with the formulation. It allows practical implementation of the algorithm. We have experimentally demonstrated the effectiveness of our estimation procedure in maintaining consistency among multiple range scans. The most expensive operation, besides pairwise scan matching, is to compute the inverse of an $3n \times 3n$ matrix. Although the number of poses n may be large for a long robot path, there are ways to limit this size to speed up the computation. The sequential procedure enables the robot to continuously maintain the optimal registration result.

Our approach assumes that the robot stops to collect a complete range scan at its current position. An alternative would be to perform continuous scanning as the robot moves. Continuous scanning would generate large amounts of data that would have to be sampled. In addition, the problem of associating measurements with the correct robot position arises, as different parts of a scan will have been obtained from different robot positions. Solving this problem would require an accurate model of the robots motion. A possible solution to the problem of excessive amounts of data is to partition the continuous scan data and transform each part to one pose on the path, based on the odometry model. These are both worthwhile problems, which we consider outside the scope of this paper.

Although we develop our method for mapping a 2D environment using 2D range scans, our formulation is general and it can be applied to the 3D case as well, by generalizing pose composition and linearization (Lu, 1995).

Acknowledgment

Funding for this work was provided by NSERC Canada and by the ARK project which receives its funding from PRECARN Associates Inc., the Department of Industry, Science and Technology, NRC Canada, the Ontario Technology Fund, Ontario Hydro, and AECL.

The authors would like to thank Steffen Gutmann, Joerg Illmann, Thomas Kaempke, Manfred Knick, Erwin Prassler, and Christian Schlegel from FAW, Ulm for collecting range scans and making the data available for our experiments. We thank Dr. Ingemar Cox, and the anonymous reviewers for many constructive comments.

Notes

1. It is possible to first convert a triangle in the network to an equivalent Y-shaped connection and then the network becomes one with serial and parallel links. However, this Delta-to-Y conversion still can not turn every network into a combination of serial and parallel connections.
2. We are grateful to Steffen Gutmann of the AI Laboratory at the Albert-Ludwigs-Universität Freiburg for providing us with these experimental results.

References

Ayache, N. and Faugeras, O.D. 1989. Maintaining representations of the environment of a mobile robot. *IEEE Transactions on Robotics and Automation*, 5(6):804–819.

Chatila, R. and Laumond, J.P. 1985. Position referencing and consistent world modeling for mobile robots. In *IEEE International Conference on Robotics and Automation*, pp. 138–145.

Cox, I.J. 1991. Blanche: An experiment in guidance and navigation of an autonomous robot vehicle. *IEEE Transactions on Robotics and Automation*, 7(2):193–204.

Crowley, J.L. 1989. World modeling and position estimation for a mobile robot using ultrasonic ranging. In *IEEE International Conference on Robotics and Automation*, pp. 674–680.

Durrant-Whyte, H.F. 1987. Consistent integration and propagation of disparate sensor observations. *International Journal of Robotics Research*, 6(3):3–24.

Durrant-Whyte, H.F. 1988a. *Integration, Coordination and Control of Multisensor Robot Systems*. Kluwer Academic Publishers: Boston, Mass.

Durrant-Whyte, H.F. 1988b. Uncertain geometry in robotics. *IEEE Journal of Robotics and Automation*, 4(1):23–31.

Gonzalez, J., Reina, A., and Ollero, A. 1994. Map building for a mobile robot equipped with a 2D laser rangefinder. In *IEEE International Conference on Robotics and Automation*, pp. 1904–1909.

Gutmann, J.-S. and Schlegel, C. 1996. AMOS: Comparison of scan matching approaches for self-localization in indoor environments. In *Eurobot '96, Kaiserslautern*, Germany, also Available in http://www.informatik.uni-freiburg.de/˜gutmann.

Kriegman, D.J., Triendl, E., and Binford, T.O. 1989. Stereo vision and navigation in buildings for mobile robots. *IEEE Transactions on Robotics and Automation*, 5(6):792–803.

Leonard, J., Durrant-Whyte, H., and Cox, I.J. 1990. Dynamic map building for an autonomous mobile robot. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Lu, F. 1995. Shape registration using optimization for mobile robot navigation. Ph.D thesis, University of Toronto, Department of Computer Science, available as ftp://ftp.cs.yorku.ca/pub/eem/FengLuPh.Dthesis.ps.gz.

Lu, F. and Milios, E. 1997. Robot pose estimation in unknown environments by matching 2D range scans. *Journal of Intelligent and Robotic Systems* (to appear in), available as ftp://ftp.cs.yorku.ca/pub/eem/matching.ps.gz.

Moutarlier, P. and Chatila, R. 1989. Stochastic multisensory data fusion for mobile robot location and environment modelling. In *5th International Symposium on Robotics Research*, pp. 85–94.

SICK Laser range scanner. http://www.sick.de.

Smith, R.C. and Cheeseman, P. 1986. On the representation and estimation of spatial uncertainty. *International Journal of Robotics Research*, 5(4):56–68.

Tang, Y.C. and Lee, C.S.G. 1992. A geometric feature relation graph formulation for consistent sensor fusion. *IEEE Transactions on System, Man, and Cybernetics*, 22(1):115–129.

Feng Lu is a senior scientist with PCI Enterprises, Richmond Hill, Ontario, Canada. He received his B.Sc. and M.Sc. degrees in Computer Science from Fudan University, China, in 1983 and 1986, respectively, and the Ph.D. degree in Computer Science from University of Toronto in 1995. He was a lecturer at the Department of Computer Science, Fudan University from 1986 to 1989. From 1995 to 1996, he conducted postdoctoral research at the Department of Computer Science, York University, where his research areas included sensor-guided autonomous robots and computer vision. Presently, he works with PCI in providing software solutions to the geomatics community. His research interests include remote sensing, digital photogrammetry, pattern classification, image analysis and visualization.

Evangelos Milios received a degree in Electrical Engineering from the National Technical University of Athens, Greece, and Master's and Ph.D. degrees in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology, Cambridge, Massachusetts. While at M.I.T., he was a member of Digital Signal Processing group and he worked on acoustic signal interpretation problems at the M.I.T. Lincoln Laboratory. He held a research faculty position at the Artificial Intelligence Group, University of Toronto (1986–1991), and was visiting professor at the Research Institute for Applied Knowledge Processing (FAW) in Ulm, Germany (1990). Since 1991, he has been Associate Professor of Computer Science at York University. His research interests are spatial reasoning for mobile robots, shape representation, and computational auditory scene analysis.